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

OECD

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

OECD
BETTER POLICIES FOR BETTER LIVES

Please cite this publication as:
OECD (2024), *OECD Handbook on the Compilation of Household Distributional Results on Income, Consumption and Saving in Line with National Accounts Totals*, OECD Publishing, Paris, https://doi.org/10.1787/5a3b9119-en.

ISBN 978-92-64-82235-1 (print)
ISBN 978-92-64-39678-4 (pdf)
ISBN 978-92-64-34311-5 (HTML)
ISBN 978-92-64-78205-1 (epub)

Photo credits: Cover © Inspiring.team/Shutterstock.com.

Corrigenda to OECD publications may be found on line at: www.oecd.org/about/publishing/corrigenda.htm.

Foreword

There is a strong need for high quality information on how different household groups in society are faring. Shocks such as the Global Financial Crisis and the COVID-19 pandemic have stressed the importance of having better information on distributions across social groups, as key input to design policies and to foster inclusive growth. Distributional results on income, consumption and saving aligned to important macroeconomic aggregates such as GDP, household disposable income and household saving are crucial in this regard, providing insight in key dimensions of material well-being across household groups.

In 2011, the OECD and Eurostat launched an Expert Group on Disparities in a National Accounts framework (EG DNA) to develop methodology for the compilation of household distributional results on income, consumption and saving in line with national accounts totals. This Handbook is the result of the work by this expert group and includes information from various questionnaires, presentations and discussions conducted since the launch of the project. It aims to assist compilers in deriving high-quality distributional results and to provide users with more insights into how the results have been derived.

The handbook was drafted by Jorrit Zwijnenburg (OECD), with Asa Johanssen providing valuable advice and editorial support. The Handbook benefited a lot from the inputs and feedback from the various EG DNA members over the years, including Tigran Baghdasaryan (Armenia), Sharon Bailey, Olivier Brunet, Nathan Chia, Derick Cullen, Jacqueline Dumpleton, Ross Harvey, Scott Jensen, Amanda Seneviratne and Michael Smedes (Australia), Anja Breitwieser, Tanja Jurasszovich and Karl Schwarz (Austria), Romain Grailet, Rutger Kemels and Laurent van Belle (Belgium), Ricardo Moraes, Katia Namir and Rebeca Palis (Brazil), Brenda Bugge, Chantal Hicks, Jackie Maisonneuve, Amanda Sinclair and Catherine van Rompaey (Canada), Gerzo Gallardo, Claudia Henríquez and Juan Radrigan (Chile), Vladimir Kermiet, Zuzana Ptackova and Jiri Vopravil (Czechia), Nina Kristiansen, Kathrine Lindeskov Johansen and Jarl Quitzau, (Denmark), Katri Soinne and Veli-Matti Törmälehto (Finland), Aliocha Accardo, Mathias André, Sylvain Billot, Maël Buron, Sébastien Durier, Fabrice Lenglart and Lucile Richet-Mastain (France), Albert Braakmann, Regina Langemann and Florian Schwahn (Germany), Brian Cahill, Justin Flannery and John Sheridan (Ireland), Yafit Alfandari, Hadassa Ben Or, Hila Dizahav, Yoel Finkel, Tali Shalem and Oz Shimony (Israel), Diego Caprara, Stefania Cuicchio, Gabriella Donatiello, Andrea Neri and Marina Sorrentino (Italy), Susumu Kuwahara, Saeko Maeda, Tetsuro Sakamaki, Ryoichi Watanabe and Tomohiro Yamazaki (Japan), Jung Yoon Choi, Woon Sun Eo, Yong Su Jeon, Cheoljae Kim, Young Tai Kim, Tae Ok Lim, Soosung Moon, Young Ji Moon, Jinho Park, and Ji Won Park, (Korea), Francisco Guillen, Fernando Pineda and Gabriela Velázquez Vera (Mexico), Arjan Bruil, Jasper Dingerink, Mélanie Koymans, Safan van der Gaauw and Wim van Nunspeet (Netherlands), Lindsay Beck, Jeff Cope, Fay Peng and Victoria Ward (New Zealand), Mikolaj Haponuik (Poland), Teresa Hilário, Cristina Ramos, Carina Rodrigues and Ana Simao (Portugal), Michal Cepela, Ludmila Ivancikova and Jozef Rosik (Slovakia), Romana Korenic, Mojca Skrlec and Jana Vajda (Slovenia), Alfredo Cristobal, Javier Orche Galindo and Sara Ruiz Gutierrez (Spain), Andreas Lennmalm, Tare Noori, Lana Omanovic and Axel Purwin (Sweden), Stephan Häni, Ueli Schiess and Lukas Schweizer (Switzerland), Angela Barry, David Matthewson, Ellys Monahan, Robert Smith, Sofiya Stoyanova, Sean White and John Wildman (United Kingdom), Mike Batty, Dennis Fixler, Kevin Furlong, Thesia Garner, Marina Gindelsky,

David Johnson, Clinton McCully and John Sabelhaus (United States), Jon Jellema (Commitment to Equity), Juha Honkkila, Ilja Kristian Kavonius, Guillaume Osier, Pierre Sola and Sofie Waltl (European Central Bank), Martha Tovar (International Monetary Fund), Paul Alkemade, Teresa Munzi, Jorg Neugschwender and Josep Espasa Reig (Luxembourg Income Study), Benson Sim (United Nations), Austin Clemens (Washington Centre for Equitable Growth), Jose Pablo Valdes Martinez (World Bank), and Matthew Fisher-Post and Marc Morgan (World Inequality Lab).

Furthermore, several current and former OECD colleagues (including trainees and secondees) provided important contributions to the work over the past years, including Carlotta Balestra, Joong Kwon Bang, Sarah Barahona, Sophie Bournot, Philip Chan, Yangsin Choi, Ole Cordes, Matthew De Queljoe, Pao Engelbrecht, Maryse Fesseau, Federico Giovannelli, David Grahn, Joseph Grilli, Emmanuelle Guidetti, Horacio Levy, Marco Mira d'Ercole, John Mitchell, Elena Tosetto, Peter van de Ven and Florence Wolff.

Special thanks go out to current and former colleagues from Eurostat for the excellent collaboration over the years and for all their contributions to the work and the Handbook, in particular Alessandra Coli, Eleni Giannopoulou, Filippo Gregorini, Sigita Grundiza, Radoslav Istatkov, Maria Liviana Mattonetti, Friderike Oehler, Francesca Tartamella, Ani Todorova, John Verrinder and Hakam Jayyousi (Eurostat).

Finally, thanks go out to Chloe Acas for formatting the final report.

Table of contents

FIGURES

BOXES

TABLES

Executive summary

This publication presents guidance for the compilation of distributional results on household income, consumption and saving consistent with national accounts totals. It is the result of the work of the Expert Group on Disparities in a National Accounts framework (EG DNA) and provides an important addition to the suite of international statistical guidance, addressing increasing policy demands for household distributional results in line with macroeconomic totals.

Economic inequality continues to be a matter of concern for policy makers and citizens, with events such as the 2008-09 Great Financial Crisis and the COVID-19 pandemic further re-emphasising the need for more detailed information on how specific household groups are faring. Evidence-based policies targeting inequality and fostering inclusive growth need to rely on systematic, robust and comprehensive data and indicators. Distributional results on income, consumption and saving aligned to important macroeconomic aggregates such as gross domestic product (GDP), household disposable income and household consumption are crucial in this regard, providing insights into key dimensions of material well-being across household groups.

The principal relevance of distributional estimates in line with national accounts totals come from the way in which they complement existing inequality measures often available from micro statistics. First of all, they provide a more comprehensive picture of economic inequality, including elements that are often not covered in inequality statistics, such as social transfers in kind. Second, the work extends distributional information from income to consumption (and eventually wealth), providing a more comprehensive overview of inequality across multiple dimensions of well-being. Thirdly, the results provide measures of inequality consistent with macroeconomic aggregates, broadening the scope for analyses, while also capturing households and transactions that are typically underrepresented in micro data. Finally, while the estimates do require a number of statistical choices and assumptions, they have a high degree of international comparability because of the common methodology and their alignment to national accounts results.

In 2011, the OECD and Eurostat launched an Expert Group on Disparities in a National Accounts framework (EG DNA) to develop methodology for the compilation of household distributional results on income, consumption and saving in line with national accounts totals. Country experts and representatives from international organisations met several times over the past decade to develop a collection template and to discuss various conceptual and methodological issues, feeding into harmonised guidance for the compilation of the relevant results. Furthermore, expert group members engaged in various compilation rounds to derive experimental estimates in line with the methodology as developed by the group.

This Handbook is the result of the work done by the expert group over the past decade and includes information from various questionnaires, presentations and discussions conducted throughout this period. It provides an overview of the conceptual framework underlying the distributional results and it discusses various aspects in relation to the compilation and presentation of the distributional results. It aims to assist compilers in deriving high-quality distributional results and to provide users with more insights into how the results have been derived.

The starting point of the work is the household sector as defined in the system of national accounts. Within the household sector, the specific focus is on private households, thus excluding any information related to institutional households, such as people living in prison, boarding schools or nursing homes. The main reason is that institutional households often comprise large groups of individuals who are not related and who may have very different income and consumption patterns. This means that their results are not comparable to those of private households, and their inclusion would hamper distributional analyses. The recommendation is to present their results separately.

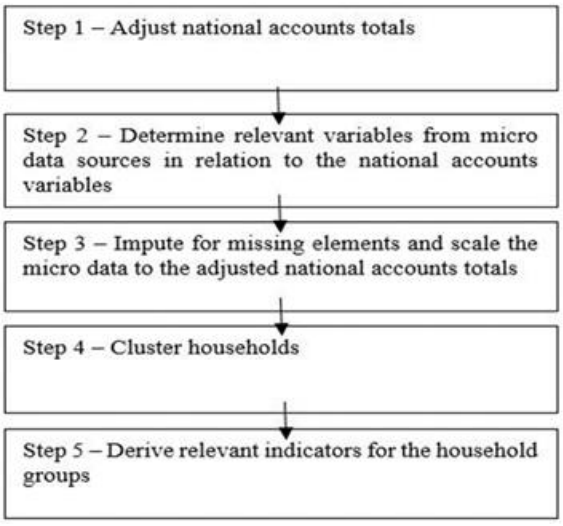
The main unit of analysis in the work is the household (and, hence, not the individual as is the case in some other studies). The main reason is that households are normally pooling some or all of their income and wealth, and consume certain types of goods and services collectively. As households may differ in size and composition, equivalence scales are used to arrive at comparable results across households, recalculating results according to the number of consumption units in each household, taking into account economies of scales of additional household members for certain consumption items.

The aim of the work is to derive distributional results for income, consumption and saving. While results are available for various underlying income and consumption items, the main focus on the income side is on adjusted disposable income. This is the most comprehensive measure of income, also including social transfers in kind, mainly relating to the provision of health, education and housing services either for free or at very low prices. As this in-kind provision is a direct alternative to cash benefits to purchase these goods and services, its inclusion leads to a more comprehensive measure of economic inequality and to more comparable results across countries. On the consumption side, the main focus is on actual final consumption, covering household spending on consumption goods and services, also including the consumption of those goods and services provided via social transfers in kind. Finally, by combining information on income and consumption results can also be derived on household saving across the distribution.

Results can be broken down into various types of household groups. The main focus in the EG DNA work has been on breakdowns according to standard of living, breaking down private households according to their equivalized disposable income. However, breakdowns are also included according to main source of income and household type, and several other breakdowns can be envisaged as well, often dependent on policy needs. The work also targets socio-demographic information to provide more insights into the background of the households included in the various household groups, for example by household type and housing status, or by individual characteristics, such as gender, age, education and employment status.

The methodology for compiling distributional information in line with national accounts totals is set up in the form of a step-by-step procedure (see Figure 1). It starts with the adjustment of national accounts totals to exclude any amounts in the macro aggregates that do not relate to resident private households. The second step involves lining up the relevant components from micro data sources to the income and consumption variables from the national accounts. These micro data provide the main underlying information to distribute income and consumption across households. This step may involve multiple data sources and specific adjustments to arrive at conceptually sound matches. In the third step, imputations are made for elements that are not covered in micro data sources, and the results are aligned to the “adjusted” national accounts totals, allocating any gaps that may exist between the micro and macro aggregates. This for example concerns adjusting for possible missing rich, the underground economy and for any under- or overreporting in the micro data. In the fourth step, households are clustered into household groups, for instance on the basis of their equivalized disposable income or on the basis of socio-demographic characteristics. In the final step, relevant indicators for the distribution of income, consumption and saving are derived, such as disparity ratios that show the degree of income and consumption inequality.

Figure 1. A step-by-step approach for the estimation of distributional information



Source: The Author.

The Handbook provides detailed descriptions of the underlying concepts and the various steps as described above. It includes various examples and guidance on how to deal with specific conceptual and practical issues in the compilation of the relevant results, assisting countries in arriving at high-quality estimates.

At the end of 2020, experimental distributional results in line with national accounts totals were included in the public databases of the OECD and Eurostat for a first time and several countries are already publishing their results on a regular basis. The expectation is that more countries will follow in the coming years, also in view of the new G20 Data Gaps Initiative (DGI-3) that includes two specific recommendations for G20 economies (and participating countries) to regularly publish distributional results in line with national accounts aggregates by the end of 2026. The Handbook will be an important tool to assist countries in setting up their compilation systems and guiding them how to deal with specific issues.

Whereas a lot of progress has been made over the past decade, much still remains to be done. This, among others, concerns improving the quality, timeliness and granularity of the data, as well as broadening the range of countries for which data become available, as also evidenced in the recommendations of the new DGI. For this reason, the work of the EG DNA will continue, further improving distributional results in line with national accounts totals in the area of income, consumption and saving. This work will feed into future updates of this Handbook.

Furthermore, in addition to income, consumption and saving, work has also started in the area of wealth. In this regard, the ECB Expert Group on Distributional Financial Accounts (EG DFA) started to develop methodology to compile distributional wealth estimates for the euro area and EU economies in 2016, and experimental results are planned for publication by early 2024. In addition, several countries have already started to regularly compile and disseminate distributional wealth estimates in line with national accounts totals. In view of the DGI-3 recommendations, the OECD launched a new international expert group early 2023, i.e. the Expert Group on Distribution of Household Wealth (EG DHW), that is going to develop internationally harmonised guidance to compile distributional wealth results in line with national accounts totals, leveraging off the work done by the ECB expert group and by various member states. The overall goal is to arrive at consistent distributional results across income, consumption, saving and wealth, providing users with a comprehensive overview of economic inequality across countries. The expectation is that this will also lead to a handbook in the coming years.

1 Overview

1.1. Aim of the Handbook

This Handbook describes how to compile distributional estimates of household income, consumption and saving consistent with national accounts, in line with the methodology as developed by the OECD (/Eurostat) Expert Group on Disparities in a National Accounts framework (EG DNA). It aims to assist compilers in deriving good quality distributional results that are comprehensive, consistent, and comparable over time and across countries. Furthermore, it aims to provide users of these data with more insight into how these results have been derived, so that they can better assess the quality of the results and understand any differences that may exist with other distributional results.

The Handbook provides an overview of the conceptual framework underlying the distributional results and discusses various aspects in relation to the methodology to compile these data on the basis of underlying micro information. This includes the selection of the relevant micro data sources, imputation techniques for items for which micro data is lacking, guidance on how to deal with gaps between micro and national accounts totals, and on how to combine all this information to arrive at consistent and reliable results for various household groups. In addition to describing the methodology, the Handbook also includes guidance on the presentation of the distributional results, including the description of indicators that provide insight in the levels of inequality in a country.

1.2. Background

The past decade has seen an increased interest in household material well-being and its distribution across households. Whereas there has always been much focus on aggregates such as Gross Domestic Product (GDP) and Net National Income (NNI), the Commission on the Measurement of Economic Performance and Social Progress (Stiglitz, Sen and Fitoussi, 2009[1]), also known as the Stiglitz report, and the G20 Data Gaps Initiative (IMF, 2023[2]) among others stressed the importance of indicators that are more directly related to households' economic well-being (such as household income, consumption and wealth), and on having more insight into how various household groups are faring. It is acknowledged that aggregates and average growth rates only provide a partial story which may conceal large discrepancies between household groups. Whereas average income of the household sector may increase, it may be the case that only a small portion is actually benefiting, while the majority hardly notices any income gain or even suffers from an income loss.

Information on the distribution of income is often available from micro data statistics, but their aggregated trends may often diverge from the national accounts data. Furthermore, due to the setup of the micro data statistics, which may often differ across countries, the results may not always be consistent over time and comparable across countries. For that reason, it was recommended to start looking into possibilities of deriving distributional results for the household sector in line with national accounts totals.

In response to the increased interest in household material well-being and its distribution, the OECD and Eurostat launched a joint Expert Group on Disparities in a National Accounts framework (EG DNA) in 2011

to investigate the possibility to develop methodology to compile distributional results in line with national accounts totals on the basis of micro data. Participants from 25 countries and two international organisations took part in this group and conducted an in-depth study of the main conceptual and practical differences between micro and macro statistics on household income, consumption and saving, discussed best ways to reconcile these differences, defined a conceptual framework, and developed a first methodology to compile internationally comparable distributional results. On the basis of this work, the expert group engaged in a first exercise to compile experimental results in 2012, the results of which were published in a working paper in 2013 (Fesseau and Mattonetti, 2013[3]).

In 2014, the work was continued by an OECD expert group, which focused on further improving the methodology by addressing specific issues, such as developing guidance on how to deal with gaps between micro and national accounts totals and on how to best link data across different data sources. On the basis of the improved methodology, this expert group conducted a second exercise in 2015, the results of which were published in a working paper early 2017 (Zwijnenburg, Bournot and Giovannelli, 2017[4]).

The work of the group continued in 2018 in the form of an OECD-Eurostat expert group, aiming to prepare the publication of the results, while also exploring possibilities to broaden the range of countries for which distributional data in line with national accounts totals would become available. This led to the inclusion of distributional results in line with national accounts totals as experimental results in the online databases of the OECD and Eurostat at the end of 2020. Several countries had also started to publish these results at the national level. A working paper was published in 2021, accompanying the results (Zwijnenburg et al., 2021[5]), highlighting the main results, as well as a final report on the work of the group in this third phase (Coli et al., 2022[6]).

The consecutive expert groups have built up a large amount of knowledge on how to compile distributional results on household income, consumption and saving in line with macroeconomic aggregates. This Handbook aims to combine this knowledge, covering results from the various questionnaires, presentations and discussions conducted as part of the work, to assist other countries to compile distributional results in accordance with the methodology as developed by the EG DNA and to help users in properly understanding the results.

In the meantime, the work of the EG DNA is continuing in the form of an OECD expert group,[1] aiming to improve the timeliness, frequency, and granularity of the results, addressing specific user demands. In this regard, the new G20 Data Gaps Initiative (IMF, 2023[2]) that was launched in 2022 includes two specific recommendations in the area of household distributional results, with the aim of G20 economies publishing annual distributional results at decile level within 18 months after the reference year by the end of 2026. In view of these recommendations, the group will work on further improving the methodology and exploring nowcasting techniques to arrive at more timely estimates, among others. It is expected that this follow-up work will feed into a future update of the Handbook.

1.3. The importance of distributional results in line with national accounts totals

In most countries, distributional results are already available from micro statistics. These provide the possibility to look at very granular levels of detail and to derive inequality results directly on the basis of the underlying data. This raises the question why there is a need for distributional results consistent with national accounts totals. The principal relevance of these data comes from the way in which they complement existing indicators on economic inequality.

First of all, they provide a more comprehensive picture of economic inequality. In that regard, the estimates include elements of income and consumption that are often not covered in micro data, but which may be very relevant in analysing inequality. An example concerns *social transfers in kind*, i.e. goods and services

provided to households by government and non-profit institutions, either free of charge or at prices that are not economically significant. As in-kind provision of these services, which often include health care and education, is a direct alternative to providing households with a cash benefit with which they may purchase these services themselves, its inclusion in distributional measures leads to a more comparable and more comprehensive measure of income. Another important example concerns the non-observed economy, which is usually absent from micro data sources, but which is accounted for in the national accounts.[2]

Secondly, the work broadens the analyses from income to consumption and saving, and eventually wealth, each with its own analytical advantages. Furthermore, the methodology ensures that these dimensions are linked in a consistent way, thus allowing for an integrated overview of economic inequality across income, consumption, saving and wealth. This provides, among others, the opportunity to derive consistent estimates on, for example, saving rates for various household groups and to analyse the joint distribution of income and wealth, e.g. assessing whether some groups may be "income poor" but "asset rich". This is usually not possible on the basis of micro data, as the results on income, consumption and wealth are often based on different underlying concepts and may suffer from specific measurement and estimation errors dependent on the underlying sources, as a consequence of which the results are seldom coherent, often leading to incorrect or even conflicting results.

Furthermore, the estimates aligned to national accounts totals provide measures on inequality consistent with macroeconomic aggregates. By construction, the results are fully consistent with economy-wide totals. This permits linking them to relevant macro-economic indicators, such as gross domestic product, total or average household income, consumption and saving figures, thereby broadening the scope for analyses. It may also assist in analysing how different household groups may be affected by specific macroeconomic trends or by specific policies.

Additionally, distributional results in line with national accounts totals ensure a high degree of international comparability. In this regard, national accounts are compiled according to internationally agreed standards. While the compilation of distributional estimates requires a number of statistical choices, assumptions and reliance on different data sources, a common methodology, developed in close collaboration with member states, helps to minimise the impact of such choices and maximise cross-country comparability of the results.

The compilation of the relevant results also has a positive impact on the quality of statistics. Increasing pressures to reduce the response burden as well as declining response rates make it more difficult to compile high quality micro statistics. Attanasio et al. (2006[7]), Garner et al. (2006[8]) and Pinkovskiy et al. (2014[9]) among others have shown an increasing gap between micro aggregates and national accounts totals over the last decades which may point to increasing measurement and estimations errors in the underlying micro data. Alignment to national accounts totals, which are the result of a process where various data sources are confronted and balanced, provides a vehicle to capture households and transactions that are typically underrepresented in micro data, while also improving comparability of results over time. Conversely, confronting national accounts totals with micro data for distributional information creates positive feedback loops for national accounts leading to improved estimates for macroeconomic aggregates.

The different underlying concepts and the alignment to national accounts totals leads to differences in inequality results. In general, the inclusion of imputed items such as *social transfers in kind* has a mitigating effect on income inequality. On the other hand, the alignment of available micro data to the relevant macro aggregates tends to increase income inequality, as the largest adjustments for the gaps between micro data and national accounts often concern items that are concentrated in higher income groups (such as property income). The overall impact on the distributional results depends on the size of the various adjustments. It is important that compilers are transparent on the main reasons for any differences. Meta data providing insight in the size of gaps between the micro and national accounts data and how they have been dealt with, as well as on the impact of the inclusion of specific items that are missing from the

micro data are very relevant. The Handbook provides guidance on the publication of this type of additional information.

1.4. Relation to other work

The work of the Expert Group on Disparities in National Accounts is closely related to other initiatives in the field of the compilation of distributional results.

First of all, together with the launch of the EG DNA, another expert group was initiated in response to the recommendations of the Stiglitz Report. The OECD Expert Group on Micro Statistics on Income, Consumption and Wealth (EG ICW) developed international guidelines for measuring the distribution of household wealth in micro statistics (OECD, 2013[10]), and a framework for the integrated analysis of micro data on household income, consumption and wealth (OECD, 2013[11]). The framework and these guidelines are at the basis of the OECD collection for the Income Distribution Database (IDD) which provides a comparable set of data on income distribution of households across OECD countries relying on micro sources, mainly household surveys. The work of this expert group has benefited the work of the EG DNA significantly, as it has led to more accurate, more complete, and more internationally comparable micro data sets on income, consumption and on wealth across countries. The work of the EG ICW was followed up in 2017 by a joint Eurostat-OECD expert group that worked on the construction of micro data sets containing consistent micro data across income, consumption and wealth (Balestra and Oehler, 2023[12]). This provides further impetus to the work on compiling distributional results in line with national accounts totals.

Another important initiative concerns the work by the ECB Expert Group on Distributional Financial Accounts (EG DFA). This group is developing distributional estimates of household (financial and non-financial) wealth for the euro area and EU economies. The work started in 2016[3] by comparing and bridging information from the financial accounts and the Household Finance and Consumption Survey (HFCS), both from a conceptual and a practical point of view. In a second stage, the group explored possibilities to further improve the coverage ratios between the micro and macro data, also assessing the possible use of administrative data. Furthermore, it started to define distributional indicators on the basis of user priorities. In response, it was decided to include non-financial assets in the work, because housing wealth and related mortgages turned out to be major components of household wealth. Their inclusion will provide for a more comprehensive overview of household distributional results and may help in cross-checking the distributional information on (the change in) wealth with the information as obtained from non-financial accounts as developed in the EG DNA. The EG DFA is currently developing methodology to compile distributional wealth estimates for the euro area and EU economies, with experimental results expected to be published by early 2024, and for this purpose, there is a close cooperation between the EG DNA and the EG DFA. This is important to harmonise the methodology across income, consumption and wealth, particularly as both work streams are likely to face similar conceptual and methodological challenges in the compilation of distributional results.

At the same time, in view of the new G20 Data Gaps Initiative (IMF, 2023[2]), the OECD launched an Expert Group on Distribution of Household Wealth early 2023, that will develop internationally harmonised templates and methodology to compile distributional results on wealth in line with national accounts totals. This work will benefit from the work of the ECB EG DFA and broaden the work to also include non-EU countries. The group will closely collaborate with both the EG DFA and the EG DNA in order to develop templates and guidelines for compiling the relevant results and to assist countries in their compilation efforts, with the aim of having regular annual distributional wealth results at the decile level within 18 months after the reference period by 2026.

The work is also closely related to the work by the WID.world team which develops so-called Distributional National Accounts (DINA). The main aim of this project is to compile annual estimates of the

distribution of income and wealth using concepts that are consistent with national accounts and using income, inheritance and wealth tax data, household income and wealth surveys, and wealth rankings provided by "rich lists" as their main inputs (see Blanchet et al. (Blanchet et al., 2021[13])). Whereas their aim is similar to that of the DNA work, both work streams differ in scope, concepts and methodology. Zwijnenburg (2017[14]) provides a detailed analysis of the differences between the two streams of work. The main differences relate to the fact that the income concept used by DINA is much broader than the one that is used in the DNA work. In DINA the focus is on national income instead of household income, thus also allocating income from other domestic sectors (e.g. primary income of financial and non-financial corporations as well as of government units) to households. Due to this and some other differences both work streams may lead to different distributional results. Therefore, it is important for both projects to be transparent on their concepts and methodologies used, so that users can understand the main reasons for possible differences in the results.

Finally, it is important to mention the launch of a new Eurostat Task Force on Household Distributional Accounts (TF HDA) in 2022 which will focus on the technical aspects of producing national distributional estimates on income and consumption in line with national accounts totals for EU countries. It will provide a forum to share experiences and knowledge across European countries which will help countries to start estimating their own distributional accounts or, for those countries who already do so, to further improve the quality of the estimates. Furthermore, it will work on further improving the quality of the experimental statistics produced through the Eurostat centralised exercise for those European countries that are not in the position to produce their own results. Because of the close links between the work of the EG DNA and the TF HDA, there will be a close collaboration between the two groups.

1.5. Overview of the Handbook

The Handbook is structured as follows. Chapter 2 presents the conceptual framework for compiling distributional results in line with national accounts totals for income, consumption and saving according to the methodology as developed by the EG DNA. It shows the link to the system of national accounts and provides an overview of the main income and consumption items that are distinguished in the process. Furthermore, it discusses the unit of analysis used in the compilation process and the main breakdowns in household groups targeted in the work. It also presents the template used by countries in compiling their distributional results.

Chapter 3 then provides a brief overview of the methodology to compile distributional results in line with this conceptual framework on the basis of the step-by-step methodology developed by the expert group. This methodology consists of five main steps which are discussed in more detailed in Chapters 4 to 9 and Chapter 13.

Chapter 4 focuses on the adjustment of the national accounts totals to exclude the amounts that do not relate to private households. As the coverage of the national accounts data is usually broader than the scope of the population included in distributional analyses, this requires some specific adjustments to arrive at the appropriate starting point for the compilation of distributional results. Subsequently, Chapter 5 discusses the micro data sources that may be used as input to derive the underlying distributions. This chapter provides a generic overview of micro data sources that may be available and discusses some of their main pros and cons. It also includes an overview of data sources that are currently used by countries in compiling their distributional results in line with the methodology as developed by the expert group.

As not all items will have a counterpart in the micro data, Chapter 6 discusses how to impute distributions for missing items, whereas Chapter 7 focuses on how to deal with gaps between micro and macro data. Both chapters provide generic discussions of these issues, whereas more detailed guidance in relation to specific underlying items is provided in separate chapters.

As multiple data sources may be used as input for the calculation of the distributional results, Chapter 8 discusses methods to link data across data sources to arrive at consistent data at the household level or at the level of household groups. After these steps, households can be clustered into household groups on the basis of the results aligned to national accounts totals. This is explained in Chapter 9.

Whereas Chapters 4 to 9 provide an overview of the basic methodological steps, Chapters 10 and 11 discuss the various income and consumption items in more detail. They provide the underlying definition for the various components, explain main common differences with items included in micro data sources, and discuss specific methodological issues that need to be considered when deriving distributional results.

As the methodology to derive distributional results in line with national accounts data will often require assumptions regarding the correct allocation of gaps between micro and macro data, and to allocate items for which micro data is lacking, Chapter 12. discusses how compilers may check the consistency and plausibility of the results. This includes analysis of the impact of the various steps in the compilation process on the overall results, the assessment of the internal consistency of the results, and analysis of the results over time.

Subsequently, Chapter 13 presents ways to publish the distributional results, also including indicators to obtain more insight in the levels of inequality in a country. The Handbook finishes with an overview of specific areas for further research in Chapter 14.

References

Attanasio, O., E. Battistin and A. Leicester (2006), "From Micro to Macro, from Poor to Rich: Consumption and Income in the UK and the US", http://www.homepages.ucl.ac.uk/~uctpjrt/Files/Attanasio-Battistin-Leicester.pdf (accessed on 26 September 2017). [7]

Balestra, C. and F. Oehler (2023), "Measuring the joint distribution of household income, consumption and wealth at the micro level", *OECD Papers on Well-being and Inequalities*, No. 11, OECD Publishing, Paris, https://doi.org/10.1787/f9d85db6-en. [12]

Blanchet, T. et al. (2021), *Distributional National Accounts guidelines - Methods and concepts used in the World Inequality Database*. [13]

Coli, A. et al. (2022), *Distributional national account estimates for household income and consumption: methodological issues and experimental results - 2022 edition*. [6]

Fesseau, M. and M. Mattonetti (2013), "Distributional measures across household groups in a national accounts framework - Results from an experimental cross-country exercise on household income, consumption and saving", OECD, http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=STD/CSTAT/WPNA(2013)10/RD&docLanguage=En (accessed on 21 September 2017). [3]

Garner, T. et al. (2006), "The CE and the PCE: a comparison", *Monthly Labor Review*, https://www.bls.gov/opub/mlr/2006/09/art3full.pdf (accessed on 30 January 2018). [8]

IMF (2023), *G20 Data Gaps Initiative*, https://www.imf.org/en/News/Seminars/Conferences/g20-data-gaps-initiative (accessed on 14 March 2023). [2]

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. [11]

OECD (2013), *OECD Guidelines for Micro Statistics on Household Wealth*, OECD Publishing, Paris, https://doi.org/10.1787/9789264194878-en. [10]

Pinkovskiy, M. et al. (2014), *Lights, Camera,…Income! Estimating Poverty Using National Accounts, Survey Means, and Lights*, Federal Reserve Bank of New York, https://www.newyorkfed.org/medialibrary/media/research/staff_reports/sr669.pdf (accessed on 26 September 2017). [9]

Stiglitz, J., A. Sen and J. Fitoussi (2009), *Report by the Commission on the Measurement of Economic Performance and Social Progress*, http://ec.europa.eu/eurostat/documents/118025/118123/Fitoussi+Commission+report (accessed on 21 September 2017). [1]

Zwijnenburg, J. (2017), "Unequal distributions? - A study on differences between the compilation of household distributional results according to DINA and EGDNA methodology", http://wordpress.wid.world/wp-content/uploads/2017/11/054-DNA_OECD.pdf (accessed on 5 December 2017). [14]

Zwijnenburg, J., S. Bournot and F. Giovannelli (2017), "OECD Expert Group on Disparities in a National Accounts framework – Results from the 2015 exercise", No. 76, OECD, http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=STD/DOC(2016)10&docLanguage=En (accessed on 21 September 2017). [4]

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, Paris, https://www.oecd-ilibrary.org/docserver/615c9eec-en.pdf?expires=1669202191&id=id&accname=ocid84004878&checksum=689A9023E81B32F43D91EF34FBDE7A7A (accessed on 23 November 2022). [5]

Notes

[1] In 2022, Eurostat also created a dedicated Task Force on Household Distributional Accounts (TF HDA) that will focus on technical aspects of producing national distributional estimates for income and consumption for EU countries.

[2] Whereas distributional results consistent with national accounts are based on a more comprehensive income concept, it has to be borne in mind that different users will have different needs and that some concepts may better suit some users than others. Micro-analysts may for example be more interested in a cash-based concept that may come closer to a household's perception of its income. In that regard, users should be aware of the conceptual differences between the various income measures used in distributional analyses, and compilers should properly explain the underlying concepts used in their compilation process, focusing on specific items that may be included or excluded in comparison with other distributional analyses.

[3] At that time, the group was called the Expert Group on Linking Micro and Macro statistics (EG LMM).

2 Conceptual framework

The aim of the work is to compile distributional estimates of household income, consumption and saving consistent with national accounts totals. This chapter describes the underlying conceptual framework, focusing on the target population, the unit of analysis, the income and consumption concepts, and the household groupings and accompanying socio-demographic information targeted in the work.

2.1. Introduction

This chapter describes the conceptual framework underlying the compilation of distributional results on household income, consumption and saving consistent with national accounts. It starts with describing how the work fits within the System of National Accounts (SNA) in Section 2.2, focusing on the position of the household sector in this framework and presenting the specific accounts covering income, consumption and saving. Then, Section 2.3 discusses the delineation of the target population in relation to the household sector as described in the SNA and defining the units of observation and analysis as used in the compilation of distributional results. Subsequently, Section 2.4 provides an overview of the various income and consumption components included in the compilation of the distributional results, also showing how saving results can be derived on the basis of this information. The chapter then discusses various breakdowns that can be used to present data for more detailed household groups in Section 2.5. This is followed by a description of socio-demographic information that can be published together with the distributional information to provide more insight in the composition of the various household groups in Section 2.6. The chapter concludes with an overview of the main templates for compiling distributional results, based on this conceptual framework, in Section 2.7.

2.2. The position of the work in the System of National Accounts

The aim of the work is to arrive at distributional results on household income, consumption and saving in line with national accounts concepts and totals. This means that the information on income, consumption and saving as described in the system of national accounts for the household sector provides the starting point for the distributional results.

The System of National Accounts (European Commission et al., 2009[1]) (hereinafter referred to as *2008 SNA*) provides a comprehensive and consistent overview of all economic activities and positions in a country. The information is presented as a sequence of accounts focusing on the various aspects of the economy (such as production, generation and re-distribution of income, consumption of goods and services, and the accumulation of assets and liabilities), also including balance sheets on the stocks of assets and liabilities at the start and the end of a recording period. Results are presented for the economy as a whole as well as for the main sectors and subsectors in the economy. These sectors are defined on the basis of similar kind of activities, with non-financial corporations, financial corporations, general government, households and non-profit institutions serving households constituting the main sectors. Furthermore, the system describes all transactions with non-resident entities, presented in the rest-of-the-world account, to arrive at a fully consistent framework.

The system reflects each transaction and position from the perspective of both parties involved and uses a quadruple-accounting system. This means that a counterpart entry is recorded in the financial accounts for each income or capital transaction recorded in the current or capital account and for each purchase/sale of a financial asset or incurrence/repayment of a liability in the financial account. This ensures that the system is always consistent, adding to the reliability and the usefulness of the results.

As explained above, the results are presented as a sequence of accounts, including balance sheets that record stocks of assets and liabilities at the start and the end of a recording period. Figure 2.1 provides a schematic overview of the main accounts distinguished in the system, broken down into the main sectors in the economy, also showing flows and positions vis-à-vis the rest-of-the world.

As presented in the figure, the sequence of accounts can be broken down into current and accumulation accounts. The current accounts provide information on production, income generated by production, the subsequent distribution and redistribution of incomes, and the use of income for consumption and saving purposes. These all relate to the upper part of the figure. The accumulation accounts record flows that affect the balance sheets and consist of the capital and the financial account, which primarily record

transactions and the other changes in assets account. Together these accounts represent the changes in the stock accounts or balance sheets. All of this is presented in the lower part of the figure.

Ideally, distributional results are available for the full set of accounts for the household sector, including their balance sheets, as this would provide a comprehensive overview of the distribution of the various types of resources available to households. For this reason, the 2008 SNA also discusses sub-sectoring of the household sector in paragraphs 4.158 to 4.165. Furthermore, it would provide the opportunity to cross-check the consistency of the data at the level of the various household groups, confronting information on saving as derived from income and consumption with saving as derived from the accumulation accounts.

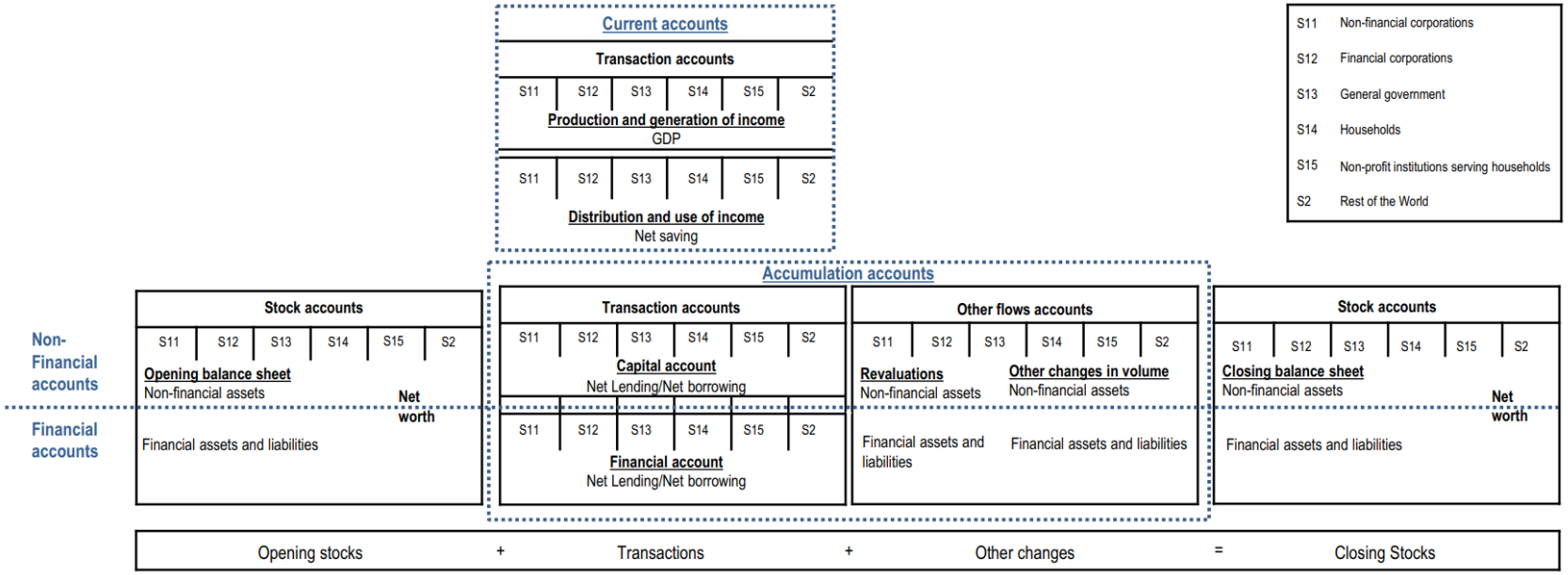
For now, the focus of the work has been on compiling distributional results on income, consumption and saving (i.e. the upper part of Figure 2.1), focusing on the allocation of primary income account, the secondary distribution of income account, the redistribution of income in kind account, and the use of disposable and of adjusted disposable income account. The decision to exclude wealth (i.e. the capital, the financial, the other changes in the volume of assets, and the revaluation accounts as well as the balance sheets) from the work up until now was mainly related to pragmatic considerations. However, the OECD launched a new expert group early 2023, to start developing templates and methodology for the compilation of distributional information on wealth, with the aim of having regular distributional results at decile level by the end of 2026 (see also Section 1.4).

In addition to the sequence of accounts as described above, the system of national accounts includes other accounts and tables that may present information in alternative ways or that may provide more detailed information on specific aspects of the economy. With regard to the compilation of distributional results on consumption, the supply-and-use tables contain very relevant information. Whereas the supply table records the supply of goods and services in an economy on the basis of domestic production and imports, the use table focuses on the use of these goods and services, either as input in the production process (i.e. intermediate consumption) or as final use, broken down into consumption, gross capital formation and exports. This means that the use table contains information on consumption of goods and services at very granular levels of detail. Although this detailed information may not always perfectly match the total consumption of resident households of these goods and services (see Section 2.4.2 for more information), it provides a good starting point in deriving distributional results on household consumption.

2.3. Delineation of the population

The household sector as defined in the SNA is the starting point for the compilation of distributional results. In the SNA, a household is defined as a group of persons who share the same living accommodation, pool some, or all, of their income and wealth, and consume certain types of goods (mainly housing and food) and services collectively (see 2008 SNA, §4.149). In general, each member of a household has some claim upon the collective resources of the household and some influence on the decisions affecting consumption or other economic activities. For these reasons, the household is regarded as institutional unit in the SNA, even though income is usually received by the individual, and the household is used as the unit of observation in compiling distributional results. Correspondingly, the Canberra Group (United Nations Economic Commission for Europe, 2011[2]) and the OECD Framework for statistics on the distribution of income, consumption and wealth (OECD, 2013[3]) also recommend focusing on the household in analysing economic well-being and its distribution.

Figure 2.1. Sequence of accounts in the system of national accounts



Source: Van de Ven and Fano (2017[4]).

Whereas the household constitutes the unit of observation in compiling distributional results, it has to be borne in mind that households may differ in size and composition, and consequently may have different consumption needs. An income level of 3 000 euros per month for a single person household is not comparable with an income level of 3 000 euros for a household consisting of two adults and three children living at home. Therefore, in analysing data at the household level, it is recommended to focus on so-called "equivalised" results to arrive at comparable figures across households. Equivalence scales are used that take into account differences in size and composition of households, recalculating results according to the number of consumption units in each household. A value is assigned to each household type in proportion to its needs, depending on its size and composition. Due to economies of scale (in particular the sharing of dwellings), the per capita requirements of larger households are lower than those of smaller households to achieve the same levels of economic well-being.

For the purpose of this study, the OECD-modified equivalence scale is used as reference method. Accordingly, the first adult[1] counts as 1 consumption unit, any additional persons aged 14 and over count as 0.5 while all children under 14 count as 0.3. However, as the most appropriate scale may depend on specific national circumstances, countries may also look for more appropriate equivalence scales to apply in deriving distributional results. Box 2.1 provides more information on the use of equivalence scales.

Not all households included in the household sector are within the scope of the work. The household sector in the SNA includes both private and institutional households. The latter concern persons living permanently in an institution or who may be expected to reside in an institution for a very long, or indefinite, period of time, with little or no autonomy of action or decision in economic matters. Examples are people living in prison, boarding schools, retirement homes, hospitals, nursing homes and religious institutions (see 2008 SNA, §4.152). These types of households may comprise large groups of individuals with very different socio-demographic backgrounds, who are not related, and who may have very different income and consumption patterns. As a consequence, they behave differently from private households and their (equivalised) results are not really comparable. Whereas it may be assumed that people in a private household have similar levels of economic well-being, this may often not be the case for people living in institutional households. For that reason, it is recommended to exclude them from the compilation of household distributional results, and to present and analyse results for this category separately.[2] Their inclusion would lead to heterogeneous results which may have a distorting effect on distributional analyses.

In addition to institutional households, micro statistics may also exclude some other types of households, such as people living in overseas or sparsely populated areas, persons with no usual place of residence, or persons illegally residing in the country. The main reason is that these people are usually hard to capture in micro statistics. However, as they are part of the target population and are assumed to be comparable with other private households, they are taken into consideration in the DNA work. Omitting them would not provide an accurate picture of inequality in an economy.

In delineating the household sector, it is also important to note how the SNA deals with unincorporated enterprises owned by households. This concerns enterprises that are owned by households and for which it is not possible to separate the assets into those that belong to the household in its capacity as a consumer from those belonging to the household in its capacity as producer (see 2008 SNA, §4.155-4.157). As a consequence, they are consolidated in the household sector in the national accounts. The SNA explains that these unincorporated enterprises can relate to all kinds of productive activities, ranging from agriculture, mining, manufacturing, construction, to retail distribution, and that it may concern single persons working as street traders or shoe cleaners to large manufacturing, construction or service enterprises with many employees.

Box 2.1. The use of equivalence scales in analysing household results

Equivalence scales are often used in distributional analyses to arrive at comparable results across households. These scales take into account that consumption needs of a household will increase with each additional household member, but not in a proportional way due to economies of scale. A value is assigned to each additional household member in proportion to its needs, often depending on their age, but possibly also taking into account other socio-demographic characteristics, such as sex, level of income, labour force status and home ownership. It may also depend on the specific delineation of the income or consumption measure that is analysed. For example, if it includes social transfers in kind, this may require a somewhat different assignment of number of consumption units to the individual household members than when these transfers are excluded. Furthermore, it may depend on the composition of consumption expenditure of various households. Equivalence scales that are appropriate for lower income households may be less appropriate for higher income households due to different consumption patterns. For that reason, equivalence scales may differ across countries, as well as within a country for households with different socio-demographic characteristics (see Radner (1994[5])) for more information).

As it is virtually impossible to derive equivalence scales that take into account all the relevant underlying factors, distributional studies often apply a simplified scale. Although this may have some caveats, it ensures consistency and transparency towards users, and also facilitates the assessment of the impact of the equivalence scale on the results. For the purpose of the DNA work, the OECD-modified equivalence scale has been chosen as reference method. Accordingly, the first adult counts as 1 consumption unit, any additional persons aged 14 and over count as 0.5 while all children under 14 count as 0.3. This means that the number of consumption units is derived according to the following formula:

$$Number\ of\ consumption\ units = 1.0 + (0.5 * additional\ adults) + (0.3 * children)$$

Household income and consumption are divided by the sum of consumption units to obtain a comparable measure across households.

The OECD-modified scale is only one of the possible scales that can be applied. As the most appropriate scale may depend on specific circumstances, countries may look for more appropriate equivalence scales to apply in deriving distributional results. The most important issue is that compilers are transparent about the equivalence scale used and the impact on the results. In that regard, it is recommended to publish information on the number of consumption units and the number of households together with the distributional results.

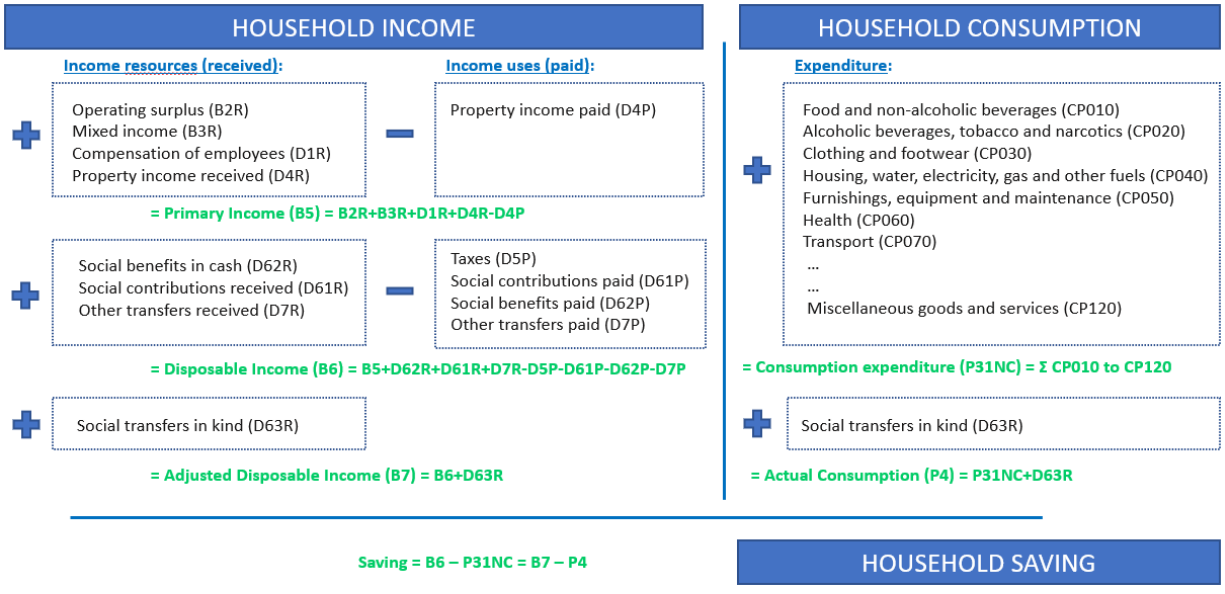
As was explained, most distributional analyses focus on equivalized results. However, the DINA project (see Section 1.3) applies a slightly different approach. In this project, the unit of observation is the individual, and results are compiled on the basis of equal split income series and individualistic series. Whereas the individualistic series focus on the actual income earned by that specific individual, the equal split series distributes the income between the adult household members. In that sense, it can be regarded as a specific application of the equivalence scale. However, where equivalence scales usually take into account economies of scale from belonging to the same household, the DINA project derives the income of the relevant individuals as a simple average of the sum of their incomes, implicitly assigning a value of 1.0 to each additional adult household member. Moreover, it does not take into account any additional consumption needs that may result from children belonging to the household. It is clear that this will lead to different numbers of consumption units for multi-person households, also affecting the distributional results and inequality measures (see Zwijnenburg (2017[6]) for more information).

In practice, it turns out to be rather difficult to come up with a clear set of criteria to describe and delineate unincorporated enterprises. Often the legal status is applied to determine whether an entity qualifies as an unincorporated enterprise. Another criterion that is often applied relates to the number of employees, but in practice other criteria may be applied as well. The difficulty in clearly defining unincorporated enterprises means that the delineation may differ across countries and that the treatment in the national accounts may differ from the treatment in micro statistics. In that regard, the Canberra Group explicitly excludes unincorporated enterprises from the definition of the household sector. As will be discussed in Chapter 7, it is important to take note of any differences in the treatment of unincorporated enterprises in aligning micro data to the national accounts totals, as they may be responsible for large gaps between the micro and the national accounts totals.

2.4. Defining Income, Consumption and Saving

The aim of the work is to derive distributional results for income, consumption and saving according to the national accounts. Figure 2.2 provides an overview of the link between these aggregates in the system of national accounts. The left-hand side of the overview presents household income, broken down into various underlying items respectively adding up to *primary income*, *disposable income* and *adjusted disposable income*. The latter is the most comprehensive measure of income which is the main measure used in the DNA work for analysing income inequality. However, for policy purposes it may also be of interest to look at inequality according to primary income[3] and disposable income. Household consumption is presented on the right-hand side which can be broken down into *final consumption expenditure* and *actual final consumption*. The bottom of the overview shows how household *saving* can be derived on the basis of these results. For that purpose, an additional item is introduced to adjust for the change in pension entitlements related to pension contributions and benefits. This adjustment is needed to reflect that changes in pension entitlements are included in the income measures but are also regarded as part of households' (dis)saving (see Section 11.16 for more information).

Figure 2.2. Income, consumption and saving in national accounts



Source: The Author.

This overview is at the basis of the work and underlies the template for compilers to derive distributional results in line with national accounts totals. The next subsections provide a more detailed overview of the various income and consumption items distinguished in the work. The template as developed by the expert group is presented in Section 2.7.

2.4.1. Income

The work identifies three income concepts in line with the system of national accounts, i.e. *primary income*, *disposable income* and *adjusted disposable income*. *Primary income* is the income that accrues to units "as a consequence of their involvement in processes of production or ownership of assets that may be needed for purposes of production" (see 2008 SNA, §7.2). The main items for the household sector concern *operating surplus*, *mixed income*, *compensation of employees* and *net property income*, which can be further broken down into more detailed underlying items, such as *interest*, *distributed income of corporations* and *rent*.

Disposable income is the income after re-distribution, i.e. after taking into account *current transfers paid* and *received*, such as *current taxes on income and wealth*, *social contributions* and *social benefits*, *non-life insurance premiums* and *claims*, and *other current transfers* like remittances. It is defined in the SNA as the maximum amount that a household can spend on consumption without having to finance it by disposing of assets or increasing its liabilities (see 2008 SNA, §8.25).

Adjusted disposable income is derived on the basis of disposable income, but also including the value of *social transfers in kind* received by households. These consist of goods and services provided to households by government and non-profit institutions either for free or at prices that are not economically significant. They are a direct alternative to receiving a social benefit in cash for the purchase of these services and therefore are included to arrive at a more comprehensive and comparable income measure. Adjusted disposable income is regarded as the most comprehensive income concept and constitutes the main income measure in the DNA work. It is regarded as providing the best insights into inequality in a country and the best measure to use for cross-country comparisons as well as analyses of dynamics of inequality over time.[4]

In the template, the income measures are further broken down into more detailed income components. Table 2.1 provides an overview of the components distinguished in the DNA work in relation to the three income measures. This more detailed breakdown provides compilers the possibility to better match the various income items to information from micro statistics and to impute for missing information at a more granular level. It is recommended to compile results at this more granular level (or even more granular levels if this provides a better match to the corresponding income item(s) in the micro data), as it is expected to lead to better quality results than directly targeting the aggregates. Furthermore, such an approach allows for a publication of distributional results at a more detailed level, which will provide users more insight in the composition of household income for the various household groups. The various items are discussed in more detail in Chapter 10.

The main focus in the template is on gross income measures. Although net measures are preferable from a conceptual point of view, gross measures are still frequently used in macroeconomic analyses given the difficulty in measuring consumption of fixed capital (i.e. depreciation). However, in the interest of giving more prominence to net measures (as will be recommended in the 2025 SNA), the template includes memorandum items for the consumption of fixed capital for the two relevant income items (i.e. operating surplus and mixed income), providing the opportunity to present net values for the main income measures in addition to the gross measures, as well as to compile net saving results (see Section 2.7).

Differences with other income measures

To have a better understanding of the income concept as used in the work, this subsection provides an overview of some other frequently used income concepts and explains the main differences with the concept used in the DNA work.

A frequently used income measure concerns the one defined by the Canberra Group (United Nations Economic Commission for Europe, 2011[2]). This income measure comes close to adjusted disposable income as defined in the 2008 SNA but differs in some important respects reflecting the different focus of the two concepts. The income measure as used by the Canberra Group is based on the conceptual definition as established by the International Conference on Labour Statisticians (ICLS) and states that "household income consists of all receipts whether monetary or in kind (goods and services) that are received by the household or individual members of the household at annual or more frequent intervals but excludes windfall gains and other such irregular and typically one-time receipts".

When comparing the Canberra definition with the income definition in the SNA, the main differences relate to the exclusion of some imputed income items (such as *investment income disbursements* (D44R)), as well as the exclusion of *non-life insurance benefits* (D72R) and *winnings from lotteries* (included in D75R) from the Canberra definition. Furthermore, whereas the conceptual definition used by the Canberra Group includes *social transfers in kind*, these are excluded from the operational definition because of measurement difficulties. Another important difference relates to the use of cash recording in the Canberra measure versus the use of accrual accounting in the SNA. This may lead to differences in time of recording when the actual payment does not coincide with the economic transaction (e.g. deferred tax payments). As the Canberra Handbook underlies a lot of the micro data sources that may be used in the compilation of distributional results in line with national accounts totals, Annex A provides a detailed overview of the differences between the income concepts as defined in the SNA and as used by the Canberra Group.

A Handbook that also underlies a lot of micro statistics is the OECD Framework for Statistics on the Distribution of ICW (OECD, 2013[3]). This framework applies the same income definition as used in the Canberra Group Handbook but distinguishes some different underlying income items. As it has been developed to maintain full consistency with the income concepts as used in the Canberra Group Handbook, the income measure as used by the ICW Framework shows more or less the same differences with the SNA concept. However, as the underlying items may sometimes differ, Annex A provides a detailed overview of the differences between the income concept as defined in the SNA and in the ICW Framework.

Thirdly, a well-known income concept is the one developed by Haig (1921[7]) and Simons (1938[8]), and by Hicks (1946[9]), measuring income as the maximum amount that can be consumed in a given period while keeping real wealth unchanged. This means that in addition to income as defined in the SNA, it also includes holding gains and losses related to the holding of non-financial and financial assets and liabilities. This type of information is included in the accumulation accounts in the framework of national accounts (see Figure 2.1) and could be taken into account once distributional information becomes available for the accumulation accounts. Although it is not part of the income definition as defined by the SNA, it would provide more insight into the economic situation of various household groups.

Finally, the DINA project also compiles distributional results in line with national accounts totals. The income measures[5] used in this project have a broader coverage than household income as defined in the SNA, as their focus is on alignment to national income, i.e. the income of the economy as a whole, so also including income of the other sectors in the domestic economy. The main underlying reason is that in their view all forms of income in the economy eventually accrue to resident individuals (see Blanchet et al. (2021[10])). In that sense, they aim to come closer to the income concepts of Haig, Simons and Hicks as explained above. However, it is questionable whether national income would in this regard lead to the best approximation of this income concept. As explained above, it may be better to focus on deriving

distributional results for the accumulation accounts to obtain more insight on how holding gains and losses are distributed across various household groups, also taking into consideration other holding gains and losses than only those resulting from (undistributed) income.

Table 2.1. Composition of main income measures

2008 SNA Code	Item	Formula
B2R	Operating surplus	
	Owner occupied dwellings	
	Leasing of dwellings	
B3R	Mixed income	
	Own account production	
	Underground production	
	Mixed income, excl. underground and own account production	
D1R	Compensation of employees	
D11R	Wages and salaries	
D121R	Employers' actual social contributions	
D122R	Employers' imputed social contributions	
D4N	Net property income	=D4R-D4P
D4R	Property income received	
D41R	Interest received (adjusted for FISIM[1])	
D42R	Distributed income of corporations	
D43R	Reinvested earnings on foreign direct investment	
D44R	Investment income disbursements	
D45R	Rent received	
D4P	Property income paid	
D41P	Interest paid (adjusted for FISIM[1])	
D45P	Rent paid	
B5	Primary income	=B2+B3+D1+D4N
D5P	Current taxes on income and wealth	
D61P	Net social contributions paid	
D611P	Employers' actual social contributions	
D612P	Employers' imputed social contributions	
D613+D614P	Households' social contributions	
D61xP	Social insurance scheme service charges	
D61R	Net social contributions received	
D611R	Employers' actual social contributions	
D612R	Employers' imputed social contributions	
D62P	Net social benefits other than social transfers in kind paid	
D62R	Net social benefits other than social transfers in kind received	
D7N	Other current transfers	=D72R-D71P+D75N
D72R-D71P	Net non-life insurance claims minus premiums	
D75N	Net miscellaneous current transfers received – paid	
B6	Disposable income	=B5-D5-D61N+D62N+D7N
D63R	Social transfers in kind	
B7	Adjusted disposable income	=B6+D63

Notes: The classification is in line with the 2008 System of National Accounts.
1. FISIM stands for "Financial Services Indirectly Measured" and is explained in more detail in Section 10.5.1.
Source: 2008 SNA (European Commission et al., 2009[1]).

The main differences between *post-tax national income* as used in DINA and *adjusted disposable income* as defined in the SNA concerns the inclusion of *net primary income of corporations*, *government surplus/deficit*, *collective consumption* and any gaps between *social insurance pension contributions* and *benefits*, and the exclusion of *net other current transfers* received in the DINA measure. This may lead to very different income levels for the household sector as a whole and may lead to quite different distributional results (see Zwijnenburg (2017[6]) for more information).

2.4.2. Consumption

With regard to consumption, the work identifies two main measures, i.e. *final consumption expenditure* and *actual final consumption*. *Final consumption expenditure* focuses on the amount that households actually spend on consumption goods and services. For distributional purposes, it is important to bear in mind that this also includes any consumption by a household which is paid for by another household (e.g. college fees paid for by a family member in another household). *Actual final consumption* also includes the consumption of *social transfers in kind*, which are goods and services that are provided to households by government and non-profit institutions and are thus not directly paid for by households themselves.

The consumption items are broken down into detailed categories on the basis of the classification of individual consumption by purpose (COICOP (United Nations Statistics Division, 2011[11])).[6] This is a functional classification which is used to classify the consumption expenditures of households. It distinguishes fourteen main categories. The first 12 items sum up to total final consumption expenditure by households. The last two items identify those parts of consumption expenditure by general government and NPISHs that are treated as social transfers in kind and that are only included in actual final consumption. The DNA work focuses on the breakdown according to COICOP as macro aggregates are usually available at this level of detail from the supply-and-use tables (as explained in Section 2.2), and as most budget surveys usually provide micro information at this level of detail, providing the opportunity to match these data at the detailed level.

Detailed information on the consumption of goods and services is available from the supply-and-use tables. However, it has to be borne in mind that this detailed information may not always concern consumption by resident households only. In some cases, it may also include consumption expenditure by non-resident households which will then have to be excluded at this detailed level. Furthermore, depending on the set-up of the supply-and-use tables, the results at the detailed level of consumption items may only concern consumption expenditure in the domestic economy, whereas any consumption expenditure abroad may only be available as an aggregate item. Depending on the coverage and the level of detail available in the micro data, this could also require a specific adjustment to arrive at total consumption expenditure of resident households only. For that reason, the template identifies specific rows to make any necessary corrections in this light. Section 4.4 explains these corrections in more detail.

Table 2.2 offers an overview of the consumption table as used in the DNA work. It shows the COICOP categories as well as some additional breakdowns for specific items which are included to provide compilers the opportunity to either better match the data with information available from micro data sources or to impute for specific elements in case of missing micro data. Countries are encouraged to compile results at these more detailed breakdowns as it is expected to lead to better results. As explained above, the table also contains two specific rows to include expenditures by resident households abroad and to correct for expenditures by non-resident households on the territory depending on how they are calculated in the supply-and-use tables and dependent on how they are covered in the micro data. The various consumption items are discussed in more detail in Chapter 11.

In addition to the various consumption items, the template also includes memorandum items to separately distinguish expenditures on consumer durables and to record information on taxes and subsidies on production and imports as paid by the various household groups. Separate information on consumer durables is of relevance as they may significantly affect saving results and might explain negative savings

for specific households in specific years. Information on taxes and subsidies on production and imports is of relevance in analysing the impact of government taxes on various household groups. For that reason, compilers are encouraged to include this information if possible.

Table 2.2. Composition of main consumption expenditure measures

Code	Item	Formula
CP010	Food and non-alcoholic beverages	
CP020	Alcoholic beverages, tobacco and narcotics	
CP030	Clothing and footwear	
CP040	Housing, water, electricity, gas and other fuels	
CP041	Actual rentals on housing	
CP042	Imputed rentals on housing	
CP043	Maintenance and repair of dwellings	
CP044	Water supply and miscellaneous	
CP045	Electricity, gas and other fuels	
CP050	Furnishings, household equipment and routine household maintenance	
CP060	Health	
CP061	Medical products, appliances and equipment	
CP062	Out-patient services	
CP063	Hospital services	
CP070	Transport	
CP071	Purchases of vehicles	
CP072	Operation of personal transport equipment	
CP073	Transport services	
CP080	Communication	
CP090	Recreation and culture	
CP100	Education	
CP110	Restaurants and hotels	
CP120	Miscellaneous goods and services	
CP12x	Miscellaneous (less FISIM, less insurance)	
CP1261	FISIM[1]	
CP125	Insurance expenditures (life and non-life)	
P33-P34	Adjustment for expenditures by resident households abroad minus expenditures by non-resident households on the territory	
P33	Resident household expenditure abroad	
P34	Non-resident household expenditure on the territory	
P3	Final consumption expenditure of resident households	= Sum of CP010 to CP120 + P33 - P34
D63	Social Transfers in Kind	
P4	Actual final consumption	= P3 + D63

Notes: The classification is in line with the 2008 System of National Accounts and COICOP 2011.
1. FISIM stands for "Financial Services Indirectly Measured" and is explained in more detail in Section 10.5.1.
Source: 2008 SNA (European Commission et al., 2009[11]).

Differences with other consumption measures

As was the case for income, micro statistics use a slightly different definition of household consumption expenditure from the SNA. The ICW Framework (OECD, 2013[3]) explains that consumption expenditure only includes the acquisition of consumption items, i.e. items that are expected to be used up immediately or in a relatively short period of time. This means that whereas *consumer durables* are included in the consumption measure of the SNA, they are excluded in the OECD Income Distribution Database (IDD) measure because they are regarded as providing services to the household over a longer period of time. The purchases of consumer durables are treated as capital expenditure in the IDD and the resulting services are treated as being consumed by households. This is not the case in the SNA (see 2008 SNA, §9.44), mainly because of the difficulty in deriving accurate measures of the value of the unpaid household services that would be produced on the basis of these durables.

The IDD database also applies a different recording of *non-life insurance premiums* and *claims*, as well as for *expenditure and gains related to gambling*. Whereas these are recorded as current transfers in the SNA, they are treated as consumption expenditure in the IDD. The expenditures on non-life insurance and gambling are recorded as consumption, whereas small windfall gains from non-life insurance and small gambling winnings are treated as negative consumption. On the other hand, large windfall gains are treated as capital transfers in the IDD. Annex A provides a detailed overview of the differences between the consumption concept as defined in the SNA and by the Canberra Group.

2.4.3. Saving

As the work focuses on both income and consumption, results can also be derived for savings across the distribution. Table 2.3 provides an overview of how saving is derived in the DNA approach.

Table 2.3. Derivation of saving

SNA item	Description	Formula
B7	Adjusted disposable income	
D8	Change in net equity of households in pension funds	On basis of the parts of D61P (+), D62R (-) and CP125 (-) that relate to pensions.
P4	Actual final consumption	
B8	Saving	= B7+D8-P4

Note: The classification is in line with the 2008 System of National Accounts.
Source: 2008 SNA (European Commission et al., 2009[1]).

In addition to income and consumption there is an additional item that is needed to derive saving for the various household groups. This concerns an *adjustment for the change in net equity of households in pension funds* (D8). This is needed because of the way contributions paid to and the benefits received from pension funds are treated in the SNA. On the one hand, they are recorded as current income and expenditure, while on the other hand they are also considered as a (dis)saving, adding to (or reducing) the value of pension entitlements. Because the related transactions are not treated as purely financial, but also as income flows, a correction has to be made when deriving saving to reflect that these flows also affect households' saving, increasing their pension entitlements as a consequence of pension contributions and decreasing their entitlements due to the receipt of pension benefits (see also Section 11.16).

2.5. Household group breakdowns as identified in the work

In the DNA approach, results can be broken down into various types of household groups. This section provides examples of breakdowns that have been identified in the DNA work so far, but depending on the available underlying information, other breakdowns can be envisaged as well.

2.5.1. Standard of living (equivalized disposable income)

First of all, data can be broken down according to standard of living, i.e. on the basis of equivalized disposable income of households. For that purpose, household disposable income has to be divided by the number of consumption units depending on the equivalence scale used (see Box 2.1), after which households can be ranked according to this equivalized disposable income. They can then be clustered accordingly, for example into income quintiles (Q1, Q2, Q3, Q4 and Q5) in such a way that each quintile represents 20% of the households.

Please note that whereas the classification is done on the basis of equivalised results, thus taking into account the number of consumption units, the distribution into quintiles is done solely on the basis of the number of households. Hence, when breaking down by income quintile, each quintile represents 20% of the total number of households and not of consumption units. Furthermore, please note that whereas the main focus for measuring income inequality is on adjusted disposable income, as it constitutes the most comprehensive income measure, the classification for standard of living is based on equivalized disposable income, as disposable income is the income that households can freely dispose of.

In the first three collection rounds, the EG DNA has focused on breakdowns by income quintile but depending on user needs and the quality of the underlying data, more granular breakdowns can be envisaged as well, for example into deciles or percentiles. In this regard, the new G20 Data Gaps Initiative (IMF, 2023[12]) includes the recommendation for G20 economies to publish annual distributional results at the decile level by the end of 2026. In view of the recommendations of the new Data Gaps Initiative and increasing user demands for more granular information, particularly for the upper and lower tail, the template already includes the possibility for recording data at the decile level, as well as for the bottom 5% and the top 5%, 1% and 0.1%.

In compiling more detailed breakdowns, compilers have to be aware of the quality of the distributional results and their sensitivity to specific assumptions in the compilation process, for example to bridge any gaps between the micro and macro data. It needs to be carefully assessed what level of detail may be opportune given the quality of the data.

2.5.2. Main source of income

A different classification that can be applied is according to main source of income. For that purpose, four categories are currently distinguished, namely a) *wages and salaries* (i.e. linked to item D11R), b) *income from self-employment* (i.e. linked to item B3R3), c) *net property income* (i.e. linked to item D4N), and d) *current transfers received* (i.e. linked to the sum of D62R (social benefits in cash received), D63R (social benefits in kind received) and D7R (other current transfers received)).[7] Households should be classified in the category which shows the highest contribution to the household income.

2.5.3. Household type

A third classification in the DNA work is according to household type. This takes into account the presence, number and age of the members of the household. In the DNA work eight categories of household types are distinguished, i.e. a) *single less than 65 years old*, b) *single 65 and older*, c) *single with children living at home*, d) *two adults less than 65 without children living at home*, e) *two adults at least one 65 or older*

without children living at home, f) *two adults with less than 3 children living at home*, g) *two adults with at least 3 children living at home*, and h) *others*. In this classification, an adult is defined as anyone 18 years or older.[8] Furthermore, the delineation of "children living at home" is based on all individuals up until the age of 16 plus the individuals whose age is between 17 and 24 and are offspring of one of the household members and are still living at home. Depending on user needs and the quality and available detail from the underlying data, more granular breakdowns can be envisaged as well.

2.5.4. Other possible breakdowns

In addition to the classifications presented here, one could also envisage other breakdowns which may be of interest to users. This may for example include breakdowns according to housing status (rental, owner-occupied with mortgage, and owner-occupied without mortgage), region (for example according to Nomenclature of territorial units for statistics (NUTS) levels in Europe) or on the basis of socio-demographic information, e.g. according to the age, gender or labour market status of the head of the household, where the head of the household is usually defined as the person with the highest income.[9] It will depend on user needs, the level of available detail in the underlying data, and on the quality of the underlying results, which level of detail could be targeted.

When the breakdown is based on characteristics of the household head, it needs to be borne in mind that this is only focusing on the characteristics of one household member and may not provide a balanced overview for the situation of the population at large. In this regard, the people that may qualify as household head may not be reflective of other people with similar socio-demographic characteristics. For example, when looking at income inequality by gender on the basis of distributional results broken down according to gender of the head of the household, this only provides insights into the difference in (equivalised) income between households with a man as household head versus those with a woman as head of the household, but not of the difference in (equivalised) income between men and women in general. Conclusions on the latter can only be derived, if the men and women acting as household head are representative for all men and women in society, which will often not be the case. This caveat needs to be made clear when presenting results on the basis of breakdowns according to household head to users.

2.6. Additional socio-demographic information accompanying the results

To obtain more background information on the various household groups, the DNA work also recommends compiling socio-demographic information to accompany the results for the various household groups. Some of these breakdowns may focus on specific socio-demographic characteristics of households, whereas others may focus on characteristics of the individuals belonging to the households in the various household groups. It will depend on the information available from micro data sources and the way in which the results are aligned to the national accounts totals, which type of socio-demographic information can be published. This section provides a description of the additional socio-demographic information that is targeted in the DNA work, although other breakdowns may be envisaged as well.

2.6.1. Household type

Background information can be provided on the size and composition of the households included in the various household groups. In the DNA work, the household types are distinguished according to the presence, number and age of the members of the household. See Section 2.5.3 for more information on the categories that are identified. Depending on user needs and the quality of the underlying data, more granular breakdowns can be targeted.

2.6.2. Housing status

An overview can also be provided on the basis of the housing status of the households included in the various household groups. This breaks down households into a) households that rent their house (*rental*), b) households that own their house (partly) financing it via a mortgage loan (*owner-occupied with mortgage*), and c) households that own their house without having a mortgage loan (*owner-occupied without mortgage*).

2.6.3. Age

The classification according to age[10] looks at how people of different ages are distributed across the various household groups. Instead of only looking at households and the age of the household head, in this case, each individual is classified according to the category in which the household he/she belongs to is classified. For this purpose, the DNA work defines six age groups, i.e. individuals between a) *0-14*, b) *15-24*, c) *25-34*, d) *35-44*, e) *45-64*, and f) *individuals above 65.*[11] Depending on user needs and the quality of the underlying data, more granular breakdowns can be envisaged.

2.6.4. Sex

A breakdown can also be envisaged according to sex, showing the number of *males* and *females* (and non-binary) in each household group. Depending on user needs and the quality of the underlying data, this could also be combined with some of the other breakdowns presented in this section, for example in combination with age.

2.6.5. Main activity

Useful information can also be obtained by breaking down the various household groups into number of individuals according to their main activity. For that purpose, the DNA work identifies nine categories, i.e. a) *unemployed*, b) *employee*, c) *employer*, d) *own-account worker*, e) *unpaid family worker*, f) *member of producer's cooperative*, g) *student*, h) *retired*, and i) *not classifiable*. These categories are partly derived on the basis of the employment status categories as defined in the ILO International Classification by Status in Employment (ICSE) 1993 (International Labour Organisation, 1993[13]), supplemented by relevant categories for individuals without any form of employment.[12]

To start with the categories related to employment status, in line with the definition of ICSE-93, *employees* should encompass persons with paid employment jobs i.e. people with explicit or implicit employment contracts "which give them a basic remuneration which is not directly dependent upon the revenue of the unit for which they work". *Own-account workers* or self-employed refer to those people for which the remuneration is directly dependent upon the profits from the production process and that have not employed staff on a continuous basis during the reference period. In case they do employ staff, they are classified as *employers*. *Members of producers' cooperatives* are workers who hold self-employment jobs in a cooperative, in which each member "takes part on an equal footing with other members in determining the organization of production, sales and/or other work of the establishment, the investments and the distribution of the proceeds of the establishment among their members". This excludes employees of producers' cooperatives who should be classified as employees. *Unpaid family workers* are workers that hold a "self-employment" job in an enterprise operated by a related person living in the same household, but that are not regarded as a partner (International Labour Organisation, 1993[13]).

In addition to these employment status categories, the DNA work also distinguishes non-employment related categories. The category "*unemployed*" concerns all persons who during the reference period were not in employment or self-employment but were available for employment or self-employment and were actually seeking work (International Labour Organisation, 1982[14]). *Students* are persons that are not

classified as usually economically active, who attend any regular educational institution, public or private, for systematic instruction at any level of education (United Nations, 2008[15]). There are several definitions that are used to delineate *retired* persons. The definition that is used in the DNA work is to distinguish those individuals that are eligible for private or public pension benefits. Finally, *individuals not classifiable* include those for whom insufficient relevant information is available, and/or who cannot be included in any of the preceding categories.

Individuals should be classified according to their main activity, which is defined as the activity which they spend most of their time on. For persons that may change status throughout the reference period, ideally information is available on the length of the activities they engaged in. In that case, individuals can be classified according to the status that they held for the longest period. Alternatively, a decision could be made on the basis of underlying micro data, e.g. tax data may provide information on how many months an individual received wages and for how many months unemployment (or other) benefits; otherwise it may perhaps be possible to derive it on the basis of the relevant amounts; if the amount of wages is still significant (e.g. in relation to wages of previous (or future) years or comparable to other persons holding the same job) and/or if unemployment benefits are still small (also in relation to other unemployed persons or in relation to previous (or future) years), it may be better to classify a person in the "employed" category than in the "unemployed" category.

2.6.6. Highest level of education achieved

Another socio-demographic breakdown that may be of interest to users is a categorisation of individuals according to their highest level of education achieved. The categories that are used in the DNA work are derived on the basis of the International Standard Classification of Education (ISCED) (UNESCO Institute for Statistics, 2012[16]) which distinguishes nine categories, i.e. 0) less than primary education, 1) primary education, 2) lower secondary education, 3) upper secondary education, 4) post-secondary non-tertiary education, 5) short-cycle tertiary education, 6) bachelor's or equivalent level, 7) master's or equivalent level, 8) doctoral or equivalent level, and 9) not elsewhere classified. In the DNA work, some of these levels are combined to arrive at a smaller level of detail, distinguishing: a) *low* (corresponding to levels 0-2 of the ISCED-A, 2011), b) *middle* (3-5), c) *high* (6-8), and d) *not elsewhere classified* (9). As individuals that never attended an education program (including small children) have to be classified in category zero according to the ISCED-A, they would end up in category "low" in the DNA template.

2.7. Presentation of the template

The conceptual framework as presented in this chapter is at the basis of the template that is recommended for the compilation of distributional results according to the methodology as developed by the EG DNA. Table 2.4 to Table 2.6 provide an overview of the general set-up of this template.

The template presents the household groups in the columns and the national accounts item in the rows. In that, it distinguishes two blocks, focusing on respectively income (Table 2.4), and consumption and saving (Table 2.5).

In the income block, the columns start with the information available from the national accounts ("*original estimates*"), targeting data on the household sector (column B1) but also providing the opportunity to first report data for the household and NPISH sector jointly (column A) if results for the household sector are not available separately. In that case, adjustments will have to be made to arrive at results that only relate to the household sector (see Section 4.2). The consumption and saving block has a similar setup, but in this case the first column (A1) relates to totals obtained from the supply-and-use tables. As these results may include amounts relating to NPISHs and non-resident households, the template includes two columns to correct for these amounts (A2 and A3, respectively). Furthermore, it includes a column (A4) to include

consumption of resident households abroad (column A4), in case this information is available at the detailed level of consumption items, to arrive at total consumption expenditure of resident households (column B1).

Both blocks then shows two columns (*"adjusted estimates"*) to distinguish institutional households (column B2) from private households (column B3), the latter constituting the target population of the DNA work. Their results are broken down into household groups on the basis of the breakdowns, as explained in Section 2.5, and the accompanying levels of detail as selected by the compiler. In the example in Table 2.4 a breakdown is presented according to equivalised disposable income levels broken down into income deciles.[13] In this case, households are categorised on the basis of their ranking according to *equivalised disposable income* as explained in Section 2.5.1.

Given the importance of the micro data in the process, the template also includes a column (C1) showing the aggregate from the micro data source used for the distribution of the specific national accounts item ("*micro source aggregate*") as well as the discrepancy between this aggregate and the adjusted national accounts totals relating to private households (column D, i.e. "*discrepancy*").

Finally, the template includes a block that focuses on socio-demographic characteristics of the various household groups (Table 2.6). This block shows breakdowns into the various socio-demographic characteristics as described in Section 2.6 for the specific household groups for which distributional results have been compiled. It starts with information for the household sector as a whole (column B1), with a breakdown into institutional households (column B2) and private households (column B3). It then targets the information by the various household groups.

Compilers are recommended to use this template in compiling distributional results according to the DNA methodology. As changes be made to the templates over time (for example in view of the G20 Data Gaps Initiative recommendations and the update of the 2008 SNA), it is advised that compilers check the latest version of the template which can be found here.

Table 2.4. Template for distributional results on income

Income		National Accounts Households S14				Micro Source	Discrepancy	Deciles D1 + D2 + … + D10 = Private households (B3)			
		Original estimates		Adjusted estimates		Aggregate		D1	D2	…	D10
		Total of S14 + S15 (A)	Total of S14 (B1 = B2 + B3)	Institutional households (B2)	Private households (B3)	(C)	D = B3 - C				
B2Rg+B3Rg	**Gross operating surplus and mixed income**										
B2Rg	Gross operating surplus										
B2R1g	Owner occupied dwellings										
B2R2g	Leasing of dwellings										
B3Rg	Gross mixed income										
B3R1g	Own account production										
B3R2g	Underground production										
B3R3g	Mixed income excluding underground and own account production										
D1R	**Compensation of employees**										
D11R	Wages and salaries										
D121R	Employers' actual social contributions										
D122R	Employers' imputed social contributions										
D4N	**Net property income received / Net property income**										
D4R	Property income received										
D41R	Interest received										
D41R'	Interest received (not adjusted for FISIM)										
D41R_FISIM	Adjustment for FISIM										
D42R	Distributed income of corporations										
D43R	Reinvested earnings on foreign direct investment										
D44R	Investment income disbursements										
D441R	Investment income attributable to insurance policy holders										
D441AR	Property income received attributed to non-life insurance policy holders										
D441BR	Property income received attributed to life insurance policy holders										

Income		National Accounts Households S14				Micro Source	Discrepancy	Deciles D1 + D2 + … + D10 = Private households (B3)			
		Original estimates		Adjusted estimates		Aggregate		D1	D2	…	D10
		Total of S14 + S15 (A)	Total of S14 (B1 = B2 + B3)	Institutional households (B2)	Private households (B3)	(C)	D = B3 - C				
D442R	Investment income payable on pension entitlements										
D443R	Investment income attributable to collective investment funds share holders										
D45R	Rent received										
D4P	Property income paid										
D41P	Interest paid										
D41P'	Interest paid (not adjusted for FISIM)										
D41P_FISIM	Adjustment for FISIM										
D45P	Rent paid										
B5g	**Balance of primary incomes**										
D5P	**Current taxes on income and wealth**										
D61P	**Net social contributions paid**										
D611P	Employers' actual social contributions paid										
D612P	Employers' imputed social contributions paid										
D613P+D614P	Households' social contributions (actual and supplements)										
D613P	Households' actual social contributions										
D614P	Households' social contributions supplements										
D61xP	Social insurance scheme service charges										
D61R	**Net social contributions received**										
D611R	Employers' actual social contributions received										
D612R	Employers' imputed social contributions received										
D62P	**Social benefits other than STiK paid**										
D62R	**Social benefits other than STiK received**										
D6211R+D6221R	Pension benefits										
D6212R+D6222R	Non-pension benefits										
D7N	**Other current transfers (net)**										

Income		National Accounts Households S14				Micro Source	Discrepancy	Deciles D1 + D2 + … + D10 = Private households (B3)			
		Original estimates		Adjusted estimates		Aggregate		D1	D2	…	D10
		Total of S14 + S15 (A)	Total of S14 (B1 = B2 + B3)	Institutional households (B2)	Private households (B3)	(C)	D = B3 - C				
D72R-D71P	Net non-life insurance claims minus premiums										
D71P	Non-life insurance premiums										
D72R	Non-life insurance claims										
D75N	Net miscellaneous current transfers										
D75R	Miscellaneous current transfers received										
D75P	Miscellaneous current transfers paid										
D75x	of which transfers between resident households (2008 SNA 8.133)										
B6g	**Disposable income**										
D63R	**STiK**										
D63R1	Education										
D63R2	Health										
D63R3	Other										
B7g	**Adjusted disposable income**										
Memorandum items											
P51c	**Consumption of fixed capital**										
	concerning: gross operating surplus										
	concerning: gross mixed income										
B2Rn+B3Rn	**Net operating surplus and mixed income**										
B2Rn	of which: net operating surplus										
B3Rn	of which: net mixed income										
B5n	**Net balance of primary incomes**										
B6n	**Net disposable income**										
B7n	**Net adjusted disposable income**										

Source: The Author.

Table 2.5. Template for distributional results on consumption and saving

Consumption & Saving		National Accounts Households S14							Micro Source	Discrepancy	Deciles D1 + D2 + … + D10 = Private households (B3)			
		Original estimates (A1)	NPISH amount (A2)	Expenditure of non-residents on territory (A3)	Adjusted estimates				Aggregate		D1	D2	…	D10
					Expenditure of residents abroad (A4)	Total of S14 (B1 = B2 + B3 = A1-A2-A3+A4)	Institutional households (B2)	Private households (B3)	(C)	D = B3 - C				
CP010	Food and non-alcoholic beverages													
CP020	Alcoholic beverages, tobacco and narcotics													
CP030	Clothing and footwear													
CP040	Housing, water, electricity, gas and other fuels													
CP041	Actual rentals on housing													
CP042	Imputed rentals on housing													
CP043	Maintenance and repair of dwellings													
CP044	Water supply and miscellaneous													
CP045	Electricity, gas and other fuels													
CP050	Furnishings, households equipment and routine maintenance of the house													
CP060	Health													
CP061	Medical products, appliances and equipment													
CP062	Out-patient services													
CP063	Hospital services													
CP070	Transport													
CP071	Purchases of vehicles													
CP072	Operation of personal transport equipment													
CP073	Transports services													
CP080	Communications													
CP090	Recreation and culture													
CP100	Education													
CP110	Restaurants and hotels													
CP120	Miscellaneous goods and services													
CP12x	Miscellaneous (less FISIM, less insurance)													

Consumption & Saving		National Accounts Households S14							Micro Source	Discrepancy	Deciles D1 + D2 + … + D10 = Private households (B3)			
		Original estimates (A1)	NPISH amount (A2)	Expenditure of non-residents on territory (A3)	Adjusted estimates				Aggregate		D1	D2	…	D10
					Expenditure of residents abroad (A4)	Total of S14 (B1 = B2 + B3 = A1-A2-A3+A4)	Institutional households (B2)	Private households (B3)	(C)	D = B3 - C				
CP1261	FISIM													
CP125	Insurances expenditures (life and non-life)													
P31DC	**Final domestic consumption expenditure**													
P33-P34	Adjustment for expenditures by resident households abroad minus expenditures by non-resident households on the territory													
P33	Final consumption expenditure of resident households abroad													
P34	Final consumption expenditure of non-resident households on the territory													
P31NC	**Final national consumption expenditure**													
D63R	STiK													
P4	**Actual final consumption**													
B7	Adjusted disposable income													
D8R	Adjustment for the change in pension entitlements													
B8	**Gross Saving**													
Memorandum items														
D2-D3	Taxes less subsidies on production and imports													
D2	Taxes on production and imports													
D21	Taxes on products													
D29	Other taxes on production													
D3	Subsidies													
D31	Subsidies on products													
D39	Other subsidies on production													
D21-D31	Taxes less subsidies on products													
D29-D39	Other taxes less subsidies on production													

Consumption & Saving		National Accounts Households S14							Micro Source	Discrepancy	Deciles D1 + D2 + … + D10 = Private households (B3)			
		Original estimates (A1)	NPISH amount (A2)	Expenditure of non-residents on territory (A3)	Adjusted estimates				Aggregate		D1	D2	…	D10
					Expenditure of residents abroad (A4)	Total of S14 (B1 = B2 + B3 = A1-A2-A3+A4)	Institutional households (B2)	Private households (B3)	(C)	D = B3 - C				
	Consumer durables													
	of which: vehicles													
	of which: others													
B8n	**Net Saving**													

Source: The Author.

Table 2.6. Template for socio-demographic background information on distributional results

Socio-demographic information	Households S14			Deciles D1 + D2 + D3 + D4 + D5 +D6 + D7 + D8 + D9 + D10 = Private households (B3)									
	Total (B1 = B2 + B3)	Institutional households (B2)	Private households (B3)	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10
Number of consumption units													
Total Number of households													
Household types													
Single less 65 year old													
Single 65 and older													
Single with children living at home													
Two adults less than 65 no child living at home													
Two adults at least one 65 or older no child living at home													
Two adults with less than 3 children living at home													
Two adults with at least 3 children living at home													
Others													
Housing status													
Rental													
Owner-occupied with mortgage													
Owner-occupied without mortgage													
Total resident population (number of persons):													
Age													
0-14													
15-24													
25-34													
35-44													
45-64													
65+													
Sex													
M													

Socio-demographic information	Households S14			Deciles D1 + D2 + D3 + D4 + D5 +D6 + D7 + D8 + D9 + D10 = Private households (B3)									
	Total (B1 = B2 + B3)	Institutional households (B2)	Private households (B3)	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10
F													
Labour market status													
Unemployed													
Employee													
Employer													
Own-account worker													
Unpaid family worker													
Member of producer's cooperative													
Student													
Retired													
Not classifiable by status													
Education													
Low													
Middle													
High													

Source: The Author.

References

Blanchet, T. et al. (2021), *Distributional National Accounts guidelines - Methods and concepts used in the World Inequality Database*. [10]

European Commission et al. (2009), *System of National Accounts 2008*, https://unstats.un.org/unsd/nationalaccount/docs/SNA2008.pdf (accessed on 29 September 2017). [1]

Haig, R. (1921), "The concept of income - Economic and legal aspects", in Haig, R. (ed.), *The federal income tax*, Columbia University Press, https://ia800208.us.archive.org/25/items/cu31924020062935/cu31924020062935.pdf (accessed on 5 December 2017). [7]

Hicks, J. (1946), *Value and Capital: An Inquiry Into Some Fundamental Principles of Economic Theory*, Clarendon Press. [9]

IMF (2023), *G20 Data Gaps Initiative*, https://www.imf.org/en/News/Seminars/Conferences/g20-data-gaps-initiative (accessed on 14 March 2023). [12]

International Labour Organisation (1993), *Resolution concerning the International Classification of Status in Employment (ICSE), adopted by the Fifteenth International Conference of Labour Statisticians (January 1993)*, http://www.ilo.org/wcmsp5/groups/public/---dgreports/---stat/documents/normativeinstrument/wcms_087562.pdf (accessed on 29 September 2017). [13]

International Labour Organisation (1982), *Resolution concerning statistics of the economically active population, employment, unemployment and underemployment, adopted by the Thirteenth International Conference of Labour Statisticians (October 1982)*, http://www.ilo.org/public/english/bureau/stat/download/res/ecacpop.pdf (accessed on 8 December 2017). [14]

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]

Radner, D. (1994), "Noncash income, equivalence scales, and the measurement of economic well-being"*, ORS Working Paper Series*, No. 63, https://www.ssa.gov/policy/docs/workingpapers/wp63.pdf (accessed on 29 September 2017). [5]

Simons, H. (1938), *Personal Income Taxation: the Definition of Income as a Problem of Fiscal Policy*, University of Chicago Press, Chicago. [8]

UNESCO Institute for Statistics (2012), *International Standard Classification of Education (ISCED-2011)*, http://uis.unesco.org/sites/default/files/documents/international-standard-classification-of-education-isced-2011-en.pdf (accessed on 29 September 2017). [16]

United Nations (2018), *Classification of Individual Consumption According to Purpose 2018*. [17]

United Nations (2008), "Principles and Recommendations for Population and Housing Censuses", https://unstats.un.org/unsd/publication/seriesm/seriesm_67rev2e.pdf (accessed on 8 December 2017). [15]

United Nations Economic Commission for Europe (2011), *Canberra Group Handbook on Household Income Statistics*, https://www.unece.org/fileadmin/DAM/stats/groups/cgh/Canbera_Handbook_2011_WEB.pdf (accessed on 27 September 2017). [2]

United Nations Statistics Division (2011), *Classification of Individual Consumption According to Purpose (COICOP) - Global Inventory of Statistical Standards*, https://unstats.un.org/unsd/iiss/Classification-of-Individual-Consumption-According-to-Purpose-COICOP.ashx (accessed on 20 October 2017). [11]

van de Ven, P. and D. Fano (eds.) (2017), *Understanding Financial Accounts*, OECD Publishing, Paris, https://doi.org/10.1787/9789264281288-en. [4]

Zwijnenburg, J. (2017), "Unequal distributions? - A study on differences between the compilation of household distributional results according to DINA and EGDNA methodology", http://wordpress.wid.world/wp-content/uploads/2017/11/054-DNA_OECD.pdf (accessed on 5 December 2017). [6]

Notes

[1] An adult is defined here as anyone 18 years or older. In line with general SNA principles, the age of a person for a given reference year should be derived on the basis of its age during the largest part of the year. This means that anyone born after the 1st of July should be assigned its age at the start of the year, whereas anyone born on or before the 1st of July should be assigned its age at the end of the year. If this is not feasible, it could also be decided to take one cut-off point in the year (e.g. at the start or at the end of the reference period), bearing in mind that this may generate slightly different results.

[2] Alternatively, they could be treated separately in the process; deriving distributional results for them, separate from the private households, and then adding these results back at the end of the exercise. Another alternative is to treat all persons within an institutional household as a separate one-person household in compiling the distributional results. The benefit of these approaches is that it will lead to results that are in line with those for the household sector as published in the national accounts (although it needs to be borne in mind that some other adjustments may be needed in the process, which may lead to deviations from the national accounts totals in any case (see also Chapter 4)). However, compilers and users need to be aware that these approaches may not do justice to the specific circumstances these individuals live in and that it may lead to heterogeneous results and distort distributional analyses.

[3] Please note that in the *standard of living* classification as applied in the DNA approach (see Section 2.5.1), households are ranked and clustered according to their equivalized disposable income. In analysing income inequality according to primary income, it may be better to rank and cluster households according to their equivalized primary income.

[4] Dependent on the specific policy purpose, some users may be more interested in alternative income measures, such as primary income and disposable income. The DNA approach also provides the

possibility to compile results according to these measures. However, for international comparisons, it is recommended to compile and analyse information on the basis of adjusted disposable income.

[5] In DINA, four income concepts are distinguished. Factor income focuses on the income before any re-distribution takes place. Pre-tax income looks at the impact of the pension system, recording pension transactions on a distribution basis instead of on a contribution basis. Post-tax income focuses on income after re-distribution, i.e. after deduction of all taxes, processing all social transfers, and after adding back all public spending, including collective consumption. In addition to post-tax national income, the project also distinguishes post-tax disposable income, which excludes the value of collective consumption and social transfers in kind, as well as any government surplus or deficit.

[6] A new version of COICOP has become available in 2018 (COICOP 2018 (United Nations, 2018[17]), with includes some minor changes in comparison to the 2011 version. Once this new COICOP will have been implemented by countries, it will also be reflected in an updated version of the DNA template.

[7] The latter category could be further broken down into *pension benefits received* and *other current transfers received*, in case the relevant information is available at that level of detail.

[8] In line with general SNA principles, the age of a person for a given reference year should be derived on the basis of his/her age during the largest part of the year. This means that anyone born on or after the 1st of July should be assigned its age at the start of the year, whereas anyone born before the 1st of July should be assigned its age at the end of the year. If this is not feasible, it could be decided to take one cut-off point in the year (e.g. at the start or at the end of the reference period), bearing in mind that this may generate slightly different results.

[9] Countries apply different rules to determine the head of the household, but most of them define it as the person with the highest income (see also (United Nations Economic Commission for Europe, 2011[2]) and (OECD, 2013[3])).

[10] See footnote 8 for more information on how to determine the age of the relevant persons.

[11] For national purposes, it may also be of interest to delineate the last two groups on the basis of the retirement age in the country. However, for international comparability, it is recommended to maintain the breakdowns as suggested here. Furthermore, in using the retirement age, it has to be borne in mind that time series analysis may be affected, when the retirement age is changing over time.

[12] In the meantime, the ISCE-93 has been updated to ISCE-18, including more granular classes based on the type of authority that the worker is able to exercise in relation to the work performed and the type of economic risk to which the worker is exposed. In the future, it may be assessed whether the categories in the DNA template could be updated to reflect (some of) these new categories.

[13] The template also includes the possibility to record results broken down by income quintile, as well as the option to provide more granular information for the tails, i.e. results for the bottom 5% and the top 5%, 1% and 0.1%. Furthermore, the template includes “optional” breakdowns which are not presented in Table 2.4 and Source: The Author.

Table 2.5 classifying households according to main source of income and household type as explained in Sections 2.5.2 and 2.5.3.

3 Overview of the methodology

This chapter provides an overview of the methodology to derive distributional estimates within the national accounts framework on the basis of micro data sources, according to the conceptual framework as laid down in the previous chapter.

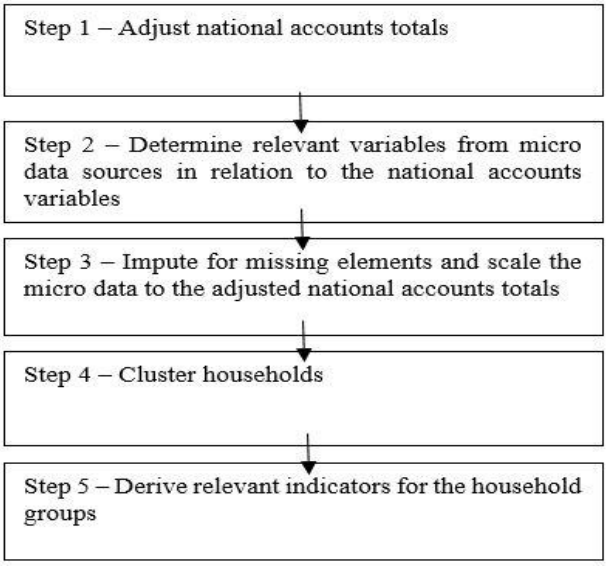
3.1. Introduction

Given the fact that the methodology uses both national accounts data and micro data sources for the compilation of the distributional results, an approach is needed to combine these two types of information in a way that addresses the important differences that may exist between them and that deals with any additional information that may be needed to arrive at robust distributional results. The overall results will depend on the inherent quality of the micro and macro data used, the quality of any necessary alignments and edits made to both sources, and of the imputations that are needed to correct for any missing elements. In consideration of these tasks, the expert group developed a methodology and practical guidance on how to deal with specific methodological challenges. This chapter provides a general description of the methodology and some explanation on the specific steps. The next chapters discuss them in more detail.

3.2. Step-by-step approach

The methodology for compiling distributional information within the national accounts framework is set up as a step-by-step procedure to clearly distinguish between the various steps that are needed to arrive at the distributional results in line with the conceptual framework as explained in the previous chapter. Figure 3.1 provides an overview of this step-by-step approach.

Figure 3.1. A step-by-step approach for the estimation of distributional information



Source: The Author.

As the distributional results only concern a specific part of the household sector in the national accounts, i.e. private households, in the first step the national accounts totals on income and consumption have to be adjusted to exclude information that does not relate to private households. This entails adjustments to exclude for the part of the national accounts data that relates to institutional households, but may also concern other adjustments, depending on the coverage of the available data. For example, in case the household sector is published in combination with non-profit institutions serving households, a correction would be needed to exclude the latter from the national accounts results. Furthermore, it may be the case that the available consumption data include expenditures of non-residents on the territory. This may also

require corrections to arrive at comprehensive results relating to private resident households only. This first step in the methodology is discussed in more detail in Chapter4.

In a second step, micro variables should be selected in accordance with the relevant national accounts variables, to provide the underlying distributional information for the relevant items. In some cases, this may be rather straightforward, but sometimes there may be conceptual and classification differences between the national accounts items and the corresponding micro variables. This may then necessitate combining multiple micro variables and/or the re-classification of certain sub-items in order to arrive at conceptually sound matches with the national accounts items. Furthermore, in selecting the relevant items from micro data, multiple data sources may be available providing information on a similar topic. In that case, it should be assessed which data source(s) and item(s) provide the best link. The selection of micro data variables is discussed in Chapter 5.

As the micro data will usually not perfectly align with the national accounts data, the third step concerns the bridging of any gaps. First of all, this may concern items that are not covered by micro data at all. Certain variables are specific to the System of National Accounts and do not have a counterpart item in the micro data. This for example concerns items such as *financial intermediation services indirectly measured* (FISIM) (see Section 10.5.1) and *investment income disbursements* (see Section 10.5.4). For these items, the distribution will have to be obtained in a different way, for example on the basis of auxiliary information or linking it to the distribution of other items. The imputation for missing items will be discussed in more detail in Chapter 6.

Secondly, for items for which a corresponding micro item is available, the aggregates will not always match perfectly. This may be due to several reasons, such as measurement and estimation errors in the micro data or quality issues with the national accounts totals or with any adjustment that was made to arrive at data for private households. As the goal is to arrive at distributional data in line with national accounts totals, these gaps will need to be bridged in the third step of the step-by-step approach. This issue is discussed in more detail in Chapter 7.

In the fourth step, on the basis of the aligned results, households can be clustered into household groups. This may for example be done based on equivalised disposable income (e.g. in income quintiles or deciles), but also on the basis of alternative classifications, such as main source of income or household composition. This step is explained in Chapter 9.

After this clustering, disparity indicators can be derived in the fifth step to provide more insight in the levels of inequality. Chapter 13 presents several indicators that can be used for this purpose.

The step-by-step procedure requires that the micro data sets that are used in the process are properly linked in order to construct coherent distributional information at the level of the household or at least at the household group. As different micro data sources may be used to obtain details on income, consumption and saving, the compiler must ensure that each source identifies households in a consistent manner. As the method of linking micro sources can significantly impact the results, special consideration is assigned to this task in Chapter 8.

4 Adjustment of National Accounts totals

The first step in the methodology concerns the adjustment of national accounts totals to exclude any amounts that do not relate to private households, i.e. the target population for the results. This chapter describes the specific adjustments that may be needed for this purpose.

4.1. Introduction

Considering that distributional results specifically relate to private households, the first step in the methodology concerns the application of any necessary adjustments to the national accounts totals to exclude those parts that do not relate to private households. This may entail three main types of adjustments. Firstly, if no separate information is available on the household sector but data are instead combined with information on non-profit institutions serving households (NPISHs), the compiler first needs to separate these two. Secondly, depending on the compilation and publication practices of the national accounts in a country, the consumption results at the detailed level may include consumption of non-residents on the territory (which is then only excluded at the aggregate level). This would require a correction at this detailed level to exclude these expenditures. Furthermore, dependent on how the relevant micro data sources cover information on expenditure of residents abroad, compilers may consider adjustments to include these expenditures at the detailed level of consumption expenditure. Finally, as the results only relate to private households, the amounts that relate to institutional households, which are also part of the household sector in the System of National Accounts, should be excluded.

This chapter discusses each of these adjustments in detail with the aim of providing guidance on their practical implementation and discussing potential challenges that may arise during the process.

4.2. Adjustment for NPISHs

The adjustment for NPISHs is necessary when households and NPISHs are compiled and/or published together in the national accounts. This may be the case when the requisite data sources to distinguish between the two are not available and/or when the NPISH sector is thought to be economically insignificant. As the distributional results only concern (private) households, any amounts related to NPISHs should be adjusted for.

Before examining the implementation of this adjustment, it is worth recalling the definition and classification of NPISHs. The System of National Accounts (European Commission et al., 2009[1]) (hereinafter referred to as *2008 SNA*) defines NPISHs as consisting of non-market non-profit institutions that are not controlled by the government. They provide goods and services to households for free or at economically insignificant prices. Most of these goods and services represent individual consumption, often provided to the members of these institutions, but it is also possible for NPISHs to provide collective services, such as medical research, which is made available to the society at large for free.

The NPISH sector covers a broad as well as a diverse range of institutions. First of all, it includes non-philanthropic organisations such as labour unions, political organisations, consumer associations and professional societies. Secondly, it covers philanthropic organisations such as charities, emergency or relief services and aid agencies. Finally, it includes institutions that provide collective services such as research institutions and environmental groups (see 2008 SNA, §4.167-170).

In order to implement the adjustment for NPISHs, the compiler will need to identify data sources that may provide distinct information on either households or NPISHs. This may concern data from satellite accounts, financial statements, tax reports and/or other types of administrative data. Depending on the type of available information, the compiler may directly target results for the household sector or derive results for the NPISH sector that can then be deducted from the totals as recorded in the national accounts. An alternative is to estimate the shares of households and NPISHs in the sum of the two. In this process, it is recommended that compilers derive results at the level of the detailed income and consumption components rather than applying a constant correction factor across all items. In this regard, it is important to realise that the shares can differ significantly across items and, moreover, that some items will by definition only relate to households. This is for example the case for *compensation of employees received* (D11) and *net social contributions paid* (D61P).

When multiple sources of information are used to determine the split for the various income and consumption items, it is also important to assess the consistency and plausibility of the adjustments across items. If these do not provide a consistent picture (for example when the adjustment factors are very different for income and consumption), it may be needed to modify some of the adjustments to bring them more in line.

In the case where no reliable data are available to derive the split between the household and the NPISH sector, an alternative is to derive a breakdown on the basis of available information from other countries. Figure 4.1 provides an overview of the share of the NPISH sector in the sum of the household sector and the NPISH sector for several countries for 2020. If the compiler is able to identify countries (or a country) where the economic role of NPISHs is thought to be more or less similar as in its own country, the ratio between households and NPISHs can be applied to those components for which a direct data source is lacking[1] (still assessing the consistency and plausibility of the adjustment factors across the various items). This approach has the advantage of allowing for expeditious item-specific adjustments. Of course, the accuracy of the adjustments depends on how well the income and consumption profiles of NPISHs in the surrogate country match with those in the target country. Any country-specific circumstances in both the proxy and/or target country should be taken into consideration, possibly deviating from the obtained ratios in specific circumstances.

Furthermore, as the weights of the various underlying items in the aggregates are likely to differ across countries and as it may be the case that countries do not share the exact same set of applicable items, decisions may have to be made at what level to apply the adjustment coefficients. Starting at different levels of aggregation may lead to different results for the aggregates. For example, starting from the underlying items of *operating surplus* (B2) (i.e. *owner-occupied dwellings* (B2R1) and *leasing of dwellings* (B2R2)) may lead to a different result than directly targeting this item at the aggregated level. In this regard, it is advised to focus on that level of aggregation that is regarded to provide the best proxy. In some cases, this may imply making assumptions on some specific (underlying) items, to make sure the results are consistent with the derived results.

4.3. Adjustment for consumption of non-resident households on the territory and where relevant of residents abroad at the detailed level

The second adjustment that may be needed pertains to separating out the consumption expenditures of non-resident households on the economic territory for each detailed consumption component in case these are included in the initial national accounts totals as derived from the supply-and-use tables and to include expenditures by residents abroad at the detailed level in case this leads to a better match with the micro data. While total household final consumption expenditure indeed only refers to resident households, this may not always be the case at the most detailed level. In that regard, the 2008 SNA explains that it may be more practical to calculate the total expenditure made by both resident and non-resident households within the economic territory and to then adjust this figure by including expenditures by residents abroad and excluding expenditures by non-residents within the economic territory at the aggregated level (see 2008 SNA, § 9.80). This approach is based on the rationale that it is easier to balance the supply and use of goods and services by considering the total domestic consumption expenditure, i.e. by residents and non-residents, instead of trying to already exclude the consumption of non-residents at this level of detail. Similarly, the consumption expenditure of residents abroad is usually not available at the detailed level of goods and services, but only added to domestic consumption expenditure at an aggregated level.

Figure 4.1. Share of NPISHs in sum of the household sector and the NPISH sector for 2020

Code	Item	AUT	BEL	CHE	CZE	DEU	DNK	ESP	EST	FIN	FRA	GBR	GRC	HUN	IRL	ITA	LTU	LUX	LVA	MEX	NLD	NOR	POL	PRT	SVK	SVN	SWE	USA
B2R+B3R	Operating surplus and mixed income	1.1%	0.5%	2.9%	0.8%	2.0%	0.7%	0.5%	1.3%	2.5%	1.5%	5.5%	0.2%	1.6%	0.8%	0.2%	0.1%	0.8%	0.8%	1.5%	0.2%	2.8%	0.1%	2.1%	0.3%	0.9%	1.8%	5.0%
B2R	Operating surplus	2.4%	1.1%	100.0%	2.1%	5.7%	1.4%	1.3%	10.8%	4.1%	2.4%	9.6%	0.8%	3.7%	1.7%	0.4%	0.4%	1.1%	1.4%	5.3%	0.8%	3.3%	1.5%	4.9%	0.6%	2.3%	4.1%	11.1%
B3R	Mixed income	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	n.a.	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
D1R	Compensation of employees	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
D11R	Wages and salaries	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
D12R	Employers' social contributions	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
D4R	Property income received	1.0%	0.5%	1.7%	0.5%	1.0%	0.8%	2.1%	0.5%	8.1%	0.8%	1.4%	0.9%	0.9%	5.2%	0.3%	0.1%	1.6%	0.1%	0.0%	0.1%	2.3%	0.7%	1.1%	1.0%	1.1%	3.5%	1.9%
D41R	Interest received	2.7%	3.6%	6.0%	0.9%	6.0%	12.1%	24.0%	5.3%	2.4%	1.4%	2.2%	1.5%	1.4%	5.8%	2.8%	0.1%	3.4%	2.1%	n.a.	0.5%	4.4%	1.7%	6.2%	2.2%	1.2%	24.4%	0.0%
D41R'	Interest received (not adjusted for FISIM)	3.2%	5.1%	3.7%	1.8%	3.8%	14.3%	10.1%	6.3%	2.3%	4.9%	2.8%	1.9%	3.0%	4.7%	3.7%	1.5%	3.9%	2.2%	0.0%	-3.7%	4.9%	3.1%	6.4%	3.7%	1.6%	10.8%	1.2%
D42R	Distributed income of corporations received	0.8%	0.0%	2.2%	0.1%	0.2%	0.9%	0.7%	0.0%	7.5%	0.0%	1.8%	0.2%	0.0%	7.9%	0.0%	0.0%	0.3%	0.0%	0.0%	0.2%	2.1%	0.0%	0.0%	0.0%	1.1%	4.7%	2.8%
D43R	Reinvested earnings on foreign direct investment	0.0%	n.a.	n.a.	0.0%	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	0.0%	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	100.0%	n.a.	100.0%	n.a.	n.a.
D44R	Investment income disbursements	0.8%	0.7%	0.8%	0.3%	2.4%	0.3%	0.0%	0.1%	11.1%	0.2%	0.2%	0.0%	0.3%	0.0%	0.0%	0.1%	0.8%	0.0%	0.0%	0.0%	1.6%	0.4%	0.1%	0.0%	0.2%	1.9%	9.4%
D45R	Rent received	n.a.	0.0%	n.a.	0.0%	5.8%	0.0%	0.0%	0.2%	17.7%	0.0%	83.5%	0.0%	0.0%	0.0%	0.0%	0.0%	33.3%	0.0%	0.0%	0.0%	3.4%	n.a.	6.3%	14.5%	0.0%	0.0%	n.a.
D4P	Property income paid (adjusted for FISIM)	1.2%	0.6%	1.6%	0.7%	0.6%	1.0%	0.9%	-6.3%	4.4%	0.7%	2.6%	1.7%	0.2%	0.6%	4.5%	0.1%	0.4%	0.9%	0.0%	-0.4%	-0.4%	1.1%	3.4%	5.4%	0.5%	0.9%	0.6%
D41P	Interest paid	1.2%	0.4%	1.7%	0.2%	1.1%	0.9%	0.8%	4.1%	1.8%	1.4%	1.3%	3.2%	0.1%	0.9%	2.3%	0.2%	0.4%	0.9%	n.a.	0.4%	0.4%	0.8%	4.3%	2.5%	0.3%	0.6%	1.0%
D41P'	Interest paid (not adjusted for FISIM)	1.2%	0.7%	1.6%	0.2%	0.7%	1.2%	1.2%	-6.3%	9.3%	0.8%	2.6%	3.0%	0.2%	0.8%	8.8%	0.2%	0.4%	0.9%	0.0%	-0.4%	-0.5%	1.1%	0.1%	6.0%	0.6%	0.9%	0.6%
D45P	Rent paid	n.a.	0.0%	n.a.	2.7%	0.0%	0.0%	0.0%	n.a.	1.5%	0.0%	0.0%	0.0%	n.a.	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.8%	0.0%	4.9%	1.7%	0.0%	0.2%	n.a.
B5	Balance of primary incomes	0.3%	0.1%	0.6%	0.2%	0.4%	0.1%	0.2%	0.1%	1.0%	0.3%	1.3%	0.1%	0.5%	0.4%	0.1%	0.0%	0.2%	0.1%	0.6%	0.0%	0.5%	0.1%	0.6%	0.1%	0.2%	0.5%	1.4%
D5P	Current taxes on income and wealth	0.1%	0.1%	0.0%	0.1%	0.0%	0.1%	0.0%	0.0%	0.3%	0.0%	0.2%	0.2%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.2%	0.1%	0.1%	0.0%
D61P	Net social contributions paid	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
D611P	Actual social contributions	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	n.a.	0.0%	0.0%	0.0%	0.0%	n.a.	0.0%	0.0%	0.0%	0.0%	0.0%
D612P	Imputed social contributions	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	n.a.	0.0%	0.0%	0.0%	0.0%	n.a.	0.0%	0.0%	0.0%	0.0%	0.0%
D61xP	Social insurance scheme service charges	n.a.	82.5%	45.8%	17.8%	100.0%	n.a.	18.8%	99.3%	100.0%	98.1%	81.2%	n.a.	100.0%	n.a.	11.0%	61.6%	0.0%	75.9%	100.0%	7.4%	76.1%	27.3%	92.1%	7.4%	24.5%	100.0%	n.a.
D61R	Net social contributions received	n.a.	82.5%	45.8%	17.8%	100.0%	n.a.	18.3%	83.3%	n.a.	0.0%	81.2%	n.a.	100.0%	n.a.	8.3%	61.6%	100.0%	75.9%	100.0%	7.4%	76.1%	27.3%	95.9%	7.4%	24.5%	n.a.	n.a.
D611R	Actual social contributions	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	7.6%	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	100.0%	n.a.	n.a.	n.a.	n.a.
D612R	Imputed social contributions	n.a.	82.5%	n.a.	17.8%	100.0%	n.a.	18.3%	83.3%	n.a.	0.0%	81.2%	n.a.	100.0%	n.a.	12.2%	61.6%	n.a.	75.9%	100.0%	7.4%	76.1%	27.3%	90.2%	7.4%	24.5%	n.a.	n.a.
D62P	Social benefits other than STiK paid	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
D62R	Social benefits other than STiK received	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
D7N	Other current transfers received	2.5%	6.5%	0.1%	1.0%	1.5%	28.8%	3.1%	1.1%	12.1%	1.9%	10.4%	0.1%	9.5%	6.3%	6.1%	8.6%	18.9%	0.0%	1.8%	15.6%	4.1%	7.4%	1.8%	4.6%	4.7%	27.3%	21.4%
D71P	Non-life insurance premiums	0.0%	0.1%	0.3%	0.8%	0.2%	0.6%	0.9%	0.0%	0.2%	0.2%	0.9%	0.0%	0.5%	0.6%	0.4%	0.7%	0.4%	0.0%	1.7%	0.4%	0.0%	1.1%	3.2%	0.1%	1.6%	0.0%	n.a.
D72R	Non-life insurance benefits	0.0%	0.1%	0.3%	1.0%	0.2%	0.9%	0.8%	0.1%	0.2%	0.2%	0.9%	0.0%	0.8%	0.5%	0.2%	0.4%	0.6%	0.0%	0.0%	0.4%	0.0%	1.0%	1.9%	0.0%	1.6%	0.0%	n.a.
D75R	Miscellaneous current transfers received	88.7%	60.5%	87.5%	13.7%	81.0%	69.7%	24.4%	48.9%	90.2%	57.6%	76.9%	34.1%	72.1%	83.7%	33.6%	11.7%	82.1%	39.9%	9.5%	54.4%	80.0%	29.6%	34.2%	11.3%	60.4%	65.3%	71.6%
D75P	Miscellaneous current transfers paid	4.1%	11.5%	0.0%	1.0%	3.7%	41.1%	3.5%	1.5%	19.3%	3.6%	16.2%	0.1%	12.8%	11.7%	8.1%	35.2%	33.8%	0.1%	1.8%	25.2%	5.4%	11.4%	1.5%	4.9%	7.0%	34.4%	21.4%
B6	Disposable income	4.4%	2.0%	3.9%	1.8%	3.0%	3.4%	2.2%	3.3%	4.2%	3.5%	3.8%	1.1%	4.8%	1.3%	0.7%	0.3%	5.4%	2.3%	2.0%	1.3%	5.6%	0.7%	3.2%	1.5%	2.1%	3.0%	2.7%
D63R	Social Transfers in Kind received	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
B7	Adjusted disposable income	0.5%	0.2%	0.4%	0.1%	0.4%	0.3%	0.4%	0.1%	0.4%	0.2%	0.4%	-1.1%	1.0%	-0.1%	0.0%	0.0%	0.0%	0.4%	0.5%	0.0%	1.0%	-0.6%	0.4%	0.0%	0.6%	0.5%	-0.2%
D8R	Adjustment for change in pension entitlements	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	n.a.	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	n.a.
P3	Final consumption expenditure	4.7%	2.2%	4.4%	2.0%	3.1%	3.1%	2.0%	3.5%	4.3%	4.0%	3.8%	2.4%	4.3%	1.9%	0.8%	0.4%	7.1%	2.1%	1.8%	1.7%	5.2%	1.6%	3.1%	1.7%	1.8%	2.7%	3.7%
B8	Gross saving	3.1%	1.0%	1.5%	0.9%	2.2%	3.5%	2.8%	1.2%	3.7%	1.5%	3.0%	-53.1%	7.6%	-0.4%	0.0%	-0.2%	0.0%	3.0%	2.0%	-0.2%	6.2%	-6.3%	3.7%	0.3%	3.2%	3.0%	-1.0%

Source: OECD.stat, Table 14A Non-financial accounts by sector.

Figure 4.2 depicts an excerpt of Table 14.12 in the 2008 SNA which shows how this works in practice. Supply and use are balanced in the use table at the detailed level with both resident and non-resident households included together in column 30. Afterwards, a correction is applied at the aggregated level to exclude *domestic purchases by non-residents* (row 13) and to include *direct purchases abroad by residents* (row 12). For this latter item, there is a direct counterpart in the supply table as an addition to imports. The domestic purchases by non-residents are added at the aggregated level to the exports of goods and services in the use table (row 13 in columns 27 and 28).

Figure 4.2. Aggregate adjustments for household final consumption expenditure in supply-and-use tables

Use of products	Total supply at purchasers' prices	Exports		Final consumption expenditure				
		Goods	Services	Sub-total final consumption expenditure	Households	NPISHs	...	...
	(1)	(27)	(28)	(29)	(30)	(31)	...	...
Products								
Total uses								
(1) Agriculture								
... ...								
... ...								
(12) Direct purchases abroad by residents								
(13) Domestic pruchases by non-residents								
(14) Total								

Source: 2008 SNA (European Commission et al., 2009[1]).

If countries have indeed followed this approach, it is necessary to determine the expenditures of non-resident households for each detailed consumption item included in the results in order to arrive at the appropriate starting point for the work, i.e. to exclude the expenditure by non-residents at this detailed level. This is crucial, as the information will be aligned to micro data that only relates to resident private households.

As a priori, the adjustment of household final consumption is only known at the aggregated level with the individual consumption components containing information for both residents and non-residents on the economic territory, the compiler will need to determine the consumption expenditure by residents for each domestic consumption component, subject to the constraint that the resident consumption expenditure items add up to the total of the final domestic resident consumption expenditure (i.e. P31NC) which is known from the supply-and-use tables.

In order to implement the necessary adjustments, the compiler can use data from tourism satellite accounts or balance of payments, among others, to determine the shares of non-residents per detailed consumption component. For example, in using tourism satellite accounts, the compiler may be able to make adjustments at the detailed consumption component level by using outbound and inbound tourism consumption.

In case no detailed information is available on the expenditure of non-residents in the domestic economy, one possible solution is to use the breakdown of expenditure of residents abroad as a proxy if this information is available. Otherwise, estimates may be made for the specific categories that these spending will most likely relate to (mainly *recreation and culture* (CP090), *restaurants and hotels* (CP110) and *transport* (CP070)). Furthermore, it may be relevant to assess whether the country may be dealing with a relatively large number of cross-border shoppers (e.g. in relation to relatively low excise duties or a relatively large number of cross-border workers) which may affect specific other consumption categories. As with the previous adjustments to exclude data related to NPISHs, it may be the case that after estimating

shares for the individual components, the compiler will need to balance the new estimates in order that they add up to the aggregate.

One could also implement the adjustment for expenditures of residents abroad at the detailed level of consumption items. In case these expenditures are covered in the micro data sources and good quality information is available to break down the aggregated amount to underlying items at the macro level (e.g. on the basis of tourism satellite accounts), this may lead to a better match and better results. For international consistency, some adjustments then need to be applied at the end of the compilation process, to move the amounts that relate to expenditure abroad at the detailed level to the aggregated correction item. In this way, results can be provided in line with the internationally agreed template. It can also be decided to already make this specific correction at the micro level, removing the expenditure of households abroad from all the underlying items at the micro level and aggregating these amounts to match it with the aggregated adjustment item in the national accounts, but this will mainly depend on what is expected to provide the most reliable results.

4.4. Adjustment for institutional households

Finally, the national accounts totals, which after the first two adjustments only relate to resident households, must also be adjusted to exclude amounts that do not relate to private households. This means that amounts related to institutional households should be removed from the totals to arrive at the correct benchmark for the distributional results. This step also further aligns the micro and macro data, as the sampling frame of micro sources also typically focuses on private households.

Persons living permanently in an institution or who may be expected to reside in an institution for a long or indefinite period of time are treated as belonging to institutional households as they have little or no autonomy of action or decision in economic matters (2008 SNA, §4.149-4.154). Examples of institutional households are people living in prison, boarding schools, retirement homes, hospitals and nursing homes, religious institutions, hotels and so on for a long period of time. They are excluded from the scope of the distributional results as they behave differently from private households and their (equivalised) results would not be comparable (see also Section 2.2). For that reason, it is recommended to present results for these institutional households separately.[2]

The implementation of the adjustment to exclude amounts relating to institutional households requires sources of information that provide details on the structure and characteristics of these households. Certain countries may compile satellite accounts on households that can be used to distinguish between these two types of households. Another practical method that can be employed concerns the use of census data. This may contain information on the number of persons living in institutional households versus persons living in private households. Furthermore, it may provide background information on people living in institutional households, for example related to socio-demographic characteristics or the type of institutional household they belong to. On the basis of that type of information, the income and consumption of individuals living in institutional households may be approximated by the income and consumption patterns of specific private households also taking into account that specific items may not be relevant for individuals living in certain types of institutional households. To arrive at results at the individual level, the assumption can be employed that each person living in an institutional household is acting as a single person household, so that his income and consumption levels can be approximated by “per household” values in relation to corresponding persons living in private households. The accuracy of these estimates will depend in part on the population breakdowns available in auxiliary data sources as well as the assumptions applied by the compiler.

As with the previous adjustment, it is recommended to apply an item-specific approach in deriving the shares of institutional households for the various income and consumption components. Institutional households may have a significant impact on specific income and consumption items such as social

protection spending and social transfers in kind related to health, whereas their impact may be negligible for other items, for example imputed rents under the assumption that most people living in institutional households will usually not own a house.

4.5. Overview of adjustments

This chapter has explained the necessary steps to go from the available national accounts totals to the amounts related to the population in scope of distributional results, i.e. resident private households.

Table 4.1 provides an overview of these steps for the income part of the work. Depending on the coverage of the available national accounts totals, a first correction needs to be made to exclude that part of the data that relates to the NPISH sector. In that case, the totals relating to the household and NPISH sector are presented in the first column (1), after which data related to the NPISH sector (S15) can be reported for each individual income item in column 2 to arrive at data that only relate to resident households (S14), reflected in column 3. Subsequently, the part that relates to institutional households should be subtracted (in column 4) to arrive at the results relating to private households (column 5).

Table 4.1. Overview of adjustments to national accounts totals to arrive at totals for the relevant income components relating to private households

Income item	S14+S15 (1)	S15 (2)	S14 (3) = (1) – (2)	Institutional households (4)	Private households (5) = (3) – (4)
Operating surplus					
Owner-occupied dwellings					
Leasing of dwellings					
Mixed income					
…					
…					

Source: The Author.

Table 4.2 provides a similar overview for the consumption part of the work. It starts with the data available from the supply-and-use tables in the first column (1). As these may include consumption by NPISHs and by non-resident households, these should be excluded (in columns 2 and 3 respectively) to arrive at data for resident households only. A correction is then needed to include any consumption expenditure by resident households abroad (in column 4), which may be done at the detailed level or at the aggregated level at the bottom of the table. Via this route, results are obtained relating to total final consumption expenditure by resident households in column 5. Subsequently, the part that relates to institutional households (column 6) should be subtracted to arrive at the results relating to private households only (column 7).

From the exercises conducted by the EG DNA, information can be obtained how the adjustment of national accounts totals to exclude those amounts that do not relate to resident private households affects the results. Of course, it has to be borne in mind that this may concern different adjustments across countries depending on the available national accounts data in countries, but it gives a sense of the magnitude of this correction across countries.

Table 4.2. Overview of adjustments to national accounts totals to arrive at totals for the relevant consumption components relating to private households

Consumption item	Sub-total final consumption expenditure from SUT (1)	Final consumption expenditure by NPISHs (2)	Final consumption expenditure by non-residents (3)	Final consumption expenditure of residents abroad (4)	Final consumption expenditure by resident households (5) = (1) – (2) – (3) + (4)	Institutional households (6)	Private households (7) = (5) – (6)
Food and non-alcoholic beverages							
Alcoholic beverages, tobacco and narcotics							
…							
…							
…							
Final domestic consumption expenditure							
Adjustment for expenditures by resident households abroad							
Final national consumption expenditure							
Social transfers in kind							
Actual final consumption							

Source: The Author.

Table 4.3 presents the overall adjustment as percentage of the initial national accounts totals for those countries providing this information in the collection round conducted in 2020. The adjusted national accounts totals are exactly equal to the original national accounts totals for the main aggregates in Mexico, the Netherlands and New Zealand which implies that their distributional results included information on institutional households. No correction was made as detailed information was missing and/or the impact of institutional households was only small. On average, the percentage difference between the adjusted and the original national accounts estimates was less than 0.4% for income and 1.1% for consumption. When looking at results over time, the adjustment coefficients showed to be relatively stable over time.

Although the overall adjustment ratios seem rather low, it has to be borne in mind that these adjustments are still very relevant when breaking down the results into underlying household groups. Without these adjustments, these amounts would have been inaccurately allocated to specific household groups, therewith affecting the overall distributional results. This stresses the importance of this specific step in the compilation process.

Table 4.3. Overview of adjustments to national accounts totals

	% difference between adjusted and original national accounts totals	
	Income (average for B5, B6 and B7)	Consumption (actual final consumption)
Canada (2015)	0.00	0.00
Czech Republic (2017)	-1.28	-1.80
France (2016)	-1.56	-2.73
Ireland (2015)	-0.39	-0.77
Israel[2] (2015)	-	-2.89
Italy[2] (2015)	0.00	-
Mexico (2016)	0.00	0.00
Netherlands (2015)	0.00	0.00
New Zealand (2015)	0.00	0.00
Slovenia (2015)	-0.15	-0.20
Sweden (2015)	-0.13	-1.34
United States (2015)	-0.36	-2.80

Note: The results for the income column show the simple average of the adjustments to primary income (B5), disposable income (B6) and adjusted disposable income (B7), as percentage of the original estimate. For Italy, only results for primary income and disposable income are included, due to missing information on adjusted disposable income. Furthermore, for Israel and Italy, information is only available for either consumption or income.
Source: Zwijnenburg et al. (2021[2]).

References

European Commission et al. (2009), *System of National Accounts 2008*, https://unstats.un.org/unsd/nationalaccount/docs/SNA2008.pdf (accessed on 29 September 2017). [1]

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, Paris, https://www.oecd-ilibrary.org/docserver/615c9eec-en.pdf?expires=1669202191&id=id&accname=ocid84004878&checksum=689A9023E81B32F43D91EF34FBDE7A7A (accessed on 23 November 2022). [2]

Notes

[1] Please note that information on final consumption expenditure is only available at an aggregated level, as the breakdown in the supply-and-use tables for the underlying consumption items does not allow for a split between the household sector and the NPISH sector. That means that assumptions will need to be made on this split for underlying items. For that purpose, it will need to be assessed what type of consumption items are most likely consumed by NPISHs.

[2] An alternative is to treat all persons within an institutional household as separate one-person households in compiling the distributional results. This may be preferred when the number of institutional households in a specific country is relatively large or has a large representation in specific household groups. However, compilers and users need to be aware that this may not do justice to the specific circumstances these individuals live in. Furthermore, they have to be transparent to users on this specific treatment of institutional households.

5 Determining relevant micro data sources

The second step in the methodology concerns the selection of relevant variables from micro data sources to link to the national accounts items. This chapter discusses the main types of micro data sources that may be available for this purpose and provides guidance on how to best approach the selection.

5.1. The importance of selecting the appropriate items from micro data sources

Once the adjustment of the national accounts totals is performed, the methodology foresees a second step to link micro data to the various national accounts items. Whereas the national accounts totals as derived in the first step provide the benchmark for the distributional results, micro data are essential for distributing these amounts across household groups.

For the purpose of obtaining the most accurate distributional results, those micro variables should be selected that best match the national accounts items in terms of scope and definition. A targeted and structured micro data sources' selection is meant to be a precondition for the success of deriving distributional results in line with the adjusted national accounts totals. For this purpose, it is essential to have a comprehensive overview of the available micro data sources with information on what variables are included and their specific definitions, their population coverage, and their timeliness and frequency. These characteristics are essential in selecting the best micro data source for each item.

Ideally, items can be found in the micro data sources that perfectly match the definition and coverage of the national accounts variables. However, because of different purposes, the classifications and definitions will not always perfectly match. This may require aggregating or disaggregating specific items from the micro side or the national accounts side, or making explicit adjustments in order to provide for a better horizontal match (i.e. matching micro and macro data at the level of the various income and consumption items). Furthermore, it has to be borne in mind that not all items will have a corresponding item in the micro data sources, especially those items that are specific to the System of National Accounts (SNA) (such as *FISIM* and *investment income disbursements*).

It is also important to note that the selection of the most appropriate items may concern the use of multiple data sources. This may necessitate the need to link the data across various data sources to arrive at coherent households' results across the various items, known as vertical consistency. These issues are dealt with in the following chapters.

In order to facilitate the selection of the relevant items from micro data sources to match the national accounts items as described in Chapter 2, this chapter discusses the main types of micro data sources that may be available in countries for the compilation of distributional results consistent with national accounts totals. In general, the main data sources can be broken down into two categories, i.e. survey data and administrative data. Sections 5.2 and 5.3 discuss these two types of data sources with their main characteristics and main pros and cons. Chapters 10 and 11 provide a more detailed overview of information that may be available at the level of the detailed income and consumption items. As the available micro data sources may differ across countries, it is not possible to state which data sources should be used by default in the compilation of distributional results. Section 5.4 presents criteria on the basis of which compilers may select the best data sources depending on their national situation. To provide some practical guidance in selecting the most appropriate data source, Section 5.5 provides an overview of the micro data sources that are used by countries involved in the DNA work.

5.2. Survey data

Traditionally, surveys are an important data source for statistics. They have provided statisticians with relevant input data for the compilation of various kinds of statistics for many years. They may be used to obtain information on specific themes (e.g. the use of different transport modes or details on the expenditures by tourists) as well as on specific groups of entities (e.g. households or corporations), the setup of the survey depending on the specific purpose of the statistics.

Survey data derive from specific questionnaires including questions related to the target variables to be published in the statistics. In some cases, the questionnaire directly targets the variables that are needed

in the statistics, but it may also be the case that information is requested on the basis of which the relevant variables can be derived. The information is usually collected for a sample of the population in which each sample unit (e.g. a person, a household or a corporation) is assigned a specific weight, indicating how many population units are represented by that specific sample unit, in order to be able to derive information for the total target population. The sample is the result of a design process that takes into account population characteristics (e.g. socio-demographic information) and non-response rates in order to arrive at representative results for the population as a whole. It also considers the response burden and other costs related to the survey, in order to arrive at good quality results while minimising the costs.

The main benefit of survey data is that statisticians have a large influence on the specific questions to be included in the questionnaire as well as on the sample design. Although some questionnaires are prescribed by international organisations to ensure international comparability, most of them are set up at the national level designed by the statistical authorities, taking into account country specific circumstances and needs. Another benefit is that surveys can include questions related to household as well as to individuals within the households, which broadens the scope for analysis and to group households according to alternative criteria based on characteristics of the individuals. This is often not possible with administrative data sources as they usually focus on either the individual or the household. Finally, survey data may be able to capture subjective variables (such as sentiments, opinions and perceptions) which can normally not be retrieved via administrative data sources.

The downside of survey results is that they may be affected by specific types of errors. The first concerns estimation errors which relate to the extrapolation of the survey results to the target population. They can be linked to the sample size, the representativeness of the sample and the magnitude of the non-response. The errors related to the sample size are referred to as standard sampling error, implying that the smaller the survey sample, the larger the margin of error surrounding the results, as less data underlie the ultimate estimates. The other two issues are referred to as coverage errors. These occur in the case of the sampling frame being different from the target population and in the case of selective non-response, both possibly causing bias to the overall results. It may for example be difficult to capture homeless people or the very rich, or it may be the case that some specific items are highly concentrated among a small set of households which may be difficult to properly capture via a sample survey. All these aspects may lead to margins of error surrounding the results.

A second type of errors concerns measurement errors, related to mistakes in the data reported in the surveys, either in the form of item non-response or the reporting of incorrect data. These errors may be due to misinterpretation of the questions, difficulty by respondents to recall the exact values, or deliberate misreporting. Meyer, Mok and Sullivan (2009[1]) show that a lot of statistics have to deal with measurement error and that these errors have increased over time, at least for some specific items. Especially questions on income are usually understood to be relatively sensitive and prone to higher non-response rates or larger measurement errors. Measurement errors may affect both survey and administrative data, but the impact on survey results may be larger due to the fact that these results are weighted to arrive at population totals, whereas this is usually not the case for administrative data (or to a much smaller extent).

Because of the possible impact of estimation and measurement errors, survey data are usually prone to substantial checks and edited in case of any errors, including imputations in case of unit non-response (if an entity does not respond to the survey) and item non-response (if a specific item is missing from the survey for a specific respondent).[1] If the unit non-response is random, the sample weights may simply be adjusted to account for the missing entity, but in case of non-random unit non-response or item non-response, more sophisticated techniques are needed to properly correct for the missing data.[2] In that regard, evidence shows that non-response is usually not random, but correlated to characteristics such as age, educational level and social status.[3]

In analysing the micro data and matching them with the national accounts totals, it is relevant to have more insight in the construction of the micro aggregates, i.e. broken down into the initial survey results, sample

weights, and any corrections that may have been made to the micro data and the sample weights. A better understanding of the composition of the micro aggregates may provide more insight into the robustness of the results and possible margins of error surrounding the results. This may be particularly important in case of large gaps between the micro aggregates and the national accounts totals.

Another important issue that should be borne in mind when analysing differences between the micro and macro results when using survey data, is that surveys often focus on a specific point in time, which implies that the target population only includes the persons or households at that specific date. Conversely, national accounts focus on results over a certain time span, including all relevant information over that period. This means that survey data may exclude households that are included in the macro totals. For example, if the reference date is the 1st of January, the survey target population will exclude any immigrants and new-borns that will enter the national accounts population throughout the year. On the other hand, if the reference date is the 31st of December, emigrants and people that may have passed away throughout the year will not be covered in the survey results. These differences should be corrected for, before aligning the micro data to the national accounts data.

The availability of survey data sources will differ across countries. Despite the fact that many statistical offices compile similar statistics and despite several initiatives to further harmonise survey designs at the international level, countries often still have their own specific survey program, with the design of underlying surveys depending on historical considerations and country-specific issues. Notwithstanding these differences, almost all countries conduct surveys to obtain information on income and consumption, sometimes collecting data on these two topics in separate surveys, sometimes combining the two in a single survey. Moreover, some countries combine them with the collection of data on wealth.

Multipurpose surveys may be a very useful tool to collect consistent data on various topics at the household level and to analyse the relationships between these various topics. On the other hand, as explained by Cifaldi and Neri (2013[2]), combining income and consumption questions in one survey may reduce the willingness of respondents to participate to the questionnaire or hamper the quality of the survey results given the high level of detailed information which is required from respondents on both topics. Despite these caveats and the different approaches applied by countries, it is common practice to include a few recall questions on consumption within the income surveys, and vice versa, with the aim of analysing household saving (computed as income minus consumption) and of having an immediate feedback on the coherence between income and consumption information, enhancing the quality of the responses. Furthermore, it may help to link information on similar types of households across different surveys.

In addition to household surveys on income and consumption, which usually target residents of private dwellings (excluding people living in institutional households and people with no usual place of residence), relevant information may also be derived from other surveys, such as business surveys which may contain information on unincorporated enterprises, and surveys targeting specific topics such as health, housing or energy use. These may provide input for the distributional of some specific items in the work. Section 5.5 provides an overview of the various data sources that are used by countries involved in the DNA work.

5.3. Administrative data

An increasing number of statisticians is using administrative data (also often referred to as register data) in the compilation of their statistics. On the one hand, this relates to the increasing pressure for statistical offices to reduce the response burden and to cut costs related to the production of statistics. On the other hand, it is understood that the use of administrative data may solve issues of decreasing quality of survey data as observed in various countries due to lower response rates and provide the possibility to publish data at much more granular levels of detail.

Administrative data usually concern large data sets that cover the whole or a large part of the population. Hence, they are less subject to estimation errors and, as stated above, provide the opportunity to publish at very granular levels of detail. This is the main benefit of administrative data over survey data. Administrative data may still suffer from under-coverage or missing data, for example due to the fact that part of the population may fall outside the scope or due to non-reporting, but the impact will be far smaller than for sample surveys. Furthermore, the risk of non-reporting is often reduced by the use of legal sanctions to enforce compliance.

Like survey data, administrative data may suffer from measurement errors. These may relate to deliberate or accidental misreporting by respondents. In this regard, Adler and Wolfson (1988[3]) explain that high gaps between administrative data and macro totals may point to large under-reporting in tax records due to tax evasion. Fioro and D'Amuri (2006[4]) even argue that survey data may have a higher reliability than administrative data, as the latter may be affected by strategic reporting to lower the tax burden. Burkhauser et al. (2010[5]) also discuss the issue of "fiscal manipulation strategies" in which taxpayers reclassify specific types of income in order to limit their tax liabilities. Moreover, it is understood that the quality of auxiliary variables in administrative data sources (such as socio-demographic information or breakdowns of certain items) may be less reliable, as this is not related to the main purpose of the data collection. On the other hand, the legal sanctions that are often related to non- or misreporting normally ensure a high level of accuracy of the information enclosed in registers (Moore, Stinson and Welniak, 1997[6]). Moreover, part of the information may be reported by companies or financial institutions. Furthermore, it has to be borne in mind that the impact of errors, at least if they are not systematic, will be smaller than for survey data, because of the fact that the data sets usually cover the population as a whole. However, in order to avoid measurement errors in administrative data to negatively affect distributional results, it is important to carefully check the data and to correct for any errors, particularly if one targets to publish at very granular levels of detail.

Another important characteristic of administrative data is that whereas statisticians often have a large influence on the items to be included in a survey questionnaire, this is usually not the case for administrative data. As these are usually collected for administrative purposes, the setup of the data collection and the items included are usually not fully tuned to statistical needs. As a consequence, the items included in administrative data may often be based on different concepts and classifications than the ones used in micro statistics or in national accounts. Administrative records on income may for example be limited to cash-based income and may exclude certain in-kind payments. It may also be the case that certain categories include benefits that are treated differently in statistical measures, such as holding gains and losses. These differences may require re-classification of items as well as adjustments to correct for conceptual differences.

Furthermore, some of the items covered in administrative data sources may change over time as a consequence of changes in policy. For example, Burkhauser et al. (2013[7]) show how the change in the tax income base in Australia to include a more detailed breakdown of income items (e.g. dividends, capital gains, etc.) led to an overstatement of the increase in the income share held by top income groups due to the fact that the newly included income sources in the tax base were disproportionately held by them.

Finally, the unit of analysis may not always align to statistical needs. Dependent on the administrative purpose, the data set may focus on persons, households and/or other combinations of individuals. This may require adjustments in order to be able to use the administrative data for specific statistical purposes.

Most administrative data sources used by countries in the DNA work concern data sets from government agencies. Examples are population census data from statistical institutes, tax data from tax authorities, information on inbound and outbound visitors from immigration authorities, data on home ownership from land registers, and information on employment and wages from social security authorities. However, other administrative data sources may be envisaged as well, such as data from pension funds, insurance companies and large energy companies. They may provide useful information on some specific items.

In assessing the usefulness of administrative data sources, it is important that compilers obtain more information on their characteristics. In addition to the items that are included in the data source, it is important to obtain meta data on the definitions of these items, the population covered by the data source, the unit of analysis applied, the degree of under-coverage of specific household groups, and the frequency and timeliness of the data. Furthermore, it is important to know whether the data have already been checked, and if so, what corrections have been made (including imputations for missing records or missing items). The latter is important for checking the robustness of the results and assessing the margins of error for various groups of households in view of matching the data with the national accounts totals.

Looking at the national accounts items that may be covered in administrative data sources, the coverage is expected to be relatively high on the income side, with tax information possibly being available for the estimation of *operating surplus from owner-occupied dwellings* (as well as land register data), *mixed income*, *compensation of employees* (as well as social security data), *property income* (also related to information on wealth), *taxes* and *social contributions and benefits* (as well as social security data). On the other hand, information will probably be lacking for income flows between households (e.g. *other current transfers*) and income items specific to the SNA (e.g. *investment income disbursements*).

On the consumption side, the amount of available administrative data sources is expected to be lower. Information may be available on consumption of *housing* (from land register data), *water, electricity, gas and other fuels* (from data obtained from energy and water suppliers), *health services* (from data from health providers), *purchases of vehicles* (from car registries), and *education* (from school registries). Furthermore, “big data” sources such as credit card data, bank statements and data from special discount cards may provide relevant information for the distribution of some parts of household consumption. However, these may not provide full coverage of all consumption expenditure and may require more research on how to use them to derive reliable estimates.

In addition to directly using administrative data to match the national accounts items, they may also be used as supplementary information to check or to complete the information obtained via surveys. In that regard, it is common practice in an increasing number of countries to combine register data with survey data in compiling income statistics. This practice may improve the quality of income estimates which may be under-reported in household surveys. As mentioned in the previous section, it may also be the case that administrative data sources include information on the number of people benefiting from a certain type of income or purchasing a specific good or service. This may then be used to assess the degree of item non-response for specific items in the survey. Combining survey and administrative data may in that regard also lead to better input in matching the micro data to the national accounts totals.

5.4. Selecting the most appropriate data sources

As survey programs and the availability of and access to administrative data sources may differ across countries, it is not possible to specify upfront which data sources should be used by countries in compiling their distributional results. This will depend on the available data sources in the countries, the variables included in the data sets with their specific definitions, the population covered, the assessment of the data quality, and the timeliness and the frequency of the data. In some cases, there may only be one data source available for a specific variable, but if information is available from multiple data sources, compilers should carefully assess the available information to see which provides the best match with the national accounts items. In some cases, this may involve combining information from multiple data sources for a specific item to complement the main data source for information that may be missing or to cross-check some of the information included in the main data source.

In selecting micro data variables in relation to the national accounts totals, there are generally two approaches, i.e. the single-source approach and the multi-source approach. In the first case, all micro variables are taken from the same micro data source, whereas in the second case, multiple micro data

sources are used in the process. Although using multiple data sources creates the need to link data across different data sets, it probably leads to the most useful micro data set underlying the distributional results. In that regard, a multi-source approach will often provide more and better links to the various national accounts items than a single-source approach. However, it also has to be borne in mind that selecting multiple micro data sources may in some cases lead to conflicting numbers on a same phenomenon (e.g. when information on a specific items is available from both the household survey and from tax records) as well as inconsistent estimates on inter-connected phenomena (e.g. when integrating data on income and consumption expenditure based on different surveys that may be based on different samples). This leads to a challenge to arrive at the best results, but it may also provide compilers with the possibility to cross-check and, if necessary, correct some of the micro data, of course after consultation of the micro experts.

In general, for each national accounts item, the micro variable should be selected that is regarded to provide the best basis for the distributional results in line with the national accounts total. On the one hand, this will depend on the conceptual fitness of the item with the national accounts variable and the difficulty to correct for any conceptual or classification differences. In this regard, items included in survey data may often provide a better match with the national accounts variables, although the concepts of the items covered in administrative data may often still come close (and auxiliary information may be available to correct for any conceptual differences). On the other hand, the selection of micro variables will depend on the quality of the underlying data to provide an accurate reflection of the actual distribution for this item for the target population (and the difficulty to impute for the part of the population that is missing). The latter relates to the possible impact of measurement and estimation errors on the underlying distribution in the micro data. It will often be difficult to assess the impact of measurement errors, as it would involve assessing the reliability of the reported data for the various data sets. Normally, it is possible to assess the impact of estimation errors, as it relates to the sample size in relation to the target population. In this respect, administrative data sources usually perform much better than survey data, as they cover the whole or a large part of the population, whereas survey data often rely on a sample, requiring weighting the data to arrive at population totals.

The selection of micro variables may also involve the combination of two or more data sources. Multiple data sets may for example provide reliable information on different parts of the population.[4] In this regard, it is important to bear in mind that survey data may often suffer from relatively low coverage in the tails of the distribution. When relying on survey data as main data source, it may therefore be relevant to use administrative data sources to obtain a better coverage in the tails. Information as included in other data sources may also be used to cross-check information as included in the selected data source. This may be particularly relevant in case of large gaps between the micro and macro totals.

In addition to looking at the best conceptual and statistical match, it is also important to look at the timeliness and frequency of the available data sources. Some data sources may be compiled on an annual basis, whereas others may only become available every couple of years. In that case, it may be preferred to choose the annual data source as it may provide the opportunity to compile distributional results more frequently or to look at ways to combine the two data sources to arrive at reliable and consistent results on an annual basis. The latter may depend on the stability of the distributions over time and the robustness of nowcasting or interpolation techniques to derive results for the intermediate years.

Furthermore, users are mostly interested to obtain information shortly after the reference period, so also the timeliness with which the data sources may become available may be an important factor in selecting the most appropriate data sources. In that regard, survey data often suffer from substantial time lags, whereas administrative data may often become available within a short period after the reference period.

5.5. Overview of data sources currently used in compilation process

As micro statistics differ across countries it is difficult to provide a comprehensive list of surveys that may be available across countries providing information on income or consumption to be used in the compilation of distributional results in line with national accounts. However, because of the importance of selecting the appropriate micro data sources and to assist compilers in obtaining a comprehensive overview of possible data sources, this section provides an overview of the data sources that are used by countries in the DNA work.

Table 5.1 provides an overview of the micro data sources used by countries for both income and consumption items. It shows that different types of surveys and administrative data sources can be distinguished that contain relevant information on the household sector to be used as input to compile distributional results in line with national accounts. Some cover a single topic, focusing on either income or consumption, whereas others may cover both topics and, in some cases, even wealth information. Furthermore, some include detailed information on various income and/or consumption items, whereas others focus on a specific item, for example health, housing, farming or fishery.

Table 5.1. Micro data sources used by countries in their DNA work (stocktake conducted in the first half of 2022)

Country	Name of the data source	Nature*	# elements where data source is used for				Frequency**
			Income	Consumption	Saving	Socio-demographic	
Australia	Survey of Income and Housing (SIH)	S	X	-	-	X	B
	Household Expenditure Survey (HES)	S	-	X	-	X	Every 6 years
	Census	C	X	X	-	X	Every 6 years
Austria	Statistics on Income and Living Conditions (SILC)	M	X	X	-	X	A
	Household Budget Survey (HBS)	S	X	X	-	-	Every 5 years
	Household Finance and Consumption survey (HFCS)	S	X	-	-	-	Every 3 years
	Education Expenditure Statistics	M	X	-	-	-	A
	School Statistics	Y	X	-	-	-	A
Belgium	Household Budget Survey (HBS)	S	X	X	-	-	B
	Statistics on Income and Living Conditions (SILC)	S	X	X	-	X	A
	Tax records - Belcotax/IPCAL data	A	X	X	-	-	A
	Demobel/CENSUS	C	-	-	-	X	A
	Data on government spending	Y	X	X	-	X	Y
	Belgian Health survey interview – HISIA	S	-	X	-	X	Every 5 years
Canada	Social Policy Simulation Database/Model (SPSDM)	M	X	X	-	X	O
	Canadian Income Survey (CIS)	M	X	X	-	X	A
	Survey of Household Spending (SHS)	M	X	X	-	X	A
	Annual Income Estimates for Census Families and Individuals (T1FF)	A	X	X	-	X	A
Czech Republic	Statistics on Income and Living Conditions (SILC)	S	X	-	-	X	A
	Household Budget Survey (HBS)	S	X	X	-	X	A
	Income tax return by individuals	A	X	-	-	X	A
	Population and Housing Census	C	-	-	-	X	Every 10 years
Finland	Statistics on Income and Living Conditions (SILC)	S	X	X	-	X	A

Country	Name of the data source	Nature*	# elements where data source is used for				Frequency**
			Income	Consumption	Saving	Socio-demographic	
	Household Finance and Consumption survey (HFCS)	S	X	-	-	-	Every 3 years
	Household Budget Survey (HBS)	S	X	X	-	-	Every 5 years
	Tax data	A	X	-	-	X	A
France	Household Budget Survey (HBS)	S	X	X	X	X	Every 5 years
	Tax and social incomes survey	M	X	-	X	X	A
	Health data set	M	X	-	-	X	O
Ireland	Statistics on Income and Living Conditions (SILC)	S	X	-	-	X	A
	Household Budget Survey (HBS)	S	X	X	-	-	Every 5 years
	Fiscal data	A	X	-	-	-	A
	Administrative data on expenditure per student per level of education	A	X	-	-	-	A
Israel	Household Expenditure Survey	S	-	X	-	-	A
Italy	Statistics on Income and Living Conditions (SILC)	M	X	-	-	X	A
	Household Budget Survey (HBS)	S	-	X	-	-	A
	Survey on Household Income and Wealth (SHIW)	S	X	-	-	-	Every 2 years
	Ministry of Economy and Finance estimates of per capita expenditure on health	Y	X	-	-	-	A
	School Statistics	M	X	-	-	-	A
Mexico	Survey of household income and expenditure (ENIGH)	S	X	X	-	X	B
Netherlands	Register for Addresses and Buildings	A	X	X	-	X	A
	Income tax data	A	X	X	-	X	A
	Wealth tax data	A	X	-	-	X	A
	Household Budget Survey (HBS)	S	X	X	-	X	Every 5 years
	Pension Claims Statistics	A	X	-	-	X	A
	Household Finance and Consumption survey (HFCS)	S	X	-	-	-	Every 3 years
	Giving in the Netherlands panel survey	S	X	-	-	X	B
	Longitudinal Internet Studies for the Social Sciences	S	X	-	-	X	B
	Wage Register	A	X	-	-	X	A
	Insurance Healthcare Act	A	X	-	-	X	A
	Long-term Healthcare Act	Y	X	-	-	X	B
	Education enrolment registration	A	X	-	-	X	A
	Legal counsel	A	X	-	-	X	A
	Population data	C	X	X	-	X	A
Portugal	Statistics on Income and Living Conditions (SILC)	S	X	-	-	X	A
	Household Budget Survey (HBS)	S	-	X	-	-	Every 5 years
	Census	C	X	-	-	X	O
Slovenia	Statistics on Income and Living Conditions (SILC)	M	X	X	-	X	A
	Real Estate Register	A	X	-	-	-	A
	Household Budget Survey (HBS)	S	X	X	-	-	Every 3 years
Sweden	Income and tax statistics	A	X	-	-	X	A
	Statistics on Income and Living Conditions (SILC)	S	X	X	-	X	O (for variables used)
	Household Budget Survey (HBS)	S		X	-	X	Every 5 years

Country	Name of the data source	Nature*	# elements where data source is used for				Frequency**
			Income	Consumption	Saving	Socio-demographic	
	Distributional analysis system for income and transfers	A	X	X	-	X	A
	Property tax register	A	X	X		X	A
	Vehicle register	A	X	X		X	A
United States	Current Population Survey (ASEC)	S	X	-	-	X	A
	Consumer Expenditure Survey	S	-	X	-	X	A
	Statistics of Income	Y	X	-	-	-	A
	Survey of Consumer Finances	S	X	-	-	-	O
	American Community Survey	S	X	-	-	-	A
	Medial Expenditure Panel Survey (MEPS)	A	X	-	-	-	A

Note: * Nature of data source: C = Census; S = Survey data; M = Combination of survey and administrative records; A = Administrative records; Y = Secondary statistics.
** Frequency: M = Monthly; Q = Quarterly; A = Annual; B = Biannual; O = Occasional.

References

Adler, H. and M. Wolfson (1988), "A prototype micro-macro link for the Canadian household sector", *Review of Income and Wealth*, Vol. 34/4, pp. 371-392, http://www.roiw.org/1988/371.pdf (accessed on 11 October 2017). [3]

Bricker, J. et al. (2015), "Measuring income and wealth at the top using administrative and survey data", *Finance and Economics Discussion Series*, No. 2015-030, Board of Governors of the Federal Reserve System, Washington, https://doi.org/10.17016/FEDS.2015.030. [12]

Burkhauser, R. et al. (2010), "Recent Trends in Top Income Shares in the USA: Reconciling Estimates from March CPS and IRS Tax Return Data", http://www.iariw.org/papers/2010/2dBurkhauser.pdf (accessed on 11 October 2017). [5]

Burkhauser, R., M. Hahn and R. Wilkins (2013), "Measuring top incomes using tax record data: A cautionary tale from Australia", *NBER Working Paper Series*, No. 19121, http://www.nber.org/papers/w19121 (accessed on 11 October 2017). [7]

Cifaldi, G. and A. Neri (2013), "Asking income and consumption questions in the same survey: what are the risks?", No. 908, Banca d'Italia, https://www.bancaditalia.it/pubblicazioni/temi-discussione/2013/2013-0908/en_tema_908.pdf?language_id=1 (accessed on 11 October 2017). [2]

D'Alessio, G. and I. Faiella (2002), "Non-response behaviour in the bank of Italy's survey of household income and wealth", No. 462, Banca d'Italia, https://www.bancaditalia.it/pubblicazioni/temi-discussione/2002/2002-0462/tema_462_02.pdf?language_id=1 (accessed on 27 October 2017). [10]

D'Alessio, G. and A. Neri (2015), "Income and wealth sample estimates consistent with macro aggregates: some experiments", No. 272, Banca d'Italia, https://www.bancaditalia.it/pubblicazioni/qef/2015-0272/QEF_272.pdf?language_id=1 (accessed on 27 October 2017). [8]

Fiorio, C. and F. D 'Amuri (2006), "Tax evasion in Italy: an analysis using a tax-benefit microsimulation model", http://fiorio.economia.unimi.it/res/tax_ev.pdf (accessed on 30 October 2017). [4]

Meyer, B., W. Mok and J. Sullivan (2009), "The under-reporting of transfers in household surveys: Its nature and consequences", *NBER Working Paper Series*, No. 15181, http://www.nber.org/papers/w15181 (accessed on 27 October 2017). [1]

Moore, J., L. Stinson and E. Welniak (1997), "Income Measurement Error in Surveys: A Review", http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.39.6539&rep=rep1&type=pdf (accessed on 11 October 2017). [6]

Sabelhaus, J. et al. (2013), "Is the consumer expenditure survey representative by income?", *NBER Working Paper Series*, No. 19589, http://www.nber.org/papers/w19589 (accessed on 27 October 2017). [11]

Törmälehto, V. (2017), "High income and affluence: Evidence from the European Union statistics on income and living conditions (EU-SILC)", *Eurostat Statistical Working papers*, http://ec.europa.eu/eurostat/documents/3888793/7882117/KS-TC-16-027-EN-N.pdf (accessed on 9 October 2017). [9]

Notes

[1] In some cases, information on the number of persons or households receiving a certain income type or purchasing a specific good or service is used for this purpose. This type of information may be available from other (mostly administrative) data sources and may be confronted with the number of persons or households reporting a number for specific items in the survey to assess the possible degree of item non-response. Dependent on the setup of the data source, it may also provide insight which records may need to be edited.

[2] For example, D'Alessio and Neri (2015[8]) discuss post-stratification in which the socio-demographic composition of the sample is aligned with known distributions from the census or other statistics. Administrative data may also be used for post-stratification in case it includes socio-demographic information. Furthermore, Törmälehto (2017[9]) explains that register data may be used to calibrate survey weights in order to reduce the estimation error in the top tail, although it has to be borne in mind that the top tail may include more heterogeneous groups of households for which it may be more difficult to correct for non-response by simply adjusting the sample weights.

[3] D'Alessio and Faiella (2002[10]) show that non-response is often more frequent among higher income and wealthier households. Furthermore, Sabelhaus et al. (2013[11]) show that high income households are likely to be underrepresented in the consumer expenditure survey in the United States. The latter issue has been addressed by applying non-interview adjustment factors to the results based on fiscal data.

[4] For example, Bricker et al. (2015[12]) explain that for income data, surveys may often represent a good instrument to collect information on low-income earners since they are not limited by any fiscal thresholds as tax registers are (e.g. households with low income that are not requested to pay taxes are not recorded in tax statistics, or income components that are not subject to taxation by law), whereas tax records are assumed to provide better estimates of top income shares, since in general surveys are characterised by under-representation of very high income households.

6 Imputation for missing items

In the third step, any gaps between the micro data and the national accounts totals need to be bridged. To this end, first, imputations need to be made for elements not covered in micro data sources, such as for missing parts of the population, informal activities and items that are specific to the national accounts. This chapter presents specific methods to apply these imputations.

6.1. Introduction

Whereas step 2 as described in Chapter 5 foresees in the selection of micro data for each national accounts item, it has to be understood that some information may not be covered in micro data sources, requiring the need for imputations to include the relevant amounts in the distributional results. Because missing information in the micro data may explain a large part of the gap between the micro aggregates and the national accounts totals, imputing for this missing information constitutes the first part of the third step of the step-by-step approach. After making the relevant imputations, compilers can assess the remaining gap and allocate the amounts to the relevant households on the basis of the most likely underlying reasons for these remaining gaps. This second part of the third step is discussed in Chapter 7.

There are four types of missing information. First of all, some items will have no counterpart in the micro data. This is for example the case for items that are specific to the System of National Accounts (SNA). Secondly, it may be the case that certain items are covered in micro data sources, but that these are not (yet) available to the compiler. This may be related to the timeliness and the frequency of the data sources. Thirdly, it may be the case that the selected micro data source may not cover the whole population, for example in case of surveys that only target specific household groups or in case of reporting thresholds in administrative data. Finally, it may concern missing information related to the underground economy and/or illegal and informal activities. These are included in the national accounts but usually not covered in micro data sources. As this missing information may concern substantial amounts that may relate to specific households or household groups, their allocation may significantly affect the distributional results. For that reason, it is important to separately impute for these types of missing information.

This chapter discusses imputations in relation to these four underlying causes, presenting basic techniques for allocating the related amounts to the relevant households or household groups. Section 6.2 discusses the imputations for items for which no counterpart information is available in micro statistics. Section 6.3 discusses the case in which items are covered in micro data sources but are not (yet) available to compilers. Subsequently, Section 6.4 discusses imputations for missing parts of the population. Section 6.5 discusses the imputation for the underground economy, and illegal and informal activities. More detailed guidance on how to impute for missing information at the level of specific income and consumption items is provided in Chapters 10 and 11.

6.2. Imputation in case an item is lacking from micro data sources

The first type of imputations concerns those for items that have no counterpart in the micro data. This often relates to items that are specific to the SNA, such as *employers' imputed social contributions* (SNA codes D122 and D612), *investment income attributable to insurance policyholders* (D441), *investment income payable on pension entitlements* (D442), *financial intermediation services indirectly measured* (FISIM), and *social transfers in kind* (D63). As these items are specific to the SNA, no direct information will be available in micro data sources and the amounts will have to be allocated in a different way.

In general, three methods are available to derive an appropriate allocation in case no micro data are available, all making use of indirect information. The first method (defined as method B[1]) proxies the missing information by using the distribution of another component, assuming that the two are distributed in a similar way. The distribution for *employers' imputed social contributions* (D122) may for example be derived on the basis of the distribution of *wages and salaries* (D11), whereas the distribution of *FISIM* may be linked to *interest paid* (D41P) and *interest received* (D41R).[2]

The second method (method C) imputes missing distributional information according to exogenous data, e.g. socio-demographic information used for the distribution of *social transfers in kind*, available at the individual or at the household level. In both cases, it is preferred to employ the imputations at a level as detailed as possible as it enables classifying households into different groupings in the remainder of the

process. When imputations are made at the group level, this will need to be done for the various classifications that are needed.

If no information is available, a third method can be used (method D) in which the distribution of one of the balancing items is used as a proxy. In that way, the inclusion or exclusion of the component does not change the distribution of that balancing item. However, this should only be done as last resort. Naturally, this can only be done at the end of the process when the distributional information has been derived on the basis of the other variables. In applying this solution, it has to be decided to which balancing item to best link the specific item. For consumption items it will be best to link it to either *final domestic consumption expenditure* (P31DC) or *final national consumption expenditure* (P31NC) (excluding the item or items for which an imputation is still needed). For income components, the distributions may be linked to the *balance of primary incomes* (B5), *disposable income* (B6) or *adjusted disposable income* (B7). It will depend on the underlying item what aggregate will provide the best proxy. It may also be the case that one would like to use the distribution of a balancing item but excluding a specific item. Compilers should assess which item or combination of items they think will provide the best proxy for the relevant item.

Chapters 10 and 11 discuss the various income and consumption items in more detail including possible imputation techniques for the items that are most likely to be missing in micro data.

6.3. Imputation in case the micro data source is not (yet) available for a specific period

The second type of imputations concerns those for items that are usually covered in micro data sources but that may not (yet) be available to compilers for a specific reference period. This may be due to the fact that the data source only becomes available with a certain time lag or is only conducted every couple of years, as a consequence of which it is not available for the specific reference period.

In the case that data are not yet available, it may be relevant to assess whether results can be obtained by extrapolating results on the basis of historic data. The most simple approach is to just apply the distribution available for the most recent year (thus assuming no change in the relative distribution across households). A more sophisticated approach would be to look whether one can spot specific trends in the historic data that may assist in deriving more accurate estimates for the reference year. Alternatively, one may assess whether the results correlate to other data for which more timely information may already be available. This may be in relation to national accounts totals but also in relation to meso-information such as labour market data or sociodemographic information. Furthermore, in case of specific policy changes, one may try to assess how these may affect specific households or household groups. In this way, one may arrive at more accurate estimates for the reference year. These may then be revised once the actual micro data become available.

If a specific micro data source only becomes available every couple of years, the above techniques may be used to derive first estimates for the missing years. These can then be revised at a later stage when results become available for a more recent year. In that case, interpolation techniques could be applied to arrive at more accurate estimates for the intermediate years, therewith overwriting the earlier results.

For both the extrapolation and interpolation techniques, it is recommended to apply them at the micro level as this will lead to the most accurate results. In this regard, it will provide the opportunity to update the clustering of households according to the interpolated or extrapolated micro data, taking into account dynamics between household groups, which may not be captured if these techniques are only applied at the level of household groups.

For both techniques, it will be important to assess their reliability on the basis of the size and direction of the revisions for the various household groups. If needed, compilers may need to further improve the techniques to arrive at more reliable results. It is also important to look at the revisions for the various

household groups to assess at what level of detail to publish the estimates. If the revisions turn out to be particularly large at a specific level of detail, it may be decided to only publish the extrapolated results at a more aggregated level of detail.

6.4. Imputation for missing parts of the population in the micro data

A third type of imputations relates to specific groups of households that may be missing from micro data sources. With regard to surveys, this may relate to people living in overseas territories or in sparsely populated areas but also to other groups that may be difficult to capture, such as very rich households or people with no usual place of residence. With regard to administrative data sources, it may be the case that these only target specific parts of the population or use thresholds, which may exclude specific groups of households from the population.

In case specific groups of households are missing, it is important to assess whether their information can be obtained in other ways. A first solution is to impute on the basis of micro data available from other micro data sources. In that regard, survey data may be complemented with administrative data and vice versa. In that case, it is important to first check whether both micro data sets are based on the same underlying concepts. If this is not the case, the micro data from the "donor" data set will first need to undergo some adjustments in order to align to the concepts of the "recipient" micro data set. These adjustments may for example be done on the basis of patterns found for households that are covered in both data sets and that are deemed comparable with households for which imputations are needed.

An alternative solution is to look whether auxiliary information may be available on the households that are missing on the basis of which their results can be approximated. For example, if no information is available on property income for a specific group of households, information may still be available on their ownership of specific types of financial and non-financial assets. In that case, this may be used to derive estimates for the missing population on the basis of assumptions of a specific rate of return. It may also be the case that another item may provide a valid proxy to derive the results for the missing households. This is similar to the technique as explained in Section 6.2 under method B, but now only being applied to a part of the population. In that regard, it is also possible to impute for the missing part of the population by linking it to exogenous information, in line with method C as explained in Section 6.2.

A third solution is to look for comparable households in the micro data set on the basis of which the missing households may be imputed. In some cases, this may concern a simple adjustment of the sample weights, but in case the missing households have very different characteristics, it may be needed to link them to specific individual households in the sample or in the register. This may be done on the basis of one-to-one linking, searching for a specific household record with similar characteristics, but it may also involve looking at a group of households with similar characteristics and taking the average amount of this group. Finally, it may involve regression analysis in which the value for a specific household is explained on the basis of a set of underlying characteristics derived on the basis of analysis of data of other households included in the data set. This is explained in more detail in the OECD Framework for Statistics on the Distribution of Household Income, Consumption and Wealth (OECD, 2013[1]).

Finally, parametric tail adjustments are often used to assess the impact of missing very wealthy households. This can be done by using external benchmark data to assess the size of the measurement error, such as done by Vermeulen (2014[2]) who uses the Forbes list of extremely wealthy to improve estimates of wealth survey micro data. This technique can also be applied to income. In this regard, Lakner and Milanovic (2013[3]) proxy the missing top incomes on the basis of the discrepancy between survey and NA consumption data and allocated this to the top using Pareto fitting.[3] The latter is, however, not preferred as multiple reasons may underlie the gaps between the micro and the macro aggregates, so taking this as a proxy for the missing top incomes may lead to incorrect distributional results.

Alternatively, Törmälehto (2017[4]) suggests, in the absence of external data on top incomes, to replace the whole tail of the outliers in survey data with estimated Pareto distributions, using hypothetical Pareto coefficients. Furthermore, Grilli et al. (2022[5]) provide a specific application of a Pareto-tail adjustment for income, using the available micro data to explore the existence of a Pareto-tail for specific items and providing guidance on how to make adjustments to the micro data in case the top-tail appears to be missing from the micro data source.

In analysing the possible need for top-tail adjustments, compilers are also encouraged to assess the distribution of the top tail in other countries and to compare survey-based results with register-based results. Furthermore, a comparison over time may provide useful insights into whether information at the top (and at the very bottom) may be missing for specific years.

As the group of households that may be missing from the micro data source may concern households with different characteristics, it may require different techniques to impute for the missing information. In that regard, it is recommended to try to derive more or less homogeneous groups of households for which a specific technique is deemed to provide the best results. This grouping can be done on the basis of socio-demographic information as well as on the basis of values obtained for these households in other parts of the work. For each household group, amounts should be derived on the basis of the technique that is deemed most reliable. This may for example imply that auxiliary information is used to impute values for unemployed persons that are not captured in the survey, an adjustment of the survey weights is applied to include households living in sparsely populated areas, and a Pareto-tail approximation is used to derive results for the very high-income households. Results on the basis of the different techniques may also be compared to see whether they show large differences and whether adjustments may be needed to some of the results before incorporating them in the distributional analysis.

It is recommended to select the appropriate imputation techniques for the relevant underlying household groups in close cooperation with the responsible experts from the relevant micro data source. They have the best overview of what is covered in the micro data source and what imputation techniques may lead to the best approximation for specific groups of missing households and to comparable results with the data included in the micro data source. Moreover, they may be able to process (some of) these imputations as part of their compilation process, providing the compilers of the distributional data with a consistent, comparable and comprehensive data set at the micro level.

6.5. Imputations for the underground economy, and illegal and informal activities

The fourth group of imputations concerns those for economic activities that are deliberately concealed to avoid tax payments (underground production) or are not captured because of their illegal or informal nature. As these activities are usually not captured in micro data sources (for that reason often referred to as the non-observed economy), the related amounts will have to be estimated indirectly in order to include them in the distributional results.

As national accountants often make explicit estimates for these activities, this will normally provide the starting point for allocating the relevant amounts to the underlying households or household groups. Ideally, the national accounts provide information on the imputed amount broken down into the three underlying types of activities (i.e. underground, illegal and informal activities), so that the amounts can be allocated accordingly. In that regard, it is not only important to obtain information on the specific values, but also on how these amounts have been derived. It may then be assessed whether the underlying assumptions for calculating these amounts may also provide input to allocate the relevant amounts to underlying households. For example, if part of the underground economy is imputed on the basis of the assumption that specific types of jobs are more likely to be involved in such types of activities, this may be used to link the amounts to specific groups of households. Of course, these assumptions can be further tuned to take into account specific characteristics that are available at the micro level on the basis of which

it can be decided which households should be assigned what amount and whether some specific groups of households should be excluded. For example, assumptions may be made with regard to the background (sex, age, employment status and living location) of drug dealers, prostitutes and traffickers. For some of these activities, information may also be available from police records.

If no specific information is available from the national accounts on the size of the underground economy, illegal and/or informal activities, it is important to separately estimate the related amounts and to separately allocate them to the relevant households, as the amounts are likely to involve (partly) different groups of households. A first step would be to look at the micro-macro gap and to assess what part may be explained by these three types of activities. In the second step, the amounts should be allocated to the households that are most likely to be involved in them. As mentioned above, in some cases information may be available on what type of households are more likely to be involved in what type of non-observed activities. In that case, the related amounts can directly be allocated to relevant households or household groups on the basis of their specific characteristics. In other cases, assumptions will need to be made, for example looking at the likelihood of households to be involved in these activities on the basis of their reported data (see below).

In looking at which households may possibly be involved in underground activities, illegal or informal activities, one may look at the plausibility of the overall results at the household or at the household group level to see whether specific amounts may be missing. For example, if for some household groups consumption by far exceeds their income, it may be the case that they are actually running a deficit and sell off assets or engage in liabilities,[4] but it may also be the case that part of their income is not covered in the micro data source. In that case, this may require an imputation, the specific item depending on what item is most likely to be underreported by the specific households. The latter may be based on the items that are most likely to be underreported in general and show the largest micro-macro gaps (e.g. mixed income, property income and social benefits) or which are most likely to be underreported for specific groups of households.

In that regard, it is also interesting to cross-check results for households with similar characteristics. It may of course be the case that they report different amounts for specific items, but in case these are much larger in a specific year or for a specific group of households, this may point to possible outliers or errors in the data. If on the basis of such analysis it is indeed concluded that the micro results are likely to be incorrect due to underreporting in relation to underground economy, illegal or informal activities, an imputation may need to be made, looking at a more plausible value in relation to previous years or comparable households.

Of course, some of these imputations will be very sensitive to assumptions on the plausibility of the micro data. For that reason, it is very important that this analysis and allocation are done by or in close cooperation with the responsible experts from the relevant micro data source. They are best equipped to assess the plausibility of the results for the various groups of households and best suited to assess where an imputation for non-observed activities may be most valid.

6.6. Conclusions

This chapter discussed general techniques how compilers may deal with elements for which micro data may be lacking. As explained, the imputation technique, which may differ across households or household groups, will depend on whether there is no micro data available at all, whether this may not (yet) be available for the specific recording period, whether only part of the population is covered, or whether information on specific activities may be missing. The micro and macro experts should discuss which technique is deemed to provide the most reliable estimates for which specific households or household groups and carefully check the results, also in relation to data that are available in the micro data sources.

Ideally, imputations are made at the micro level. This provides the opportunity to check the reliability of the results at the micro level and also ensures that the next steps in the process can start from underlying micro data. In that regard, it has to be borne in mind that the alignment of the micro data to the national accounts totals should also be done on the basis of the micro data, after which the households can be ranked at the micro level according to their income levels including the imputed amounts. This also ensures that results can be aggregated into multiple household groupings, all arriving at consistent results in line with national accounts totals.

References

Grilli, J., P. Engelbrecht and J. Zwijnenburg (2022), *Pareto tail estimation in the presence of missing rich in compiling distributional national accounts*. [5]

Lakner, C. and B. Milanovic (2013), "Global Income Distribution - From the fall of the Berlin Wall to the Great Recession"*, Policy Research Working Paper*, No. 6719, World Bank, https://openknowledge.worldbank.org/bitstream/handle/10986/16935/WPS6719.pdf?sequence=1&isAllowed=y (accessed on 27 October 2017). [3]

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. [1]

Törmälehto, V. (2017), "High income and affluence: Evidence from the European Union statistics on income and living conditions (EU-SILC)"*, Eurostat Statistical Working papers*, http://ec.europa.eu/eurostat/documents/3888793/7882117/KS-TC-16-027-EN-N.pdf (accessed on 9 October 2017). [4]

Vermeulen, P. (2014), "How fat is the top tail of the wealth distribution?"*, ECB Working Paper*, No. 1692, ECB, http://www.ecb.europa.eu/pub/scientific/wps/date/html/index.en.html (accessed on 27 October 2017). [2]

Notes

[1] Please note that method A is reserved for deriving distributional results on the basis of actual underlying micro information.

[2] Please note that ideally this should only include interest paid to and received from banks, but the total amounts can still provide a good proxy.

[3] Pareto-tails are based on the observation that in many populations the income distribution at the top is distributed in a similar way. As explained by Lubrano (2017[6]) it assumes that the number of individuals whose income exceeds a given level x can be approximated by Cx^{α} for some choice of C and α. This approximation seems particularly accurate for large incomes, i.e. for x above a certain threshold. Therefore, Pareto tails approximations are often used to check the plausibility of survey results for higher income households. In that regard, they can also be used to derive estimates in case very high-income households are deemed to be missing. For more information on Pareto tails, please see Vermeulen (2014[2]), Lakner and Milanovic (2013[3]), Armour et al. (2014[7]) and Chakraborty and Waltl (2018[8]).

[4] For this purpose, it would be very useful if the information could be combined with information from the capital and the financial accounts.

7 Aligning micro data with national accounts totals

The second part of the third step concerns the alignment of the micro data to the national accounts totals. Dependent on the size of the gaps, this may have an important impact on the results. This chapter explores potential reasons for the gaps and presents a framework that may assist in their allocation to the relevant households.

7.1. Introduction

In the second part of the third step, after imputations have been made for items for which micro data are lacking, micro data have to be aligned to the national accounts totals. How this step may affect distributional results will mainly depend on the size of the gaps between the micro data and the national accounts totals, and on the available information on how to allocate the amounts to the relevant households.

The alignment of the micro data to the national accounts aggregates may have a large impact on the results. That is why it is important to look at the most likely reasons for these gaps and to allocate them to the underlying households accordingly. A simple proportional or equal allocation across the distribution may otherwise only add bias to the distributional results instead of providing a more comprehensive and consequently a more accurate overview of inequality.

This chapter presents a framework that may be used to properly allocate micro-macro gaps to the relevant households. It distinguishes the various reasons that can be causing any gaps between the micro and macro data and provides compilers the possibility to allocate the gaps to underlying households on the basis of the most likely underlying reasons. Compilers are encouraged to use this framework as it is assumed to lead to better results than simply applying a proportional allocation. Whereas this chapter presents the basic framework, Chapters 10 and 11 discuss the various income and consumption items in detail, also providing more background information on the most likely causes for gaps between the micro and macro data.

The chapter first discusses the possible impact of gaps between micro and macro data on distributional results and how compilers may assess this impact for their results in Section 7.2. It then provides an overview of the items that have shown the largest micro-macro gaps in countries in Section 7.3. This is followed by an explanation of the main reasons that may be causing the gaps between the micro and macro aggregates, which forms the first part of the framework in Section 7.4. In Section 7.5, it is explained how the gaps can be allocated to the relevant households on the basis of the most likely underlying reasons. The chapter concludes with an overview of the framework in Section 7.6.

7.2. The impact of micro-macro gaps in compiling distributional results

To assess the possible impact of gaps between the micro data used in the compilation process and the macro totals derived after step 1, one can look at their share in the overall results. In that regard, it has to be understood that the distributional outcomes are the result of underlying micro data,[1] imputations for items or part of the population for which micro data is lacking,[2] and the alignment of these micro data to the macro aggregates. A method to acquire more insight into the role of these three components in the compilation of distributional results is by deriving their coverage rates. This can be done by looking at the relative shares of micro data, imputations and alignments in the absolute flows that constitute adjusted disposable income and actual consumption expenditure.

As adjusted disposable income consists of positive and negative items, the absolute flows should be considered to get a correct view of the contributions of each of the three factors to this aggregate. The relative shares for the imputations and alignments should then be calculated by dividing the sum of their absolute values by the sum of the absolute flows constituting adjusted disposable income and actual final consumption. In deriving the share of the micro data, it has to be borne in mind that micro totals can exceed the macro aggregates, so simply looking at the sum of the micro totals as percentage of the absolute flows would not provide a correct picture. Therefore, the share of the micro total in the balancing items should be derived as a residual, i.e. after deduction of the shares of the imputations and the alignments.

Table 7.1 provides an example of how this works in practice. It shows an aggregate that is composed of five underlying items, the first three items positively contributing to the aggregate while the last two

contributing negatively. For three of the items micro data underlying the distributional results is available. For two items no micro data is available ("NAV"), so distributions will have to be derived via imputations.

Table 7.1. Example of deriving contributions of alignment and imputation

	Adjusted NA aggregate (1)	Micro aggregate (2)	Micro-Macro gap (3) = (2) – (1)	Imputation (4) = (1) if (2) is NAV
Item A	90	70	20	-
Item B	50	NAV	-	50
Item C	60	85	-25	-
Item D (-)	35	30	5	-
Item E (-)	15	NAV	-	15
Total (= A + B + C – D – E)	150	125	-10	35

Source: The Author.

To derive the contribution of the alignment of the micro-macro gaps in the example above, the sum of the absolute amounts of the micro-macro gaps should be divided by the absolute amounts that constitute the total, i.e.

$$Contribution_{alignment} = \frac{(|20| + |-25| + |5|)}{(|90| + |50| + |60| + |35| + |15|)} = \frac{50}{250} = 20.0\%$$

The contribution of the imputation can be computed by dividing the sum of the absolute amounts of the imputations by the absolute amounts that constitute the total, i.e.

$$Contribution_{imputation} = \frac{(|50| + |15|)}{(|90| + |50| + |60| + |35| + |15|)} = \frac{65}{250} = 26.0\%$$

The contribution of the micro data can now be derived as a residual, i.e.

$$Contribution_{micro\ data} = 1 - contribution_{alignment} - contribution_{imputation} = 1 - 0.20 - 0.26 = 54.0\%$$

Figure 7.1 and Figure 7.2 show the share of each of these three components in the total absolute flows that constitute adjusted disposable income and actual final consumption expenditure for the household sector as a whole, derived from the EG DNA exercise conducted in 2015.

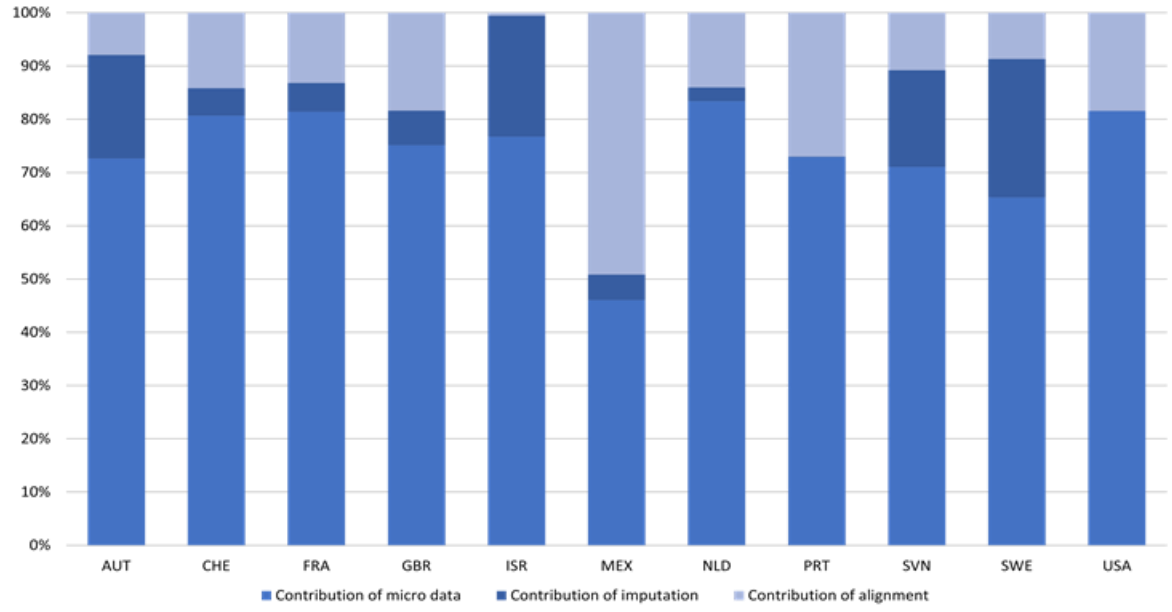
It can be concluded that the distributional results are to the largest extent based on micro data, but that the impact of imputations and alignments on the distributional results is often significant. In the 2015 exercise, micro data sources covered more than 70% of the underlying flows on average for adjusted disposable income, whereas this was more than 60% for actual final consumption expenditure.

The impact of the imputations and alignments on the distributional results can best be reviewed by presenting the size of these adjustments in absolute terms as percentage of the balancing items. This provides insight into the maximum amount that has to be allocated to the various households. Whereas positive and negative adjustments may (partly) cancel out at the level of the household sector as a whole, their overall impact on distributional results may still be significant, especially when they are allocated differently to the various household groups.

Figure 7.3 and Figure 7.4 present the impact of imputations and alignments on adjusted disposable income and actual consumption expenditure respectively, for the household sector as a whole, on the basis of the results of the 2015 exercise. For most countries, the size of the alignments is larger than that of imputations which relates to the fact that most countries have micro data available for the majority of items and only need to rely on imputations for few of them. In that regard, it should also be mentioned that the number of items for which countries report imputations and alignments varies across countries.

Figure 7.1. Contribution of micro data, alignment and imputation to adjusted disposable income

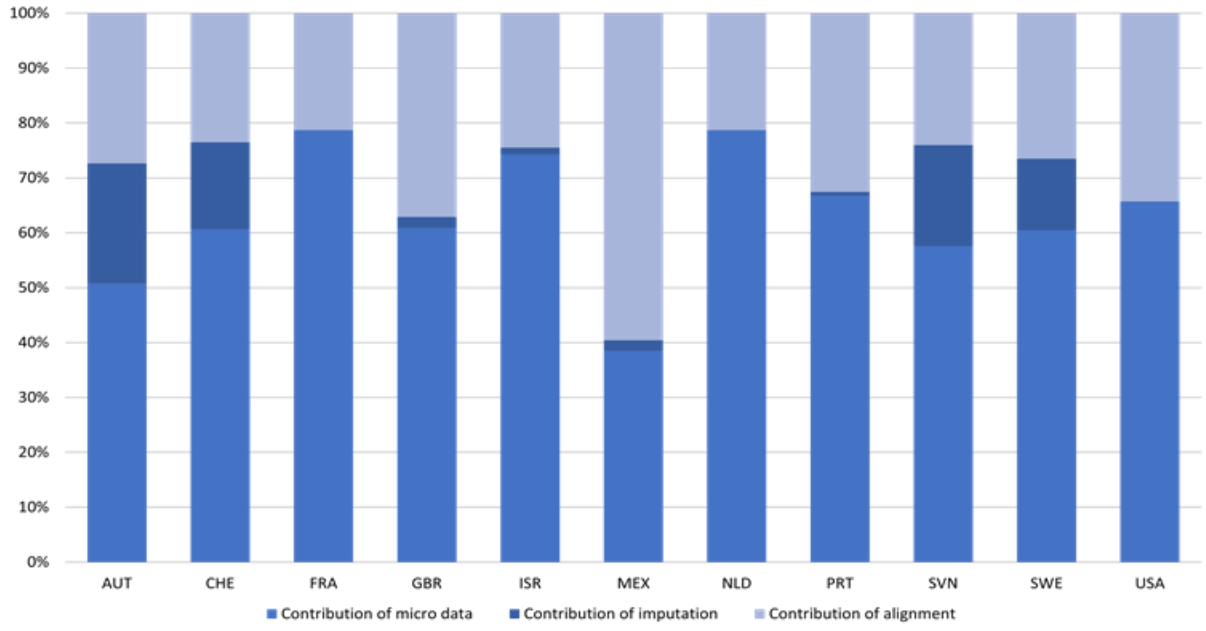
Share of absolute micro-macro gaps (alignment), imputations and micro data (derived as residual) in absolute amounts that constitute adjusted disposable income, EG DNA exercise 2015



Source: The Author.

Figure 7.2. Contribution of micro data, alignment and imputations to actual final consumption expenditure

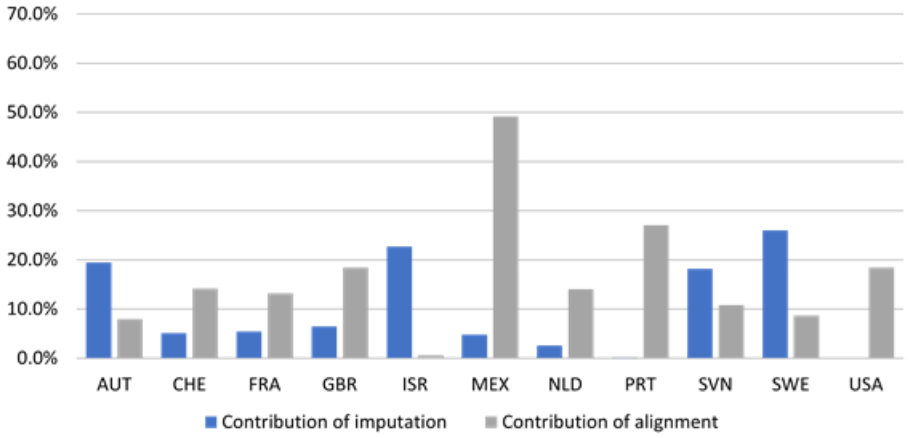
Share of absolute micro-macro gaps (alignment), imputations and micro data (derived as residual) in absolute amounts that constitute actual final consumption expenditure, EG DNA exercise 2015



Source: The Author.

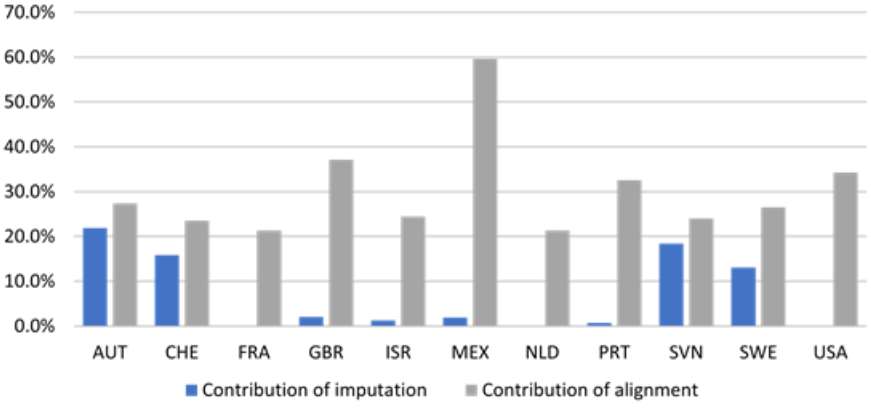
The size of the alignments appears to be significant in all countries. When looking at the average for all countries, the impact of alignments was 26.5% for adjusted disposable income and 30.2% for actual consumption expenditure in the 2015 exercise. The method for allocating alignments to underlying individuals or households may significantly affect distributional results. Ideally, information is available for the correct allocation of the amounts to the relevant individuals or households, but often (part of) the allocation may need to be done on the basis of some assumptions. It will depend on the degree of information and on the robustness of the assumptions how this will affect distributional results and to what extent this may add to the margins of error surrounding the results (see also Chapter 12).

Figure 7.3. Size of the absolute alignments and imputations as percentage of adjusted disposable income for the household sector as a whole



Source: Zwijnenburg (2016[1]).

Figure 7.4. Size of the absolute alignments and imputations as percentage of actual consumption expenditure for the household sector as a whole



Source: Zwijnenburg (2016[1]).

7.3. Items that show largest gaps

The previous section showed that the alignment of gaps between micro and macro data has a large impact on the distributional results in most of the countries. This implies that for some items large gaps exist

between the micro aggregates and the national accounts totals. Table 7.2 shows the adjustment coefficients for the main income components on the basis of the exercise conducted in 2015. The adjustment coefficient shows by how much the micro results need to be adjusted to align them with the adjusted national accounts totals. It is calculated as the adjusted national accounts aggregate divided by the micro aggregate. The table shows the number of countries for which an adjustment coefficient could be calculated (i.e. micro data was available to compile the distributional results), the average value of the coefficient, the median value and the minimum and maximum values in the exercise. Table 7.3 presents results for the main expenditure components.

Table 7.2. Adjustment coefficient for the main income components

National accounts aggregate divided by the corresponding micro aggregate, EG DNA exercise 2015.

Code	Instrument	Number of countries	Average		Median		Minimum		Maximum	
			most recent year	second most recent year	most recent year	second most recent year	most recent year	second most recent year	most recent year	second most recent year
B2	Operating surplus	4	1.79	1.27	1.90	1.27	0.94	1.12	2.43	1.42
B3	Mixed income	4	2.20	1.79	2.00	1.79	1.30	1.67	3.50	1.91
D1R	Compensation of employees	3	1.19	…	1.20	…	1.16	…	1.20	…
D41R	Interest (not adjusted for FISIM), received	8	2.08	1.90	1.56	1.05	0.66	0.72	6.40	4.77
D42R	Distributed income of corporations	7	5.06	10.67	1.88	5.53	0.70	3.00	17.76	23.50
D41P	Interest (not adjusted for FISIM), paid	9	3.58	2.47	2.94	1.50	1.02	1.01	11.31	4.65
D5P	Current taxes on income and wealth	10	1.18	1.19	1.18	1.15	0.78	0.74	1.54	1.78
D61P	Net social contributions	2	1.23	2.01	1.23	2.01	1.19	1.28	1.27	2.73
D62R	Social benefits other than STiK	10	1.22	1.30	1.15	1.26	0.97	0.98	1.55	1.65
D63R1	Education	3	0.94	0.88	0.95	0.88	0.72	0.78	1.13	0.98
D63R2	Health	3	1.36	1.37	1.18	1.37	1.16	0.99	1.73	1.75

Source: Zwijnenburg (2016[1]).

It can be observed that the gaps between the micro and macro data are often quite substantial. An adjustment coefficient that is close to 1 implies good alignment, but the tables show that the average values substantially differ from 1 for most of the income and consumption components. For the majority of the items the coefficient is above 1, meaning that the micro aggregates are lower than the macro aggregates. Only in a few cases, it is the other way around.

In the 2015 exercise, *distributed income of corporations* (D42R) turned out to have the highest adjustment coefficient on average (5.06), followed by *alcoholic beverages, tobacco and narcotics* (CP020) (3.60), *interest paid* (D41P) (3.58), *mixed income* (B3R) (2.69), *health* (CP060) (2.47), and *interest received* (D41R) (2.08). *Distributed income of corporations* (D42R) also records the highest maximum values. The consumption components generally have smaller differences across components than income. The average coefficients are between 1.09 and 2.47 in the most recent year, and between 0.92 and 2.72 in the second most recent year, when excluding the item *alcoholic beverages, tobacco and narcotics* (CP020).

Table 7.3. Adjustment coefficient for the main expenditure components

National accounts aggregate divided by the corresponding micro aggregate, EG DNA exercise 2015.

Code	Instrument	Number of countries	Average		Median		Minimum		Maximum	
			most recent year	second most recent year	most recent year	second most recent year	most recent year	second most recent year	most recent year	second most recent year
CP010	Food and non-alcoholic beverages	10	1.48	1.53	1.34	1.40	1.06	0.95	2.87	2.76
CP020	Alcoholic beverages, tobacco and narcotics	9	3.60	5.37	2.51	2.52	1.68	1.13	12.00	21.03
CP030	Clothing and footwear	10	1.57	1.70	1.25	1.40	1.09	1.03	2.90	2.80
CP040	Housing, water, electricity, gas and other fuels	9	1.23	1.16	1.06	0.94	0.84	0.87	2.47	2.30
CP050	Furnishings, households equipment & house maint.	10	1.60	1.71	1.41	1.59	1.15	0.96	2.93	2.93
CP060	Health	9	2.47	2.72	2.16	2.27	1.22	1.15	4.78	4.74
CP070	Transport	8	1.56	1.59	1.34	1.36	0.98	0.95	3.18	2.87
CP080	Communications	10	1.25	1.53	1.26	1.34	0.71	1.08	2.28	2.50
CP090	Recreation and culture	10	1.90	1.85	1.45	1.65	1.14	1.01	4.05	3.50
CP100	Education	10	1.09	0.92	1.08	1.05	0.19	0.09	1.87	1.51
CP110	Restaurants and hotels	10	1.54	1.32	1.52	1.29	0.97	1.06	2.20	1.64
CP120	Miscellaneous goods and services	7	1.89	1.88	2.06	1.78	0.97	1.13	2.63	2.85

Source: Zwijnenburg (2016[1]).

7.4. Possible reasons for micro-macro gaps

Possible reasons for the differences between the micro results and the adjusted national accounts totals are related to the first three steps of the step-by-step approach presented in Chapter 3, related to the quality of the data and of the assumptions used in the process:

- Step 1: Adjustment of the national accounts totals:
  - The quality of the national accounts totals
  - The quality of the adjustments to the national account totals
- Step 2: Linking micro data source variables to the national accounts variables:
  - Assumptions regarding the conceptual and classification differences
- Step 3: Imputation for missing elements and aligning data to national accounts totals
  - The quality of the correction for missing elements
  - The quality of the micro data – Estimation and measurement errors

The reasons for the gaps are discussed below, in accordance with the above categorisation.

7.4.1. The quality of the national accounts totals

A first possible reason for the gap between the micro and macro results may be quality issues related to the national account totals. The national accounts totals are the product of a balancing framework in which data from various data sources are combined and confronted. Often source data need to be adjusted to arrive at consistency and comprehensiveness. In that process, choices have to be made that may cause differences from the direct data sources. The quality of the data that are used in the system and the strength of the assumptions made in the balancing process will determine the quality of the final results. Gaps between micro and macro data may point to possible quality issues in this process.

In compiling the national accounts, the data for the household sector may be derived in three ways, i.e. directly based on household micro data sources (i.e. independent of other sectors in the accounts), estimated using counterpart information (from other sectors, e.g. banking statistics, pension data or government statistics), or as a residual after combining all other data sources in the system of national accounts. The robustness of the results for the household sector will depend on the quality of the various data sources used in constructing the national accounts and the amount of detail they provide. Furthermore, it will depend on how good these data sources align.

In analysing the gaps between the micro and macro results, it is important to have a more detailed look at how the national accounts results have been derived and whether this may contain any inconsistencies. Process table information that describes the various steps to get from the basic information to the final national accounts totals may be very relevant for that purpose. This may include information on adjustments made to correct for conceptual differences, to impute for missing elements, and to reach internal consistency within the framework of national accounts.

7.4.2. The adjustment of the national accounts totals

In the compilation process to arrive at distributional results, national accounts totals may have to be adjusted to exclude NPISHs, institutional households and consumption expenditure by non-residents, and to include expenditure of residents abroad at the detailed level of consumption items if this is deemed to lead to better matching of the micro and macro data. In some cases, specific information will be available to make these adjustments, but in other cases, these adjustments will have to be based on assumptions. Gaps between micro and macro results may be due to quality issues in making these adjustments. Therefore, it is also important to be transparent about the specific adjustments made in this step and in case of micro-macro gaps to discuss whether these may be partly due to incorrect adjustments or underlying assumptions in this step. For more information on this specific step, please refer to Chapter 4.

7.4.3. Conceptual differences and classification issues

Gaps may also appear as a consequence of conceptual differences and classification issues between micro and macro data. Sometimes the definition of the national accounts may vary from the one used in the survey or administrative data source, and (part of the) transactions may be classified differently. Chapters 10 and 11 provide more insight into possible conceptual differences between micro and macro results for various items.

Also, the time of recording may differ between the national accounts totals and the micro results. The latter often focus on a certain point in time (e.g. end of the quarter or end of the year), whereas the national accounts aim to capture all transactions within a certain time frame. This may give rise to differences between the micro and macro results, for instance related to changes in the population or for specific economic events that may have occurred during the period.

Moreover, data may refer to different time periods, for example if a specific survey is only conducted every other year, the information of a previous year may be used to derive results for a more recent year. This may also cause gaps between the micro and macro results.

It is important to carefully assess the reliability of any adjustments that may have been made to the micro data to adjust for any conceptual and classification differences when assessing possible reasons for micro-macro gaps.

7.4.4. Correction for missing elements

For some (sub)items or parts of the population, information may be lacking from micro data sources. In those cases, it is expected that compilers come up with imputations to correct for the missing information (see also Chapter 6). Micro-macro gaps may point to the fact that the missing part is actually smaller or larger than initially assessed. Furthermore, it may point to additional elements that may be missing from the micro data. In that regard, in case of a gap between the micro and the macro data, it is important to assess whether this can indeed be related to missing information in the micro data and, in case an imputation has already been made, whether the imputation is deemed to be correct or whether part of the remaining micro-macro gap may still relate to the need for additional imputations.

7.4.5. The quality of the micro data

Just as the macro aggregates may turn out to be incorrect, micro estimates may also be subject to quality issues. This may be increasingly the case, as many statistical offices struggle with the quality of their household surveys due to increased unwillingness to take part in surveys, to respond to specific questions, as well as increasing inaccuracy in filling out the surveys (see for example Meyer et al. (2015[2]) and Pinkovskiy et al. (2014[3])). The increased use of administrative data may partly overcome this issue, although it has to be borne in mind that these data sources come with their own downsides, not always providing matching concepts, and not always having full coverage of all parts of the population.

In general, micro data can suffer from two types of errors, i.e. estimation errors and measurement errors. Estimation errors relate to the extrapolation of the micro results to the target population and can be linked to the sample size, the representativeness of the sample and the magnitude of the non-response. The errors related to the sample size are referred to as standard sampling error, implying that the smaller the sample, the larger the variance surrounding the results, as less data underlie the ultimate estimates. The other two issues are referred to as coverage errors. These occur in the case of the sampling frame being different from or non-representative of the target population, and in the case of selective non-response, both causing bias to the results. As discussed in Chapter 5 all these aspects may cause gaps with the macro results. Especially survey data may suffer from estimation errors. Administrative data sources tend to have broad coverage and are therefore less prone to these kinds of errors. In this regard, Törmälehto (2017[4]) reports a striking example of France that changed from interview to register-based incomes in their EU-SILC results for 2008 which led to “a conspicuous increase in the share of property income: for the top 5 per cent, it jumped from 7.1 to 32.6% from 2006 to 2007”, possibly indicating the existence of estimation errors in the survey results.

In analysing the gaps between the micro data and the adjusted national accounts totals as derived after the first step in the compilation process, it is important to assess the likeliness of the gaps being influenced by estimation errors. For that purpose, it would be helpful to have information on the survey results in terms of underlying micro data, survey weights and standard errors. With regard to estimation errors, this may provide more insight into the margins of error surrounding the results. Furthermore, as this type of error will most likely affect more target variables at the same time, similar micro-macro gaps across multiple items may indeed be an indication of the existence of estimation errors.

Errors may also occur when the recorded values depart from the actual true values. These are referred to as measurement errors and may relate to item non-response or the reporting of incorrect data. These may be due to misinterpretation of the questions, difficulty by respondents to re-call the exact values, and deliberately reporting incorrect data. As was explained in Chapter 5, a lot of statistics have to deal with measurement error and these kinds of errors seemed to have increased over time, at least for specific items. Especially questions on income are usually understood to be relatively sensitive and prone to higher non-response rates or larger measurement errors, both in survey and administrative data. Particularly *income from self-employment* (see Johns and Slemrod (2008[5]) and Neri and Zizza (2010[6])), *property income* (see Neri and Zizza (2010[6])) and *social benefits* (see Meyer et al. (2009[7])) are prone to underreporting. The same goes for specific consumption items, such as illegal goods and services (e.g. *illegal drugs* and *prostitution*) or for socially unacceptable goods or services (e.g. *alcohol* and *gambling*).

As measurement errors may also be responsible for gaps between the micro and macro aggregates, it is important to assess to what extent the micro data may have been liable to these kinds of errors. This may be done by looking at the consistency and plausibility of the micro results, for example by confronting information on income, consumption and wealth at the micro level, checking the information with data from other data sources, looking at consistency of the data over time, and comparing data with results for comparable households. It is important to do this in close collaboration with the micro experts responsible for the specific statistics as they will have the best knowledge of the underlying data. If any errors are detected, it is recommended to correct for these in the micro data, so that an updated micro data set can be used as new input in the compilation process of the distributional results.

7.5. Allocation of gaps to relevant households

After the gaps have been attributed to the most likely causes, the related estimates have to be allocated to the relevant households or household groups. As the allocation may differ per cause, the allocation on the basis of these underlying causes will lead to more accurate results than allocating the full gap in one go.[3] For all causes that concern the micro data underlying the distributional results, specific solutions will have to be found.[4] This implies addressing the issue of possible measurement errors and of possible estimations errors.

When looking at estimation errors, research has shown that non-response is often correlated to specific household characteristics. D'Alessio and Faiella (2002[8]) for example show that it is often more frequent among higher income and wealthier households.[5] This is confirmed by Sabelhaus et al. (2013[9]) who analysed that high income households are underrepresented in the consumer expenditure survey in the US as they are less likely to participate. Pareto-analyses may be helpful in analysing whether the top tail of the distribution is covered in the micro data and to correct for this if needed (see Grilli et al. (2022[10])). Another option is to impute for the missing information by looking at administrative data, as done by the US Bureau of Labor Statistics in applying non-interview adjustment factors to the results of the consumer expenditure survey based on fiscal data (see U.S. Bureau of Labor Statistics (2022[11])). On the other hand, D'Alessio and Neri (2015[12]) found that in the Survey of Household Income and Wealth (SHIW) non-reporting was more frequent among the poorer and less educated. If that is the case, specific adjustments may be needed to correct for that.

Regarding the issue of measurement errors, it may in some cases be straightforward which micro data to adjust (for example in case of confrontation with data from other micro data sources)[6], but in other cases, this may require specific assumptions. In that case, the analysis of the plausibility of the underlying micro results may for example provide some direction where to best allocate specific amounts in relation to the various causes for micro-macro gaps. This can for example be done on the basis of constructing full sets of accounts at the micro level, i.e. comparing information on income, consumption and wealth. As saving

derived from the non-financial accounts should match the saving from the financial accounts, it may be checked whether there are inconsistencies or implausibilities at the micro level when comparing data from various sources.

In the case when not all information is available at the micro level one can try to derive results for groups of households and check the consistency and plausibility at the most detailed level possible.[7] Looking at outliers and the distribution of the data within the household group may also be helpful for this purpose. Furthermore, it may be useful to look at the development of micro-macro gaps over time, also in relation to trends in micro and meso results for specific groups of households. This may also reveal insight into less plausible trends in some of the elements.

Furthermore, information from research may provide insight into what type of households are most likely involved in specific types of activities or affected by specific types of errors which may provide the underlying rationale for adjustment of these records in this specific step of the process. In relation to the non-observed economy, Coli and Tartamella (2014[13]), for example, show that non-registered workers are not equally distributed across the household sector, but show to be concentrated in specific subgroups. Furthermore, Accardo et al. (2009[14]) made specific adjustments for income "from fraud and undeclared work", which mainly affected "self-employed, the most well-off senior managers, salaried workers in the first half of the income distribution and non-active persons, excluding retired people". Carson (1984[15]) provides information on which household types are more likely to be involved in the underground economy or illegal activities.

A lot of research is also available on the impact of measurement errors in various statistics. Several studies confirm that misreporting often depends on socio-demographic characteristics (such as age, family type, ethnicity, income level, region and education), i.e. some groups are more likely to misreport for some items than others. Neri and Zizza (2010[6]) found that in the Survey of Household Income and Wealth (SHIW) misreporting tends to be more diffuse among males, the older, the self-employed and respondents at the higher end of the earnings distribution. They also found some regional differences in the likelihood of misreporting. Sabelhaus et al. (2013[9]) and Cifaldi and Neri (2013[16]) show that there is large underreporting at the top of the distribution. They did not find similar evidence for consumption, explaining that consumption is a less sensitive topic and more difficult to hide from an interviewer. Lohmann (2010[17]) and Romanov and Gubman (2012[18]) explain that there is also evidence that part-time and irregular employees are more likely to incorrectly report their earnings (e.g. reporting income levels for a full month that may not be representative of their average income). Furthermore, whereas some groups tend to underreport their income, it has also been the case that some other socio-demographic groups tend to over-report their income. For example, Bound and Krueger (1989[19]) found that women have a slight tendency to underreport their earnings.[8] In that regard, even if the micro-macro gap is zero, there may be the need to adjust some of the underlying data at the household level.

It is clear that the allocation of the micro-macro gaps to the relevant households on the basis of the most likely underlying causes will often involve subjective decisions. The examples provided above may provide some insights into how to approach the allocation question and what groups may be more prone to what specific types of measurement errors, but it will depend on the items, the data sources and the country characteristics which approach will work best to solve and allocate any micro-macro gap for a particular item at the country level. Furthermore, it is important that any decisions on how to allocate the gap is done in close cooperation between the micro and the macro experts, as they both have specific knowledge of the underlying micro and macro data which is relevant to come up with the best possible solution.

The allocation of the amounts to the underlying households should ideally be done at the level of the micro statistics, i.e. by making adjustments to the survey or administrative data, applying imputations at the micro level, or by adjusting the survey weights to arrive at the relevant aggregates. This will lead to improved micro data that underlie the new distributional measures and will make sure that the income group classification is re-adjusted on the basis of these improved data. An alternative is to allocate the amounts

at the aggregated level. In that case, the quintile (or other household group) allocation on the basis of the "unadjusted" micro data is taken as starting point and the amounts that have been attributed to the various causes are allocated to the quintiles. It is clear that the distributional results on the basis of this aggregated approach will not be as accurate as in the case of processing the corrections at the micro level, but in the end may lead to better results than simply applying a proportional allocation.

7.6. A framework to allocate the micro-macro gaps

To assist compilers in discussing possible reasons for the gaps and to allocate them to the relevant households (or household groups), a framework has been developed on the basis of the reasons expressed in the previous section. This framework consists of two parts. The first part focuses on assigning parts of the gap to possible underlying causes. This part is presented in Table 7.4. The first block (block I) in the table focuses on the derivation of the adjusted national accounts estimate for a specific item, starting from the national accounts total and adjusting for NPISHs, institutional households and expenditures of non-resident households on the territory respectively. The first column in this block shows the original estimates that were used to derive the adjusted national accounts figure. The second column provides the possibility to correct any of these original figures to close part of the gap between the micro and the macro results. The final result is presented in the third column.

The second block of the framework (block II) confronts the adjusted national accounts result with the micro aggregate, showing the gap between the two. The initial macro-micro gap is presented in the first column. The third column shows the gap that still remains after corrections have been made to the adjusted national accounts aggregate. This remaining gap still needs to be attributed to other reasons. This is done in block III. This block lists possible causes related to conceptual or classification differences, missing items or errors with regard to the micro data. In addition to the reasons presented in the previous section, it also contains an item for reasons that are not covered by the other categories. The block ends with the gap that still remains after attributing parts of the gap to the underlying reasons. Ideally, the amount of this remaining gap is zero, which would imply that the complete gap is explained by the various causes.

After the attribution of the macro-micro gap to the underlying causes, the related amounts should be allocated to the relevant household groups. Table 7.5 presents a framework for this step. Block IV focuses on the allocation on the basis of revised micro data, which, as was explained in the previous section, is the preferred option. In that case, corrections are processed at the micro level and new results are derived following the standard step-by-step approach. However, in some cases, this may be deemed too time-consuming or too complex. In those cases, corrections may be allocated at an aggregated level.[9] This can be done in block V which provides the opportunity to allocate the remaining gaps at the quintile level (or other household groups depending on the targeted breakdown). Finally, block VI deals with allocating the remaining gap that could not be linked to any of the possible causes. The sum of the corrected micro data and the consecutive meso-corrections leads to the distributional results for the quintiles (or other household groups).

Results from two studies in which EG DNA members applied the framework to the five items that appear to be most relevant for their country showed that the allocation across quintiles indeed differs across the various reasons and that in most cases they differ from the distributions according to the micro data. The differences turned out to be particularly large for "measurement errors" and "underground activities". The latter seemingly relates to the non-inclusion of underground economy in initial estimates and shows the importance of a separate estimation of these transactions. Furthermore, the results showed that estimation errors may significantly alter the distribution across households for specific items (e.g. *food and non-alcoholic beverages*; *alcoholic beverages, tobacco and narcotics*; *transport*; *restaurants and hotels*; and *miscellaneous goods and services*).[10]

Table 7.4. Framework for attributing micro-macro gaps to underlying causes

	Item xx.	Original estimate	Correction	Ultimate Estimate
I	National account total (A)	…	…	…
	- Adjustment for NPISHs (B1)	…	…	…
	- Adjustment for institutional households (B2)	…	…	…
	- Adjustment for expenditures of non-resident households on the territory (B3)	…	…	…
	+ Adjustment for expenditure of resident households abroad (B4)	…	…	…
	= Adjusted NA total (C=A-B1-B2-B3+B4)	…	…	…
II	Micro total (D)	…		…
	= Macro-Micro gap (E=C-D)	…	…	…
III	Conceptual or classification issues (F)	…	…	…
	Underground and illegal activities (G)	…	…	…
	Other elements missing in micro data (H)	…	…	…
	Estimation errors (under-/overcoverage) (I)			…
	Measurement errors (under-/overreporting) (J)			…
	Reasons n.e.c. (K)			…
	= Remaining gap (L=E-F-G-H-I-J-K)			…

Source: Zwijnenburg (2016[1]).

Table 7.5. Framework for allocating gaps to household groups

	Item xx.	Estimate	Q1	Q2	Q3	Q4	Q5
IV	Allocation on the basis of micro data						
	Original micro aggregate (P)	…	…	…	…	…	…
	Revised micro aggregate (Q)	…	…	…	…	…	…
V	Allocation on the basis of meso corrections						
	Conceptual or classification issues (R)	…	…	…	…	…	…
	Underground and illegal activities (S)	…	…	…	…	…	…
	Other elements missing in micro data (T)	…	…	…	…	…	…
	Estimation errors (under-/overcoverage) (U)	…	…	…	…	…	…
	Measurement errors (under-/overreporting) (V)	…	…	…	…	…	…
	Reasons n.e.c. (W)	…	…	…	…	…	…
VI	Alignment of remaining gap (X=C-Q-R-S-T-U-V-W)	…	…	…	…	…	…
	Final estimate (Y=Q+R+S+T+U+V+W+X)	…	…	…	…	…	…

Source: Zwijnenburg (2016[1]).

It is recommended that micro and macro experts regularly discuss the gaps between micro and macro results, particularly for the items that show the largest gaps, to find the most likely underlying reason(s), possibly reduce the gaps, and decide to which households these gaps most likely relate. Regular discussions will add to the awareness of these gaps and exchange of expertise may provide useful insights in how to deal with them. This will not only be relevant for projects in which micro and macro results are combined but would also be beneficial to properly explain to users why the results of micro and macro statistics on similar subjects may deviate.

References

Accardo, J. et al. (2009), "Inequality between households in the national accounts", No. 1265, INSEE, https://www.insee.fr/en/statistiques/1280710 (accessed on 21 September 2017). [14]

Bound, J. and A. Krueger (1989), "The extent of measurement error in longitudinal earnings data: do two wrongs make a right?", *NBER Working Paper Series*, No. 2885, National Bureau of Economic Research, http://www.nber.org/papers/w2885.pdf (accessed on 30 October 2017). [19]

Carson, C. (1984), "The underground economy: An introduction", *Survey of current business*, Vol. 64/5, pp. 21-37, https://bea.gov/scb/pdf/1984/0584cont.pdf (accessed on 9 January 2018). [15]

Cifaldi, G. and A. Neri (2013), "Asking income and consumption questions in the same survey: what are the risks?", No. 908, Banca d'Italia, https://www.bancaditalia.it/pubblicazioni/temi-discussione/2013/2013-0908/en_tema_908.pdf?language_id=1 (accessed on 11 October 2017). [16]

Coli, A. and F. Tartamella (2014), "Using Administrative and Survey Data to Analyse Tax Evasion from Unregistered Labour", http://www.iariw.org/papers/2014/ColiPaper.pdf (accessed on 24 October 2017). [13]

D'Alessio, G. and I. Faiella (2002), "Non-response behaviour in the bank of Italy's survey of household income and wealth", No. 462, Banca d'Italia, https://www.bancaditalia.it/pubblicazioni/temi-discussione/2002/2002-0462/tema_462_02.pdf?language_id=1 (accessed on 27 October 2017). [8]

D'Alessio, G. and A. Neri (2015), "Income and wealth sample estimates consistent with macro aggregates: some experiments", No. 272, Banca d'Italia, https://www.bancaditalia.it/pubblicazioni/qef/2015-0272/QEF_272.pdf?language_id=1 (accessed on 27 October 2017). [12]

Gottschalk, P. and M. Huynh (2007), "Are earnings inequality and mobility overstated? The impact of non-classical measurement error", http://www.sem.tsinghua.edu.cn/semcms_com_www/upload/home/store/2008/5/12/2767.pdf (accessed on 30 October 2017). [20]

Grilli, J., P. Engelbrecht and J. Zwijnenburg (2022), *Pareto tail estimation in the presence of missing rich in compiling distributional national accounts*. [10]

Johns, A. and J. Slemrod (2008), "The Distribution of Income Tax Noncompliance", http://www.bus.umich.edu/OTPR/DITN%20091308.pdf (accessed on 30 October 2017). [5]

Lohmann, H. (2010), "Comparability of EU-SILC survey and register data: The relationship among employment, earnings, and poverty", http://www.iariw.org/papers/2010/6bLohmann.pdf (accessed on 20 October 2017). [17]

Meyer, B., W. Mok and J. Sullivan (2015), "Household Surveys in Crisis", *NBER Working Paper Series*, No. 21399, http://www.nber.org/papers/w21399 (accessed on 24 October 2017). [2]

Meyer, B., W. Mok and J. Sullivan (2009), "The under-reporting of transfers in household surveys: Its nature and consequences", *NBER Working Paper Series*, No. 15181, http://www.nber.org/papers/w15181 (accessed on 27 October 2017). [7]

Neri, A. and R. Zizza (2010), "Income reporting behaviour in sample surveys", No. 777, Banca d'Italia, https://www.bancaditalia.it/pubblicazioni/temi-discussione/2010/2010-0777/en_tema_777.pdf?language_id=1 (accessed on 11 October 2017). [6]

Pinkovskiy, M. et al. (2014), *Lights, Camera,…Income! Estimating Poverty Using National Accounts, Survey Means, and Lights*, Federal Reserve Bank of New York, https://www.newyorkfed.org/medialibrary/media/research/staff_reports/sr669.pdf (accessed on 26 September 2017). [3]

Romanov, D. and Y. Gubman (2012), "Well-Being and Measurement Error of Income Reported in a Social Survey versus Income Recorded by a Tax Administration", http://www.iariw.org/papers/2012/RomanovPaper.pdf (accessed on 30 October 2017). [18]

Sabelhaus, J. et al. (2013), "Is the consumer expenditure survey representative by income?", *NBER Working Paper Series*, No. 19589, http://www.nber.org/papers/w19589 (accessed on 27 October 2017). [9]

Törmälehto, V. (2017), "High income and affluence: Evidence from the European Union statistics on income and living conditions (EU-SILC)", *Eurostat Statistical Working papers*, http://ec.europa.eu/eurostat/documents/3888793/7882117/KS-TC-16-027-EN-N.pdf (accessed on 9 October 2017). [4]

U.S. Bureau of Labor Statistics (2022), *BLS Handbook of Methods: Consumer expenditures and income*, https://www.bls.gov/opub/hom/cex/calculation.htm#calculation-methodology (accessed on 12 September 2023). [11]

Zwijnenburg, J. (2016), "Further Enhancing The Work On Household Distributional Data – Techniques For Bridging Gaps Between Micro And Macro Results And Nowcasting Methodologies For Compiling More Timely Results", OECD, http://www.iariw.org/dresden/zwijnenburg.pdf (accessed on 21 September 2017). [1]

Notes

[1] This relates to the micro data for the items for which Method A (as described in Section 6.2) has been applied.

[2] This relates to the items for which Method B, C or D (as described in Section 6.2) has been applied.

[3] An alternative is to apply a proportional allocation of the gap, i.e. simply multiplying all micro data by the same factor to arrive at the macro aggregates. This would assume that all households misreport to the same degree. Whereas this may be a valid assumption if no other information is available (see also Section 12.3.1), it should only be applied as a last resort, i.e. after trying to allocate the majority of the gap on the basis of the most likely underlying reasons and to the most likely households concerned.

[4] Corrections that relate to the adjusted national accounts totals will only affect the benchmark totals so only having an indirect impact on the distributional results.

[5] In the Survey of Household Income and Wealth (SHIW) conducted by the Bank of Italy, it was found that respondents that are persuaded to participate after an initial refusal have average income and wealth that is 20% and 30% higher than the overall average. This was confirmed by a study in which data for a sample of 2000 households were matched with banking information. This also showed that non-response was not random, but more frequent among the wealthiest households.

[6] See for example D'Alessio and Faiella (2002[8]) and D'Alessio and Neri (2015[12]) who have done research in which consistency of micro results within the same survey is checked.

[7] A good example of such a consistency check is the way in which the French statistical office checks the data. They ask for information on income, consumption and financial well-being in their Household Budget survey, on the basis of which it is possible to adjust incomes on the basis of a coherence filter between income and consumption. Accardo et al. (2009[14]) explain that "when households declared an income which was very much lower than their everyday consumption expenditure (defined as consumption excluding major or exceptional purchases), yet without indicating that they felt they were in any financial difficulty, their income was aligned with the level of their consumption expenditure."

[8] Furthermore, Gottschalk and Huynh (2007[20]) explain that measurement error may be mean reverting, "in the sense that persons with low earnings tend to overstate their earnings and persons with high earnings understate their earnings" (see also Lohmann (2010[17])).

[9] Bearing in mind that this is a sub-optimal solution as it does not take into account possible reclassification of households across household groups on the basis of corrected micro-data. Furthermore, it does not provide the possibility to take into account specific characteristics at the household level that may lead to more nuanced adjustments.

[10] See for more information, Zwijnenburg (2016[1]).

8 Linking or matching data across data sources

As multiple micro data sources may be used in the compilation process, linking data across these datasets in a proper way to arrive at coherent and consistent sets of accounts for underlying households is of crucial importance. This chapter describes four methods to achieve this objective, with their main pros and cons.

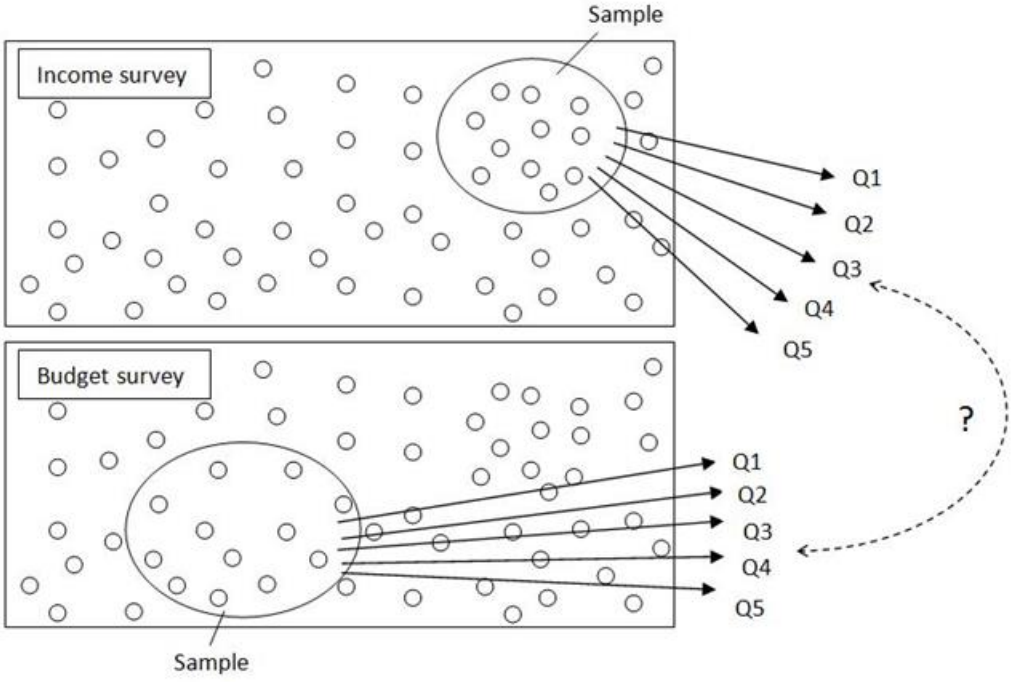
8.1. Introduction

A very important step in the compilation process is the linking or matching of data across various data sets to construct coherent data on income, consumption and saving for the various household groups. In many cases, data from different micro data sources are used, obtained via micro data surveys or administrative data, and the way in which they are combined may seriously impact the overall results.

Sometimes the various data sources may describe exactly the same households, in which case it will be easy to link the data, but in many cases, it will concern different samples of households. The question then arises how these data should be matched to create complete sets of accounts for similar types of households, and to arrive at coherent distributional results for income, consumption and saving for various household groups.

Figure 8.1 provides a simplified example of the issue, showing a country that uses two different sample surveys for its income and its consumption items. As it concerns sample surveys, different households may be selected in the samples. Furthermore, the samples may differ in size.

Figure 8.1. The issue of linking data across surveys



Source: The Author.

In order to arrive at reliable and consistent distributional results for income, consumption and saving across household groups (e.g. income quintiles as shown in Figure 8.1), income and consumption data from different micro data sets need to be matched in a coherent way. Generally, there are four methods to achieve this objective, the first two aiming to link or match data at the micro level[1] and the latter two processing results separately for each data source and only matching results at the aggregated level:

1. Link records on the basis of common household IDs or identifiers present in the data sources (record linking). Results can then be clustered on the basis of these matched micro data.
2. Merge data from different data sets into a single micro data set via statistical matching and modelling. This approach uses matching variables available in all data sets to impute missing

variable(s) available in one specific data set (the donor set) into the data set where this/these variable(s) is/are missing (the recipient set). Results can then be clustered on the basis of information from this "new" synthetic data set.

3. Construct household groupings for each data source separately in case the variable needed for clustering is available in all of them and link these results at the aggregated level.
4. Construct household groupings for each data source separately on the basis of an imputed variable in case this variable is not available in all data sources and link these results at the aggregated level. This is a variant of option c but relying on imputations instead of direct observations to match the data at the aggregated level.

These methods are explained in more detail in the following sections, explaining the basic technique as well as the main pros and cons of each of the methods.

8.2. Linking records on the basis of identifiers

In the first approach, records from different data sources are linked on the basis of unique identifiers that enable the direct linking of records across different data sources. This can for example be done on the basis of social security numbers, fiscal numbers or addresses. This option will often be available in case countries use administrative data as one of their main data sources. Data from these administrative data sources may then be linked to data available from surveys.[2]

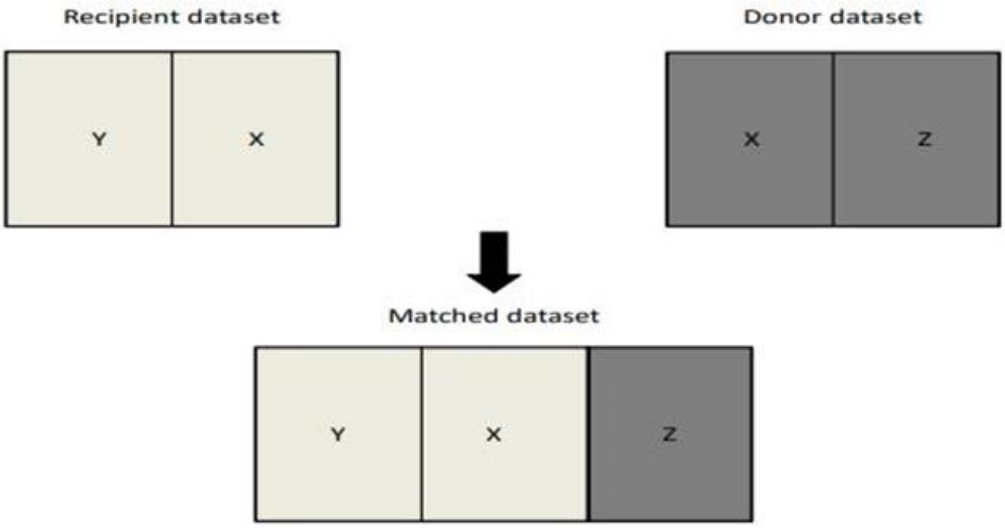
Linking on the basis of identifiers is the preferred method to link data across multiple data sources, as it ensures that the data on income, consumption and saving are fully consistent at the micro level, without the need to rely on any assumptions to link the data. This means that there is no margin of error feeding into the results as a result of the matching exercise (except in case of any errors in the identifiers themselves).

8.3. Integrating data sets through statistical matching

In the second approach, information from different data sources is fused on the basis of statistical matching. In this technique, a specific variable that is missing from one data set (the recipient data set) is imputed from another data set (the donor data set) by looking at common variables available in both. These may concern information on income (group), age, gender, marital status, region, household size, main source of income, occupation, type of labour contract, country of birth, education level, etc.

For example, a compiler may have data from two surveys, i.e. an income survey and a budget survey, in which disposable income (variable Y) is missing from the budget survey and consumption expenditure (variable Z) from the income survey. In order to obtain a data set that includes data on both income and consumption, both data sets can be fused with the help of common variables X available in both data sets. For that purpose, the relation between these common variables with the target variable need to be assessed on the basis of data from the donor data set, assessing the specific matching variables that will be used to conduct the matching. This relation can then be used to impute the target variable in the recipient data set. This will lead to a synthetic (or "matched" or "fused") file, [3] containing records that include both X, Y and Z (see Figure 8.2).

Figure 8.2. Integrating data sets through statistical matching



Source: Balestra and Oehler (2023[1]).

After the data fusion, results are processed according to the step-by-step approach and allocated to the relevant household groups on the basis of the underlying information in the synthetic data file.

A prerequisite for statistical matching is that the populations match across the various data sets and that the matching variables are identical in terms of concepts and reporting.[4] To the extent that this is not the case, adjustments will be needed to ensure a good alignment between the data sources.

The quality of the matching will largely depend on the selection of the matching variables. Balestra and Oehler (2023[1]) explain that these should meet two essential criteria. First of all, they should show homogeneous distributions across the relevant data sources, ensuring that the data sets cover similar types of households with coherent information on the distribution of the matching variables. Secondly, they should have a significant correlation to the target variable(s), i.e. they should behave as good predictor of the target variable(s) to be imputed in the recipient data set.

Ideally, the target variables (i.e. Y and Z) are independent of each other and the full relationship between the two is explained by the common variables (i.e. X). This is known as the conditional independence assumption. However, this assumption rarely holds and is difficult to test in practice (see Eurostat (2013[2])). Balestra and Oehler (2023[1]) explain that auxiliary information may help in increasing the likelihood of meeting this assumption. This may for example be in the form of having a proxy variable for Z in the recipient data set or a proxy for Y in the donor set (e.g. a reported income variable in the budget survey). Normally, the more detail that can be used in the matching, the more accurate the results. A more detailed description of statistical matching is available in Eurostat (2013[2]).

The main advantage of this approach is that households are fused at the micro level on the basis of common characteristics. This is expected to lead to relatively good matches (to the extent they meet the criteria as explained above) and provides the opportunity to assess the plausibility of the results at the micro level at the start of the process. If some records show implausible results for the combinations of income and consumption, edits may be performed before further processing the data. This could be done by correcting either income or consumption results, or by changing some of the characteristics that are at the basis of the matching. Editing the underlying micro data may be particularly relevant in case of large micro-macro gaps for specific items. Instead of applying a proportional allocation to close the gaps for the various items, one could edit those items at the micro level for which the gaps between the micro data and

the relevant national accounts data are most significant and for which the matching may show implausible results for some specific households.

The downside of this approach is that it requires assumptions for the matching which may lead to some degree of uncertainty surrounding the matched results. This will largely depend on the coherence of distributions of the common variables across the various data sets and of the explanatory power of the common variables to explain the target variable. As the statistical matching may not perfectly capture the full relationship between the common variables and the target variable, it is sub-optimal to direct matching on the basis of identifiers (Balestra and Oehler (2023[1])), but preferable over matching at the aggregated level only (as described in the next two sections).

8.4. Construct household groupings for each data source separately on the basis of a common variable

Data from different sources can also be processed and clustered into household groups independently, with matching only taking place at the aggregated level. In that case, the various steps will be processed separately for the various data sets and distributional results will only be linked in a final step. Figure 8.3 presents an example of how this works, showing a country that uses different sample surveys for its income and its consumption items, with different households included in both. The data are processed separately for the two surveys and then combined at an aggregated level on the basis of the targeted household groups.

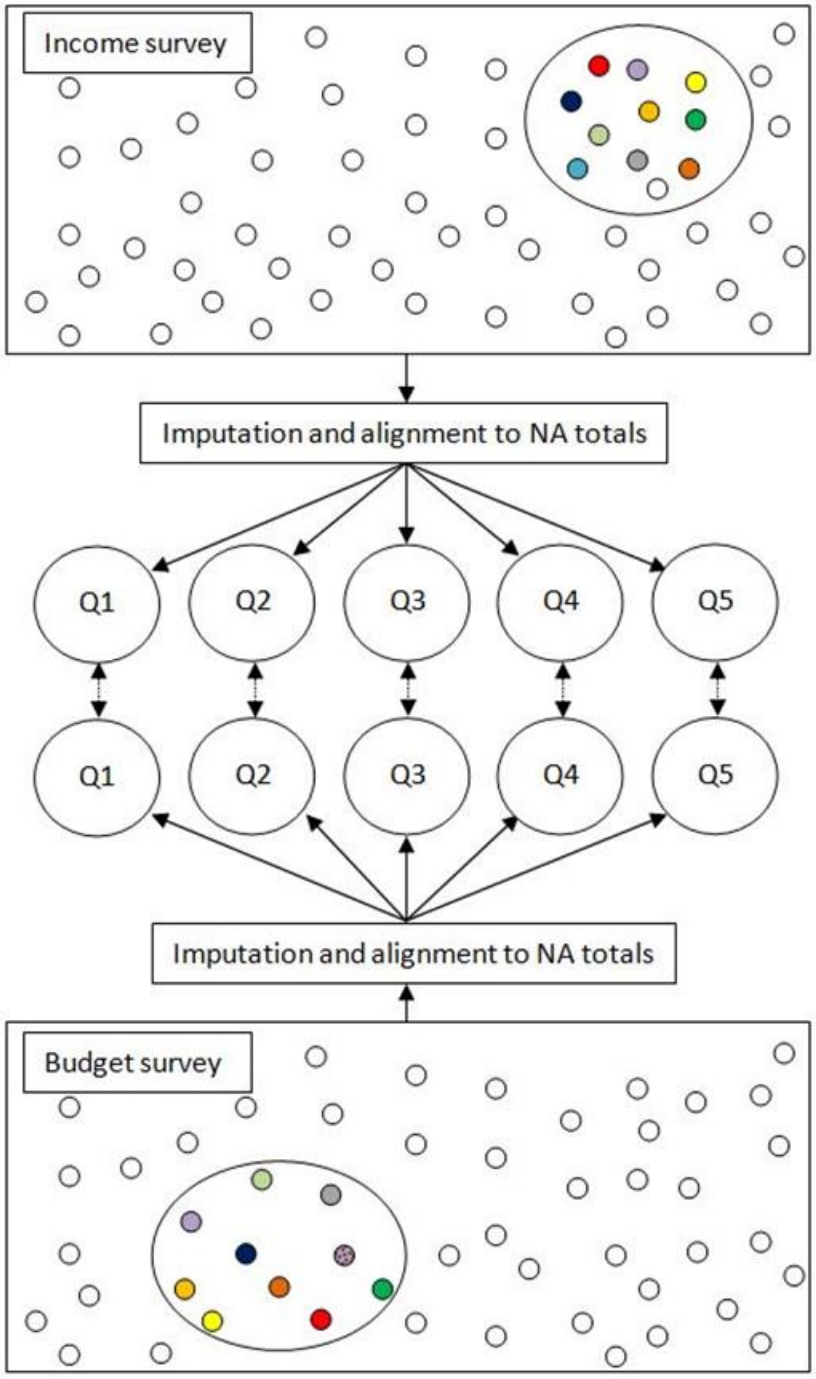
This approach can be applied when the variable necessary for clustering into household groups is available in all relevant data sets and is comparable in terms of concepts and reporting. If that is not the case, this would require specific adjustments and edits to ensure good alignment between the variables. Alternatively, one could opt to impute the relevant variable(s) in the missing data sets (see Section 8.5).

When applying this approach, it is important that the data sources describe the same population and that they show similar distributions for the common items, i.e. the same prerequisites for applying statistical matching. This will increase the likelihood of starting from similar data sets and it will help avoiding incoherent matching results. The latter may occur, for example, with regard to the level of income reported for each of the income groups, the number of households included in each of the household groups, and with regard to the socio-demographic characteristics of the various household groups.

With regard to the first issue, the income levels when clustering households according to their equivalised disposable income may be different when clustering results for each data set separately. It will be important to apply similar kinds of adjustments to income as reported in the various data sources, to ensure similar income concepts across data sources, bearing in mind that the specific adjustments needed may differ across data sources dependent on their concept and coverage of income items.[5] Ideally, compilers would then arrive at similar income levels, upper and lower bounds, and distributions for the various household groups across the different data sources. If this is not the case, compilers should investigate the main underlying reasons for any differences and try to make informed adjustments to bring the results closer in line.

When looking at clustering according to other characteristics than income, it is important to ensure consistency in the number of households for each household group across the data sources used.[6] For example, when clustering according to main source of income, a different number of households may end up in the group “income from self-employment” according to income survey results than according to budget survey data. In case of large differences in numbers of households for specific household groups, compilers should investigate the main reasons for these differences. These may for example relate to differences in concepts, differences in weights and/or differences in reported values. Dependent on the most likely issues, compilers need to make adjustments to ensure closer alignment between the results.

Figure 8.3. The issue of linking data at an aggregated level



Source: The Author.

Finally, differences may show up in the sociodemographic composition of various household groups. For example, the first income decile clustered on the basis of budget survey data may show a much larger number of single households and people below 25 that results according to the income survey. This may point to differences in (sample) populations and/or income definitions. The same may apply for other household groupings. If this occurs, compilers should investigate the main underlying reasons and make necessary adjustments to better align the results.

The benefit of this approach is that it is less complicated and probably less time-consuming than the first two approaches. The downside is that it may lead to less reliable results as the plausibility of the results can only be checked at an aggregated level and is normally only done on the basis of the variable relevant for the clustering. This is different when applying statistical matching, which relies on matching at the micro level and is done on the basis of multiple common variables, leading to closer matches and to one synthetic data set underlying the distributional results, ensuring consistency in income levels, number of households and socio-demographic characteristics for the various household groups across income and consumption.

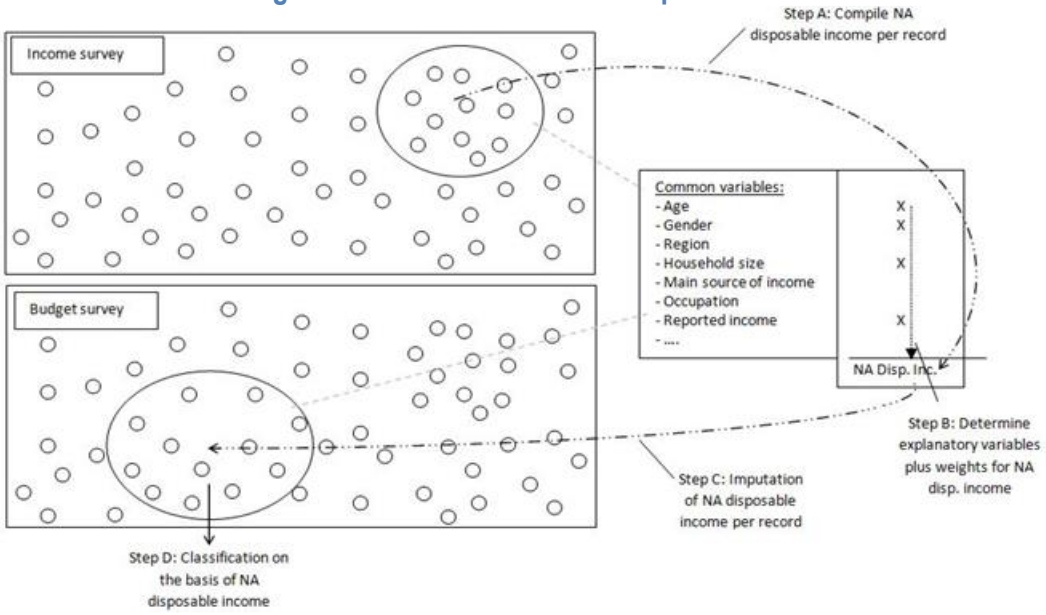
Furthermore, statistical matching enables analysing the plausibility of the results across income and consumption at the micro level. Plausibility checks are much more complex when linking at the aggregated level, as inconsistencies may be due to a larger number of factors. This also means that it is more difficult and may take more time to make accurate adjustments in case of any observed inconsistencies. In that regard, statistical matching is preferable over linking at the aggregated level.

8.5. Construct household groupings for each data source separately on the basis of an imputed variable

The last option to arrive at a coherent distribution of households across household groups is a variant of option 3, when the common variable to cluster households into household groups may not be available for all data sets. In that case, one may consider imputing the common variable in the relevant data sets and then cluster households accordingly for the various data sets separately.

For example, if disposable income is not available in all data sets, a disposable income variable could be imputed on the basis of common characteristics, via which households can be classified consistently into deciles. This method has some similarities to statistical matching in the sense that households are matched on the basis of similar characteristics, but it differs as in this option records are not matched at the micro level but at the aggregated level on the basis of an imputed disposable income item. As records are not fused individually, the various steps in the methodology can be processed independently, and at the final stage households can be classified on the basis of this imputed disposable income. Figure 8.4 presents a simplified example of how this technique works.

Figure 8.4. The issue of linking records on the basis of an imputed income variable



Source: The Author.

Looking at the specific steps, first, in the income survey, an income item has to be created according to national accounts definitions and in line with the national accounts totals (step A). This requires linking and aligning the relevant items from the micro survey to national accounts and imputing for any missing items. As a result, one arrives at an "*NA aligned disposable income*" per record. Subsequently, a regression analysis can be run on the basis of common variables in the various data sets to find explanatory variables (e.g. relating to households' characteristics, reported income and/or consumption, etc.) to explain these disposable income levels (step B). As these variables will be used to impute an "NA aligned disposable income" in all data sets, it is important to look at common characteristics available in all data sets. This may include "age", "gender", "marital status", "region", "household size", "main source of income", "occupation", "income", "type of labour contract", "country of birth", "level of education" etc. The regression analysis will lead to a model that can be used to assign NA aligned disposable income levels to micro records in the other micro data sets (step C).

In the final step, households in the other data sets can be classified into income groups on the basis of these imputed income levels (step D). The latter may be done using income boundaries defined on the basis of the imputed income results in the respective data sets or boundaries determined on the basis of the income part of the work. In the former case, one can make sure that the ten deciles consist of 10% of the households according to the results of the specific data set. However, income levels may deviate from the ones used for the classification of households in the income part, also implying that households with similar characteristics would not necessarily end up in the same deciles across all data sets. In the second option, the boundaries will match those used for allocating households in the income part (probably leading to a better match between income and consumption results), but as this may lead to different numbers of households per decile for the consumption part, this may require adjustment of weights for the underlying micro data. In adjusting the weights, one has to make sure that all deciles consist of 10% of the households and that the sum of the deciles still adds up to the national accounts totals.

As was the case with statistical matching, the approach depends on the coherence of information as reported for the common variables in the relevant data sets (i.e. homogeneous distributions) and that the common variables provide a good predictor of the target variable. Furthermore, it is important that the common variables are identical in terms of concepts and reporting, possibly requiring specific adjustments and/or edits if this is not the case.

This approach generally has the same benefits and downsides as the previous approach. The main additional benefit in view of clustering according to income groups, is that it ensures that the income concept for clustering households is consistent across different data sets and that – as it relies on multiple common variables to derive this variable - similar types of households will be assigned similar types of income. However, it may not be straightforward to run the regression analysis. Furthermore, given the downsides of only linking at the aggregated level (as explained in Section 8.4), the approach is still sub-optimal in comparison to linking data at the micro level as is done in the first two approaches.

References

Balestra, C. and F. Oehler (2023), "Measuring the joint distribution of household income, consumption and wealth at the micro level", *OECD Papers on Well-being and Inequalities*, No. 11, OECD Publishing, Paris, https://doi.org/10.1787/f9d85db6-en. [1]

Eurostat (2013), "Statistical matching: a model based approach for data integration", https://ec.europa.eu/eurostat/documents/3888793/5855821/KS-RA-13-020-EN.PDF.pdf/477dd541-92ee-4259-95d4-1c42fcf2ef34?t=1414780333000 (accessed on 30 August 2023). [2]

Notes

[1] These methods are discussed in detail in Balestra and Oehler (2023[1]).

[2] Please note that this may not always be possible (even if micro data may include unique identifiers) due to legal constraints (e.g. in view of general data protection regulations).

[3] As explained in Balestra and Oehler (2023[1]), the term "synthetic" is used as not all data for each household as included in the resulting data set have been directly observed but may have been obtained by combining information from different data sources.

[4] For example, if common variables in a specific survey are deemed more liable to reporting errors than in others, this may affect the quality of the matching. This may be particularly relevant in combing survey data with administrative data. In order to avoid incorrect matching results, it is important to first edit the micro data in the various data sets before applying the statistical matching.

[5] Compilers should avoid taking reported income from other surveys as a direct proxy for disposable income as defined in the System of National Accounts. Imputation for missing items and alignment to national accounts totals is often affecting different types of households in different ways, affecting income levels in different ways and altering the ranking. For that reason, using the reported income as a direct proxy will likely lead to incorrect matching.

[6] This will not be an issue for grouping by income as the relevant groupings are defined by number of households, e.g. 10% of households in each decile when breaking down by equivalised disposable income deciles.

9 Clustering households into household groups

In the fourth step, households can be clustered into household groups. This chapter describes how this can be done for the main household groupings targeted in the work.

9.1. Introduction

In the fourth step, households can be clustered into household groups. This may be done on the basis of equivalised disposable income, but the clustering can also be based on alternative characteristics such as main source of income, household type or age of the head of the household. It will depend on the available information and on the quality of the distributional results what household groups can be targeted.

This chapter describes how distributional results can be derived for the various household groups. For breakdowns based on socio-demographic characteristics, this will be relatively straightforward, by simply allocating households on the basis of their underlying characteristics. This is explained in Section 9.4. However, for classification according to standard of living (i.e. equivalised disposable income) or main source of income, more guidance may be needed. Section 9.2 describes the classification according to standard of living and Section 9.3 according to main source of income.

9.2. Clustering according to standard of living

In the classification according to standard of living, households are clustered on the basis of their equivalised disposable income. For this purpose, household disposable income in line with national accounts totals (i.e. after imputation for missing elements and alignment of the micro data to the macro results) is recalculated into per consumption unit results, taking into account the size and composition of the household, to arrive at comparable results across households. This is known as *equivalised disposable income*. The Handbook uses the OECD-modified equivalence scale as reference method, but compilers may also decide to apply a different scale if this is deemed more appropriate in relation to country specific circumstances. The OECD-modified scale assigns a value of 1 consumption unit to the first adult in the household, a value of 0.5 for each additional person aged 14 and over, and 0.3 for all children under 14 (see also Box 2.1).

After disposable income for each household has been divided by its number of consumption units, households can be ranked on the basis of this equivalised disposable income. On the basis of this ranking, households can then be clustered into income groups, for example into income deciles. In that case, the clustering should be done in such a way that each decile represents 10% of the households. Hence, each decile represents 10% of the total number of households (not consumption units). Depending on the reliability of the distributional results, more detailed breakdowns may be envisaged as well, such as into income percentiles, which may be particularly relevant for the top end of the distribution. However, this would require a careful assessment of the robustness of the results at these levels of detail (see also Chapter 12).

9.3. Clustering according to main source of income

Households can also be clustered into household groups on the basis of their main source of income. For that purpose, households should be clustered in the category which shows the highest contribution to their adjusted disposable income. It should be borne in mind that this should be based on the adjusted disposable income in line with national accounts, thus after imputation for missing elements and alignment of the micro data to the national accounts totals.

The DNA work distinguishes four main sources of income, namely a) *wages and salaries* (based on item D11R), b) *income from self-employment* (based on item B3R3), c) *net property income* (based on item D4N), and d) *current transfers received* (based on items D62R (social benefits in cash received), D63R (social benefits in kind received) and D7R (other current transfers received)).[1] As explained above,

households should be ranked in the category which shows the highest contribution to its national accounts aligned adjusted disposable income.

9.4. Clustering according to socio-demographic information

Households can also be clustered on the basis of socio-demographic characteristics. One specific classification distinguished in the DNA work is according to household type. This takes into account the presence, number and age of the members of the household. In the DNA approach, eight categories of household types are distinguished, i.e. a) *single less than 65 years old*, b) *single 65 and older*, c) *single with children living at home*, d) *two adults less than 65 without children living at home*, e) *two adults at least one 65 or older without children living at home*, f) *two adults with less than 3 children living at home*, g) *two adults with at least 3 children living at home*, and h) *others*. In this classification, an adult is defined as anyone 18 years or older.[2] Furthermore, the delineation of "children living at home" is based on all individuals up until the age of 16 plus the individuals whose age is between 17 and 24 and are offspring of one of the household members and are still living at home. Depending on user needs and the quality and available detail from the underlying data, more granular breakdowns can be envisaged as well.

With regard to socio-demographic information, several alternative classifications can be envisaged. This will depend on user needs, availability of the necessary information to cluster the households accordingly, and the quality of the underlying distributional results. Examples of breakdowns that could be envisaged are according to the *age of the head of the household* (e.g. into a) 0-24, b) 25-34, c) 35-44, d) 45-54, e) 55-64, and f) 65 and above),[3] *housing status* (e.g. a) rental, b) owner-occupied with mortgage, and c) owner-occupied without mortgage), or *main activity of the head of the household* (e.g. a) unemployed, b) employee, c) employer, d) own-account worker, e) unpaid family worker, f) member of producer's cooperative, g) student, h) retired, and i) not classifiable (see Section 2.6.5 for more information)). Countries apply different rules to determine the head of the household, but most of them define it as the person with the highest income (see also (United Nations Economic Commission for Europe, 2011[1]) and (OECD, 2013[2])). In assessing these alternative breakdowns, compilers should check the availability of the relevant information needed to cluster the information on the basis of these specific characteristics. However, they should also carefully assess the robustness of the results at these levels of detail (see also Chapter 12).

The clustering according to socio-demographic information is relatively straightforward in the sense that households should simply be allocated on the basis of their underlying characteristics. In case of changes throughout the year, compilers should look at the duration of the different situations and allocate households according to the situation it was in for the longest period of time.

References

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. [2]

United Nations Economic Commission for Europe (2011), *Canberra Group Handbook on Household Income Statistics*, https://www.unece.org/fileadmin/DAM/stats/groups/cgh/Canbera_Handbook_2011_WEB.pdf (accessed on 27 September 2017). [1]

Notes

[1] As mentioned in Section 2.5.2, the latter category could be further broken down into pension benefits received and other current transfers received, in case the relevant information is available at that level of detail.

[2] In line with general principles of the System of National Accounts, the age of a person for a given reference year should be derived on the basis of his/her age during the largest part of the year. This means that anyone born after the 1st of July should be assigned its age at the start of the year, whereas anyone born on or before the 1st of July should be assigned its age at the end of the year. If this is not feasible, it could be decided to take one cut-off point in the year (e.g. at the start or at the end of the reference period), bearing in mind that this may generate slightly different results.

[3] For national purposes, it may also be of interest to delineate the last two groups on the basis of the retirement age in the country. However, for international comparability, it is recommended to maintain the breakdowns as suggested here. Furthermore, in using the retirement age, it has to be borne in mind that time series analysis may be affected, when the retirement age is changing over time.

10 Overview of the income items

This chapter provides an overview of the national accounts components distinguished on the income side, describing their main characteristics and highlighting items from micro data sources that may provide the best possible link. It also explores possible reasons for gaps between the micro aggregates and the national accounts totals and provides guidance on how to arrive at underlying distributions in case micro data is lacking.

10.1. Introduction

Three main aggregates are distinguished on the income side, i.e. *primary income* (B5), *disposable income* (B6) and *adjusted disposable income* (B7), as defined in the System of National accounts (European Commission et al., 2009[1]) (hereinafter referred to as *2008 SNA*). Primary income is the income that accrues to units as a result of their involvement in the production process or because of ownership of assets that may be used for purposes of production. The main items for the household sector concern *operating surplus*, *mixed income*, *compensation of employees* and *net property income*. Disposable income is the income after re-distribution, involving transactions such as *current taxes on income and wealth*, *social contributions and benefits*, *non-life insurance premiums and claims*, and *other current transfers* like remittances. Adjusted disposable income is derived on the basis of disposable income, but also includes the value of *social transfers in kind* received by households. As mentioned before, these consist of goods and services provided to households by government and non-profit institutions either free or at prices that are not economically significant. They are a direct alternative to receiving a social benefit in cash for the purchase of these services and therefore are included to arrive at a more comprehensive and comparable income measure.

Adjusted disposable income is regarded as the most comprehensive income concept and therefore constitutes the main income measure in the DNA work. It is regarded to provide the best insight into the inequality in a country and the best measure to use for cross-country comparisons as well as analysis of dynamics of inequality over time.

10.2. Operating surplus from actual and imputed rentals (B2R)

The surplus accruing from the production process (before deducting any property income) is reflected by *operating surplus* and *mixed income*. For the household sector, only the surplus generated by home owner-occupiers in their capacity as producers of housing services for own final consumption and by households leasing dwellings is recorded as operating surplus (see 2008 SNA, §7.9 and §24.55). The surplus of unincorporated enterprises is recorded as mixed income. The main reason for the latter is that this surplus implicitly contains an element of remuneration of work done by the owner(s) of the enterprise that cannot be separately identified from the return to the owner as entrepreneur (see 2008 SNA, §7.9).

10.2.1. Owner-occupied dwelling services (B2R1)

In the 2008 SNA, households that own the dwellings in which they live are treated as owners of unincorporated enterprises that produce housing services consumed by these same households (see 2008 SNA, §6.117). The rationale is that the ratio of owner-occupied to rented dwellings can vary significantly between countries or regions and over short periods of time which may hamper international and inter-temporal comparisons of the production and consumption of housing services in case no imputation was made for the value of own-account housing services (see 2008 SNA, §6.34). Furthermore, in distributional analyses, not imputing for own-account housing services may lead to misleading results in which a house-owner and a household that is renting a house may seem to have similar levels of income, but in which the house-owner may be far better off with not having to pay any explicit rent. By imputing housing services produced by the house-owner, this benefit is reflected in its income, with a corresponding imputation on the consumption side to also reflect its consumption of these services.

The housing services are recorded at market prices and valued on the basis as the estimated rental that a tenant would pay for accommodation of the same size, quality and type. This is also the value that is recorded as consumption expenditure by the household. To arrive at operating surplus related to owner-occupied dwellings, the related intermediate costs should be deducted from the output of these housing services. This usually concerns regular maintenance costs (excluding major repairs, which should be

treated as gross fixed capital formation) and the payments for financial services in case that the house purchase has been financed by a mortgage loan (i.e. financial intermediation services indirectly measured (FISIM) related to interest payments on the mortgage loan (see also Section 10.5.1)). The resulting gross operating surplus is treated as being earned by the household in its capacity as the owner of the unincorporated enterprise owning the house.

The principle as described above applies to both main residences and second homes. However, in case the dwelling is located in another country, it is being treated as belonging to a notional resident unit in that country. The legal owner then has a financial claim on this notional unit, regarded as a foreign direct investment relation. The operating surplus from renting out the house is then treated as being withdrawn from the notional unit and fully remitted to the owner in the form of property income (i.e. D42). In this way, there are no retained or reinvested earnings as recorded under item D43 (see also Section 10.5.3). Furthermore, the consumption of the relevant housing services shows up in the consumption of the legal owner, being imported from the country where the dwelling is located.

Both survey and administrative data sources may provide relevant information for the allocation of the amounts to the underlying households. For example, surveys may ask whether respondents own the house they live in and/or about the value of the dwelling. In this regard, various countries use information from income and budget surveys for deriving the distributional breakdown. Furthermore, information may be available on the mortgage loan and on the related interest payments, as well as on the maintenance costs. Information may also be available on characteristics of the dwelling (e.g. the number of square meters, type of residence, neighbourhood) that may be used to estimate the imputed rent (see for example Tsakloglou et al. (2010[2])). Some countries, for example, use cadastral information or census data. Furthermore, in some countries the imputed value of the income generated by production of housing services is taxed, so information may be available from fiscal records. All this information may cover both main residences and second homes. However, if (part of the) second homes are missing, this would require an explicit imputation.

As *operating surplus from owner-occupied dwellings* (B2R1) is the result of the production of housing services minus the costs of maintenance and repairs, FISIM and taxes (less subsidies) on production, it is recommended to also compile the distributional results for this item on the basis of these underlying components. This means that the national accounts item should first be broken down into these underlying components, after which each of these components should be linked to a corresponding item from the micro data sources. In case no direct information is available for one of the underlying items, its distribution can best be obtained by linking it to one of the other subcomponents. It is expected that this calculation will lead to better results than directly targeting the balancing item.

Operating surplus can be derived on a gross or on a net basis, i.e. before or after deducting *consumption of fixed capital* (i.e. depreciation) related to the dwelling. From a conceptual point of view, the net measure would be preferred, but because of challenges in arriving at accurate and comparable estimates of consumption of fixed capital, compilers may also decide to compile results on a gross basis. This is the approach that has been applied in the DNA work so far. In case results are presented on a net basis, it is recommended to separately show the results for consumption of fixed capital (see Section 10.2.2), so that users can derive both gross and net results. Furthermore, the same approach should be applied for both operating surplus and mixed income. The collection template includes a specific block to account for these estimates.

10.2.2. Leasing of dwellings (B2R2)

In accordance with gross operating surplus from owner-occupied dwellings, this category records the *operating surplus from leasing of dwellings* by households. In this case, the value of the output of the rental service is equal to the rental paid by the tenant, after which operating surplus can be derived by deducting the costs for the maintenance and repair of the dwelling, FISIM related to the mortgage interest payments

and taxes (less subsidies) on production. Furthermore, the costs may also include service charges paid to a rental agency.

Relevant micro information for the calculation of the distributional results may be available from both survey and administrative data sources. These may contain information on rent received by households and on costs related to renting out dwellings, although the latter may sometimes be combined with costs related to owner-occupied dwellings. It may also be the case that micro data sources include direct information on the profits made from leasing dwellings. However, it has to be borne in mind that the underlying concepts of the items in the micro data sources may not always match the national accounts concept for operating surplus from leasing of dwellings. The micro data source may for example combine income from renting dwellings with income from renting other fixed assets and natural resources. In the SNA, these relate to different items, i.e. *operating surplus* (B2) related to the rental from leasing dwellings, *mixed income* (B3) related to the rental from leasing other fixed assets, and *rent* (D45) related to the rent from leasing natural resources. If the item in the micro data source indeed combines some of these other components, a reclassification should be performed in order to align micro data to national accounts totals.

The best way to derive distributional results for *operating surplus from leasing of dwellings* will depend on the level of detail available from micro data sources. If micro information is available on the underlying components (i.e. output from rental services, maintenance and repairs, FISIM, taxes (less subsidies) on production, and service charges from a rental agency), it is recommended to derive the distributional results on the basis of the distributions of these underlying items. However, if only information is available on the profits made from leasing out dwellings, this can also be used to directly derive the distribution for this balancing item.

10.2.3. Consumption of fixed capital (memorandum item)

In theory, *operating surplus* and *mixed income* should be corrected for the *consumption of fixed capital*, as the latter reflects the reduction in the value of fixed assets due to their use in the production process. Its deduction would thus lead to measures that account for all costs related to production, providing net measures instead of gross measures. The concept of consumption of fixed capital is closely related to the concept of depreciation, but whereas in commercial accounting depreciation is often used in the context of writing off historic costs, the consumption of fixed capital in the SNA often depends on the current value of the assets.

Although conceptually it may be preferable to focus on net measures, gross measures are often used in national accounts, because of the difficulty of measuring consumption of fixed capital. As explained above, the underlying concept in commercial accounting often differs from the national accounts concept and reported information may often be derived on the basis of arbitrary assumptions which may lead to results that are not comparable across households.

To arrive at economically meaningful results in line with national accounts concepts, statisticians should estimate the present value of the stock of fixed assets, the lifetime of the various underlying assets and the appropriate patterns of depreciation (see also 2008 SNA, §6.240-257). Depending on the available information, compilers may try to come up with estimates as input for the distributional analyses, but if reliable input data is missing, compilers may also decide to publish gross measures. These are generally considered to be more comparable across countries (see 2008 SNA, §2.142), although the 2025 SNA will put more emphasis on net measures, encouraging countries to further invest in improving the relevant estimates.

10.3. Mixed income (B3)

Mixed income is the balancing item from the generation of income account for unincorporated enterprises owned by households. It measures the surplus or deficit accruing from the production by unincorporated enterprises owned by households after deducting compensation of employees, taxes on production and intermediate consumption (including FISIM), but before deducting any payment of property income. It is called mixed income as it implicitly contains an element of remuneration for work done by the owner, or other members of the household, that cannot be separately identified from the return to the owner as entrepreneur.

In addition to income from unincorporated enterprises which is usually reported in surveys or in administrative data, mixed income also covers the surplus from own account production and from underground production. As these components may differ in size across countries and may rely on different techniques for their allocation to the relevant households, it is recommended to treat them separately in deriving distributional results.

As was the case for operating surplus, mixed income can also be derived on a gross or on a net basis, i.e. before or after deducting *consumption of fixed capital* (i.e. depreciation). As explained before, the net measure is preferable from a conceptual perspective, but compilers may focus on gross measures from practical feasibility considerations. The latter is also the approach that has been applied in the DNA work so far. In case results are presented on a net basis, it is recommended to separately show the results for consumption of fixed capital (see Section 10.3.4), so that users can derive both gross and net results. The collection template includes a specific block to account for these estimates.

10.3.1. Own account production (B3R1)

The production boundary of the SNA includes the own-account production of all goods that are retained by their producers for their own final consumption or gross fixed capital formation. This may concern, for example, the production and processing of agricultural products, dairy products, beer and wine, weaving cloth, wood-cutting, hunting and fishing, and the supply of water (see 2008 SNA, §6.32). On the other hand, it does not include the own-account production of services within households except for housing services by owner occupiers (see Section 10.2.1) and the production of services by employing paid domestic staff for example to wash, cook or to look after children (see Section 10.3.3).

Output for own final use should be valued at the basic prices at which the goods and services could be sold if offered for sale on the market. When reliable market prices cannot be obtained, a second-best procedure must be used in which the value of the output of the goods or services produced for own final use is deemed to be equal to the sum of their costs of production. The goods produced for own-account production are treated as being consumed immediately by the relevant household and are recorded as part of consumption expenditure in the relevant COICOP (Classification of Individual Consumption according to Purpose) item.

Often no micro information will be available on own account production, although some countries report to have information available from income or budget surveys. In case no information is available, it may be relevant to assess the underlying assumptions that are used by the national accounts to impute for own account production and whether this may be linked to specific household characteristics that are available in micro data sources. For example, farmers may be assumed to produce and process more agricultural and dairy products for own final consumption than other households. In a similar way, other types of own account production may also be linked to people or households with specific characteristics. By combining this with information available in micro data sources, the amounts related to own-account production may be allocated to underlying households.

As there is a direct link between the own account production and consumption of goods, it has to be borne in mind that the same distribution should be applied to both income and consumption.

10.3.2. Underground production (B3R2)

In the national accounts, adjustments are also made to correct for the non-observed economy, i.e. economic activities that are illegal, underground or informal, or otherwise missed by the statistical system (see also Section 6.5). Depending on country practices, the adjustment for the non-observed economy may affect several components of the household accounts, in particular *compensation of employees*, *mixed income* and *property income received*.

As described in Tartamella and Coli (2010[3]), *mixed income* is one of the items that is most heavily affected by the non-observed economy. For that reason, it is included as a separate sub-item (labelled "underground production") under mixed income to separately allocate the relevant amounts to the underlying households. This category includes the deliberate concealment of legal production activities to avoid tax payments by registered and unregistered units and any illegal production activities.

Since information on underground production is not available in micro data sources, imputations will have to be made by modelling the likelihood of households to benefit from concealed mixed income. As explained in Section 6.5, it is important to first assess the underlying assumptions that are used by the national accounts to impute for underground activities and whether these assumptions may also be used to allocate the relevant amounts to underlying households. For example, if part of the underground economy is imputed on the basis of the assumption that specific types of jobs are more likely to be involved in such types of activity, this may be used to link the amounts to specific groups of households. Then, it is also important to assess which households may be more likely to be involved in underground activities, for example by looking at the plausibility of their overall results. If for some households or household groups consumption by far exceeds their income as reported in the micro data sources, this may be an indication that part of their income derives from underground activities which may not have been covered in the micro data.

10.3.3. Mixed income excluding underground and own account production (B3R3)

This item covers mixed income excluding underground and own account production. This part of mixed income relates to the production by unincorporated enterprises owned by households for which the accounts are not sufficiently detailed to treat the activity as that of a quasi-corporation. In this regard, according to the SNA, unincorporated enterprises owned by households should be treated as quasi corporations, included in one of the corporations sectors when a full set of accounts, including balance sheet entries and information about withdrawals of income from the quasi-corporation, is available (see 2008 SNA, §4.42-4.46). Although frequently information may be available on the production activities, it may not always be possible to separate out other income flows, transfers and financial transactions relating to the production activity from those for the household in general. In that case, as well as in ones where even the information on the production activity is incomplete, the unincorporated enterprise is included in the household sector (see 2008 SNA, § 24.6). As their surplus implicitly contains an element of remuneration for work done by the owner that cannot be separately identified from the return to the owner as entrepreneur, the full amount is recorded as mixed income.

In most countries, mixed income is computed on the basis of administrative records, business surveys or a mix of surveys and administrative data. These usually cover information on self-employment income which could form a good proxy for mixed income, dependent on the exact definition of the income and the delineation of the self-employed. In micro surveys income from self-employment often includes the profit or loss that accrues to owners of, or partners in, unincorporated enterprises who work in these enterprises, after deduction of charges such as interest, dividends and rents payable that are related to the production

activities (see OECD (2013[4])). Mixed income as defined in the national accounts corresponds to the value of output less operating costs (such as intermediate consumption, payment of compensation of employees and net taxes) and before any deduction and receipt of property income (i.e. interests, dividends and rents). Therefore, the main difference will often relate to the treatment of property income received/paid by the enterprises. A correction may be needed to align the micro concept with the national accounts concept.

Moreover, profits or losses from partners who do not work in these enterprises (i.e. "silent" or "sleeping" partners) may be treated differently. They may be included in dividend income in survey results, whereas they are included in *mixed income* (B3) or in *withdrawals from shareholders* (D422) in the national accounts. Also, for this issue, there may be a need to conduct a specific correction in order to better align the concepts. Finally, differences may also occur due to a different treatment of *consumption of fixed capital*. It is recommended to start from gross figures, i.e. before deducting consumption of fixed capital, but in case the micro data are based on net figures, a correction will have to be applied to arrive at similar measures. In that regard, it also has to be considered that the consumption of fixed capital as included in micro data results may often be based on tax and accounting rules and may thus deviate from the concept as applied in the national accounts, where it is based on current replacement cost, not historic cost, and on estimates of actual prices of capital consumption (see also Section 10.3.4).

Looking at the delineation of self-employed for which the production surplus should be included in mixed income, the SNA defines them according to whether they keep separate accounts or not (see above). This delineation may differ from households' self-perception and may lead to divergences between what is recorded as income from unincorporated businesses owned by households in the national accounts and what people declare as income generated as a self-employed business activity in surveys. Further complexity is added by the fact that, as legal arrangements vary across countries, even the compilers' interpretation of national accounts rules may differ across countries. Although international guidelines have been developed on how to classify unincorporated enterprises, countries' experiences still demonstrate difficulties in estimating the share of self-employment by institutional sector and pointing out issues with regard to comparability across countries (see Pionnier and Guidetti (2015[5])). For these reasons, it is important that national experts try to harmonise the concepts used in micro data and in national accounts as much as possible (ex-ante), and to make corrections ex post in case of any remaining conceptual differences. The latter may be done by confronting the household data with information from the business register. This may provide insight for which households the comparability between micro and macro results may be hampered due to a different classification of activities related to unincorporated enterprises. On the basis of that information specific corrections may be applied at the micro level to better align the data.

In addition to conceptual differences, gaps between micro aggregates and national accounts totals may also occur due to other causes, such as misreporting and/or under-or over-coverage of specific household groups. Johns and Slemrod (2008[6]) and Neri and Zizza (2010[7]) have shown that self-employment income is one of the items that is most liable to underreporting in survey data. If part of a possible gap between micro data and national accounts totals is related to underreporting, it has to be assessed which households this most likely relates to.

The gap between micro and national aggregates may also come from households that do not report any information on this item (item non-response). This may call for different imputations. To arrive at appropriate imputations, it would be relevant to assess whether national accountants have already applied specific corrections in relation to possible underreporting and whether this provides information on how to allocate the amounts to the relevant households. Furthermore, it is important to assess which households may be more likely to have underreported their mixed income, for example by looking at the plausibility of their overall results. If for some household groups consumption by far exceeds their income as reported in the micro data sources, this may be an indication that part of their income has not bene reported in the micro data. As it was the case for income from underground activities, this requires careful analysis of the data.

10.3.4. Consumption of fixed capital (memorandum item)

As was the case with operating surplus, from a conceptual point of view mixed income should exclude consumption of fixed capital, to arrive at measures that account for all costs related to production. However, as explained in Section 10.2.2, because of the difficulty of measuring consumption of fixed capital, compilers may decide to only compile and publish gross figures. To arrive at economically meaningful results for consumption of fixed capital, statisticians should estimate the present value of the stock of fixed assets used in production, the lifetime of the various assets and the appropriate patterns of depreciation. Depending on the available information, compilers may try to come up with estimates as input for the distributional analyses. Otherwise, compilers may also decide to publish results at gross measures. These results are generally considered to be more comparable across countries (see 2008 SNA, §2.142), although the 2025 SNA will put more emphasis on net measures, encouraging countries to further invest in improving the relevant estimates.

If compilers aim to arrive at net measures, it is important to look for appropriate micro data to distribute the amount of consumption of fixed capital as recorded in the national accounts. In that regard, it has to be understood that depreciation as used to derive business profits in surveys is usually based on tax and accounting rules, based on historic cost. This may not reflect the actual value at which fixed capital is used up in the production process and may deviate from consumption of fixed capital as defined in the SNA, which is based on current replacement cost. In that regard, it may be better to derive estimates of consumption of fixed capital for the relevant households on the basis of estimates of the present value of their stock of fixed assets used in production, including assumptions on the lifetime of these assets and the appropriate patterns of depreciation. This may be done on the basis on assumptions of the type and amount of capital stock that would be needed in the production of goods and services in specific industries. This may then provide ratios between the amount of consumption of fixed capital and mixed income for specific industries, on the basis of which appropriate values can be allocated to relevant households in proportion to their mixed income in a specific industry.

10.4. Compensation of employees (D1R)

Compensation of employees is the total remuneration, in cash or in kind, payable by an enterprise to an employee in return for work done by the latter during the accounting period. It has two components, i.e. wages and salaries, and social insurance contributions payable by employers, which includes both actual and imputed contributions to social insurance schemes.

10.4.1. Wages and salaries (D11R)

Wages and salaries include any social contributions, income taxes, etc., payable by the employee even if they are actually withheld and paid directly by the employer on behalf of the employee. Wages and salaries in cash include wages or salaries that are paid regularly; enhanced payments or special allowances, for example for working abroad, to cover the costs of travel to and from work; ad hoc bonuses; and commissions, gratuities and tips. They do not include reimbursements for expenditures made by employees to take up their jobs or to carry out their work, and payments to workers absent from work because of illness, accidental injury, maternity leave, etc. Wages and salaries in kind should include the value of goods and services that employers provide to their employees, either for free or at reduced prices, such as meals and drinks; housing services or accommodation; services of vehicles or other durables provided for the personal use; transportation to and from work; and childcare (see 2008 SNA, §7.43-7.55).

Micro information will usually be available from survey data and/or from administrative data sources. However, conceptual differences may exist that require adjustments. One important issue in this regard is the recording of wages and salaries while an employee is on sick, injury or maternity leave. These amounts

are not recorded as *wages and salaries* (D11R) in the SNA, but as *social benefits other than social transfers in kind* (D62R) (see Section 10.9), whereas they are usually included in wages and salaries in micro data sources. This means that a correction will be needed in matching the micro and macro data.

Another issue is that survey data may not always cover wages and salaries from secondary jobs, which may lead to under-coverage in comparison with national accounts totals. If that is the case, imputations should be made for those persons who are expected to have secondary jobs. On the other hand, survey results may sometimes lead to over-coverage due to the fact that they relate to a specific period in time that may not be representative of a full reference period. This would also require specific adjustments for the households involved.

Furthermore, wages and salaries in kind may be treated differently in micro data sources and in the national accounts. Almost any kind of consumption good or service may be provided as remuneration in kind, with the most common examples being meals and drinks, housing services, the services of vehicles, goods and services produced as outputs from the employers' own processes of production, transportation to and from work, sports, recreation or holiday facilities for employees and their families, childcare for the children of employees. Another form of income in kind that has become more popular over the past decade results from the practice of an employer giving an employee the option to buy stocks (shares) at some future date. It is often the case that the related amounts are not included in micro data sources. In that case, the relevant amounts should be allocated on the basis of assumptions on who is most likely to benefit from these forms of remuneration.

10.4.2. Employers' actual social contributions (D121R)

Employers' social contributions are social contributions actually paid by employers to social security funds or other employment-related social insurance schemes to secure social benefits for their employees, former employees or dependants. As they are made in relation to employment for the benefit of these specific groups, their value is recorded as one of the components of compensation of employees. Subsequently, the contributions are recorded as being paid by the employees as current transfers into the social security schemes or other employment-related social insurance schemes (see also Section 10.7.1). By definition, these amounts received as part of compensation of employees and as paid into social insurance schemes are identical.

The contributions are divided into actual and imputed contributions. The *employers' actual social contributions* consist of the contributions actually paid by employers to both social security and other employment-related schemes. The *employers' imputed social contributions* relate to social benefits that are provided by employers directly to their employees, former employees or dependants without involving an insurance enterprise or autonomous pension fund and without creating a special fund or segregated reserve for the purpose (see Section 10.4.3). Further distinctions can be made into pension and non-pension contributions (see 2008 SNA, §7.56-7.70).

Information on employers' actual social contributions will usually be available from micro data sources and from administrative data sources. In case no information is available, a solution would be to use the distribution of wages and salaries as a proxy for the distribution of the employers' actual social contributions.

10.4.3. Employers' imputed social contributions (D122R)

In addition to the actual contributions, the SNA also distinguishes imputed social contributions, reflecting those contributions that are not directly recognisable as being paid by employers, but that are still benefiting households as they fund a social security scheme or accrue a social security entitlement for the employees. The SNA distinguishes two types of imputed employers' social contributions, i.e. imputed pension contributions and non-pension contributions.

When looking at employers' imputed social contributions related to pension schemes, a distinction should be made between defined contribution schemes and defined benefit schemes. In the case the employer runs the former scheme him-/herself, the value of the costs of operating the scheme should be treated as an imputed contribution payable to the employee as part of compensation of employees (with a counterpart recording as final consumption expenditure by households of financial services) (see 2008 SNA, §7.64). Although this is not an actual flow from the employer to the employee, it makes sure that the accounts properly reflect that the employees are benefiting from obtaining this service from their employer. All other contributions made to a defined contribution scheme will concern actual contributions that will be included under D121R, so they will not require any additional imputations.

In case of defined benefit schemes, the imputed social contributions are calculated such that the sum of the employer's actual contribution plus the sum of any contributions by the employee plus the imputed contribution by the employer is equal to the increase in benefit due to current period employment plus the costs of operating the scheme (see 2008 SNA, §7.65). This imputation ensures that the full increase in the pension entitlement due to the current period employment is accounted for in the accounts, with the employer normally being responsible for any shortfall between the accrual and the contributions received.

The employers' imputed non-pension contributions relate to the situations where social benefits are provided by employers directly to their employees, former employees or dependants without involving an insurance enterprise or autonomous pension fund and without creating a special fund or segregated reserve for the purpose. In this situation, the 2008 SNA (§7.68, 8.83 and 8.84) considers existing employees as being protected against various specified needs or circumstances, even though no reserves are built up to provide future entitlements. Remuneration is therefore imputed for such employees equal in value to the amount of social contributions that would be needed to secure the de facto entitlements to the social benefits they accumulate.

Although the amounts should in principle not only depend on the levels of the benefits currently payable, but also on how the future benefits are likely to evolve (as a result of factors such as expected changes in the number, age distribution and life expectancies of their present and previous employees), in practice, due to difficulty in deriving these actual amounts, the unfunded non-pension benefits payable by the enterprise during the same accounting period are often used as an estimate of the imputed remuneration that would be needed to cover the imputed contributions.

These imputed social contributions are specific to the national accounts framework and are usually missing from micro data sources (e.g. they are not part of the household income definition as defined by the Canberra Group Handbook). If no information is available, it is recommended to use the distribution of *wages and salaries* (D11) or of *employers' actual social contributions* (D121) as a proxy. Although the inclusion of these items does not affect disposable income (as the amount received as part of *compensation of employees* (D11) is equal to the amount paid as part of *social contributions* (D61)), the availability of breakdowns for these items is deemed valuable, as these affect the distribution of primary income and provide insight on how re-distributional transactions change the incomes of households.

10.5. Net property income (D4N)

Property income accrues when the owners of financial assets and natural resources put them at the disposal of other institutional units (see 2008 SNA, §7.107). It is usually broken down into underlying items that provide more information on the type of income or payments and on the related financial instrument. The focus in the DNA work is on *net property income* which is the result of property income received (D4R) and paid (D4P). The following sub-sections discuss the various components in detail.

10.5.1. Interest received (D41R) and paid (D41P)

Interest is a form of income that is receivable by the owners of certain kinds of financial assets (i.e. deposits, debt securities, loans and (possibly) other accounts receivable) for putting their financial assets at the disposal of another institutional unit (see 2008 SNA, §7.113-7.126). It may be a predetermined sum of money or a fixed or variable percentage of the principal outstanding. It is recorded on an accrual basis, i.e. continuously accruing over time. This may differ from actual amounts paid in a specific reference period which may often be what is included in micro data sources.

The amounts of interest on loans and deposits payable to and receivable from financial corporations include a margin that represents an implicit payment for the services provided by financial corporations in providing loans and accepting deposits (see 2008 SNA, §7.116). As these amounts constitute payments for the intermediation services provided by the financial corporations, the actual interest payments and receipts to or from financial corporations are corrected for these service charges to arrive at interest as defined in the SNA.

This service charge, which is known as *Financial Intermediation Services Indirectly Measured* (FISIM), is calculated on the basis of both the difference between the interest rate received on loans and a reference rate, and the difference between a reference rate and the interest rate paid on deposits (see 2008 SNA, §A3.24-A.3.27). The impact for the household is that if it receives interest on a deposit, the actual amount will be lower than the notional amount as recorded in the national accounts, as the bank deducts an amount related to the service it provides (i.e. a higher notional interest receipt for the household sector with an accompanying FISIM payment). If a household pays interest on a loan, the actual payment is higher than the notional amount recorded in the national accounts as the bank adds a service fee for their intermediary service (i.e. a lower notional interest payment by the household sector with an additional FISIM payment). Offsetting adjustments are applied to consumption, increasing intermediate consumption in the case of interest payments related to business deposits and to mortgage and business loans, and increasing final consumption in case of other deposits and loans (e.g. loans for purchasing final consumption goods).

As FISIM is a specific item in the SNA that has no specific counterpart in micro data sources, the template starts from the unadjusted “actual” interest flows, i.e. not adjusting for FISIM, and includes a specific item for the allocation of the FISIM correction, both on the uses and the resources side.

Interest (not adjusted for FISIM) received (D41R’) and paid (D41P’)

As explained above, it is recommended to start from unadjusted “actual” interest flows at the macro level, as it will provide a better link to the data as recorded in the micro data sources, and to only then apply the correction for FISIM. This means that the national accounts items of interest received and paid have to be recalculated into the “actual” interest flows by removing the adjustments that were made to correct for the financial intermediate services (FISIM). On the receipt side, this will lead to a higher amount, whereas on the payment side, this will lead to a lower amount. As FISIM is specific to the SNA, it should be relatively easy to retrieve the unadjusted flows.

Micro information on interest payments and receipts by households is usually available from survey data as well as from administrative data. This may concern fiscal data obtained from tax files or data obtained from financial corporations directly. The definition in the micro data will usually be in line with the definition as used in the template for interest payments and receipts before the adjustment to correct for FISIM, except for a possible difference between the accrual recording and cash payments.

In case no separate information is available on interest payments and/or receipts, an alternative would be to derive estimates on the basis of micro information on financial assets and liabilities of households, provided that those are available. By linking interest rates to the amounts for the various financial instruments held and owed by households, interest payments and receipts can be derived. These

estimates may also be used to check the plausibility of the results, especially in case of large gaps between the micro and macro aggregates.

In the EG DNA collection rounds it became clear that micro-macro gaps for countries relying on survey data for interest receipts were often related to under-coverage of the very rich and/or underreporting for specific household groups. In case administrative data are available (especially in case these are provided by financial corporations) these caveats may be overcome.

Adjustments for FISIM (D41R_FISIM and D41P_FISIM)

The item *Financial Intermediation Services Indirectly Measured* (FISIM) reflects the indirect service charge for the service by banks of bringing together borrowers and lenders. This leads to a higher notional interest received for the household sector as recorded in the national accounts than the actual amount received (as it is assumed that the bank already deducted an amount related to the service it provides) and to a lower notional interest paid as recorded in the national accounts than the actual amount paid (as it is assumed that this payment also includes a service fee for the intermediary service provided by the bank).

To arrive at amounts in line with the national accounts concepts, it is proposed to compile the FISIM correction separately for both interest received and interest paid. When these corrections are combined with the "actual" interest receipts and payments, this will lead to the balance of the notional interest flows as recorded in the SNA. The correction regarding interest received should be recorded with a positive sign, reflecting that the notional amount is higher than the actual amount received. The correction regarding interest paid should be recorded with a negative sign, reflecting that the notional amount is lower than the actual amount paid. Table 10.1 provides a schematic overview of the related flows.

Table 10.1. Recording of SNA interest receipts and payments in the EG DNA template

Receipts (R)	D41R	SNA interest received
	D41R'	Actual amount received (not adjusted for FISIM)
	D41R_FISIM	+ Adjustment for FISIM
Payments (P)	D41P	SNA interest paid
	D41P'	Actual amount paid (not adjusted for FISIM)
	D41P_FISIM	- Adjustment for FISIM

Source: The Author.

As explained above, as FISIM is a specific national accounts concept, no counterpart variable will be available from micro data sources. To derive an appropriate distribution, compilers are recommended to make a link to the actual interest payments and receipts by households. If detailed information is available on the different actual interest rates for various types of deposits and loans and these can be linked to the various households groups, such information can be used to allocate FISIM at a very detailed level. However, if such information is lacking, one could also assume equal margins for all types of deposits and loans, allocating FISIM proportionally to the relevant amounts of interest payments and receipts by households or household groups.

When allocating FISIM to households, one should bear in mind that a corresponding correction should be made for the consumption of FISIM. Depending on the type of deposit or loan, it should be recorded as intermediate consumption or as final consumption. If FISIM relates to mortgage loans, it should be recorded as intermediate consumption in the production of housing services related to owner-occupied dwellings and the leasing of dwellings. This is reflected in a lower value of *operating surplus* (B2) (see also Section 10.2). If FISIM relates to business loans owed by households or deposits held by household businesses, it should also be recorded as intermediate consumption, but in this case reflected in a lower value of *mixed*

income (B3) (see also Section 10.3.3). For all other deposits and loans, the related FISIM should be recorded as part of final consumption (CP1261).

As there is a direct link between the consumption of FISIM and the correction items recorded in the property income block, one should make sure that in allocating FISIM to households (or household groups), this link is maintained. This means that the sum of FISIM recorded as final consumption (CP1261) and as intermediate consumption (reflected in lower values of B2 and B3) should equal the amount of FISIM recorded as correction items in property income (discussed above) at the level of the household (or household group). The breakdown into type of consumption should ideally be made on the basis of information on the types of deposits and loans or type of interest receipt and payments. If that type of information is not available, assumptions should be made to break it down into these three types of consumption.

10.5.2. Distributed income of corporations received (D42R)

Distributed income of corporations consists of two underlying items, i.e. dividends and withdrawals of income from quasi-corporations (see 2008 SNA, §7.127-7.135). *Dividends* are a form of investment income to which shareholders, i.e. the collective owners of corporations holding the shares in their equity, become entitled as a result of placing funds at the disposal of corporations. *Withdrawal of income from quasi-corporations* consists of that part of distributable income that the owner withdraws from the quasi-corporation. These are unincorporated enterprises that have sufficient information to compile a complete set of accounts and are operated as if they were separate corporations. The amount of income withdrawn from a quasi-corporation is decided by the owner and in that regard such a withdrawal is equivalent to the payment of dividends by corporations to their shareholder(s).

As was explained in Section 10.3.3, the distinction between unincorporated enterprises that should be treated as quasi-corporations and the ones that should be included in the household sector is not always straightforward. This may differ across countries and may also lead to different treatments in the national accounts and in micro data sources. This has to be borne in mind when comparing micro and macro data for both mixed income and distributed income of corporations. Harmonisation of the concepts used in micro data and in national accounts on how to treat unincorporated enterprises is very relevant in this regard. In combining the micro and macro data, it may also be useful to assess how specific entities are recorded in the business register, to check whether the way in which data are recorded in micro data sources indeed aligns with the business register which normally serves as the basis for the national accounts recording.

Furthermore, in combining the micro and macro data, it has to be borne in mind that withdrawal from quasi-corporations in the SNA also includes operating surplus related to ownership of dwellings abroad. As was explained in Section 10.2.1, this dwelling is treated as belonging to a notional resident unit in that country, with the legal owner having a financial claim towards this unit. The operating surplus from renting out the house is then treated as being withdrawn from the notional unit and fully remitted to the owner in the other country the form of property income.

When looking at micro data sources on distributed income of corporations, most income surveys and fiscal data sources will include this type of information. The definition will also be more or less the same, except (as was explained above) for the delineation between unincorporated enterprises to be included in the household sector (for which the production surplus will be recorded under mixed income) and unincorporated enterprises that should be recorded as quasi-corporations (for which the distributed income will be recorded under distributed income of corporations). Furthermore, as explained in the ICW Framework (OECD, 2013[4]), micro data sources may often not include income from family trusts. These are discretionary trusts set up to hold families' assets or to conduct family businesses. Generally, they are established for asset protection or tax purposes. While the income from these trusts is recorded as distributed income of corporations in the SNA, it is usually not covered in micro survey data.

Furthermore, the distributed income related to dwellings owned abroad may often not be covered in micro data sources either.

As was the case with interest, if no separate information is available on distributed income of corporations received, an alternative would be to derive estimates on the basis of micro information on financial assets and liabilities of households, provided that those are available. By assuming a certain rate of return in relation to equity held by households, distributed income of corporations as received by households can be derived. These estimates may also be used to check the plausibility of the results and be of help in allocating some of the amounts to underlying households in case of large gaps between the micro and macro aggregates. However, it must be borne in mind that in comparison with interest flows, dividends may show larger dispersions. Corporations usually show large differences in terms of profits and will also differ in dividend policies, with some corporations distributing all profits while others retaining all or part of it. These differences may perhaps partly cancel out when looking at results at more aggregated levels but may distort results when looking at more granular levels of detail.

In the DNA work it became clear that this specific item shows the largest gaps between the micro and macro data. Therefore, it is important to assess the main underlying reasons for these gaps and try to allocate these accordingly. First of all, gaps may relate to different treatment of unincorporated enterprises in the micro data and in the national accounts, as explained above. This will affect both mixed income and distributed income of corporations. If a profit-making unincorporated enterprise is recorded as a quasi-corporation in the national accounts whereas it is treated as part of the household sector in the micro data, this will lead to a higher mixed income and lower distributed income of corporations in the micro data in comparison with the national accounts. Therefore, it is important that both items are analysed in conjunction when analysing micro-macro gaps. Secondly, micro-macro gaps may often relate to under-coverage of the very rich and/or underreporting for specific household groups, mainly in relation to survey data. Pareto-tail adjustments may assist in overcoming this issue.

10.5.3. Reinvested earnings on foreign direct investment (D43R)

This item specifically relates to the retained earnings of (quasi-)corporations that are part of a foreign direct investment relation. Foreign direct investment is defined as a cross-border investment relation in which a resident in one country (the direct investor) has control or a significant degree of influence on the management of an enterprise (the direct investment enterprise) resident in another economy (see 2008 SNA, § 21.34). Any earnings that are not actually distributed to the direct investor are treated as being distributed implicitly and reinvested, as the decision to retain some of the earnings is seen as representing a deliberate investment decision on the part of the foreign direct investor (see 2008 SNA, §7.137 and 7.138).

Although foreign direct investment usually takes place between corporations or quasi-corporations, a couple of countries record reinvested earnings on foreign direct investment as being received by households. This would imply that the household is regarded as the direct investor in a foreign direct investment relation. As it is not very likely that this would normally be done in the form of an unincorporated enterprise for which it is not possible to set up separate accounts, the amounts recorded under this item are expected to only exist for a very small number of countries and for very small amounts.

In case a country records reinvested earnings received by the household sector, its allocation to the relevant households will depend on available underlying information. It is not expected that these amounts will be covered by income surveys, but perhaps information may be available on specific equity holdings of certain households that may give rise to these reinvested earnings. Alternative would be to link it to equity holdings in general, although it is assumed that this will be suboptimal, as the reinvested earnings will only concern very specific equity holdings.

10.5.4. Investment income disbursements (D44R)

Investment income disbursements concern property income flows that are not explicitly paid out, but that do accrue to the owners of the underlying assets, normally directly feeding into an increase in their wealth. For that reason, they are imputed in the SNA as part of property income.

The SNA recognizes three types of investment income disbursements: (1) investment income attributed to insurance policy holders (D441R), (2) investment income payable on pension entitlements (D442R) and (3) investment income attributed to investment fund shareholders (D443R). As these components concern rather different forms of income and may have very different distributions, it is recommended to compile the results for this item on the basis of these underlying components instead of directly targeting the aggregate. In that regard, it has also to be borne in mind that investment income disbursements are normally not recorded in micro data sources, as a consequence of which imputations will be needed to arrive at distributional results, for which different types of auxiliary data will be needed for the three components.

Another reason to focus on the underlying components is that parts of the investment income disbursements are treated as premium supplements that need to be recorded in the distribution of income account as part of *social contributions* in D61 (i.e. the investment income payable on pension entitlements) and *non-life insurance premiums* in D71 (i.e. the part of investment income attributed to insurance policy holders that relates to non-life insurance).[1] In that regard, one should make sure that the breakdown into household groups is identical for the corresponding parts. The calculation of distributional results for the various underlying components may help compilers in attributing the right amounts in the remainder of the accounts.

Investment income attributed to insurance policy holders (D441R)

For non-life insurance policies, the insurance corporation holds technical reserves that are seen as a liability towards the insurance policy holders. The investment income on these reserves is treated as income attributable to the policyholders (as it is assumed to feed into these technical reserves), which is distributed to policyholders in the allocation of primary income account (as part of D441R) and paid back to the insurance corporation as a premium supplement in the secondary distribution of income account (as part of D71P, i.e. *net non-life insurance premiums*). Net non-life insurance premiums comprise both the actual premiums payable by policyholders to obtain insurance coverage during the accounting period and the premium supplements payable out of the investment income attributed to insurance policyholders, less the service charges payable to the insurance corporation.

For life insurance policies and annuities, the insurance corporations have liabilities towards the policyholders and annuitants equal to the present value of expected claims. Bonuses declared in connection with life policies are treated as being distributed to policyholders and as premium supplements recorded in the financial account as payable by households and receivable by insurance corporations as changes in life insurance and annuities entitlements. As the recording of this item in the remainder of the accounts differs from the supplements on non-life policies, which are also treated as flowing back to policyholders but have a counterpart in the re-distribution account as premium supplements payable by households and receivable by insurance corporations (D71P), the template uses two codes to make a clear distinction between the two (D441A versus D441B).

Investment income attributed to insurance policyholders is usually not covered in micro data sources, as a consequence of which the distribution across households should be derived on the basis of alternative information.[2] Because of the link with the insurance technical reserves (that are at the basis of the investments of the insurance corporation on which it receives the investment income that has to be attributed to the policyholders), it would make more sense to use the distribution of these reserves as a proxy for the distribution of the investment income. However, this would require reliable information at the

micro level on both life insurance and non-life insurance technical reserves. An alternative would be to take the premium payments for both life and non-life insurance as a proxy for the income attributed to insurance policyholders. This information may be available from survey data. A last resort would be to derive the distribution for this item on the basis of one of the aggregates in a way that the inclusion or exclusion of the component does not affect the distribution of the main aggregates. In that sense, it would be best to link it to the distribution of disposable or adjusted disposable income. However, this should only be done as a last resort.

Investment income payable on pension entitlements (D442R)

The second category concerns investment income payable on pension entitlements. The exact calculation of this item depends on the underlying type of pension scheme. For defined contribution pension schemes, contributions are invested on behalf of the employees as future pensioners and the investment income receivable by the pension funds is therefore recorded as property income for the households. The investment income payable on defined contribution entitlements is equal to the investment income on the funds plus any net operating surplus earned by renting land or buildings owned by the fund. For defined benefit schemes the increase in the present value of the entitlements due to the unwinding of the discount rate represents the investment income distributed to the employees and should be recorded under this specific item.

For both types of schemes, the investment income is attributed to the policyholders which are then treated as paying an equal amount back to the funds as premium or contribution supplements in the secondary distribution of income accounts (as part of D61P, i.e. social contributions paid). It also forms part of the adjustment item for the change in pension entitlements (D8), which will be explained in more detail in Section 11.16.

Investment income payable on pension entitlements is usually not covered in micro data sources, which means that its distribution should be based on alternative information. As the calculation of investment income differs between defined benefit and defined contribution pension schemes, ideally information is available on which households accrue pension entitlements according to what type of scheme. The investment income can then be derived for both types of schemes on the basis of the accrued entitlements (from the balance sheet) in combination with a rate of return. For defined benefit schemes this rate of return will be determined by the discount rate, whereas for defined contribution pension schemes the rate of return can be derived by dividing the actual property income of the relevant pension funds by its pension liabilities.

If no micro information is available on the pension entitlements accrued by households, an alternative may be to estimate these entitlements on the basis of auxiliary information, such as the pension premium payments (also taking into account that retired persons will no longer contribute but still need to be assigned investment income on their pension entitlement accrued on previous pension contributions), the number of years that households have contributed to a scheme (which may also be estimated on the basis of age information), and/or the benefit formula of the relevant pension schemes.

It is clear that arriving at accurate estimates of the pension entitlement for the various households is a very complex and time-consuming task which requires a lot of actuarial assumptions. However, as the related amounts may be significant, compilers are encouraged to try to come up with distributional estimates that at least take into account age groups to arrive at an appropriate distribution of the investment income payable on pension entitlements. In that regard, it has to be borne in mind that older persons (up to retirement) will have accrued more pension entitlements and therefore may benefit to a larger extent from investment income payable on these entitlements. Furthermore, it may also be expected that entitlements of retired people will decline over time (the decline depending on longevity tables) which also needs to be taken into account when deriving the results. Simply applying an equal distribution or linking it to one of the aggregates is deemed to lead to suboptimal results.

Investment income attributable to collective investment funds shareholders (D443R)

A third category of investment disbursements is the investment income attributed to investment fund shareholders. Also for these funds, it is reasoned that the shareholders are actually the owners of the investments and therefore should receive all the earnings on the investments. As only part of the earnings of investment funds is actually distributed to the shareholders in the form of dividends, the remainder of the earnings is also recorded in the SNA as being distributed to the shareholders (leaving the investment funds with no saving) and being reinvested into the funds via a transaction recorded in the financial accounts.

As is the case with the other two investment income disbursements, micro information will usually be lacking for this item. In that case, it is recommended to derive the distribution on the basis of information on holdings of investment fund shares by households. If that information is not available, an alternative would be to use the distribution of distributed income of corporations as a proxy. A last resort would be to derive the distribution for this item on the basis of one of the aggregates in a way that the inclusion or exclusion of the component does not affect the distribution of the main aggregates. In that sense, it would be best to link it to the distribution of primary, disposable or adjusted disposable income. However, this should only be done as a last resort.

10.5.5. Rent received (D45R) and paid (D45P)

Rent is the income receivable by the owner of a natural resource (the lessor or landlord) for putting natural resources such as land or subsoil assets at the disposal of another institutional unit (a lessee or tenant) for use in production (see 2008 SNA, §7.109). The terms under which rent on a natural resource is payable are expressed in a resource lease. Rents differ from rentals, which are payments under an operating lease to use fixed assets such as dwellings or machines belonging to another unit which remains responsible for maintenance and replacement of the asset if necessary. These payments are treated as sales or purchase of services. This means that not only the type of underlying asset is different between rent and rentals, but also the nature of the lease.

Information on rent received and paid may be available from survey data and from administrative data sources. However, when linking the micro data to the national accounts totals, it has to be borne in mind that the concepts may not perfectly match. For example, it may be the case that the micro data source combines the income from renting natural resources and fixed assets in one category, whereas according to the SNA they should be broken down into a part that feeds into *operating surplus* (B2) (i.e. the rental from leasing fixed assets) and a part that is recorded as *rent* (i.e. the rent from leasing natural resources). In that case, a reclassification should be performed in order to align micro data to national accounts totals, by reclassifying income received from renting fixed assets.

10.6. Current taxes on income and wealth

This category includes all current taxes on income and wealth. Taxes are compulsory, unrequited payments, in cash or in kind, made by institutional units to government units regularly every tax period (see 2008 SNA, §8.15 and §8.52-8.64). They are recorded on an accrual basis, i.e. when the activities, transactions or events occur that create the liabilities to pay taxes. Taxes on income consist of taxes on incomes, profits and capital gains, whereas current taxes on capital consist of taxes that are payable periodically, usually annually, on the property or net wealth of institutional units (with the exception of taxes on land or other assets that are used in production which are treated as other taxes on production (lowering operating surplus and mixed income). Furthermore, this category includes miscellaneous current taxes such as poll taxes, payments by households to obtain certain licences, and taxes on international transactions.

Information on taxes on income and wealth are usually available from income surveys and administrative data. It will depend on the exact definitions whether the coverage is similar to that of the national accounts. In that regard, micro surveys may exclude taxes on holding gains and/or taxes on wealth which may lead to a gap with the national accounts totals. This should then be allocated to the relevant households, preferably on the basis of underlying information on the distribution of the holding gains and/or the wealth underlying these taxes. Furthermore, the time of recording in micro surveys may be different from the accrual recording in the national accounts. This should also be borne in mind when linking micro data to the national accounts totals.

As there is a link between taxes on income and the primary income as recorded in the SNA, their distributions may be compared to check the plausibility of the results. As various items may be liable to different tax rates and as some items may be exempted or subject to a threshold, the distributions will probably not be identical, but some correlation may be expected. The same goes for taxes on holding gains and taxes on wealth. If these could be linked to distributional information on holding gains and wealth, this would provide more insight into the plausibility of the related tax results. Furthermore, these techniques may be used to estimate the distribution of taxes on income and wealth in case micro data is missing.

10.7. Net social contributions paid (D61P)

Social contributions are actual or imputed payments to social insurance schemes to provide for the payments of social insurance benefits. A social insurance scheme is a specific type of insurance scheme where the following two conditions are satisfied: (a) the benefits received are conditional on participation in the scheme and constitute social benefits as defined in the SNA (see Section 10.9); and (b) at least one of the three conditions following is met: (i) participation in the scheme is obligatory either by law or under the terms and conditions of employment of an employee, or group of employees; (ii) the scheme is a collective one operated for the benefit of a designated group of workers, whether employed or non-employed, participation being restricted to members of that group; (iii) an employer makes a contribution (actual or imputed) to the scheme on behalf of an employee, whether or not the employee also makes a contribution (see 2008 SNA, §8.65).

In the SNA, all contributions to social insurance schemes are shown as made by households. They consist of employers' and households' social contributions. The former comprise *employers' actual social contributions* (D611P) and *employers' imputed social contributions* (D612P). Both items are exactly the same as those recorded in the primary income account as part of compensation of employees, respectively D121R and D122R. Households' contributions consist of *households' actual social contributions* (D613P) and *households' social contribution supplements* (D614P). The actual contributions reflect the contributions payable by employees on their own behalf, by self-employed and by non-employed persons. The contribution supplements consist of the property income earned during the accounting period on the stock of pension and non-pension entitlements, as recorded in the primary income account respectively under items D442R and D441R. Set against these contributions is the service fee charged by the unit administering the social security scheme which should be deducted to arrive at the net social contributions paid.

As some of the underlying components link to specific income items and as the components may have different underlying distributions, the template distinguishes them separately. Ideally information is compiled at this detailed level as it is expected to lead to more accurate results than simply focusing on the aggregate.

10.7.1. Employers' actual social contributions paid (D611P)

This item reflects the actual contributions made by employers to social insurance schemes. The item is exactly the same as the one recorded under item D121R and compilers should make sure that the reported amounts match. Please see Section 10.4.2 for more information on this item.

10.7.2. Employers' imputed social contributions paid (D612P)

This item reflects the imputed social contributions by employers related to the unpaid part of the accrual of pension entitlements related to the current service period and to non-pension social benefits provided by employers directly to their employees, former employees or dependants. It is exactly the same as that recorded in the allocation of primary income account under item D122R and compilers should make sure that the reported amounts match. Please see Section 10.4.3 for more information on this item.

10.7.3. Households' actual social contribution paid (D613P)

This category records all contributions to social insurance schemes payable on their own behalf by employees, self-employed or non-employed persons. The amounts are recorded on an accrual basis, which for those in work implies the times when the work that gives rise to the liability to pay the contributions is carried out.

Micro data will often be available from income surveys or administrative data, although it has to be checked whether the concepts match those as used in the national accounts. It may be the case that the micro data also include some of the other components reported as social contributions. In that case, it is recommended to correct for the part that does not relate to households' actual social contributions.

10.7.4. Households' social contribution supplements paid (D614P)

This category consists of the property income earned during the accounting period on the stock of pension and non-pension entitlements. The former amount is equal to the property income item D442R, i.e. *investment income payable on pension entitlements*, whereas the second item is part of D441R, i.e. *investment income attributable to insurance policy holders*. However, this last item also includes income on life insurance policies which should not be recorded as part of households' social contribution supplements (as it is fully reflected in the financial accounts). In compiling the distribution for this item, compilers should be aware of these links, trying to derive the distribution on the basis of the relevant matching items.

10.7.5. Social insurance scheme service charges paid

Set against the social contributions is the service fee charged by the unit administering the social security scheme. This may be an explicit or an implicit charge (e.g. equal to the sum of costs incurred by the employer administering the scheme) and this amount should be deducted to arrive at the net social contributions. It is presented as a separate item in the template, although some countries may already reflect the service charge in a lower amount of imputed social contributions or premium supplements in which case no additional correction is needed.

In case the unit administering the social security scheme applies an explicit charge, micro information may be available that can be used to derive its distribution. Otherwise, the allocation to the relevant households should be done on the basis of other information. As the service charge will likely depend on the premium payments, the distribution of the premiums can be used as a proxy for the distribution of the service charge. If information is available that the service charge is equal across participants and information is available on who is participating, this can also be used to allocate the service charge equally across participants.

10.8. Net social contributions received (D61R)

Social contributions may be received by dedicated units running social insurance schemes or by employers in case they provide a social insurance scheme to their employees directly. The latter may also concern households in their role as unincorporated enterprises included in the household sector (see 2008 SNA, §8.16).

When an unincorporated enterprise operates its own employer-related social insurance scheme, any actual social contributions (as paid by the unincorporated enterprise to its employees and then paid back by the employees into the social insurance schemes as operated by the enterprise) is recorded under category D611R. When it provides social insurance benefits directly to its employees a social contribution is imputed under category D612R, equal to the amount of social contributions that would be needed to secure for the same social benefits. In theory, it may also involve *households' actual social contributions* (D613R), although the related amounts are expected to be small. Furthermore, *households' social contributions supplements received* (D614R) are deemed to be irrelevant for the household sector.

10.8.1. Employers' actual social contributions (D611R)

In case of an unincorporated enterprise running its own social insurance scheme for its employees, actual social contributions may be received by the household sector.

It will depend on the set up of income surveys whether specific information is collected on households in their role as unincorporated enterprise. In that case, it may provide insight which households run their own social insurance scheme and it may contain information on the related amounts. It may also be the case that this kind of information is available from business surveys. Furthermore, administrative data sources may provide relevant information.

Ideally, information is available on social contributions, but if that is not the case, information on social benefits paid may be used as a proxy to derive the distribution of actual social contributions received. An alternative would be to look at wages and salaries paid by households in their role as unincorporated enterprise or at mixed income. However, it has to be borne in mind that this will also include unincorporated enterprises that do not operate their own social insurance scheme. Furthermore, it is not expected that there will be a perfect correlation between mixed income and social contributions.

10.8.2. Employers' imputed social contributions (D612R)

In case of unincorporated enterprises providing social benefits directly to their employees, social contributions should be imputed, equal to the amount of social contributions that would be needed to secure for the same social benefits.

The distribution of this item should ideally be based on underlying micro information. As it concerns an imputed item, no information will be available in micro data sources, but it may be the case that information is available on the social benefits paid by households in their role as unincorporated enterprise. This may be available from household or business surveys or from administrative data. This information can be used as a proxy for the distribution of the imputed social contributions. An alternative would be to look at wages and salaries paid by households in their role as unincorporated enterprise or at mixed income. However, as explained above, this will also include unincorporated enterprises that do not operate their own social insurance scheme, and mixed income and imputed social contributions will not be perfectly correlated.

10.8.3. Households' actual social contributions (D613R)

Households may in theory receive actual social contributions in their role as owner of an unincorporated enterprise running its own social insurance scheme for its employees. However, the amounts will usually

be relatively small. The distribution should ideally be based on underlying micro information which may be available from household or business surveys or from administrative data, but if such information is not available the allocation could be based on the distributions of the employers' actual and imputed social contributions received (D611R and D612R) or on the basis of wages and salaries paid by households in their role as unincorporated enterprise or on the basis of mixed income (B3), bearing in mind the caveats as expressed in Sections 10.8.1 and 10.8.2.

10.8.4. Households' social contribution supplements (D614R)

As this item relates to property income earned during the accounting period on the stock of pension and non-pension entitlements, this generally does not appear as part of social contributions received by the household sector in their role as owners of unincorporated enterprises.

10.9. Social benefits other than STiK received (D62R)

Social benefits are current transfers received by households which are intended to provide for the needs that arise from certain events or circumstances, such as sickness, unemployment, retirement, housing, education or family circumstances. This will often be in the form of regular payments but may also be in the form of a lump sum (see 2008 SNA, §8.68).[3] The main social benefits in cash concern pension provision for retirees or widows and permanently disabled. Social benefits may be provided under social insurance schemes or via social assistance. Whereas social insurance schemes require formal participation by the beneficiaries, often linked to employment and usually evidenced by the payment of contributions, this is not the case for social assistance schemes. Eligibility to receive social assistance benefits is not dependent on those kinds of criteria and they are often paid for via general funds such as taxes. Social insurance benefits in kind provided by employers are treated as if they were paid in cash. However, if they are provided under general social security schemes or social assistance, they are recorded under item D63, i.e. *social transfers in kind*.

Due to the increasing importance of pension benefits in the income distribution, also in view of ageing societies as experienced by a lot of countries, the template distinguishes between pension benefits and other social benefits in cash. Compilers are encouraged to compile results according to this breakdown, particularly when there is a lot of user interest for this type of information in the country.

Micro information on social benefits may be available from household surveys and from administrative data. However, with regard to household surveys it has to be borne in mind that the underlying items sometimes suffer from unit or item non-response and from underreporting (see for example Meyer et al. (2009[8])). In that regard, the availability of administrative information on what types of social benefits are received by which households may assist in checking in the plausibility of the underlying micro data and in correcting for any missing information or underreporting of benefits. Although information may not always be available on the exact amounts received by households, register information may indeed be available with information on who is benefiting from what types of benefits.

10.10. Social benefits other than STiK paid (D62P)

As the household sector may include unincorporated enterprises with paid employees, social benefits may also appear as a use for the household sector. This is the case when an unincorporated enterprise operates an employer-related social insurance scheme itself or provides social insurance benefits directly to its employees.

Information for the allocation of this item may be available from household or business surveys or from administrative data. If micro information is lacking, wages and salaries paid by households in their role as

unincorporated enterprise or *mixed income* (B3) could be used as proxy. However, it has to be borne in mind that this will also include unincorporated enterprises that do not operate their own social insurance scheme. Furthermore, it is not expected that there will be a perfect correlation between mixed income and social contributions.

10.11. Net other current transfers (D7N)

Other current transfers are provisions of goods or services from one unit to another without receiving any goods, services or assets in return, other than *current taxes on income and wealth* (covered under D51) and *social contributions and benefits* (covered under D61 and D62). Current transfers are different from capital transfers in the sense that the latter are linked to the acquisition or disposal of a financial or non-financial asset. As they involve a transfer of assets, they are assumed not to directly affect disposable income. In practice, capital transfers tend to be large, infrequent and irregular, whereas current transfers tend to be relatively small and are often made frequently, on a regular basis.

It is not always possible to clearly distinguish between current and capital transfers, and it may be the case that some cash transfers are regarded as a capital transfer by one party involved in the transaction and as a current transfer by the other party. As in the SNA only one recording can be applied to ensure consistency within the framework, this sometimes implies that transfers that may be large and irregular from the viewpoint of a household are still recorded as current transfers, as they are regarded as regular and frequent from the viewpoint of the other party involved. This for example relates to specific insurance benefits that may be received by households. This also gives rise to differences between micro statistics and the national accounts. Whereas the SNA treats all transfers as current as long as it does not involve the disposal or acquisition of an asset, the ICW Framework treats transfers of cash as capital if they are large and irregular, regardless of whether they involve the sale or purchase of an asset (2013[4]). This has to be borne in mind when linking micro data to national accounts results.

In the template *net other current transfers* are further broken down into non-life insurance transactions and miscellaneous current transfers, because of the different nature of these transactions, the fact that the data availability may differ, and due to the fact that their distribution across households may be quite different.

10.11.1. Net non-life insurance claims minus premiums (D72R-D71P)

Non-life insurance policies provide cover against various events or accidents resulting in damage of goods or property, harm to persons, or against financial losses resulting from events such as sickness, unemployment, accidents, etc. (see 2008 SNA, §8.117). For the household sector, this item covers policies taken out by households on their own initiative and for their own benefit, independently of their employers or government and outside any social insurance scheme.

At the aggregated level, e.g. for the household sector as a whole, premiums and claims often cancel out. However, this will normally not be the case at the individual household level and will most likely also not hold for less aggregated groups of households. Therefore, it is important to separately derive distributional results on the premiums and claims to arrive at the net result of non-life insurance transactions. Only if information on the underlying flows is missing, one may directly target the net impact of claims minus premiums, but it has to be carefully checked whether this leads to plausible results. Otherwise, it may be opted to set the benefits equal to the premiums, cancelling out the impact of this transaction on disposable and adjusted disposable income. Alternative is to derive the distribution on the basis of the distribution of disposable income, as a last resort.

As mentioned above, the inclusion of net non-life insurance claims minus premiums as part of disposable income differs from the approach applied by the ICW Framework (OECD, 2013[4]) and by the Canberra Handbook (United Nations Economic Commission for Europe, 2011[9]). They treat premiums

to insure dwellings as intermediate consumption and related claims as capital transfers offsetting the capital loss in the stock of wealth, whereas premiums paid for other types of insurance (e.g. protecting against unemployment, illness, disruption of travel) are recorded as consumption expenditure with claims being treated as negative consumption expenditure affecting the relevant expenditure components in relation to the risk covered (e.g. for housing, health, transport). This means that on the micro side, the information collected corresponds to household expenditure net of private insurance claims. The main reason is that whereas non-life insurance benefits can be regarded as regular flows from the perspective of the insurance company, this is not the case from the viewpoint of individual policy holders. When matching micro and macro variables, adjustments to address the above differences should be considered.

Net non-life insurance premiums

Net non-life insurance premiums comprise both actual premiums and premium supplements payable out of the *investment income attributed to insurance policyholders* (D441). They are recorded on a net basis, thus already deducting the service charge related to the insurance services provided by the insurance company.

Micro information on actual premium payments may be available from household income or budget surveys or from administrative data from insurance companies. This can be used to derive the allocation of the actual premiums across households. If information on premiums is lacking, but information is available on non-life insurance benefits, this may be used as a proxy. The distribution of the premium supplements should be equal to the part of *investment income attributed to insurance policy holders* that relates to non-life insurance policyholders (D441R). See Section 10.5.4 for more information on this specific item.

Net non-life insurance claims

Non-life insurance claims are the amounts payable in settlement of damages that result from an event covered by a non-life insurance policy during the current accounting period. They normally become due at the moment when the eventuality occurs that gives rise to a valid claim under the terms of the policy. Claims are usually treated as current transfers, even when large sums may be involved as a result of the accidental destruction of a fixed asset or serious personal injury to an individual (see 2008 SNA, §8.115-8.121).

Micro information on non-life insurance claims may be available from household income surveys or from administrative data from insurance companies. This can be used to derive the allocation of the benefits across households. If information on benefits is lacking, but information on actual premiums is available, this could be used as a proxy.

10.11.2. Net miscellaneous current transfers (D75N)

Miscellaneous current transfers consist of all other current transfers than recorded under D5, D6, D71 and D72. This may concern current transfers paid to and received from other sectors, both in cash and in kind, such as membership fees, subscriptions, voluntary donations to non-profit institutions serving households, and fines and penalties, as well as current transfers between households. Examples of the latter include regular remittances between members of the same family that do not belong to the same household (e.g. parents supporting children no longer living at home) and transactions related to lotteries and gambling. With regard to the latter, the amounts paid for lottery tickets or placed bets (minus the service charge)[4] and the related winnings are in the SNA regarded as taking place directly between those participating in the lottery or gambling (see 2008 SNA, §8.136).

Whereas current transfers vis-à-vis other sectors will usually be reflected in the national accounts totals, this may not be the case for transfers between households. When focusing on results for the household

sector as a whole this may not be problematic as the related amounts will cancel out. However, this need not be the case when breaking down the household sector into more granular household groups, as the payments and receipts may be expected to concern different types of households. Some household groups may turn out to be net contributors, while other household groups may turn out to be net receivers. In that regard, the DNA results show that the impact of transfers within the household sector may significantly affect income results for specific household groups in some countries. For that reason, it is important that, in case transfers between households are not reflected in the national accounts totals, compilers try to come up with estimates for the purposes of compiling distributional results. The template includes a separate row for other current transfers between households to encourage compilers to explicitly look into this issue.

When looking at underlying data sources for the allocation of the relevant amounts across households, several components are probably well-covered in household surveys. Before matching the micro data with the national accounts totals, it is important to check which specific items are covered in the micro data sources and which items are likely to be missing. This may assist compilers in distributing specific underlying amounts and to impute for specific items that may be missing. This is particularly important if auxiliary information may be available to distribute (some of) the missing items.

Information on other current transfers between households may also be (partly) available from household surveys. If information is missing, it would be important to assess what items are most likely covered under these specific transfers and who are the most likely households or household groups to benefit and the ones most likely to pay. For example, if it concerns transfers from parents to their children that no longer live at their parents' home, the related amounts may be assigned on the basis of socio-demographic information. If it concerns other types of transfers, it may also be possible to allocate the payments and receipts to specific households on the basis of specific socio-demographic information, depending on the type of transfer.

A specific case concerns the transactions related to betting and gambling. As explained above, the related transactions are assumed to take place directly between those participating, except for the service charge which is paid to the unit organising the lottery or gambling. This differs from the treatment in micro statistics where placed bets and purchases of lottery tickets are recorded as consumption expenditure and minor lottery prizes and other winnings are recorded as negative consumption expenditure, whereas large winnings are treated as capital transfers received. Despite the different recording, micro information on the purchase of lottery tickets and placed bets, as well as on the winnings, will usually be covered in micro data sources, with the exception of the part that relates to the service charge. An adjustment is therefore needed when linking micro and macro data on betting and gambling, but the underlying micro data will provide a good proxy for distributing the corresponding amounts from national accounts.

10.12. Social Transfers in Kind (D63R)

Social transfers in kind are goods and services that are provided to households by government and non-profit institutions[5] either free of charge or at subsidised rates (see 2008 SNA, §8.141). Health care and education are the most well-known examples of social transfers in kind, but in-kind goods and services also cover housing, childcare and elderly care. As social transfers in kind can be regarded as a direct alternative to providing households with a cash benefit to purchase associated goods and services themselves, their inclusion in income measures leads to a more comprehensive measure and fairer comparison across countries and over time. In that regard, Tonkin et al. (2014[10]) explain that if in a country A certain services are largely provided by the state, whereas in country B households need to pay for those services directly, all other factors being equal, someone with the same disposable income in country A would have a higher standard of living than in country B. For this reason, social transfers in kind are included in adjusted disposable income as defined in the SNA. The inclusion of social transfer in kind also

leads to a better assessment of income inequality and of the impact of re-distributional policies. For that reason, the DNA work covers both disposable and adjusted disposable income.

Direct information for the distribution of social transfers in kind is usually not available from micro data sources. Therefore, estimates are often based on secondary information. As the distributions and underlying auxiliary data sources may differ across the various types of social transfers in kind, the template distinguishes three categories, i.e. social transfers in kind on health, on education, and other social transfers in kind.

10.12.1. STIK on Health (D63R1)

The first category that is distinguished concerns *social transfers in kind on health*. As direct micro information will usually be lacking, it is recommended to distribute the amounts on the basis of the insurance value approach, according to which an insurance premium equivalence is allocated to the households.[6] In the absence of further information, the latter approach basically comes down to allocating the average per capita STiK for health to each individual. However, it can be refined by segmenting the population based on socio-demographic information and allocating STiK in line with the various needs/provision costs related to each population segment. For example, it is demonstrated that health related spending is highly age dependent, therefore by allocating to each individual health STiK proportionally to the STiK spending by age, the results would be closer to the actual value approach. In that respect, it is recommended to implement a basic scenario following the insurance value approach by relying on as much socio-demographic information as possible to refine individual allocations.

To implement the basic scenario following the insurance value approach, it is suggested to apply the following (minimum) procedure:

- Step 1: Adjust the national accounts total for the part received by institutional households.
- Step 2: Try to find a source providing an estimate of average public health spending by age, and perhaps other categories (e.g. gender).
- Step 3: Impute to each individual the average health care cost of a person with the corresponding age. Each individual is thus assumed to receive a public benefit determined by the average public spending of his/her group, irrespective of whether or not actual use of health care services has been made.
- Step 4: Scale up or down the imputations so that they match the adjusted national accounts totals (as determined in step 1 above).

10.12.2. STiK on Education (D63R2)

The second category concerns social transfers in kind on education. An actual value approach or a modelled approach using socio-demographic information can be used to allocate spending on education to individuals and households. For example, if socio-demographic information is available on age or schooling status/level of education (and whether or not the relevant students are in public education), and STiK spending per capita for all these sub-groups is available, then education related STiK allocations can be made fairly close to the actual value.

10.12.3. Other STiK and other allocation approaches (D63R3)

The remaining category includes all other social transfers in kind such as housing, childcare and elderly care. Although the related amounts will usually be smaller than social transfers in kind on health and education, it is still important that compilers try to find the best auxiliary information to allocate the amounts to the relevant households. As was the case with the other two categories, this could be done on the basis of the actual value approach and the insurance value approach, depending on the available information.

It would be ideal to work on a finer decomposition of other STiK, such as for housing, early childhood education and childcare services, or long-term elderly services, and make imputations separately for each STiK category using all the available socio-demographic information. If for some categories no information is available, it could be opted to allocate the related amounts flatly to all households or individuals. This is to be preferred over a proportional allocation in proportion to (adjusted) disposable income as these transfers in kind do not seem related to income.

References

Coli, A. and F. Tartamella (2010), "Micro-macro integration: survey data on household income for the estimate of the Italian GDP", http://www.ec.unipi.it/documents/Ricerca/papers/dsm/2010/Report-335.pdf (accessed on 24 October 2017). [3]

European Commission et al. (2009), *System of National Accounts 2008*, https://unstats.un.org/unsd/nationalaccount/docs/SNA2008.pdf (accessed on 29 September 2017). [1]

Johns, A. and J. Slemrod (2008), "The Distribution of Income Tax Noncompliance", http://www.bus.umich.edu/OTPR/DITN%20091308.pdf (accessed on 30 October 2017). [6]

Meyer, B., W. Mok and J. Sullivan (2009), "The under-reporting of transfers in household surveys: Its nature and consequences", *NBER Working Paper Series*, No. 15181, http://www.nber.org/papers/w15181 (accessed on 27 October 2017). [8]

Neri, A. and R. Zizza (2010), "Income reporting behaviour in sample surveys", No. 777, Banca d'Italia, https://www.bancaditalia.it/pubblicazioni/temi-discussione/2010/2010-0777/en_tema_777.pdf?language_id=1 (accessed on 11 October 2017). [7]

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. [4]

Pionnier, P. and E. Guidetti (2015), "Comparing profit shares in value-added in four OECD countries: Towards more harmonised national accounts", *OECD Statistics Working Papers*, No. 2015/3, OECD Publishing, Paris, https://doi.org/10.1787/5js0bsm2g0lt-en. [5]

Tonkin, R. et al. (2014), "Social Transfers in Kind in the United Kingdom and Finland: Micro-level measurement and distributional impact Social transfers in-kind in the United Kingdom and Finland: Micro-level measurement and distributional impact 1", http://www.iariw.org/papers/2014/tonkinPaper.pdf (accessed on 29 September 2017). [10]

Tsakloglou, P. et al. (2010), "Distributional effects of non-cash incomes in seven European countries", http://www.iariw.org/papers/2010/7dTsakloglou.pdf (accessed on 29 September 2017). [2]

United Nations Economic Commission for Europe (2011), *Canberra Group Handbook on Household Income Statistics*, https://www.unece.org/fileadmin/DAM/stats/groups/cgh/Canbera_Handbook_2011_WEB.pdf (accessed on 27 September 2017). [9]

Notes

[1] The other elements of investment income disbursements (i.e. the part of investment income attributed to insurance policyholders that relates to life insurance and the investment income attributed to investment

fund shareholders) are also flowing back to the insurance corporation and the investment funds, but these flows are recorded entirely in the financial accounts.

[2] Please note that in the ICW Framework (OECD, 2013[4]) investment earnings received on invested life insurance funds but not immediately distributed to life insurance and pension scheme participants are included in element K03, i.e. *Adjustments to life insurance, annuity and private pension entitlements*.

[3] In case of pension benefits in the form of a lump sum, it depends on the specific requirements of the scheme how this should be treated in the national accounts. If there is a requirement to immediately convert these funds to an annuity, the lump sum should not be recorded as pension benefits receivable immediately upon retirement but as annuity benefits receipts in the relevant recording periods (see 2008 SNA, §17.138).

[4] The service charge is recorded as final consumption expenditure under category CP090. It is defined as the difference between the amount paid for lottery tickets or placed in bets and the amounts paid out to winners.

[5] Social insurance benefits in kind provided by employers are treated as if they were paid in cash.

[6] The actual value approach, according to which the actual values of health benefits are allocated to the various households, is not recommended, although the results at more aggregated levels of detail will come close to the results according to the insurance value approach.

11 Overview of all consumption items

This chapter provides an overview of the national accounts components distinguished on the consumption side, describing their main characteristics and highlighting items from micro data sources that may provide the best possible link. It also explores possible reasons for gaps between the micro aggregates and the national accounts totals and provides guidance on how to arrive at underlying distributions in case micro data is lacking.

11.1. Introduction

This chapter provides an overview of the various national accounts components that are distinguished on the consumption side, based on the Classification of Individual Consumption according to Purpose (COICOP). Two main aggregates are distinguished, i.e. *final consumption expenditure* (P31) and *actual final consumption* (P4), as defined in the System of National accounts (European Commission et al., 2009[1]) (hereinafter referred to as *2008 SNA*). The former measures the amount of household expenditure on consumption goods and services, whereas the latter records the amount of consumption goods and services acquired by households. The difference relates to the treatment of social transfers in kind. These are goods and services provided to households by government and non-profit institutions serving households (NPISHs) either for free or at prices that are not economically significant. To reflect that they are acquired by households, although they are not actually purchased by households, they are included in actual final consumption, which is the counterpart of adjusted disposable income as recorded on the income side.

11.2. Food and non-alcoholic beverages (CP010)

This category covers all food products and non-alcoholic beverages purchased for consumption at home. It excludes food and beverages sold for immediate consumption away from home by hotels, restaurant, cafés, bars etc., as well as cooked dishes prepared by restaurants for consumption off their premises or by catering contractors which are all to be included in *restaurant and hotels* (CP110). It also excludes pet foods which are to be included in *recreation and culture* (CP090). Examples of products that are included in this category are bread, meat, fish, milk, eggs, oils, fruit, vegetables, chocolate, coffee, tea, mineral waters, soft drinks and juices (see for more information category 01 of the COICOP classification).

In addition to products purchased on the market, this category also includes food products and non-alcoholic beverages that are produced by households for own consumption. These are part of household production and the surplus that derives from this production forms part of mixed income. The goods are treated as either being consumed immediately by the household or stored in inventories for later use.

Household budget surveys will normally cover the consumption of food and non-alcoholic beverages, which can be used to allocate the amounts to underlying households. However, this will usually only concern the products as purchased on the market and may not cover the consumption of food and non-alcoholic beverages that are produced for own consumption. In that case, it is recommended to allocate the amounts related to the latter separately, as its distribution may deviate from the one for purchased foods and non-alcoholic beverages. Information may be available on households that are likely to be involved in the production of these goods for own consumption, for example because they own a farm or an allotment. In allocating the related amounts, compilers should be aware of the link between the distribution of the consumption of these goods and of the mixed income related to their own account production. Although their distributions need not be identical (e.g. due to the impact of intermediate consumption), it may be assumed that these will be closely aligned.

11.3. Alcoholic beverages, tobacco and narcotics (CP020)

This consumption category includes alcoholic beverages (including the non-alcoholic variant such as non-alcoholic beer) that are purchased for consumption at home, thus excluding those beverages sold for immediate consumption away from home by hotels, restaurant, cafés, bars etc. Alcoholic beverages that are distinguished in this category are spirits, wine and beer. Furthermore, this category includes all purchases of tobacco, including purchases in restaurants, cafés, bars, et cetera. Finally, it also includes

narcotics, such as marijuana, opium and cocaine (see for more information category 02 of the COICOP classification).

As was the case with food and non-alcoholic beverages, this category includes consumption of own account production, which may not be covered in micro data sources. In that case, the related amounts should be allocated to the relevant households separately. Information may be available on households that are likely to be involved in the production of alcoholic beverages, tobacco and/or narcotics for own consumption, which in that case may be used for its distribution across households. It is recommended to also keep a close link between the distribution of the consumption of these goods and the distribution of the mixed income related to their production, as was mentioned in the previous section.

In addition to consumption of own account produced goods, this category may also include consumption of goods that are part of the non-observed economy. In that regard, the production and consumption of narcotics will often be illegal, and production and consumption of alcoholic beverages and tobacco may partly be hidden from authorities to avoid excise duties. As it is unlikely that the related consumption expenditure will be reported in household surveys, the related amounts will largely have to be allocated separately on the basis of auxiliary information. Assumptions will have to be made which households are most likely to be involved in the consumption of these goods.

The work conducted by the expert group showed relatively large gaps between the micro aggregates and the national accounts totals for this specific category. Whereas this may be partly due to the lack of information of consumption of own account produced goods and of goods produced and consumed as part of the underground economy, this could also be due to underreporting. This is related to the fact that large consumption of these goods is usually regarded as socially unacceptable (see OECD (2013[2])). If information is available on which types of respondents are more likely to underreport, this could be used in better aligning the micro data to the national accounts totals.

11.4. Clothing and footwear (CP030)

This category comprises clothing and footwear, which in addition to clothing materials, garments, shoes and other articles of clothing, clothing accessories and footwear, also includes their cleaning, repair and hiring (see for more information category 03 of the COICOP classification). In addition to newly bought goods, it may also include purchases of second-hand goods, which will normally not be visible in the national accounts (at least for the largest part) as it mainly concerns transactions between households.[1] Box 11.1 explains the role of trade in second-hand goods in the system of national accounts and how it should be dealt with in compiling distributional results.

Household budget surveys usually include information on the consumption expenditure on clothing and footwear. This can be used for the allocation of the amounts to underlying households. It may be assumed that the budget surveys will also include purchases of second-hand goods, but if these are not included, these should be added separately before aligning the results to the national accounts totals. Furthermore, the sale of these products should also be included as negative consumption. It is likely that this is not treated in this way in budget surveys or other micro data sources, so corrections will most likely be needed for these sales. Information will be needed on which households are more likely to involve in the sale of second-hand products to make the appropriate corrections. As explained in Box 11.1 properly accounting for these second-hand sales is very relevant in order to arrive at accurate consumption and saving levels for the various household groups, especially in countries where second-hand trade forms an important part of the economy.

Box 11.1. Treatment of second-hand trade

Most consumption goods and services are normally used up immediately and completely by households, but in some cases, they may be used or re-used over a longer period of time. In those cases, goods can be resold (or transferred) to other households. This is for example the case with clothing and footwear (CP030), certain household equipment (CP050), and vehicles or other means of transportation (CP070). In the national accounts, such re-sales (and transfers) of consumption goods are treated as negative consumption expenditure for the household selling (or transferring) the good, whereas it is recorded as positive consumption expenditure for the household buying (or receiving) the good (see European System of Account 2010 (Eurostat, 2013[3]), §9.47). The purchase (or transfers) should be valued at market price (or original purchase price if no information is available on the market price). In case of a transfer, the same amount should be recorded as current transfer paid and received.

Whereas the trade in second-hand goods between households does not affect the national accounts totals for the household sector (except for any service charge that may be related to the sale) as the related transactions (negative and positive consumption expenditure) cancel out at the aggregated level, this will not be the case when breaking out the household sector in more granular household groups. The sales (or transfers) and purchases (or receipts) of second-hand goods may concern different types of households and therefore no longer cancel out at disaggregated levels. As this may affect household groups in different ways and as the amounts involved may be significant (also in relation to the increased use of digital platforms such as eBay to facilitate second-hand trade), it is important that second-hand trade is accounted for explicitly within the compilation of distributional results.

To explicitly address the issue of second-hand trade, estimates will be needed on the amounts related to second-hand trade for the relevant products involved. These should be added to the national accounts totals that currently only reflect the net purchases of these products (i.e. excluding any trade in second-hand goods) by the household sector. By explicitly showing the sales and purchases (or transfers) of second-hand goods, they can be attributed separately to the relevant households, also depending on how both sales and purchases of second-hand goods are covered in micro data sources.

An example is provided below on how to derive the distribution of the consumption expenditure for a specific good when considering the explicit treatment of the trade in second-hand goods. First, assume that the national accounts aggregate for the household consumption expenditure of good A amounts to 900. This concerns the net purchases of the household sector in which second-hand trade in this good (which in this example is set equal to 200) has been netted out. Furthermore, assume that the household budget survey provides the following information on the consumption of good A by three household groups.

	Aggregate	HH group 1	HH group 2	HH group 3
Consumption expenditure on good A	+1,000	+250	+350	+400

In the example, it is assumed that this information in the budget survey concerns all purchases of good A, so covering both purchases of new and of second-hand goods, but that it does not include a correction for any sales of goods. That means that the numbers reflect the gross purchases of the various second-hand household groups, adding up to 1,000 in this example.

Without any correction for second-hand trade, the distribution of the consumption of this good can be derived by proportionally allocating the gap between the micro and macro aggregates to the three household groups. This would give the following result.

	Aggregate	HH group 1	HH group 2	HH group 3
Consumption expenditure on good A	+900	+225	+315	+360

However, as was explained above, it is assumed that good A is subject to second-hand trade within the household sector for an amount of 200. In order to properly account for the involvement of the three household groups in this second-hand trade, separate information is needed on the sales and purchases involved in the second-hand trade. The table below provides the relevant information for each household group.

	Aggregate	HH group 1	HH group 2	HH group 3
Second-hand sales of good A	-200		-100	-100
Second-hand purchases of good A	+200	+150	+50	0

Household group 1 only purchases second-hand goods, whereas group 3 only sells them. On the other hand, group 2 both sells and purchases second-hand goods.

To properly account for the second-hand trade in the overall distributional results for good A, both the trade in goods with other sectors and the second-hand trade within the household sector should be accounted for. This can be done by distinguishing between the relevant underlying flows. As the micro data are assumed to not only cover purchases from other sectors, but also purchases of second-hand goods, the micro data can be linked to gross purchases in line with national accounts totals. These are equal to the net purchases as recorded in the national accounts (900) plus the purchases of second-hand goods (200), i.e. 1 100. In this case, a proportional allocation of the relevant amounts would lead to the following result for gross purchases.

	Aggregate	HH group 1	HH group 2	HH group 3
Gross consumption expenditure on good A	+1,100	+275	+385	+440

To arrive at the correct distributional results for this item, the sales of the second-hand goods (-200) should also be allocated to the relevant households, as they are treated as negative consumption. Combining the gross consumption expenditure on good A as derived above with the distribution of the sales of second-hand goods, leads to the following results for the overall distribution of the consumption of good A.

	Aggregate	HH group 1	HH group 2	HH group 3
Gross consumption expenditure on good A	+1100	+275	+385	+440
Second-hand sales of good A	-200	0	-100	-100
Net purchases of good A	+900	+275	+285	+340

This treatment has led to significantly different results. The consumption of good A by household group 1 comes out more than 20% higher (275 versus 225), whereas for group 2 and 3 it comes out respectively 8% and 9% lower.

The example deals with a situation in which the micro data are assumed to include purchases of second-hand goods. In that case, these kinds of purchases can be added to the national accounts totals before aligning the micro and macro data (if separate micro information is available on the purchases of new goods and of second-hand goods, it would be better to separately derive their distributions). If the micro data do not cover purchases of second-hand goods, the impact of the second-hand trade should be assessed separately and only added to the results after aligning the micro data to the national accounts totals. On the other hand, if both the impact of purchases and sales of second-hand goods are already reflected in the micro data, these underlying flows should first be distinguished separately in the micro data to arrive at separate distributions for all three underlying components.

11.5. Housing, water, electricity, gas and other fuels (CP040)

This consumption category consists of different consumption items related to housing. As these components may differ in size, link to different micro variables, and may have different distributions across households, the template separately distinguishes them, and compilers are advised to derive results at this detailed level to arrive at best possible results. See category 04 of the COICOP classification for more information on the coverage of this specific item and its underlying components.

11.5.1. Actual rentals on housing (CP041)

This category includes the actual payments for the use of the dwelling including the land on which the property stands. In case of the rental of a furnished dwelling, it also includes the payments for using the furniture. Furthermore, it includes payments for the use of a garage to provide parking in connection with the dwelling. On the other hand, it excludes income receivable by the owner of a natural resource which may be put at the disposal of another institutional unit for use in production. The latter is recorded as rent as covered under item D45 (see also Section 10.5.5). The amounts recorded under this category reflect the actual rents, so before deduction of any allowances as may be received by some households. These allowances should be recorded as current transfer received by the relevant households.

Information on actual rentals is usually available from micro data sources. This can be used as underlying information to derive the distribution in line with the national accounts total. It should, however, be borne in mind that the underlying concept of the related micro variable may differ from the national accounts concept, in that the amount may already have been corrected for any allowances received. In that case, a correction should be made at the micro level for the households that received allowances before aligning the micro data to the national accounts totals. For this purpose, information may be available from administrative data sources.

11.5.2. Imputed rentals on housing (CP042)

This item relates to the own-account production of housing services by owner occupiers (for both main residences and second homes in the domestic economy). As was explained in Section 10.2.1 the rationale is to arrive at more comparable data on the production and consumption of housing services across countries with different ratios of owner-occupied to rented dwellings and for countries in which this ratio may fluctuate a lot over time. Furthermore, the imputation of the production and consumption of housing services by owner-occupiers leads to fairer inequality measures. Without the imputation house-owners and households renting their house may arrive at similar income levels, whereas the house-owners may be far better off with not having to pay for any explicit rent. By imputing for the housing services produced by the house-owner both its income and consumption are corrected for this difference.

The housing services are recorded at market prices, usually based on estimated rental that a tenant would pay for accommodation of the same size, quality and type. This full output value is recorded as produced and consumed by the same household. It also constitutes the input for deriving *operating surplus from owner-occupied dwellings* (B2R1) which is obtained by deducting intermediate costs from these produced housing services. As a consequence, the distribution of the consumption of imputed rentals on housing is usually closely aligned to the distribution of operating surplus from owner-occupied dwellings.

Micro data will usually be available to derive the value of the own-account production of housing services by owner occupiers. For example, surveys may ask whether respondents own the house they live in and/or about the value of the dwelling (often both covering main residences and second homes). Furthermore, information may be available on characteristics of the dwelling (e.g. the number of square meters, type of residence, neighbourhood) that may be used to estimate the imputed rent (see for example Tsakloglou et al. (2010[4])). Also, in some countries the imputed value of the income generated by production of

housing services is taxed, so information may be available from fiscal records. As mentioned above, there is a close relation between operating surplus from owner-occupied dwellings and consumption of imputed rentals on housing. This should be kept in mind when deriving the distribution for both items.

11.5.3. Maintenance and repair of dwellings (CP043)

This item refers to activities that have to be undertaken regularly in order to maintain the dwelling in a good working order and that do not change the dwelling's performance, capacity or expected service life. They can be further broken down into minor and major maintenance and repairs. Only expenditures on materials and services for minor maintenance and repair that are normally the responsibility of a tenant, such as interior decoration and repairs to fittings, are part of this consumption category. Expenditures that owner-occupiers incur on the decoration, minor repairs and maintenance of the dwelling, which would normally be seen as the responsibility of a landlord, are not treated as household final consumption expenditure but as intermediate expenditure in the production of housing services. The same goes for major maintenance and repairs, such as re-plastering walls or repairing roofs. Furthermore, major renovations or extensions to dwellings are recorded as fixed capital formation (see 2008 SNA, §9.66-9.68).

Micro data will often be available on expenditure on maintenance and repair of dwellings, but the coverage may differ from the national accounts concept. The micro variable may for example include expenditure on repairs that according to the national accounts should be recorded as intermediate consumption in the production of housing services. If possible, a correction should be made to align the micro data with the national account concept, possibly also involving a correction to the calculation of operating surplus on owner-occupied dwellings, to which part of the micro concept may relate. In that sense, the correction to the micro data on maintenance and repair of dwellings to exclude the part that relates to intermediate consumption should be reflected in a lower value of operating surplus from owner-occupied dwellings for the relevant households.

11.5.4. Water supply and miscellaneous (CP044)

This category covers expenditure associated to water supply, including hire of meters, reading of meters and standing charges. Furthermore, it includes refuse and sewage collection and disposal, and co-proprietor charges for caretaking, gardening, heating and lighting etc. in multi-occupied buildings.

Underlying distributions may be obtained from budget survey information or from administrative data that may be available from water suppliers. If no micro data is available, the distribution could be derived by using the sum of actual and imputed rentals on housing (i.e. CP041 and CP042) as a proxy.

11.5.5. Electricity, gas and other fuels (CP045)

This category covers expenditures related to electricity, including hire of meters, reading of meters, and standing charges. Furthermore, it includes expenditures related to gas, liquid fuels, solid fuels, and heat energy.

Underlying distributions may be obtained from budget survey information or from administrative data that may be available from energy suppliers. If no micro data is available, the distribution could be derived by using the sum of actual and imputed rentals on housing (i.e. CP041 and CP042) as a proxy.

11.6. Furnishings, household equipment and routine household maintenance (CP050)

This category includes furniture and furnishings, carpets and other floor coverings, including their delivery, installation and repair, as well as household textiles, appliances, glassware, tableware and household

utensils, tools and equipment for house and garden, and goods and services for routine household maintenance. Part of the items may concern consumer durables which are goods that may be used for purposes of consumption repeatedly or continuously over a period of more than a year (see 2008 SNA, §9.42). Section 11.18 provides more information on the treatment of consumer durables in the SNA.

Information on this specific category is usually available from micro data sources. However, the treatment of purchases of consumer durables may differ. In this regard, the ICW Framework (OECD, 2013[2]) treats consumer durables as assets that provide services to their owners, in analogy to the treatment of owner-occupied dwellings. For that reason, they may be excluded from the corresponding micro variables what would necessitate an explicit imputation before aligning the micro data to the national accounts totals. It may be the case that purchases of consumer durables are covered in another item in the micro data source, which may then be used to make the necessary correction. However, it has to be borne in mind that this may also include consumer durables that relate to one of the other consumption items in the DNA approach.

Furthermore, it has to be noted that this category may also include trade in second-hand goods. That also requires specific treatment of the related flows (see Box 11.1).

11.7. Health (CP060)

This category consists of different items related to health care expenditure by households. As these components may differ in size, link to different micro variables, and may have different distributions across households, the template separately distinguishes them, and compilers are advised to derive results at this detailed level to arrive at best possible results. See category 06 of the COICOP classification for more information on the coverage of this specific item and its underlying components.

11.7.1. Medical products, appliances and equipment (CP061)

This category covers medicaments, prostheses, medical appliances and equipment, and other health-related products, intended for use outside a health facility or institution. It includes pharmaceutical products, other medical products (such as clinical thermometers, bandages, first-aid kits and pregnancy tests) and therapeutic appliances and equipment (such as eyeglasses and contact lenses, hearing aids and prosthetic devices), including their repair. It includes products that are paid for on the basis of health insurance benefits received by households but excludes any health care that has been acquired by government or non-profit institutions serving households provided to households for free or at prices that are not economically significant. These are treated as social transfers in kind, which are not part of the *consumption expenditure of households* (P31) but are added to this consumption expenditure to arrive at *actual final consumption* (P4).

Micro information may be available from budget surveys or from other data sources. In linking the data to the national accounts totals, it has to be assessed whether the coverage of the items is similar. The micro concept may for example also include health insurance premiums which are excluded in the national accounts concept (these are covered under item D71P) or amounts related to social transfers in kind. If that is the case, these have to be corrected for, before linking the micro and macro data. It also has to be borne in mind that, as it concerns sensitive information, respondents may not always provide honest answers to questions relating to their health situation. This may lead to gaps between the micro aggregates and the national accounts totals. In those cases, it has to be carefully assessed which households this most likely concerns.

11.7.2. Out-patient services (CP062)

This category covers medical, dental and paramedical services delivered to outpatients by medical, dental and paramedical practitioners and auxiliaries. These may be delivered at home, in facilities, dispensaries or outpatient clinics of hospitals. As was the case with the previous category, this also includes products that are paid for by health insurance benefits but excludes health care that has been provided via social transfers in kind.

Micro information may be available on these expenses from budget surveys or from other data sources. In linking the micro and macro data, it has to be assessed whether the coverage of the items is similar. If the micro concept also includes health insurance premiums or amounts related to social transfers in kind, corrections will need to be made before linking the micro and macro data. Furthermore, also for this category it may the case that micro-macro gaps may be related to reluctance of part of the household population to provide accurate information on their expenses on health care. In that case, it has to be considered which households are most likely to underreport and/or are most likely to consume out-patient services.

11.7.3. Hospital services (CP063)

This category covers the services of general and specialist hospitals, medical centres, maternity centres, nursing homes and convalescent homes that chiefly provide in-patient health care. It also includes the services of institutions serving old people in which medical monitoring is an essential component and the services of rehabilitation centres providing in-patient health care and rehabilitative therapy where the objective is to treat the patient rather than to provide long-term support. The services related to the latter are covered under category *miscellaneous goods and services* (CP120). As was the case with the previous two, this category also includes products that are paid for on the basis of health insurance benefits and excludes health care that has been provided via social transfers in kind.

As for the other health expenditure categories, micro information may be available from budget surveys or from other data sources, but it has to be carefully assessed whether the coverage is the same. If the micro concept also includes health insurance premiums or amounts related to social transfers in kind, corrections will be needed before linking the micro and macro data. Furthermore, micro-macro gaps for this category may be related to reluctance of some households to provide accurate information on their health care expenses. In that case, it has to be considered which households are most likely to underreport and/or are most likely to purchase out-patient services.

11.8. Transport (CP070)

This category consists of different items related to transport expenditure by households. As these components may differ in size, link to different micro variables, and may have different distributions across households, the template separately distinguishes them, and compilers are advised to derive results at this detailed level to arrive at best possible results. See category 07 of the COICOP classification for more information on the coverage of this specific item and its underlying components.

11.8.1. Purchases of vehicles (CP071)

This group covers the purchases of motor cars, motorcycles, bicycles and animal-drawn vehicles, but it excludes recreational vehicles such as camper vans, caravans, trailers and boats, which are covered under *recreation and culture* (CP090). Although all these goods may be considered as assets because they have a life span of over a year and may render services to their owning households over a longer period of time, they are treated as consumption items in the SNA (as long as they are not purchased by households in

their role as unincorporated enterprise) as the services they render are not covered by the production boundary (see Section 11.18 for more information). This may differ from treatment in micro statistics and has to be borne in mind when linking micro data to the national accounts totals. Furthermore, as was explained in Box 11.1, second-hand trade may play an important role in this category. Compilers have to be aware of the impact of this second-hand trade in compiling the distributional results.

Micro data on the purchases of cars will usually be available from budget surveys, even though the treatment may often be different. In linking these micro data to national accounts totals, it has to be borne in mind that the reported micro data may also include purchases of second-hand cars. As explained in Box 11.1, a specific treatment is needed to avoid any errors feeding into the distributional results, as these sales and purchases of second-hand cars within the household sector are not reflected in the national accounts totals. Compilers are therefore encouraged to estimate the gross estimates of car purchases by the household sector and distribute the purchases on the basis of those results and only then process the impact of sales of second-hand cars within the household sector. Directly linking micro data on car purchases to the net purchases of cars by the household sector in the national accounts is expected to lead to inaccurate distributional results.

11.8.2. Operation of personal transport equipment (CP072)

This group covers the purchases of spare parts, accessories, fuels and lubricants for, as well as services for the maintenance and repair of personal transport equipment. It also covers hire of parking space not providing parking in connection with the dwelling (the latter is included in *actual rentals on dwellings* (CP041)), toll and parking meters, and driving lessons and licenses.

Micro data may be available on purchases related to the operation of personal transport equipment, although not all items may be covered in the micro concept. In case the micro concept differs from the macro concept, adjustments may be made to correct for these differences and allocated to the relevant households separately from the part that provides a conceptual match. If no additional information is available on the items that do not correspond between the micro and macro data, a proportional allocation may be appropriate.

11.8.3. Transport services (CP073)

This category covers purchases of services for passenger transport by railway, road, air, sea and inland waterway, also including services of other modes of transportation such as cable car, removal and storage services, and travel agents' commissions. Transport that is part of all-inclusive holidays or tours is not included under this item, but under *recreation and culture* (CP090).

Micro data on purchases of transport services may be available from budget surveys. If no information is available, it is not recommended to use the other two categories as a proxy, as no direct link is assumed between owners of vehicles and the use of passenger transport. In that case, it would be better to look for other types of information that may be available on the use of transport services by households or alternatively link it to aggregated consumption expenditure by households.

11.9. Communication (CP080)

This group covers expenditures related to postal services, telephone and telefax equipment and their services (see for more information category 08 of the COICOP classification).

Micro data may be available from budget surveys, but in case information is lacking, aggregated consumption expenditure by households can be used as a proxy.

11.10. Recreation and culture (CP090)

This category comprises multiple items. First of all, it includes audio-visual, photographic and information processing equipment, such as television and radio sets, cameras, personal computers, records and discs. Secondly, it covers major durables for recreation and culture, including camper vans, caravans, boats, horses and ponies, and musical instruments. Thirdly, it includes other recreational items and equipment (such as games and toys, sport equipment and camping gear), gardens, plants and flowers, and pets and related products (including veterinary and other services for pets). A fourth category comprises recreational and sporting services, such as provided by sport stadiums, fitness centres, amusement parks, cinemas, theatres, museums and national parks. It also includes the hire of equipment and accessories for sport, recreation and culture. Furthermore, it includes costs related to games of chance, such as service charges for lotteries and casinos. The fifth category concerns newspapers, books and stationery, and the final item relates to package holidays. These are all-inclusive holidays or tours which provide for travel, food, accommodation, guides etc. See category 09 of the COICOP classification for more information on the coverage of this specific item and the underlying components.

Micro data are usually available on various aspects of recreation and culture. In case micro data is available at more granular levels of detail, it may be best to derive distributional results at these more detailed levels. In case of large gaps between the micro and macro aggregates, and in case of sub-items for which micro data is lacking, this is expected to gain better results than simply aligning the results at the aggregated level. For example, it may be the case that only one item shows large micro-macro gaps, and that the underlying distribution is completely different from the items that show perfect alignment. Proportionally allocating on the basis of aggregated results would in that case lead to significantly different results from allocating on the basis of the underlying items. Furthermore, analysing gaps between the micro and macro data at this more detailed level may also provide better insight into possible underlying reasons for these gaps and for their proper allocation across households.

11.11. Education (CP100)

This category comprises educational services, but does not include expenditures on educational materials, such as books and stationery. These latter are covered in *recreation and culture* (CP090). Furthermore, it does not include education that has been acquired by government or non-profit institutions serving households to provide to households for free or at prices that are not economically significant. These are treated as social transfers in kind which are not part of the consumption expenditure of households but are added to arrive at actual final consumption (see for more information category 10 of the COICOP classification).

Micro information is usually available from budget surveys or from data sources on education. In linking the data to the national accounts totals, it has to be assessed whether the coverage of the micro items correspond to the national accounts concepts. If the micro data for example include amounts related to social transfers in kind, this has to be corrected for before linking the micro and macro data.

11.12. Restaurant and hotels (CP110)

This category covers catering services provided by restaurants, cafés, bars, canteens etc., as well as accommodation services provided by hotels, boarding houses, motels, holiday villages, boarding schools, universities and hostels for young workers or immigrants, etc. However, with regard to the latter, it excludes rentals paid by households for secondary residences, which should be recorded as rentals for housing (CP041 and CP042). See category 11 of the COICOP classification for more information on the coverage of this item.

Micro information is usually available from budget surveys, although the underlying concepts may differ from the COICOP definition. The micro data may for example include rent for secondary residences or costs related to package holidays. In that case, adjustments will be needed to arrive at a better conceptual alignment before linking the micro data to the national accounts totals.

11.13. Miscellaneous goods and services (CP120)

This category consists of all consumption items not covered by all previous categories and includes a wide range of categories. As these underlying items may differ in size, link to different micro variables, and may have different distributions across households, the template separately distinguishes them and compilers are advised to derive results at this detailed level to arrive at best possible results. See category 12 of the COICOP classification for more information on the coverage of this item and its underlying components.

11.13.1. Insurance expenditures (life and non-life) (CP125)

This category covers the service charges for insurance, such as life insurance and insurances connected with the dwelling, health and transport. It does not include the actual insurance premiums or benefits, which should be recorded on the income side, under net non-life insurance premiums (D71P) and net non-life insurance benefits (D72R) respectively.

As the payment for the service is normally included in the insurance premium payment, no separate information will usually be available from micro data sources on the actual service charge. In that case, the national accounts total should be allocated to the relevant households on the basis of auxiliary information. As there is a direct link between the service charge and the insurance premiums, it is recommended to use the sum of the life and non-life insurance premiums (if possible, including the premium supplements related to the investment income attributable to insurance policy holders) as a proxy for the distribution of the service charge. If separate information is available on the service charges for the different types of insurance as well as on premiums paid by households at this more detailed level, this can be used to arrive at more accurate distributions.

11.13.2. FISIM (CP1261)

This category concerns the part of *financial intermediation services indirectly measured* (FISIM) that is consumed by household as part of their individual consumption expenditure. As was explained in Section 10.5.1, in addition to charging customers directly for financial intermediation services, financial corporations also generate service income by lending at a rate of interest higher than the one at which they borrow. This is an indirect charge for the service of bringing together borrowers and lenders. That means that the interest that is charged by financial corporations on loans also includes a service charge and that a service charge is already withheld on interest paid out on deposits. This service charge is calculated on the basis of both the difference between the interest rate received on loans and a reference rate, and the difference between a reference rate and the interest rate paid on deposits (see also 2008 SNA, §A3.24-A.3.27). The actual interest flows are corrected to exclude these indirect service charges, which have to be recorded as part of consumption reflecting the payment for the indirect intermediary service provided by financial corporations. Depending on the underlying interest payment, it will be recorded as intermediate consumption or as final consumption. If FISIM relates to mortgage loans, it should be recorded as intermediate consumption in the production of housing services related to owner-occupied dwellings, reflected in a lower value of *operating surplus* (B2). If FISIM relates to business loans owed by households or deposits held by unincorporated enterprises, it should be recorded as intermediate consumption reflected in a lower value of *mixed income* (B3). For all other deposits and loans, the related FISIM should be recorded as part of final consumption (CP1261).

As FISIM is a specific SNA concept, no direct micro data will be available on its distribution. Therefore, as was explained in Section 10.5.1, it is recommended to derive the distribution of FISIM on the basis of the actual interest payments and receipts by households. If detailed information is available on the different actual interest rates for various types of deposits and loans and if these can be linked to underlying households or specific household groups, this can be used to allocate FISIM at a detailed level. If such information is lacking, one could assume equal margins for all types of deposits and loans, allocating FISIM proportionally to the aggregated absolute amount of interest payments and receipts by households or household groups.

The calculation of the FISIM correction at the income side provide the starting point for the calculation of the related FISIM items at the consumption side: as explained above, the FISIM that relates to mortgage loans should be recorded as intermediate consumption in the production of housing services related to owner-occupied dwellings and be reflected in a lower value of *operating surplus* (B2); the FISIM that relates to business loans taken out by households or deposits held by household in their role as unincorporated enterprise should also be recorded as intermediate consumption but now reflected in a lower value of *mixed income* (B3); and the FISIM related to all other deposits and loans should be recorded as part of final consumption (CP1261). As there is a direct link between the consumption of FISIM and the correction item as recorded in the property income block, one should make sure that this link is maintained in the distributional results. This means that the sum of FISIM recorded as final consumption (CP1261) and as intermediate consumption (reflected in lower values of B2 and B3) should equal the total amount of FISIM recorded as correction items in property income (FISIM_R and FISIM_P) at the level of the household (or household group). The breakdown into type of consumption should ideally be made on the basis of information on the types of deposits and loans or type of interest receipt and payments. If that type of information is not available, assumptions should be made to break it down into these three types of consumption.

11.13.3. Miscellaneous (less FISIM, less insurance) (CP12x)

This category includes consumption expenditure that is not covered by any of the other categories. Items that are covered in this category are: personal care (e.g. hairdressing salons and appliances and articles for personal care); prostitution; jewellery, clocks and watches; other personal effects (e.g. travel goods, articles for babies, miscellaneous personal articles); social protection (i.e. assistance and support services provided to persons who are elderly, disabled, having occupational injuries and diseases, etc.); and other services not classified elsewhere such as fees for legal services, charges for undertaking, payments for the services of estate agents, and payments for advertisements.

Micro data may be available on various underlying categories. In case micro data is available at more granular levels of detail, it may be best to derive distributional results at these more detailed levels when they can be matched to corresponding national account results. This will be particularly relevant if this category turns out to be substantial. In case of large gaps between the micro and macro aggregates, and in case of sub-items for which micro data is lacking, it is expected to gain better results than simply aligning the results at the aggregate level. It may for example be the case that only one underlying item shows a large gap between the micro and macro results, and that the underlying distribution is completely different from the items that show perfect alignment. Proportional allocation on the basis of aggregated results would in that case lead to significantly different results from allocation on the basis of the underlying items. Furthermore, analysing gaps between the micro and macro data at this more detailed level may also provide better insight into possible underlying reasons for these gaps and for their allocation across households.

When looking at the coverage of the various items in micro data sources, it is expected that no information will be available with regard to prostitution which is usually part of the non-observed economy. For that specific item, the distribution will have to be estimated on the basis of other information or in proportion to

aggregated final consumption expenditure. In case information is also missing for other items, compilers are recommended to look for relevant information that can be used for the allocation of the related amounts.

11.14. Adjustment for expenditures by resident households abroad (and expenditures by non-residents on the territory) (P33 and P34)

As was explained in Chapter 2, in most countries, national accounts data for household final consumption expenditure by type of goods and services refer to all final consumption on the territory. This means that they include the amount of final consumption expenditure of non-resident households on the territory and exclude expenditure of resident households abroad. In that case, corrections are needed to arrive at results for total final consumption of resident households,[2] which should be the starting point for the DNA calculations (see also Section 4.4). A first correction concerns separating out the consumption expenditures of non-resident households on the economic territory. A second correction concerns the inclusion of consumption expenditure of residents abroad. Both corrections can be done at the detailed level of consumption components, in which case no additional corrections will be needed at the aggregated level. Alternatively, one or both of the corrections can be applied at an aggregated level, in which case aggregated correction items will be needed in the distributional template to reflect the impact of these corrections on the distributional results.

As budget survey data only apply to resident households, it is recommended to correct for the expenditure of non-residents at the detailed level of consumption items. Alternatively, this can be done at the aggregated level via item P34, but there will be no good approximation of how to allocate this correction to underlying households, for which the detailed consumption amounts will then incorrectly include expenditure by non-residents.

With regard to the correction for the inclusion of the expenditure of resident households abroad, it is recommended to include this at an aggregated level. However, if the micro data include these consumption expenditures at the detailed level, it may be preferable to also correct the national accounts totals for these expenditures at the detailed level, as this may lead to a better match and to better results. This may be done on the basis of tourism satellite accounts or on the basis of available micro data. For international consistency, some adjustments will then be needed at the end of the compilation process, to move the amounts that relate to expenditure abroad at the detailed level to the aggregated correction item. In this way, results can be provided in the line with the internationally agreed template. It can also be decided to already make this correction at the micro level, but this will mainly depend on what is expected to provide the most reliable results.

11.15. Social transfers in kind (D63P)

As was explained in Section 10.12, social transfers in kind are goods and services that are provided to households by government and non-profit institutions either free of charge or at subsidised rates. Health care and education are the most well-known examples of social transfers in kind, but in-kind goods and services may also cover housing, childcare and elderly care. As social transfers in kind can be regarded as a direct alternative to providing households with a cash benefit to purchase associated goods and services themselves, they are included in the measure of adjusted disposable income to arrive at a more comprehensive income measure and to provide for a fairer comparison across countries and over time. In line with this more comprehensive income measure, the national accounts also include a consumption measure that reflects that these goods, although the expenditure is actually undertaken by government and NPISHs, are consumed by households. In this regard, actual final consumption of households reflects the goods and services acquired by households for the satisfaction of their needs regardless of whether they are purchased by households or obtained by transfer from government units or NPISHs.

The item is equal to the one recorded as part of adjusted disposable income (see Section 10.12) and the distribution should be equal to the one derived in that context.

11.16. Change in net equity of households in pension funds (D8)

In order to arrive at household saving, consumption expenditure has to be subtracted from income. However, because of the way pension contributions and benefits are treated in the SNA, an adjustment is needed for the change in pension entitlements. This is to reflect that whereas these transactions are included in household income, they also affect the saving of households, increasing their pension entitlements as a consequence of pension premium payments and decreasing their entitlements via the receipt of pension benefits.

The adjustment item D8 covers those parts of social insurance schemes for which the liabilities are recognised in the SNA (i.e. employment-related social insurance schemes). Pensions due under social assistance and social security schemes are excluded, as they do not lead to an accrual of pension entitlements for households (2008 SNA, §9.20-9.25). Furthermore, the item does not cover individual pension plans (e.g. life insurance), as these are not regarded as social insurance (thus not affecting disposable income) and are only recorded in the financial accounts.

The adjustment for the change in pension entitlements is equal to:

- the total value of the actual and imputed social contribution payable into pension schemes (parts of D611P, D612P and D613P that relate to pensions),
- plus the total value of contribution supplements payable out of the property income attributed to pension fund beneficiaries (part of D614P that relates to pensions),
- minus the value of the associated service charges (part of CP125 that relates to pension schemes),
- minus the total value of the pensions paid out as social insurance benefits by pension schemes (part of D62R that relates to pensions).

As the adjustment for the change in pension entitlements is the result of various components with their own specific distributions across households, one should make sure that the distribution for this component is derived on the basis of these underlying items. Ideally, this is done at the level of the household or at the level of detailed household groups, but as a minimum it is recommended to distinguish between the active population (for which actual pension contributions (part of D61P), related service charges (part of CP125) and their part of the premium supplements are relevant), and the retired population (for which pension benefits (part of D62), related service charges (part of CP125) and their part of the premium supplements are relevant). This allows for a better link with the main underlying flows.

11.17. Taxes less subsidies on production and imports (memorandum item)

Although taxes and subsidies on production and imports as paid or received by households are already reflected in the relevant income and consumption amounts, there is a lot of user interest in separate information on these items, in order to obtain more insights in the impact of government policies on various household groups. For that reason, the relevant items are included in the template in the form of memorandum items.

Taxes are “compulsory, unrequited payments, in cash or in kind, made by institutional units to government units” (2008 SNA, §7.71). Unrequited refers to the fact that the government provides nothing in return to the individual unit making the payment. *Taxes on production and imports* (D2) consist of *taxes on products* (D21), which are payable per unit of the product, and *other taxes on production* (D29), which are imposed on the producer and do not apply to the products nor to the profits of the producer (see 2008 SNA, §6.50).

Subsidies are “current unrequited payments that government units, including non-resident government units, make to enterprises on the basis of the levels of their production activities or the quantities or values of the goods or services that they produce, sell or import” (2008 SNA, §7.98). Similar to taxes, *subsidies* (D3) consist of *subsidies on products* (D31) and *other subsidies on production* (D39).

Taxes and subsidies on products will be reflected in the purchasers’ prices of the relevant products and, hence, already be recorded in the template under the relevant items. For example, any taxes on products borne by households in their capacity as final consumers will be reflected in the amounts of the relevant consumption categories. Other taxes and subsidies on production are not related to the purchase of products but rather to the production of products and will be reflected in measures of gross operating surplus (in case the taxes and subsidies relate to owner-occupied dwelling services) and/or mixed income. For example, any other taxes on production borne by the household in its role as producer in a specific industry will have led to a lower value of mixed income.

For the purposes of the DNA work, only taxes (and subsidies) on products and production that are explicitly paid (and received) by the households sector are considered. While it is possible that part of the taxes (less subsidies) on products and production paid (and received) by other sectors may be passed on and, in the end, get paid (or be received) by households (as final consumers), to assume this in its entirety not only ignores the components of final demand by the government sector, non-profit institutions serving households and non-residents, but it also removes the ability for the statistical outputs to properly reflect the behaviour of organisations in response to changes in taxation policy.

The total amount of taxes less subsidies on products as paid by the household sector as part of their final consumption expenditure is normally available from input-output tables.[3] However, a direct estimate of other taxes less subsidies on production is normally not available from the national accounts, as this information is only available by industry in input-output tables and supply-and-use tables. In that case, an estimate may be derived by multiplying the taxes less subsidies for each industry by the share of mixed income in the sum of gross operating surplus and mixed income for the relevant industry (except for the industry *imputed rents of owner-occupied dwellings* for which the full amount of other taxes less subsidies on production relates to the household sector).

Information on taxes and subsidies on products will normally be lacking from household surveys as the amounts are concealed in the overall consumption expenditure amounts. Information may be available from tax records, but it may be difficult to link this to purchases made by specific (groups of) households. This means that a proxy will be needed to arrive at distributional results for taxes and subsidies on products. On the other hand, micro data sources may cover information on other taxes and subsidies on production, as these may be explicitly paid by households as part of their production activities. If this information is lacking, a proxy will be needed to arrive at a proper allocation of the amounts to the underlying households.

With regard to a proxy for the allocation of taxes and subsidies on products, the recommendation is to allocate the amounts on the basis of consumption expenditure, ideally at the level of detailed product categories as different rates may apply to different product groups. The latter would require calculating the total amount of taxes less subsidies paid by the household sector per COICOP category, which would imply breaking down the total from the input-output tables (as mentioned above) into underlying consumption categories. This may be done by using information from supply table on the total of taxes (less subsidies) by products and deriving the part paid by the household sector on the basis of the share of household final consumption in the total supply for each product category, where necessary adjusting the amounts to align to the total as obtained from the input-output tables (for example adjusting the results via a proportional adjustment).[4] The results at product level may be transferred into results at COICOP level (e.g. via a concordance table) and, subsequently, these amounts may be allocated to the underlying households on the basis of their share in total consumption for each of the categories.

If no micro data is available to allocate the amount of other taxes and subsidies on production, this amount may be allocated to households in proportion to their gross mixed income and, where relevant, gross operating surplus. If more granular information is available on the specific production activity of the various household, this may be combined with information of the specific amount of taxes (less subsidies) paid per industry to arrive at more accurate results.

As explained, the information on taxes less subsidies on production and imports has been included in the form of memorandum items on the consumption sheet. It is broken down into specific subcategories and aggregations, so that compilers can choose at what level of detail to report this information, dependent on the available information and the methodology chosen to derive the results.

11.18. Consumer durables (memorandum item)

Whereas expenditure on consumer durables is already captured under the relevant consumption items, it may be useful to show it separately as it may significantly affect saving results and might explain negative savings for specific households in certain years. For that reason, the template includes a memorandum item for *consumer durables*, broken down into *vehicles* and *other consumer durables*.

Consumer durables concern goods that may be used repeatedly or continuously over a year or more. Examples are vehicles, furniture, and electrical equipment such as washing machines. In case that these goods are purchased for business purposes (by households in their role as unincorporated enterprise or by other units engaged in production) they are classified as gross fixed capital formation. However, when they are purchased for personal use of household members, they are classified as final consumption expenditure. This relates to the asset boundary of the SNA that states that the coverage of assets is limited to those assets that are used in economic activity as defined in the SNA. As the production of domestic services by households for their own final consumption (with the exception of own-account production of housing services by owner-occupiers) are excluded from the production boundary, durables that are used in the production of these services are also excluded from the asset boundary and their purchases are recorded as part of consumption expenditure (see 2008 SNA, §3.46-3.47).

While the treatment of purchases of consumer durables looks relatively straightforward, it often turns out to be more complicated in practice, especially when certain durable goods are used partly for business purposes and partly for personal benefit (e.g. a car that is used both as a taxi and for personal rides by the household). In those cases, the expenditure on the purchase of the durable should, in theory, be split between gross fixed capital formation and household consumption expenditure in proportion to its usage for business and personal purposes (see 2008 SNA, §9.60). However, whereas this may seem easy from a theoretical point of view, it often turns out to be very complicated from a practical point of view.

The purchase of consumer durables is often treated differently in micro statistics. The ICW Framework (OECD, 2013[2]) treats consumer durables as assets that are used in the production of household services for their own final use. That also implies that according to this framework the surplus from unpaid domestic household services is regarded as an element of household income. As a consequence, it recommends recording purchases of consumer durables as gross fixed capital formation by households instead of as consumption expenditure.

The different treatment of consumer durables in micro statistics and in national accounts may lead to differences in coverage of the national accounts consumption items and related micro variables, mainly affecting furnishings, household equipment and routine household maintenance (CP050) and purchases of vehicles (CP071). Information on consumer durables may be covered in other items in the micro data, but if no micro data is available imputations will be needed to arrive at an appropriate distribution of the related amounts. As these may be significant, ideally auxiliary information can be used for a proper allocation.

It also has to be borne in mind that consumer durables are often goods that may be subject to second-hand trade, because of their relatively longer life span. For that reason, it is important to also keep track of any second-hand trade that may take place for these goods. Again, as the related amounts may be substantial, a correct recording of the sales and purchases of second-hand trade is very relevant. For some of these consumer durables, second-hand trade may take place via well-organised markets or via websites (e.g. trade in second-hand cars), from which relevant information may be derived that can be used as input to derive the relevant distributions. See Box 11.1 for more information on how to deal with second-hand trade.

References

European Commission et al. (2009), *System of National Accounts 2008*, https://unstats.un.org/unsd/nationalaccount/docs/SNA2008.pdf (accessed on 29 September 2017). [1]

Eurostat (2013), *European system of accounts 2010*, https://doi.org/10.2785/16644. [3]

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. [2]

Tsakloglou, P. et al. (2010), "Distributional effects of non-cash incomes in seven European countries", http://www.iariw.org/papers/2010/7dTsakloglou.pdf (accessed on 29 September 2017). [4]

Notes

[1] In some cases, it may concern transactions involving other sectors or the rest of the world. In those cases, the transactions would indeed show up in the accounts, although it is expected that the amounts will normally be very small.

[2] If the national accounts data already exclude consumption of non-residents on the territory and include consumption of residents abroad, no corrections will be needed. If one of the two conditions does not hold, this would require adjustments to bring the data in line with final consumption of resident households.

[3] If compilers would also like to include an estimate of any taxes less subsidies as paid by the household sector as part of their gross fixed capital formation (GFCF), this may be obtained by multiplying the total amount of taxes less subsidies on products on GFCG by the share of household GFCF in total GFCF for the economy as a whole. Although GFCF is currently not included in the DNA work, the inclusion of any taxes paid in relation to GFCF would provide a more complete picture of the impact of government policies on various household groups. Ideally, distributional estimates would be available for GFCF to use for distributing the relevant amounts to underlying households. Alternatively, estimates may need to be made to distribute the amounts, bearing in mind that most GFCF will relate to housing and to machinery that may be needed in production.

[4] While the amounts may differ, they should normally come relatively close. Significant differences may be due to country specific peculiarities and may point to the need for a slightly different approach.

12 Consistency and plausibility checks

The quality of the distributional results largely depends on the quality and alignment of the micro data and the national accounts totals. Furthermore, assumptions may play an important role, for example in imputing for any missing elements, closing gaps between the micro and macro data, and linking data across different data sets. In this light, it is important to closely assess the consistency and plausibility of the results at the end of the compilation process. This chapter provides an overview of checks that may be conducted in this regard.

12.1. Introduction

The previous chapters have explained how national accounts data may be combined with micro data to arrive at distributional results on income, consumption and saving in line with national accounts. It is clear that the quality of these distributional results will be highly dependent on the quality and the alignment of these two sources of information, the size of the items for which micro data is lacking and for which compilers have to rely on assumptions, and the quality of the matching of data across different data sets to arrive at homogenous results for households or household groups. With this in mind, it is important to properly assess the consistency and plausibility of the results at the end of the compilation process. This chapter provides an overview of checks that may be conducted in this respect.

Section 12.2 discusses general approaches to check the consistency and plausibility of the distributional results, both at the level of underlying items and at the level of balancing items. As the distributional results are the outcome of a step-by-step approach, many issues may underlie possible inconsistencies and implausibilities in the results. For that reason, it is important to have insight into how these results have been constructed and how different assumptions in the compilation process (e.g. regarding the allocation of micro-macro gaps or the allocation of imputed items across the distribution) may have led to different results. Section 12.3 presents reconciliation tables that may provide this type of insight. The chapter ends with some conclusions in Section 12.4.

12.2. Checking consistency and plausibility of distributional results

This section discusses approaches that may be used to check the consistency and plausibility of the distributional results. This concerns specific checks that can be applied to certain items due to direct or indirect links with other items as well as general checks that can be applied to all items and aggregates in the DNA work.

12.2.1. Checking internal consistency of distributional results

A first check is to look at the consistency of the distributional results for specific items that have a direct or indirect link to other items in the DNA work. Direct links may result from the setup of the System of National Accounts (SNA) in which some items should be identical by definition (e.g. *employers' actual social contributions* received by households as part of their *compensation of employees* (D121R) and paid by households as part of their *social contributions* (D611P) are equal by definition) and in which other items may constitute an input in another component in the framework (e.g. *property income attributed to non-life insurance policyholders* as received by households (D441AR) forms part of *insurance premiums paid* by households (D71P)).

Table 12.1 provides an overview of direct conceptual links between income and consumption items inherent to the setup of the SNA. The table presents the conceptual link and includes a short explanation. Please refer to the relevant sections on the related items in Chapters 10 and 11 for more information on these conceptual links. Compilers should make sure that these links are respected in compiling distributional results for the relevant items. If they encounter any inconsistencies, they should check where this inconsistency stems from (see also Section 12.3) and try to solve it on the basis of the underlying information.

Table 12.1. Direct links between variables in the system of national accounts

Overview of the conceptual links between income and consumption items in the system of national accounts.

	Conceptual link	Explanation
Employers' actual social contributions	D121R = D611P	These contributions are presented as being paid by employers to their employees as part of compensation of employees which are then paid by these employees into social insurance schemes (see §8.83 of the 2008 SNA).
Employers' imputed social contributions	D122R = D612P	These contributions are presented as being paid by employers to their employees as part of compensation of employees which are then paid by these employees into social insurance schemes (see §8.84 of the 2008 SNA).
Property income received, attributed to non-life insurance policyholders	D71P ≥ D441AR	The investment income on reserves related to non-life insurance policies is treated as income to policyholders which is then treated as paid back to the insurance corporation as premium supplement (see §7.142 of the 2008 SNA.
Investment income payable on pension entitlements	D614P ≥ D442R	The investment income payable on pension entitlements is treated as being received by policy holders in the form of property income which is then treated as paid back to the pension fund as social contribution supplement (see §8.86 of the 2008 SNA).
FISIM	CP1261 ≤ D41R_FISIM + |D41P_FISIM|	The amount of FISIM as recorded as correction item for property income should be allocated to intermediate consumption and final consumption for the relevant households. This means that the sum of FISIM recorded as final consumption (CP1261) and as intermediate consumption (reflected in lower values of B2 and B3) should equal the amount of FISIM recorded as correction item in property income at the level of the household (or household group). As a consequence, the consumption item cannot exceed the total amount of FISIM at the household level.
Change in net equity of households in pension funds	$D8 = D61P_{pensions} - D62R_{pensions}$	The adjustment for the change in pension entitlements is derived on the basis of a formula which should hold at the level of the household and consequently also at the level of household groups (see §9.24 of the 2008 SNA).
Social insurance scheme service charges	CP125 ≥ D61xP	Social insurance scheme service charges are presented as a correction item in D61 as they are not part of the social contribution but concern (indirect) payments insurance services. As this consumption item also includes services charges for other types of insurance, the consumption item should at least be equal to the social insurance scheme service charge (see also §8.82 of the 2008 SNA).
Net lending/net borrowing	Saving + net capital transfers + net capital accumulation = net lending/net borrowing from financial accounts	Net lending/net borrowing as derived from the income and capital account should equal net lending/net borrowing as derived from the financial account (see §11.1 of the 2008 SNA).

In addition to links between specific items, the setup of the system of national accounts also ensures consistency between resources and uses at the level of the various sectors and subsectors in the system. This identity is usually checked by confronting *net lending/net borrowing* as derived from the current and capital account with that from the financial account. These results should be identical but may differ in practice due to the use of different data sources and different compilation techniques. Any inconsistencies will be reflected in "statistical discrepancies". The smaller these discrepancies, the better the alignment between the current, capital and financial accounts, and the higher the plausibility of the underlying results.

For this reason, if distributional information is also available for the capital[1] and the financial accounts, it is recommended to confront these results with the results derived for income and consumption at the level of the household or the household group. If these show discrepancies, this may point to issues in the underlying data. This check is particularly important as results for various countries showed large negative saving rates for low-income household groups. Whereas this may reflect economic reality, it may also point to weaknesses in the compilation process. Confronting these results with distributional information from the capital and financial accounts may provide more insight into the plausibility of these results. Box 12.1 describes the issue of negative saving in more detail.

Finally, direct checks can be performed on the fact that several items cannot become negative,[2] while others can assume negative values by definition. More specifically, in the DNA template, the following items on the income side can assume negative values as a result of the compilation process: *operating surplus* (B2R),[3] *mixed income* (B3R),[4] *net property income* (D4N), *reinvested earnings on foreign direct investment* (D43R), *balance of primary incomes* (B5), *net other current transfers* (D7N), *net non-life insurance claims minus premiums* (D72R-D71P), *net miscellaneous current transfers* (D75N), *transfers between resident households* (D75x), *disposable income* (B6) and *adjusted disposable income* (B7). On the consumption side, it concerns items that may be affected by second-hand trade, such as *purchases of vehicles* (CP071). Finally, in deriving results for saving, it has to be borne in mind that the *adjustment for the change in pension entitlements* (D8) and *gross saving* (B8) may also be negative. When looking at the template, it is also important to note that, as discussed in Section 10.5.1, *FISIM* related to interest paid (D41P) should be reported with a negative sign. All other items in the template should be reported with a positive sign, regardless of the fact whether they positively or negatively contribute to the balancing items.

In addition to direct links, compilers may also focus on the consistency and plausibility of the distributional results by looking at indirect links between certain items. In that regard, some items may be expected to correlate with others from an economic point of view. Current taxes on income may for example be expected to correlate with primary income. Furthermore, property income items will usually be closely linked to the assets owned and liabilities owed by the specific households or household groups. Table 12.2 provides several examples of indirect links. It will depend on country-specific characteristics how strong these links are and on data availability (e.g. distributional information on wealth will not be available in all countries) and quality (e.g. some types of assets or liabilities may also suffer from large gaps between micro and macro results) which checks may be performed. Furthermore, compilers may also identify other indirect links that may be relevant at the national level.

Although no immediate conclusions can be derived with regard to (in)correctness of some of the distributional results on the basis of these indirect links, they may often provide insight into their plausibility. Particularly in case historical distributional results show strong correlation between certain items, which is no longer present (or less strong) in the current results, this may point to an issue in the data. The same goes if two related items move in opposite directions. Furthermore, if other countries show a strong correlation between specific items whereas this is not the case for the country at stake, this may also raise questions whether there is a specific explanation for this. It does not necessarily imply an error in the data, but it may trigger the compiler to try to assess the rationale underlying this difference.

Table 12.2. Indirect links between variables

Overview of indirect links between items that may provide insight into possible implausibilities in the distributional results.

	Economic link	Explanation
Link between wages and salaries and employers' actual social contributions	D121R versus D11R	Often social contributions will be linked to wages and salaries of employees. It may not always concern a fixed percentage, but at least a correlation may be expected. Information may be available at the country-level on the exact link.
Link between financial intermediation services and interest receipts and payments	D41R_FISIM versus D41R and D41P_FISIM versus D41P	FISIM implies that the actual interest received from and actual interest paid to financial corporations includes a correction for the service provided by the financial corporation. As a consequence, there is a link between the actual interest flows and the amount of FISIM.
Link between taxes paid and primary income	D5P versus B5	Current taxes on income and wealth will relate to the income received by households. It may depend on specific tax rules in a country how the two relate and which specific items should be taken into consideration, but at least it is expected that this item will relate to primary income received by households.
Link between operating surplus from housing services produced and consumption of these housing services	B2R1 versus CP042	Housing services produced by homeowners are consumed by the same households. They also constitute the main input for deriving operating surplus from owner-occupied dwellings (B2R1). This is obtained by deducting intermediate costs from these housing services. As a consequence, the distribution of the consumption of imputed rentals on housing is usually closely aligned with the distribution of operating surplus from owner-occupied dwellings.
Link between property income flows and financial stock data	D41R versus interest-bearing assets (AF.2, AF.3, AF.4 and AF.8) D41P versus interest-bearing liabilities (AF.4 and AF.8) D42R versus equity holdings (AF.51) D441AR versus non-life insurance technical reserves (AF.61) D441BR versus life insurance technical reserves (AF.62) D442R versus pension entitlements (AF.63) D443R versus investment fund shares (AF.52)	Property income is the income receivable by the owner of a financial asset or a natural resource in return for providing this asset to another unit. The return may often depend on the type of underlying asset, but it is expected that there is also a strong correlation between the amount of the assets and the income received on these assets. This holds for a range of property income flows received and paid by households.
Link between net social contributions received, social benefits paid, and mixed income	D61R and D62P versus B3R3	Net social contributions received and social benefits paid by households relate to their role as unincorporated enterprises which run their own social insurance scheme for their employees. As a consequence, there is a link between these amounts and the households that include an unincorporated enterprise. An indication of the latter is formed by the amount of mixed income recorded for a specific household or household group, although there will not be a perfect correlation between the two.
Link between non-life insurance premiums and non-life insurance technical reserves	D71P versus non-life insurance technical reserves (AF.61)	It may be expected that the amount of non-life insurance technical reserves will depend on the amount of non-life insurance premiums that are paid by policy holders. As a consequence, the two may be expected to be correlated.

Box 12.1. Plausibility of negative saving rates

Negative saving rates have proved to be a common area of focus when evaluating the plausibility of distributional results. In that regard, several countries have reported negative saving rates for the lowest income quintile and some even for the second and/or third income quintile. These results may be viewed with some scepticism as such a situation cannot be maintained over a longer period of time at the individual household level. In that respect, these outcomes may be regarded as statistical artefacts, resulting from the methodology to derive the distributional results, possibly relating to errors in the underlying micro or macro data, incorrect allocation of imputed items (including the imputation for transactions between households), and/or incorrect linking of data across data sources. On the other hand, there may also be plausible economic explanations for negative saving rates for specific groups of households.

First of all, it is important to note that a negative saving rate for a specific household group does not necessarily mean that each household in that group reports negative saving. It just reflects the average for the group as a whole. In that regard, it would be interesting to have more information on the number of households within the group that report negative saving. Also information on the median saving rate would provide useful information. Furthermore, more granular breakdowns into other socio-demographic characteristics (such as household type, education level and age of the head of the household) may provide more insight in specific groups that report negative saving.

Secondly, it is important to keep in mind that household groups, particularly when clustering according to income, may often not consist of the same households over time. For example, if self-employed households experience an unproductive business year and do not earn their normal annual income, they will be grouped in a lower income group in a given period. In future periods they may be re-grouped in a higher income group when their income situation returns to normal. This will also be the case for households that experience temporary unemployment. Furthermore, one would expect upwards movement in income group over time for certain types of households that may be found in the lower income groups in the current recording period such as single-student households. Thus, the fact that the income groups are not fixed over time could provide an explanation for the existence of negative saving for certain income groups for a longer period of time.

Looking at the plausibility of negative saving at the level of individual households, there may also be a number of possible explanations. First of all, as mentioned above, they may concern temporary negative saving for households that became (partly) unemployed or self-employed households that experienced an unproductive business year. Furthermore, a temporary negative saving may be explained by a large purchase for a household which may be financed by disposing assets or by obtaining consumer credit. Information from the financial accounts may provide more insight in that regard. There is also a wealth of economic literature on households' saving and consumption behaviour related to their position in the life cycle which can provide explanation for negative saving results. This field of study has been guided by two main macro-economic theories, i.e. the Life Cycle Hypothesis (Modigliani and Brumberg, 1954[1]) and the Permanent Income Hypothesis (Friedman, 1957[2]).

The Life Cycle Hypothesis states that individuals plan their consumption on the basis of their entire lifetime such that consumption can remain relatively stable across all periods of their lives. When individuals are young, they accumulate debt as they consume more than they earn and then afterwards when they find employment, they can begin to reduce their debts and accumulate savings for retirement where they will use up their accumulated assets. In this manner, saving over a lifetime would follow an inverted "U"-shaped pattern. Negative saving rates for specific household groups can thus be explained by the presence of single-student households or households drawing on pensions.

The Permanent Income Hypothesis explains that a person's consumption is based both on their current income as well as their expected future income. This means that when there is a temporary change in income, the adjustment to consumption will be smoothed out over time. According to this hypothesis one would predict that a household suffering from a large negative income shock in a given period would maintain the same level of consumption as in previous periods, possibly leading to temporary negative saving.

In response to the impact of the position of a person or household in its lifecycle on its income and saving levels, it is often argued that consumption may provide more insight into the real level of inequality in a country. Only looking at income inequality on the basis of the results of a specific year would not incorporate the position of households in their life cycle and not reflect inequality on the basis of the permanent income of households, therewith very likely overstating inequality (Fisher, Johnson and Smeeding, 2012[3]).

12.2.2. Analysing changes over time

In analysing the plausibility of distributional results, compilers can also look at changes over time. For this purpose, one may focus on changes in the absolute values or in the distribution of specific items or aggregates. Both may show remarkable changes which may point to issues in the data. For example, a sharp increase or decrease in the absolute value of a specific item for a household group will probably raise questions, whereas this will also be the case when the relative distribution shows a significant change in comparison with previous years. The plausibility of such changes will also depend on the specific item. Some items may be expected to show more fluctuation over time (e.g. mixed income and distributed income of corporations) than others (e.g. operating surplus and compensation of employees).

When analysing the plausibility of changes over time at the level of household groups, one has to be aware that changes may be caused by changes at the level of the underlying households or by changes in the composition of these groups. For example, households that were classified in the lowest income group in a specific year need not necessarily be classified in the same group in the following year. Ideally, more information is available on dynamics between household groups over time. This may for example be done on the basis of panel information which may enable following the same household over time. An alternative is to look at more granular breakdowns of household groups, for example further breaking down income groups into household type. Also the underlying socio-demographic information may provide more insight into changes in the stability of the composition of specific household groups over time.

As was the case with the indirect links (as explained in the previous subsection), no immediate conclusions can be drawn on the basis of this analysis with regard to inconsistencies in the distributional results. It will depend on the underlying reasons for these changes, whether it may indeed point to implausible results. For that reason, it is important to look at dynamics between household groups over time (as explained above), but also to have more insight into (changes in) the computation of these results over time. Analysis of changes in the contributions of the underlying micro data, the imputations for missing elements and alignment of the micro data to the national accounts totals over time may show the main underlying reason for the change in the distributional results at the aggregated level, also providing more insight into the plausibility of these results. This is explained in more detail in the next section.

12.2.3. Making cross-country comparisons

A third approach to obtain more insight into the plausibility of distributional results is to compare them with results from other countries. Although results are likely to differ across countries because of country-specific characteristics, such a comparison may provide useful insight in how comparable or different distributions are for specific items from those in other countries. Particularly for items for which no micro

data is available or for which gaps between the micro and macro data show to be relatively large and for which compilers may have had to rely on several assumptions to derive the distributional results, such a comparison may prove very helpful. In making these cross-country comparisons, it is recommended to look for countries that are deemed most comparable. Furthermore, in case of any differences it may also be important to assess how the results in other countries have been derived.

In making cross-country comparisons, one can look at the distributional results for the various items and aggregates. It is recommended to compare results at the most granular level of detail possible as this is expected to provide the most reliable comparison (bearing in mind that the share of the various items in the aggregates may differ across countries).

In addition to comparing the distributional results, one may also compare the socio-demographic composition of the various household groups across countries. On the one hand, this may provide insight into the plausibility of the allocation of households to the various household groups and, on the other hand, it will provide more information on whether distributional results for specific household groups may expected to be comparable across countries. If the socio-demographic compositions of household groups show to be very different, their results are less likely to be comparable.

Finally, one may also compare the composition of (adjusted) disposable income and actual final consumption in terms of their underlying components of household groups across countries. It may for example show the share of compensation of employees in disposable income for the various income quintiles across countries or the percentage of income that is spent on food and non-alcoholic beverages by various income groups across countries. It may be interesting to see how comparable or different these compositions are across countries, particularly for those countries for which household groups may be deemed comparable.

Again, cross-country comparisons may not directly point to incorrect distributional results but may show where additional analysis may be needed to check the plausibility of the results. It may trigger the compiler to try to assess the underlying rationale for these differences.

12.3. Obtaining insight into possible reasons for implausible results

If questions arise about the plausibility of some of the distributional results, it is important to look at the underlying composition of these results. This will provide insight to what extent these results have been driven by underlying micro data (including any corrections for conceptual and classification differences), imputations for missing items, and the alignment of the micro data to the national accounts totals. The robustness of the distributional outcomes will then depend on the size and the reliability of each of these factors. This will also provide insight to what extent different decisions in the process may lead to different distributional results. For example, if the assumptions to allocate imputed amounts to underlying households are deemed to be relatively weak, it may be assessed whether a different allocation may lead to more plausible results. When looking at results for the aggregates, it also has to be borne in mind that these may have been influenced by linking of data across different data sources. Therefore, the quality of this matching should also be assessed in checking the robustness of the distributional results.

12.3.1. Assessing the composition of the distributional results

A first step in obtaining more insight into the robustness of the distributional results is to look at their composition in terms of underlying micro data (including corrections for conceptual and classification differences), imputation for missing items, and the alignment of the micro data to the national accounts totals. A general overview as presented in Section 7.2 may already provide insight to what extent these different components have contributed to the overall distributional results and how this compares to other countries and how it develops over time. More detailed overviews focusing on the composition of the results

for the underlying household groups will provide more insight into how the allocation of these components affect the distributional results and how different decisions in the compilation process may lead to different outcomes. These more detailed overviews are the focus of this subsection.

Table 12.3 presents a template to show the composition of the results for a specific item across household groups, broken down into the contributions by micro data (including corrections for conceptual and classification differences), the imputation for missing items, and the alignment of the micro data to the national accounts totals. It has to be borne in mind that most of these contributions will be positive (e.g. if an item is missing and should be imputed), but that they may also be negative (e.g. if the micro aggregate exceeds the national accounts aggregate). For that reason, the template separately distinguishes positive and negative contributions. Depending on whether the micro data fall short or exceed the national accounts total and on how the gap is allocated to underlying households, the alignment of the micro-macro gap will be recorded as positive or negative contribution. Please note that the template only shows the contributions of micro data, imputations and the alignment of micro and macro data, but depending on the amount of underlying information further details may be included in the template, such as separate corrections for conceptual and classification differences and further breakdowns of the imputations made for various missing elements.

Table 12.3. Template to show the composition of a specific item

	Q1	Q2	Q3	Q4	Q5
NA aligned total	…	…	…	…	…
Positive contribution					
Micro data	…	…	…	…	…
Imputation	…	…	…	…	…
Positive alignment of micro-macro gap	…	…	…	…	…
Negative contribution					
Negative alignment of micro-macro gap	…	…	…	…	…

Whereas these tables can be used to assess the composition of the various items in terms of micro data, imputations and alignments, it can also be used to assess the composition of the main aggregates. In that case, the results for the underlying items should be added up to arrive at the compositional results for the totals, bearing in mind that some items will positively contribute to the aggregates whereas others may contribute negatively.

Table 12.4 provides an example of the composition of adjusted disposable income on the basis of its underlying contributions. A distinction has been made into positive and negative contributions, both broken down into micro data, imputations and alignment of the micro and macro data. The micro data reported under the positive contributions show the sum of all the micro data that positively contribute to adjusted disposable income, such as the micro data underlying compensation of employees, interest received, distributed income of corporations received and current transfers received. The micro data under the negative contributions show all the micro data that negatively contribute to adjusted disposable income, such as the micro data underlying interest paid, current on income and wealth, and current transfers paid. Whereas this will usually be relatively straightforward for micro data and for imputed items, this may be less straightforward for the alignment of micro data to national accounts totals. Whether this should be recorded as positive or negative contribution will depend on whether the micro aggregate exceeds or falls short of the national accounts total and whether the item itself positively or negatively contributes to the aggregate. For example, if the micro aggregate for interest received exceeds the national accounts total, the alignment of the micro-macro gap will negatively contribute to adjusted disposable income. But the same will be the case when the micro data underlying taxes paid fall short of the national accounts totals. Compilers should be aware of this issue to properly account for the positive and negative contributions of the various underlying components.

Table 12.4. Example of the composition of an item on the basis of its underlying contributors

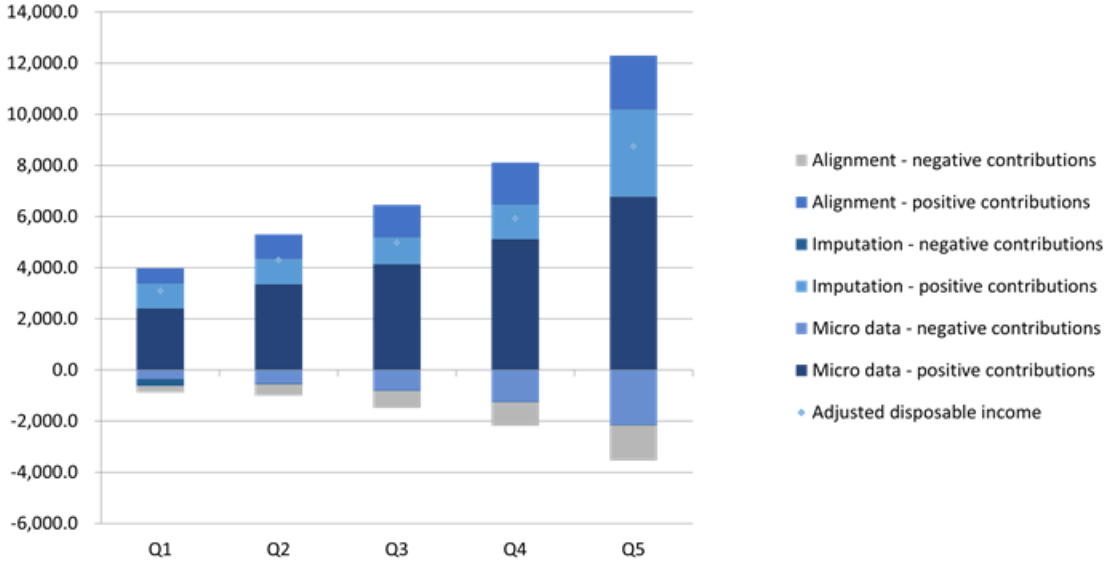
Composition of adjusted disposable income as result of underlying micro data, imputations and alignment of micro-macro gaps.

Adjusted disposable income	Q1	Q2	Q3	Q4	Q5
NA aligned total	165	185	245	250	295
Positive contributions (+)					
Micro data	200	250	300	350	400
Imputation	40	35	40	28	30
Alignment of micro-macro gap	15	20	30	22	25
Negative contributions (-)					
Micro data	60	80	90	110	120
Imputation	20	20	20	20	20
Alignment of micro-macro gap	10	20	15	20	20

The composition of (aggregated) items may also be presented graphically. Figure 12.1 provides an example for the composition of adjusted disposable income. It presents distributional results for the five quintiles, showing the positive and negative components, broken down into underlying micro data, imputations and alignments.[5]

Figure 12.1. Example of composition of distributional results

Composition of adjusted disposable income for five income quintiles



When analysing the results, the contributions can be assessed at an aggregated level, such as presented in Table 12.4 and Figure 12.1. The positive and negative contributions may be further broken down into underlying components. This may provide further insight into which items have the largest contributions to aggregates for specific household groups. Furthermore, it may provide more insight in how different allocations for some specific items for which the allocations are deemed to be less reliable, may affect the distributional results. For example, allocating a larger part of a micro-macro gap to lower income households will change the overall distribution and may at the same time increase the plausibility of the overall results.

By carefully assessing the margins of error surrounding the components contributing to the distributional results, one may derive the margins surrounding these outcomes and assess how these may be affected by different allocations. The initial allocation may already be the result of a careful analysis considering all available (direct or indirect) information, but there may still be some margin surrounding these allocations. This will be particularly true for allocations of micro-macro gaps in case no information is available on what may have been causing the remaining gap.

In assessing possible margins of error in relation to the allocation of the remaining micro-macro gaps, one has to look at sensible alternatives for allocating the gap. In this regard, the proportional allocation may be assumed as a rather reliable way of closing any final micro-macro gap under the assumption that any remaining gap is likely related to more generic issues that will be affecting households across the distribution in rather similar ways.[6] However, slightly different allocations could be considered. For that reason, it may be useful to conduct sensitivity analyses, starting from the underlying micro data and slightly adjusting the weights in comparison to a proportional allocation (i.e. assigning relatively larger amounts to some household groups at the expense of others), which would respect the initial distribution to a large extent and be in line with the assumption that any remaining gap is likely due to more generic factors.[7]

For any other components, compilers may assess the reliability of their applied allocation and assess what plausible alternatives could be envisaged. For example, for items for which direct micro data is missing, compilers may assess the reliability of the proxy applied to arrive at the relevant distributions. In some cases, the allocation has been based on an item that closely links to the target variable, leaving very small margins of error. In other cases, the proxy may have been much weaker, leaving more uncertainty around the results.

The same applies to the matching of the results across data sets. Compilers should try to conduct sensitivity analyses to check how the results may change when applying slightly different assumptions for the linking or matching. In some cases, this may have a very small impact on the results (i.e. when applying direct linking on the basis of identifiers), but the impact may be larger when the matching relies on specific assumptions (i.e. in case of statistical matching or matching at the aggregated level).[8]

12.3.2. Checking plausibility of the national accounts data

Although it is not mentioned as one of the direct contributors to the distributional results, national accounts totals may have a large impact on the data, as they constitute the benchmark for the distributional results. In that regard, they feed into the alignment of the micro data to the national accounts totals. Especially in case of large micro-macro gaps, it is important to assess the quality of the national accounts totals, as was explained in Chapter 7 For that purpose, it would be helpful to have more insight in how the national accounts result has been derived, obtaining more information on the main underlying data source(s) (e.g. whether it has been based on direct information on the household sector, on counterparty sector information or whether it has been derived as a residual) and on changes that may have been applied in the course of the compilation to arrive at the final national accounts result. This information would help in assessing the reliability and possible margins of error surrounding the results.

12.3.3. Checking plausibility of the micro data

In analysing the consistency and plausibility of the results, it is also important to assess the quality of the underlying micro data. Usually, these data will already have undergone extensive checks as part of the compilation of the relevant micro statistics, but additional checks may be performed in the process of compiling distributional results in line with national accounts totals. For example, some of the checks as described in Section 12.2 may not have been performed yet at the micro level, as it concerns confronting data from multiple data sources that may not have been available at the level of the individual statistics. A specific example is the confrontation of income and consumption results at the micro level, which often

can only be done by combining information from different data sets. Furthermore, compilers can check the consistency of micro data over time, look at comparisons of results within and across household groups, and look at the distribution of the results within the micro source. Such checks may provide insight into the plausibility of the data included in the various data sources. However, it has to be borne in mind that inconsistencies or implausibilities after linking data across multiple data sources may also be due to incorrect linking of the data. Thus, in case inconsistencies or implausible results are found at the micro level after linking data from multiple data sets, both the quality of the linking procedure and the underlying micro data should be carefully assessed.

12.3.4. Checking plausibility of imputations

It is also important to assess the reliability of the amounts and the allocation for items that are missing in the micro data. In some cases, these may be derived on the basis of auxiliary information, but in other cases, they may have been derived to a large extent on the basis of assumptions. In the latter cases, it is important to assess how different assumptions would affect the results and whether that would lead to more plausible results, either for the item itself or for the aggregated items. As imputations may relate to various elements that may be missing in the micro data, such as missing parts of the population, missing elements in relation to the national accounts items, underground economy and illegal and informal activities, ideally the amounts and their allocations should be assessed separately, as they may have different distributions and may rely on different types of underlying information.

12.3.5. Checking plausibility of aligning micro-macro gaps

The alignment of the gaps between micro and macro data may account for a large part of the distributional results. For that reason, it is important to check the plausibility of this allocation. In some cases, auxiliary information may be available on the basis of which these gaps can be allocated to the most probable underlying causes and for their allocation to the relevant households, but if this is not the case, it should be assessed whether the allocation that has been applied indeed leads to the most plausible results for the relevant item and for the aggregates. In case different allocations lead to more plausible results and also make sense from a technical and economic point of view, compilers may decide to apply an alternative allocation. In that regard, it is also important to take note of underlying reasons distinguished by other countries to explain any micro-macro gaps for specific items and how they have dealt with these. Furthermore, it is important to monitor the size and allocation of these gaps over time.

12.3.6. Checking plausibility of linking or matching data across data sources

Finally, the quality of the matching of data across different data sources should be checked. It will depend on the specific technique used (see Chapter 8) how one may best check for this plausibility. If linking or matching is done at the micro level, the plausibility of the matching results can be checked at the micro level, for example by looking at consistency between items from different data sets that have direct or indirect links or by looking at the consistency between income and consumption (and possibly wealth). If these checks show implausible results, this could point to errors in the micro data or to errors in matching the data. In that case, both the quality of the matching procedure as the underlying micro data should be carefully assessed. Such a confrontation may also be applied when matching data at the aggregated level. Furthermore, in that case, one may also look at the fitness of underlying socio-demographic characteristics of the groups after linking. For example, if the characteristics of the first income group according to data sources on income are completely different from those derived on the basis of data sources on consumption, this may point either to errors in the underlying data (i.e. in the income and/or consumption data or in the reported socio-demographic information) or in the matching across data sources. It would in that case be best to liaise with the micro experts to retrieve the most likely cause for the implausible results and to solve this in the best possible way.

12.4. Conclusions

This chapter has presented several ways to check the consistency and plausibility of the distributional results and provided guidance to assess possible causes for any inconsistencies or implausible results. As has become clear from this chapter, it will often be difficult to draw firm conclusions on errors in the (processing of the) data, but these checks may point to possible issues in the compilation process. This should help compilers to carefully assess the inputs and assumptions used to arrive at these results and to check whether alternative decisions would be possible that may lead to more plausible results. Of course, this does not imply that compilers may just work towards any specific outcome, as they should carefully assess to what extent different assumptions may indeed be possible and valid from a technical, economic and historic point of view. In that regard, compilers should always account for any changes in the data or in the assumptions that may change the distributional results and to explain their rationale from a technical and economic perspective.

References

Fisher, J., D. Johnson and T. Smeeding (2012), "Inequality of Income and Consumption: Measuring the Trends in Inequality from 1985-2010 for the Same Individuals", http://www.iariw.org/papers/2012/FisherPaper.pdf (accessed on 29 September 2017). [3]

Friedman, M. (1957), *A Theory of the Consumption Function*, Princeton University Press, Princeton. [2]

Modigliani, F. and R. Brumberg (1954), "Utility analysis and the consumption function: An interpretation of cross-section data", in *Post-Keynesian Economics*, Ruthers University Press New Brunswick New Jersey, London. [1]

Notes

[1] The EG DNA has been investigating possibilities to include information on *capital transfers received* (D9R) and *paid* (D9P), *gross capital formation* (P5) and *acquisitions less disposals of non-produced assets* (NP) to arrive at results up until net lending/net borrowing.

[2] In some cases, negative values may still show up for these items in the micro data, for example related to repayments of wages and taxes. Theoretically, these repayments should lead to a correction to the relevant item in the period in which the "incorrect" amount was recorded, but in practice, this may not always be feasible.

[3] This also holds for its underlying components, *operating surplus from owner occupied dwellings* (B2R1) and *operating surplus from leasing of dwellings* (B2R2).

[4] This also holds for *mixed income excluding underground and own account production* (B3R3). Furthermore, in theory it may also hold for *own account production* (B3R1) and *underground production* (B3R2), although this is not very likely in practice.

[5] The breakdown into micro data and alignments is derived on the basis of the assumption of proportional adjustment of the gaps between the micro and macro aggregates for each individual component.

[6] One could also apply assumptions that the gap is due to very specific reasons only involving a limited group of households, but this may be less likely. For example, it could be assumed that the gap for a specific item is fully due to underreporting by households at the lower end of the income distribution. However, as these may consist of very different types of households, it would seem very unlikely that the underreporting would only apply to this specific group of households and that there would be no issues for similar types of households elsewhere in the distribution. Furthermore, assigning the full micro-macro gap to a very limited group of households may imply assigning relatively large values to specific households in comparison to their initially reported data (possibly also assigning large values to households that did not report any values for a specific item to start with), with possibly large impacts on their ranking in the overall distribution. This will be much smaller when keeping closer to the initially reported data and assuming more generic causes for the remaining gaps.

[7] The EG DNA is currently developing a more specific proposal on how to conduct sensible sensitivity analyses on the basis of these assumptions, as part of the development of a more elaborated quality framework.

[8] The EG DNA is looking into the development of standard sensitivity analyses to accompany the results as part of the new quality framework.

13 Indicators to present distributional results

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

13.1. Introduction

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

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

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

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

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

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

using the notation:

X: income/consumption component

i: {1, 2, …, I} to identify household groups

n_i: total number of households in group i

$X_i^{NA\_adj}$: adjusted national accounts subtotal for group i

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

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

using the notation:

X: income/consumption component

i: {1, 2, …, I} to identify household groups

cu_i: total number of consumption units in group i

$X_i^{NA\_adj}$: adjusted national accounts subtotal for group i

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

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

13.3. Measures of disparity

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

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

$$Ratio\ to\ average_i = \frac{\bar{X}_i^{NA\_adj}}{\bar{X}^{NA\_adj}}$$

using the notation:

X: income/consumption component

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

i: {1, 2, …, I} to identify household groups

n_i: total number of households in group i

N: total number of households in the population

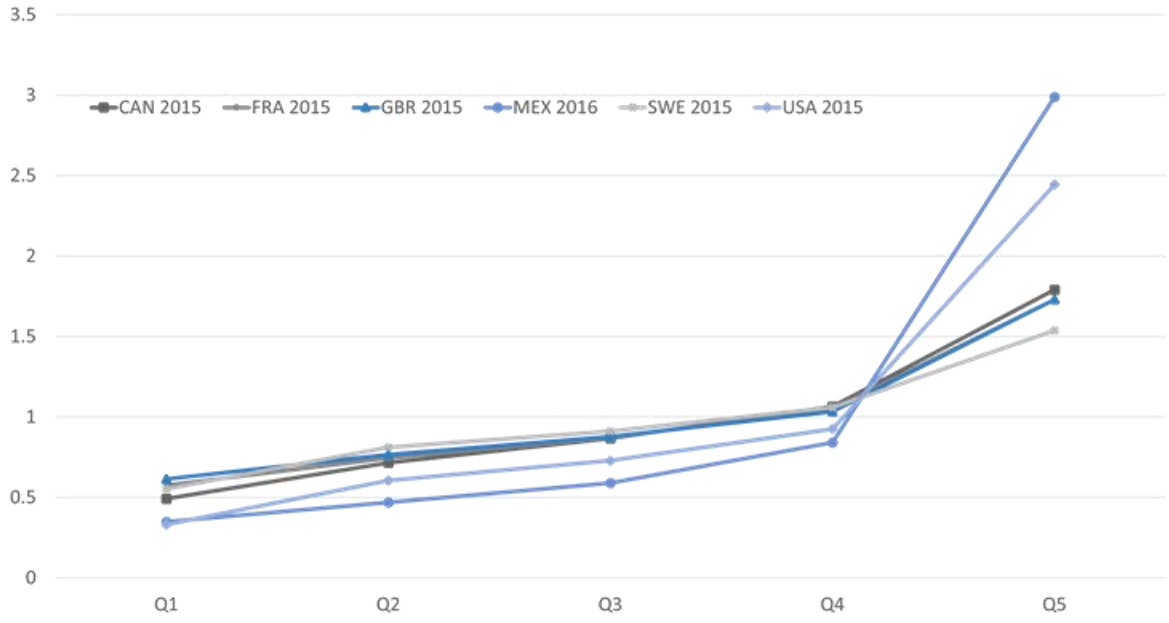
$\bar{X}_i^{NA\_adj}$: per household or per consumption unit adjusted national accounts for group i

$\bar{X}^{NA\_adj}$: per household or per consumption unit adjusted national accounts

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

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

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



Note: The results show adjusted disposable income per consumption unit for each group to the average adjusted disposable income per consumption unit for the private household sector as a whole.
Source: Zwijnenburg et al. (2021[1]).

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

$$Ratio\ highest\ to\ lowest_z = \frac{Max_{i \in z}\{\bar{X}_i^{NA\_adj}\}}{Min_{i \in z}\{\bar{X}_i^{NA\_adj}\}}$$

using the notation:

X: income/consumption component

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

i: {1, 2, …, I} to identify household groups

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

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

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

$$Coefficient\ of\ variation_z = \frac{\sqrt{\frac{1}{N} \times \sum_{i \in z} \left[n_i \times \left(\bar{X}_i^{NA\_adj} - \bar{X}^{NA\_adj} \right)^2 \right]}}{\bar{X}^{NA\_adj}} \times 100$$

using the notation:

X: income/consumption component

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

i: {1, 2, …, I} to identify household groups

n_i: total number of households in group i

N: total number of households in the population

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

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

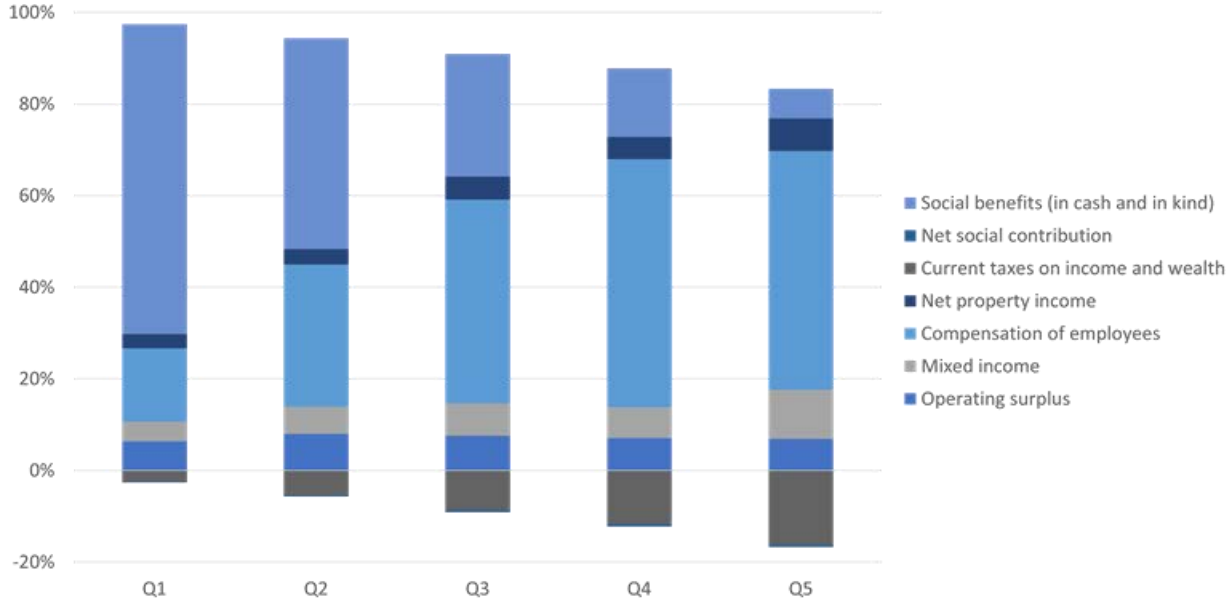
As was the case with the ratio highest to lowest, this ratio may be used to make cross-country comparisons as well as to monitor changes over time within a country. When analysing the disparity index, two properties need to be taken into account. First, the above calculation assumes that each household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

13.4. Composition of household income and consumption

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

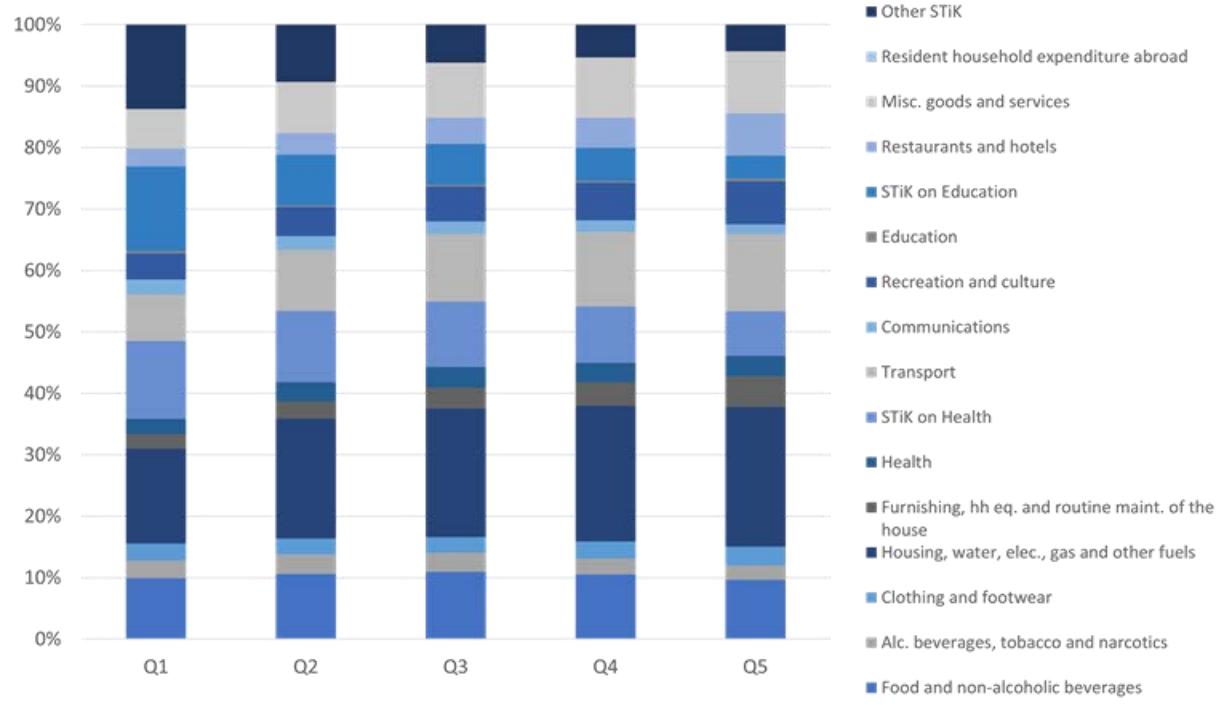
The most common way to present these results is to look at the share of the various items in total disposable or adjusted disposable income, and in total final consumption expenditure or actual final consumption. It will depend on the available underlying information what level of detail could be provided. Figure 13.2 and Figure 13.3 show examples on how to present this information on the basis of DNA results collected in 2020 (Zwijnenburg et al., 2021[1]).

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



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

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



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

13.5. Indicators based on underlying micro data

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

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

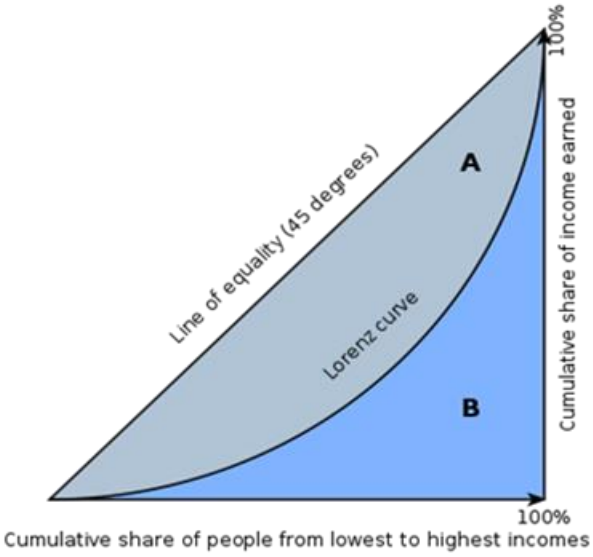
A first indicator that could be derived from underlying micro data is the household participation rate. This ratio provides insight into the share of households that report a value for a specific item. For example, the participation rate for distributed income of corporations shows how many households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

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

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

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

Figure 13.4. Example of Lorenz curve



Source: Wikipedia (2023[3]).

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

References

Cowell, F. (2011), *Measuring Inequality*, Oxford University Press, https://doi.org/10.1093/ACPROF:OSOBL/9780199594030.001.0001. [2]

Wikipedia (2023), *Lorenz curve*, https://en.wikipedia.org/wiki/Lorenz_curve (accessed on 21 November 2023). [3]

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, Paris, https://www.oecd-ilibrary.org/docserver/615c9eec-en.pdf?expires=1669202191&id=id&accname=ocid84004878&checksum=689A9023E81B32F43D91EF34FBDE7A7A (accessed on 23 November 2022). [1]

14 Further work

A lot of work has been done in developing methodology to compile distributional results in line with national accounts totals, as presented in this handbook. However, more work is needed. This chapter provides an overview of the main areas for further work.

14.1. Introduction

The EG DNA was launched in 2011 and has done a lot of work to develop a harmonised methodology and to explore specific methodological issues that may pose challenges in the compilation of distributional results in line with national accounts. However, there are still some areas where more work is needed. This chapter provides an overview of the main areas for further work.

14.2. Broadening the range of countries

Many countries are already compiling results according to the DNA methodology, but data are still missing for specific countries. In this regard, it is expected that more countries will start compiling the relevant results in the coming years in view of the new G20 Data Gaps Initiative (IMF, 2023[1]). Furthermore, the OECD and Eurostat have started working on the development of so-called centralised approaches to compile results for missing countries. These approaches rely on either publicly available micro data or micro data available to the international organisations, which are then combined with publicly available national accounts totals to arrive at DNA results. Of course, the results are sub-optimal in comparison with results that could be compiled by national statistical offices, as these would normally have more data sets at their disposal, have better knowledge of the underlying data, and have better insights in country specific circumstances that may be of relevance in imputing for missing items and in allocating any micro-macro gaps. However, the approaches would still prove useful to compile estimates for those countries for which data is not (yet) available.

In the work, Eurostat is focusing on EU countries, using data from the EU Survey on Income and Living Conditions (EU-SILC)[1] and the EU Household Budget Survey (HBS)[2] as their main inputs (Eurostat (2022[2])). The OECD is focusing on non-EU OECD member countries, relying on data available from the Luxembourg Income Study (LIS)[3] (Zwijnenburg et al., forthcoming (2024[3])). Results for EU countries are already available and have been included in the database of Eurostat and the OECD. Results for non-EU countries are expected to become available in the course of 2023. At the same time, both institutions continue to further improve their centralised approaches, by obtaining feedback on the results and by finetuning specific elements in the methodology.

14.3. Improving the timeliness of the distributional results

Timeliness of data is an important quality characteristic and comes at a premium in periods of rapid and important changes in the economy (such as the Great Financial Crisis and the Covid pandemic). Given the complexity of deriving DNA estimates and the time lag for many of the underlying micro data sources, DNA estimates currently suffer from relatively long time lags, often only becoming available a couple of years after the reference year. Given the huge user demand for timely data, it is important to develop nowcasting techniques to reduce the existing time lags and to ensure full relevance of the DNA results. In this regard, the new G20 Data Gaps Initiative (IMF, 2023[1]) includes the ambition for G20 economies to publish annual distributional results within 18 months after the reference period, by the end of 2026. It may be explored whether even more timely estimates may be feasible, dependent on the development of reliable nowcasting techniques.

There are already various initiatives, both at the national and international level, to nowcast more timely distributional information (see for example Office for National Statistics (2020[4]), Statistics Canada (2021[5]), Blanchet, Saez and Zucman (2022[6]) and Eurostat (2022[7])). These efforts will provide a useful starting point to explore how nowcasting techniques may be used to compile more timely DNA results.

14.4. Increasing the granularity of the results

Whereas the DNA work currently focuses on breakdowns by income quintiles, there is a large user demand for more granular breakdowns, e.g. results broken down by income decile (see the ambition for distributional results in the new G20 Data Gaps Initiative (IMF, 2023[1])) and/or percentile. The latter is deemed particularly relevant to obtain more insights into the bottom and top end of the distribution. Furthermore, in addition to breakdowns by household type and by main source of income, there is demand for breakdowns according to other types of socio-demographic characteristics, such as age and gender. It is important for the work to explore possibilities to publish at these more granular levels of detail, with the templates already including the possibilities to report data at these more granular levels.

The possibility to publish more detailed results will largely depend on the quality of the results. In this regard, it needs to be borne in mind that the allocation of micro-macro gaps and of any items for which micro data is lacking may lead to some margins of error surrounding the results, dependent on the information available to properly allocate the amounts to the relevant underlying households. These margins of error may prevent publishing at more granular levels of detail, particularly when they hamper a proper analysis of the trends. This means that more work will be needed to improve the linking between micro and macro items, to reduce gaps between the micro and macro results, to further finetune the guidance to impute for missing items and/or parts of the population, and to improve the linking across different data sources. This will reduce the margins of error surrounding the results, providing the possibility to publish results at more granular levels of detail. Furthermore, it may involve the development of sensitivity analyses to assess the sensitivity of the results to alternative assumptions for bridging micro-macro gaps and to deal with missing elements.

14.5. Increasing the frequency of the results

There is also a clear user request for more frequent results. In that regard, several countries currently only compile results every couple of years, often dependent on the availability of the underlying micro data. On the other hand, some countries are already compiling the results on an annual basis, with some even exploring publication at a quarterly frequency. It will be relevant to assess whether specific methodological guidance can be developed to assist countries in compiling results at a higher frequency. This may involve the development of interpolation techniques to assist countries to compile results for those years for which micro data may be missing. Furthermore, it may involve the development of nowcasting techniques for those reporting periods for which micro data may not yet be available.

14.6. Exploring distributional results on wealth

Whereas distributional information on income, consumption and saving aligned to macroeconomic totals are essential to obtain a better understanding of how household groups are faring, it is important to broaden this work to include the wealth dimension, to provide insights into the three main dimensions of material well-being, i.e. income, consumption, and wealth. This allows policymakers to have a comprehensive overview of the economic situation of different household groups and to better attune policies to their specific needs. Furthermore, it provides compilers with the opportunity to cross-check results across income, consumption and wealth, adding to the quality of the overall results.

The Expert Group on Distributional Financial Accounts (EG DFA) has already done extensive work on developing distributional financial and non-financial balance sheets for the household sector in the euro area and EU economies, and several countries have started to develop distributional results on wealth at the national level. Furthermore, in view of the specific recommendation on distributional wealth results in the new G20 Data Gaps Initiative (IMF, 2023[1]), the OECD launched a new Expert Group on Distribution

of Household Wealth (EG DHW) early 2023. This group will develop internationally harmonised templates and methodology for the compilation of household distributional wealth results on wealth in line with national accounts totals, leveraging off the work already done in this area by the EG DFA. The ambition is to have G20 economies compile distributional wealth results at decile level on an annual frequency by the end of 2026.

14.7. Conclusions

The EG DNA will continue its efforts in the coming years to work on these specific issues (as well as on other issues that may arise in the coming years). It is expected that this will feed into future updates of this Handbook.

References

Blanchet, T., E. Saez and G. Zucman (2022), "Real-Time Inequality", *NBER Working Paper Series*, No. 30229, https://github.com/thomasblanchet/real-time-inequality (accessed on 30 November 2022). [6]

Eurostat (2022), *Early estimates of income inequalities*, https://ec.europa.eu/eurostat/statisticsexplained/ (accessed on 30 November 2022). [7]

Eurostat (2022), *Eurostat Centralised Exercise - Methodological note*. [2]

IMF (2023), *G20 Data Gaps Initiative*, https://www.imf.org/en/News/Seminars/Conferences/g20-data-gaps-initiative (accessed on 14 March 2023). [1]

Office for National Statistics, U. (2020), *Provisional estimates of income inequality in the UK for the financial year ending 2020*, https://www.ons.gov.uk/peoplepopulationandcommunity/personalandhouseholdfinances/incomeandwealth/bulletins/householdincomeinequalityfinancial/financialyearending2020provisional (accessed on 30 November 2022). [4]

Statistics Canada (2021), *Distributions of household economic accounts for income, consumption and saving of Canadian households, fourth quarter 2020*, https://www150.statcan.gc.ca/n1/daily-quotidien/210907/dq210907a-eng.htm (accessed on 30 November 2022). [5]

Zwijnenburg, J. et al. (2024), "OECD centralised approach to calculate distributional results in line with national accounts totals - Methodology and results". [3]

Notes

[1] https://ec.europa.eu/eurostat/web/income-and-living-conditions/overview.

[2] https://ec.europa.eu/eurostat/web/household-budget-surveys/overview.

[3] https://www.lisdatacenter.org/.

Annex A. Comparison between macro and micro concepts

The Canberra Group Handbook provides a conceptual definition of household income. This definition is in line with that as established by the International Conference of Labour Statisticians (ICLS) and reads as follows:

> *"Household income consists of all receipts whether monetary or in kind (goods and services) that are received by the household or by individual members of the household at annual or more frequent intervals but excludes windfall gains and other such irregular and typically one-time receipts. Household income receipts are available for current consumption and do not reduce the net worth of the household through a reduction of its cash, the disposable of its other financial or non-financial assets or an increase in its liabilities" (United Nations Economic Commission for Europe, 2011[1]).*

This conceptual income definition is, as much as possible, consistent with the income definition as used in the SNA. The operational definition is to a large degree consistent with the conceptual definition, apart from the exclusion of the value of unpaid domestic services, the value of consumer durables and social transfers in kind, due to the difficulty in valuing these components.

Appendix 2 of the Canberra Group Handbook provides an overview of the relationship between the income concept as used in micro statistics (according to the definition used in the Canberra Group Handbook) and the one used in macro statistics (as defined in the 2008 System of National Accounts (European Commission et al., 2009[2])). The table below provides an overview of the main differences on the basis of that appendix.

Table A.1. Comparison of income according to Canberra and SNA framework

	Description	Correspondence in Canberra definition	Comments
B2R1	Operating surplus from owner occupied dwellings	Net value of housing services provided by owner-occupied dwellings and subsidised rentals	The SNA item is recorded on a gross basis, thus after deducting intermediate consumption costs (including FISIM), but before deducting consumption of fixed capital and interest paid.
B2R2	Operating surplus from leasing of dwellings	Rental income from residential properties net of operating expenses, depreciation and interest	The SNA item is recorded on a gross basis, thus after deducting intermediate consumption costs (including FISIM), but before deducting consumption of fixed capital and interest paid.
		Rental income from non-residential properties net of operating expenses, depreciation and interest	The SNA item is recorded on a gross basis, thus after deducting intermediate consumption costs (including FISIM), but before deducting consumption of fixed capital and interest paid.
B3R1	Mixed income from own account production	Goods produced for own consumption, less cost of inputs	
B3R2	Mixed income from underground production	-	
B3R3	Mixed income excluding underground and own account production	Profit/loss from unincorporated enterprises, measured net of operating costs and after deduction for the depreciation of assets used in the production, and net interest	The SNA item is recorded on a gross basis, thus after deducting employee costs and intermediate consumption costs (including FISIM), but before deducting consumption of fixed capital and interest paid. Furthermore, the micro data may include amounts related to profits by units that are treated as quasi-corporations in the SNA. The withdrawal of income of these quasi-corporations is included in distributed income of corporations in the SNA.
		Goods produced for barter, less cost of inputs	
		Royalties, i.e. from intellectual property rights, etc.	
D11R	Wages and salaries received	Direct wages and salaries for time worked and work done	In concept both the SNA and micro income measures do not include social insurance benefits paid by employers, such as sick leave or maternity leave, in wages and salaries. The Canberra Handbook explains that in practice, it may be difficult to separate these payments.
		Cash bonuses and gratuities	
		Commissions and tips	
		Directors' fees	
		Profit-sharing bonuses and other forms of profit-related pay	
		Shares offered as part of employee remuneration	
		Free or subsidised goods and services from an employer	
		Severance and termination pay	
D121R	Employers' actual social contributions received	Employers' social insurance contributions	
D122R	Employers' imputed social contributions received	-	
D41R'	Interest received (not adjusted for FISIM)	Interest from financial institutions	Micro statistics reflect amounts net of expenses, e.g. interest paid on borrowings for investment purposes
D42R	Distributed income of corporations received	Dividends, including income from own incorporated business	Micro statistics may include dividends as distributed by mutual funds or other investment funds, which are included in Investment income attributable to collective investment funds shareholders (D443R) in the SNA. Micro statistics reflect amounts net of expenses, e.g. interest paid on borrowings for investment purposes Furthermore, the micro data include withdrawal of income of quasi-corporations in income from self-employment (see B3R3).

	Description	Correspondence in Canberra definition	Comments
D43R	Reinvested earnings received on foreign direct investment	-	
D441AR	Property income received attributed to non-life insurance policy holders	-	
D441BR	Property income received attributed to life insurance policy holders	-	
D442R	Investment income payable on pension entitlements received	-	
D443R	Investment income attributable to collective investment funds shareholders received	-	
D45R	Rent received	Rental income from non-produced assets (land and subsoil assets)	Micro statistics reflect amounts net of expenses, e.g. interest paid on borrowings for investment purposes
D41'P	Interest paid (not adjusted for FISIM)	-	
D45P	Rent paid	-	
FISIM_P	Adjustment for FISIM	-	
D5P	Current taxes on income and wealth	Direct taxes (net of refunds)	
		Compulsory fees and fines	
D611P	Employers' actual social contributions paid	Employee and employers' social insurance contributions	Employers' social insurance contributions may be separately available (see D121R)
D613P	Households' actual social contributions paid		
D612P	Employers' imputed social contributions paid	-	
D614P	Households' social contributions supplements paid	-	
D61xP	Social insurance scheme service charges paid	-	
D611R	Employers' actual social contributions received	-	
D612R	Employers' imputed social contributions received	-	
D62P	Social benefits other than STiK paid	-	
D62R	Social benefits other than STiK received	Social security pensions	
		Pensions and other insurance benefits	Lump sum retirement benefits are not included in the micro measure
		Social assistance benefits (excluding social transfers in kind)	
D72R	Non-life insurance claims	-	
D71P	Non-life insurance premiums	-	
D75R	Miscellaneous current transfers received	Current transfers from non-profit institutions	
		Compulsory and quasi-compulsory inter-household transfers received	The micro items specifically relate to transfers between resident households
			Gambling wins and losses (net of service charge) are not part of the income concept as used in the micro measure
	Miscellaneous current	Current transfers to non-profit	

	Description	Correspondence in Canberra definition	Comments
D75P	transfers paid	institutions	
		Compulsory and quasi-compulsory inter-household transfers received	The micro items specifically relate to transfers between resident households
			Gambling wins and losses (net of service charge) are not part of the income concept as used in the micro measure
D75x	Net miscellaneous transfers received between resident households	Compulsory and quasi-compulsory inter-household transfers received	
D63R1	Social transfers in kind on education received	Social transfers in kind received	
D63R2	Social transfers in kind on health received		
D63R3	Other social transfers in kind received		
		Value of unpaid domestic services	Excluded from the SNA
		Value of services from household consumer durables	Excluded from the SNA

Source: Appendix 2 of the Canberra Group Handbook (United Nations Economic Commission for Europe, 2011[1]).

Table A.2. Comparison of income according to ICW and SNA framework

	Description	Correspondence in ICW Framework definition	Comments
B2R1	Operating surplus from owner occupied dwellings	Net value of housing services provided by owner-occupied dwellings	The SNA item is recorded on a gross basis, thus after deducting intermediate consumption costs (including FISIM), but before deducting consumption of fixed capital and interest paid.
B2R2	Operating surplus from leasing of dwellings	Rent from real-estate other than owner-occupied dwellings, net of expenses (partly)	The SNA item is recorded on a gross basis, thus after deducting employee costs and intermediate consumption costs (including FISIM), but before deducting consumption of fixed capital and interest paid. The SNA only includes the part that relates to renting of dwelling under B2R2. The other part is recorded under B3R3.
B3R1	Mixed income from own account production		
B3R2	Mixed income from underground production		
B3R3	Mixed income excluding underground and own account production	Income from self-employment	The SNA item is recorded on a gross basis, thus after deducting employee costs and intermediate consumption costs (including FISIM), but before deducting consumption of fixed capital and interest paid. The income of “sleeping” or “silent” partners of unincorporated enterprises will be included under mixed income in the SNA, while it will be recorded as Income from shares and other equity, net of expenses in the ICW Framework. However, most household enterprises with “sleeping” or “silent” partners will concern quasi-corporations for which any withdrawal is recorded as distributed income of corporations in the SNA (see D42).
		Rent from real-estate other than owner-occupied dwellings, net of expenses (partly)	The SNA item is recorded on a gross basis, thus after deducting employee costs and intermediate consumption costs (including FISIM), but before deducting consumption of fixed capital and interest paid.

	Description	Correspondence in ICW Framework definition	Comments
			The SNA only includes the part that relates to renting of non-dwelling buildings and structures under B3R3. The other part is recorded under B2R2.
		Royalties and other income from other non-financial assets, net of expenses (partly)	The SNA item is recorded on a gross basis, thus after deducting employee costs and intermediate consumption costs (including FISIM), but before deducting consumption of fixed capital and interest paid. Rent earned from land or other natural resources (as well as rent paid) are recorded as Rent in the SNA (see D45).
D11R	Wages and salaries received	Cash wages and salaries	Wages and salaries paid while employee is on sick, injury or maternity leave are included in employee income items in the ICW Framework while recorded as social benefits other than STiK received (D62R) in the SNA.
		Cash commission and piece-work payments	
		Cash tips and gratuities	
		Directors' fees	
		Shared offered as part of employee remuneration	
		Profit-sharing bonuses and other forms of profit-related pay	
		Other cash bonuses	
		Free or subsidised goods and services from employers	
		Severance and termination pay	
D121R	Employers' actual social contributions received	Employers' social insurance contributions	
D122R	Employers' imputed social contributions received	-	
D41R'	Interest received (not adjusted for FISIM)	Interest from deposits, net of expenses	Micro statistics reflect amounts net of expenses, e.g. interest paid on borrowings for investment purposes.
		Income from bonds and other debt securities, net of expenses	
		Income from other financial assets, net of expenses (partly)	Micro statistics reflect amounts net of expenses, e.g, interest paid on borrowings for investment purposes. This may concern interest payments on other assets, but also other types of property income as defined in the SNA.
D42R	Distributed income of corporations received	Income from shares and other equity, net of expenses	Micro statistics reflect amounts net of expenses, e.g. interest paid on borrowings for investment purposes. The income of "sleeping" or "silent" partners of unincorporated enterprises will be included under mixed income in the SNA, while it will be recorded as Income from shares and other equity, net of expenses in the ICW Framework. However, most household enterprises with "sleeping" or "silent" partners will concern quasi-corporations for which any withdrawal is recorded as distributed income of corporations in the SNA (see D42).
D43R	Reinvested earnings received on foreign direct investment	-	
D441AR	Property income received attributed to non-life insurance policy holders	-	
D441BR	Property income received attributed to life insurance policy holders	-	
D442R	Investment income payable on pension entitlements received	-	

	Description	Correspondence in ICW Framework definition	Comments
D443R	Investment income attributable to collective investment funds shareholders received	Income from mutual funds and other investment funds, net of expenses	Micro statistics reflect the actual dividends as distributed by the funds but excludes the investment earnings made by the funds but not distributed as payments to their shareholders which is included in the SNA item. Micro statistics reflect amounts net of expenses, e.g. interest paid on borrowings for investment purposes
D45R	Rent received	Royalties and other income from other non-financial assets, net of expenses (partly)	Rent earned from land or other natural resources (as well as rent paid) are recorded as Rent in the SNA.
D41'P	Interest paid (not adjusted for FISIM)	Interest paid on consumer credit	In the ICW Framework the interest paid on non-consumer credit is subtracted from the income earned from the assets associated with the loans concerned.
D45P	Rent paid		
FISIM_P	Adjustment for FISIM		
D5P	Current taxes on income and wealth	Direct taxes, net of refunds	
		Irregular taxes on wealth, including taxes on holding gains	Taxes on holding gains are not regarded as current transfer in the ICW Framework but as capital transfer
		Compulsory fees and fines (partly)	Fines imposed by courts of quasi-judicial bodies are recorded as part of miscellaneous current transfers (D75) in the SNA
D611P	Employers' actual social contributions paid	Employee and employers' social insurance contributions	Employers' social insurance contributions may be separately available (see D121R)
D613P	Households' actual social contributions paid		
D612P	Employers' imputed social contributions paid	-	
D614P	Households' social contributions supplements paid	-	
D61xP	Social insurance scheme service charges paid	-	
D611R	Employers' actual social contributions received	-	
D612R	Employers' imputed social contributions received	-	
D62P	Social benefits other than STiK paid	-	
D62R	Social benefits other than STiK received	Pensions and other cash benefits from social security	The ICW Framework treats transfers as capital if they are large and irregular.
		Pensions and other benefits from employment-related social insurance	
		Social assistance benefits in cash from government	
		Current transfers in cash received from non-profit organisations (partly)	This may relate to social benefits as well as to miscellaneous current transfers (see D75R)
D72R	Non-life insurance claims		In the ICW Framework term insurance claims are treated as part of other capital transfers received, whereas accident insurance claims are treated as negative consumption expenditure
D71P	Non-life insurance premiums		Premiums actually paid (thus excluding the premium supplements related to property income received attributed to non-life insurance policy holders (D441AR)) are recorded as part of consumption expenditure in the

	Description	Correspondence in ICW Framework definition	Comments
			ICW Framework
D75R	Miscellaneous current transfers received	Current transfers received from other households	The SNA treats transfers as current if they do not involve a disposal or acquisition of an asset (excluding cash), whereas the ICW Framework treats transfers as capital if they are large and irregular.
		Other current transfers received, excluding STiK	
		Current transfers in cash received from non-profit organisations (partly)	This may relate to social benefits as well as to miscellaneous current transfers (see D75R)
D75P	Miscellaneous current transfers paid	Current transfers paid to other households	The SNA treats transfers as current if they do not involve a disposal or acquisition of an asset (excluding cash), whereas the ICW Framework treats transfers as capital if they are large and irregular.
		Current transfers paid to non-profit organisations	
		Other current transfers paid	
		Compulsory fees and fines (partly)	Fines imposed by courts of quasi-judicial bodies are recorded as part of miscellaneous current transfers (D75) in the SNA. The rest is part of current taxes on income and wealth (D5).
D75x	Net miscellaneous transfers received between resident households	Current transfers received from other households	
D63R1	Social transfers in kind on education received	Social transfers in kind	
D63R2	Social transfers in kind on health received		
D63R3	Other social transfers in kind received		
		Value of unpaid domestic services	Excluded from the SNA
		Value of services from household consumer durables	Excluded from the SNA
		Annuity and other regular payments from life insurance funds	These are included in the ICW Framework as part of property income. In the SNA these benefits are regarded as dissaving and only recorded in the financial accounts.
		Regular payments from private pension funds	Excluded from the SNA

Source: Annex B of OECD Framework for Statistics on the Distribution of Household Income, Consumption and Wealth (OECD, 2013[3]).

Table A A.3. Comparison of consumption according to ICW and SNA framework

	Description	Correspondence in ICW Framework definition	Comments
CP010	Food and non-alcoholic beverages		
CP020	Alcoholic beverages, tobacco and narcotics		
CP030	Clothing and footwear		
CP040	Housing, water, electricity, gas and other fuels		
CP041	Actual rentals on housing		
CP042	Imputed rentals on housing		
CP043	Maintenance and repair of dwellings		
CP044	Water supply and miscellaneous		
CP045	Electricity, gas and other fuels		
CP050	Furnishings, household equipment and routine household maintenance		

	Description	Correspondence in ICW Framework definition	Comments
CP060	Health		
CP061	Medical products, appliances and equipment		
CP062	Out-patient services		
CP063	Hospital services		
CP070	Transport		
CP071	Purchases of vehicles		Purchases of durables are treated as purchases of assets in the ICW Framework
CP072	Operation of personal transport equipment		
CP073	Transport services		
CP080	Communication		
CP090	Recreation and culture		In the ICW Framework all payments on gambling are treated as consumption expenditure with winnings as either negative consumption expenditure or capital transfers received. The SNA only includes the service charge component imputed to have been taken by the promotor as part of consumption expenditure.
CP100	Education		
CP110	Restaurants and hotels		
CP120	Miscellaneous goods and services		
CP12x	Miscellaneous (less FISIM, less insurance)		In the ICW Framework cost related to the repair and maintenance of durables are treated as costs to income from these durables. In the ICW Framework implicit service charges by financial institutions are excluded. These charges (except FISIM and insurance expenditure) are covered under this item in the SNA.
CP1261	FISIM		FISIM charges are not included under consumption expenditure in the ICW Framework but recorded as part of interest receipts and payments.
CP125	Insurance expenditures (life and non-life)		The ICW Framework records actual insurance premiums (excluding the premium supplements related to property income received attributed to non-life insurance policy holders (D441AR)) as part of consumption expenditure, whereas accident insurance claims are treated as negative consumption expenditure (term insurance claims are treated as part of other capital transfers). In the ICW Framework costs related to the insurance of durables are treated as costs to income from these durables. For the SNA this only relates to insurance related to activities that feed into operating surplus or mixed income.
D63	Social transfers in kind		
D8	Change in net equity of households in pension funds	-	Not included in the ICW Framework as social insurance pension entitlements are not treated as financial assets.

Source: Annex B of OECD Framework for Statistics on the Distribution of Household Income, Consumption and Wealth (OECD, 2013[3]).

References

European Commission et al. (2009), *System of National Accounts 2008*, https://unstats.un.org/unsd/nationalaccount/docs/SNA2008.pdf (accessed on 29 September 2017). [2]

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]

United Nations Economic Commission for Europe (2011), *Canberra Group Handbook on Household Income Statistics*, https://www.unece.org/fileadmin/DAM/stats/groups/cgh/Canbera_Handbook_2011_WEB.pdf (accessed on 27 September 2017). [1]

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

This publication presents guidance for the compilation of distributional results on household income, consumption and saving consistent with national accounts totals. These results are a key input for evidence-based policies targeting inequality and fostering inclusive growth, providing insights into key dimensions of material well-being across household groups. The results complement existing inequality measures by including elements that are often lacking from micro statistics and by providing inequality measures consistent with macroeconomic aggregates, broadening the scope for analyses, while also capturing households and transactions that are typically underrepresented in micro data. Moreover, while the estimates do require a number of statistical choices and assumptions, they have a high degree of international comparability because of the common methodology and their alignment to national accounts results. The handbook provides an overview of the conceptual framework underlying the distributional results and discusses various aspects in relation to the compilation and presentation of the distributional results. It aims to assist compilers in deriving high-quality distributional results and to provide users with more insights into the main benefits of these results as well as into the way that the results have been derived.

PRINT ISBN 978-92-64-82235-1
PDF ISBN 978-92-64-39678-4

9 789264 822351