WIFO WORKING PAPERS
599/2020

Distributional National Accounts (DINA) with Household Survey Data.

Methodology and Results for European Countries

Stefan Ederer
Stefan Humer
Stefan Jestl
Emanuel List

Distributional National Accounts (DINA) with Household Survey Data. Methodology and Results for European Countries

Stefan Ederer, Stefan Humer, Stefan Jestl, Emanuel List

WIFO Working Papers 599/2020

April 2020

Abstract

The paper builds Distributional National Accounts (DINA) using household survey data. We present a transparent and reproducible methodology to construct DINA whenever administrative tax data are not available for research and apply it to various European countries. By doing so, we build synthetic microdata files which cover the entire distribution, include all income components individually aligned to national accounts, and preserve the detailed socioeconomic information available in the surveys. The methodology uses harmonised and publicly available data sources (SILC, HFCS) and provides highly comparable results. We discuss the methodological steps and their impact on the income distribution. In particular, we highlight the effects of imputations and the adjustment of the variables to national accounts totals. Furthermore, we compare different income concepts of both the DINA and EG-DNA approach of the OECD in a consistent way. Our results confirm that constructing DINA is crucial to get a better picture of the income distribution. Our methodology is well suited to build synthetic microdata files which can be used for policy evaluation like social impact analysis and microsimulation.

E-mail: stefan.ederer@wifo.ac.at

2020/087/W/0

Medieninhaber (Verleger), Hersteller: Österreichisches Institut für Wirtschaftsforschung
1030 Wien, Arsenal, Objekt 20 | Tel. (43 1) 798 26 01-0 | https://www.wifo.ac.at
Verlags- und Herstellungsort: Wien

Kostenloser Download: https://www.wifo.ac.at/wwa/pubid/65907

Distributional National Accounts (DINA) with Household Survey Data: Methodology and Results for European Countries*

Stefan Ederer[a,b,**] · Stefan Humer[b,**] · Stefan Jestl[c] · Emanuel List[b,d]

[a] *Austrian Institute of Economic Research (WIFO)*
[b] *Vienna University of Economics and Business (WU)*
[c] *The Vienna Institute for International Economic Studies (wiiw)*
[d] *The Macroeconomic Policy Institute (IMK)*

[**] *Corresponding authors:* stefan.ederer@wifo.ac.at, shumer@wu.ac.at

April 15, 2020

Abstract

The paper builds Distributional National Accounts (DINA) using household survey data. We present a transparent and reproducible methodology to construct DINA whenever administrative tax data are not available for research and apply it to various European countries. By doing so, we build synthetic microdata files which cover the entire distribution, include all income components individually aligned to national accounts, and preserve the detailed socioeconomic information available in the surveys. The methodology uses harmonized and publicly available data sources (SILC, HFCS) and provides highly comparable results. We discuss the methodological steps and their impact on the income distribution. In particular, we highlight the effects of imputations and the adjustment of the variables to national accounts totals. Furthermore, we compare different income concepts of both the DINA and EG-DNA approach of the OECD in a consistent way. Our results confirm that constructing DINA is crucial to get a better picture of the income distribution. Our methodology is well suited to build synthetic microdata files which can be used for policy evaluation like social impact analysis and microsimulation.

Keywords: Distributional National Accounts; survey data; income inequality
JEL Classifications: C55, D31, E01

*This research project is supported by funds from the Oesterreichische Nationalbank (Oesterreichische Nationalbank anniversary fund, project 16728).

1 Introduction

More than two decades have passed since Anthony B. Atkinson reminded the economics profession of the importance of "bringing income distribution in from the cold" (Atkinson, 1997). Since then, matters of income and wealth distribution and their determinants have enormously gained in importance in the field of economics and the public debate. A recent milestone in this line of research is the work on Distributional National Accounts (DINA), which aim at reconciling micro- and macroeconomic data. The need to build DINA arises from the fact that, on the one hand, national accounts are limited in their representation of the socio-economic heterogeneity of households and individuals within an economy. Survey and tax data, on the other hand, offer information on distributions in great detail but are often inconsistent with macroeconomic variables. The main contribution of DINA is to allow for a distributional analysis of core macroeconomic concepts (i.e., national income or GDP) along the entire spectrum of the population.

So far, this endeavour has been pursued by two almost separate initiatives. On the one hand, a group of prominent scholars such as Thomas Piketty, Emmanuel Saez and Gabriel Zucman (see Piketty, Saez, & Zucman, 2018) collect their work under the umbrella of the World Wealth and Income Database (WID). On the other hand, national statistical agencies cooperate closely with the Organisation for Economic Co-operation and Development (OECD) in compiling distributional measures of household income, consumption and saving within the framework of the national accounts. While both aim to align distributional statistics to the economic aggregates, they differ in scope, concepts and methodology. Against this backdrop, it is crucial to analyse how these differences shape our understanding of the developments of inequality.

Even more importantly, both approaches have different goals. WID focuses on historical time series in order to track the long-run developments of inequality. They emphasise the role of the top tail for distributional statistics and therefore heavily rely on tax data and Pareto imputation techniques. While this has become the de facto standard procedure to derive inequality indicators for total income, it usually comes at the cost of losing information on separate income components and their relation to socioeconomic characteristics at the individual and household level. EG-DNA builds its analysis on sample surveys that are collected and maintained within the national statistical offices. Based on their longtime experience with both survey data and national accounts, their goal is to fuse their expertise in both branches to enhance their standard reporting on the development of living conditions. As a result, they are planning to publish updated inequality indicators and information on income, consumption and saving at the quintile level.

Obviously, both approaches have their virtues and are well suited to answer specific research questions. Unfortunately, from the perspective of this paper, they also share a common weakness. They lack either the granularity of income components or the information on the joint distribution of income with policy-relevant socioeconomic characteristics at the household level, which is necessary to conduct thorough social impact analysis. Such analysis has become a vital tool to inform policymakers about the various consequences of reform proposals and support their decision-making process.

This paper builds on the insights of both the DINA and the EG-DNA initiative. We propose a transpar-

ent and reproducible methodology, which uses only harmonised and publicly available data sources. It facilitates comparable cross-country analysis and can be applied whenever tax data are not available for research. As a result, we provide enhanced micro datasets that cover the entire distribution, include all income components separately aligned to national accounts, and preserve the detailed socioeconomic information available in the surveys. Furthermore, we compare the methodological choices and assumptions of the DINA and the EG-DNA approach and analyse their impact on the measured income distribution. In particular, we highlight the effects of imputations and the adjustment of the variables to national accounts totals. Moreover, we apply the methodology to calculate distributions for various income definitions of both the DINA and EG-DNA approach in a consistent way. Thus, using the same unit of analysis and the same target population, we can directly compare the results of these two, so far entirely separated, approaches.

Distributional National Accounts form the most recent endeavour in compiling data on the distribution of income and wealth. The origins of this line of work go back to Kuznets (1955), who was not only one of the first scholars compiling inequality statistics but was also very prominent in the process of developing the system of national accounts. Only several decades later, Atkinson (1971), Piketty (2003) and Piketty and Saez (2003) undertook comparable exercises and compiled time series of top income and wealth shares for the United Kingdom, France and the United States. These contributions breathed new life into inequality research and brought the topic back centre stage in economics (Piketty, 2015). They, however, came under some criticism for focussing only at the top of the distribution and not representing the developments in the bottom- and middle-income segments of the population. Around the same time and from a somewhat different angle, the Commission on the Measurement of Economic Performance and Social Progress (Stiglitz, Sen, & Fitoussi, 2009) concluded that aggregate measures such as GDP do not adequately picture the progress of well-being. As these statistics conceal differential outcomes for different groups of the population, more prominence should be given to the distribution of macroeconomic variables. These distributions, however, are traditionally not included in national accounts. Survey and tax data, which provide information on distributions are, on the other hand, usually not consistent with national accounts.

DINA can be interpreted as an attempt to address these shortcomings. In contrast to previous research, the DINA initiative aims at providing statistics on the entire distribution that are consistent with the macroeconomic aggregates. Since the guidelines on the general methodology (Alvaredo et al., 2016) were published, DINA have been compiled for various countries (e.g. Alvaredo, Chancel, Piketty, Saez, & Zucman, 2017; Blanchet, Chancel, & Gethin, 2019; Garbinti, Goupille-Lebret, & Piketty, 2018; Piketty et al., 2018; Piketty, Saez, & Zucman, 2019). Most of these datasets are built from administrative tax data, on which the DINA methodology relies as its primary source of distributional information. However, in a variety of countries tax data are not available to researchers, at least not at the necessary level of detail, which makes it even more difficult to apply the core DINA methodology. Even if tax data are available, the definition of taxable income and their coverage in official statistics differ in countries because of the tax law, so that an idiosyncratic approach is necessary for each country to construct DINA. Furthermore, access to data often hinges on special arrangements with local authorities that impede comparable cross-country analysis with administrative data.

The other prominent approach compiling distributional results in line with national accounts goes back to the OECD, Eurostat and national statistical institutes from various EU countries, which formed the Expert Group on Disparities in National Accounts (EG-DNA) in 2011. The EG-DNA has so far issued several reports on the progress of developing a standard methodology for building distributional measures for household income, consumption and saving that are consistent with national accounts (Fesseau & Mattonetti, 2013; Fesseau, Wolff, & Mattonetti, 2013; Zwijnenburg, Bournot, & Giovannelli, 2017). The approach primarily works with survey data and aims to arrive at breakdowns of the household sector at a quite aggregated level (usually quintiles). It differs from the DINA approach in that it focuses on households instead of individual adults as their unit of analysis and uses different income concepts, which do not include parts of national income outside the household sector, such as retained earnings or collective consumption (see section 2). A variety of countries, such as Australia, have already published distributional results following the guidelines of the EG-DNA (Seneviratne, 2016). As of today, the expert group is planning to release its final report with methodical guidelines any time soon.

The paper contributes to both strands of the literature in three ways: First, we build DINA datasets from survey data, using harmonized and publicly available sources of information (SILC, HFCS) and applying a simple and reproducible methodology. This "one-method-fits-all" approach can be applied to all European countries for which these sources are available without relying on any idiosyncratic knowledge about tax and social security systems. Furthermore, the uniform approach makes the results highly comparable across countries. Second, our results consist of synthetic microdata files, which cover the entire distribution, include all income components separately at a granular level, and are consistent with national accounts. Such data files are in principle usable for the impact analysis of policy measures, such as a tax or social security reform. Third, we compare the DINA and EG-DNA approach, using the same target population and unit of analysis. By doing so, we can analyze the implications of the use of different income concepts for the distribution, controlling for other sources of differences in outcomes.

We find that realigning the different income flows to national account totals and including the income of other sectors increase inequality in most countries. This generally confirms the need to construct DINA to get an adequate picture of the income distribution. Furthermore, the impact of a specific variable on inequality depends (1) on its distribution in the original dataset, (2) how well it is covered in the surveys (compared to national account totals), and (3) its contribution to national income. All three factors vary substantially between countries and data sources. Our methodology is in general very useful to build synthetic datasets which are better suited for policy analyses than the original survey data. It should, however, be applied with some caution when building a longer-term time series of inequality.

The paper is structured as follows: Section 2 presents a detailed description of our methodology and compares it with those of the two approaches mentioned above. In particular, it discusses the matching between household surveys and national accounts as well as the imputations of variables that are not included in the microdata. Section 3 discusses the results for the distribution of income both for the household sector and for national aggregates. Section 4 concludes.

2 Methodology

2.1 Data and income concepts

The aim of the paper is to build Distributional National Accounts from household surveys. We generally follow the methodology of Piketty et al. (2018) but use the insights of the EG-DNA initiative to reconcile survey data with national accounts. Since we take a "one-method-fits-all" approach for European countries, we only use information which is provided by common sources. To check for the robustness of our results and the sensitivity of the methodology, we build DINA based on two different data sources: the European Survey on Income and Living Conditions (SILC), provided by Eurostat, and the Household Finance and Consumption Survey (HFCS) of the European Central Bank, both of which are extensively used for distributional analyses in the literature.

The SILC provides yearly data and includes several types of incomes as well as socio-economic data for around 230,000 households and 570,000 individuals in 29 European countries. The most recently published wave of the HFCS was conducted in 2014, including more than 84,000 households and 210,000 individuals in 20 countries. Detailed descriptions of both datasets and their methodology are provided by Eurostat (2019) and the Household Finance and Consumption Network (2016).

To re-align survey data to national accounts' totals, we use annual data for non-financial sectoral accounts, provided by Eurostat. The data include incomes and expenditures for institutional sectors such as private households, non-financial and financial corporations, and the government. They are harmonized across countries and are commonly used as key indicators for economic development and well-being. A detailed description of the data is found in Eurostat (2013).

Following Piketty et al. (2018), we use national income as our key income concept. National income includes all institutional sectors of the economy, not only private households, which are the primary focus of surveys. The basic idea of using national income in DINA is that all income ultimately accrues to households and individuals in one form or another.

The DINA literature focuses on four different definitions of national income. Pre-tax factor income (PRTFI) is the primary income of all domestic sectors of the economy, net of depreciation and including net income flows with other countries. Pre-tax national income (PRTNI) corrects pre-tax factor income for effects of the social securtiy system by adding benefits and substracting contributions. For the latter, the DINA guidelines propose two alternative definitions: First, taking into account the redistributive of the pension system only, and second, considering the effects of the entire social security system. We choose the second approach because pension and other social contributions cannot be separated due to limitations in the Eurostat data.[1] Since beneficiaries and contributors are usually different households and individuals, and the money flows do not add up, we furthermore include the balance between these two items in order to make pre-tax national income equal to pre-tax factor income. Pre-tax national income is

[1] Due to the limitaions in the data, we deviate from the DINA guidelines in that we include all monetary transfers (D62) instead of pension and other social benefits (D621+D622). The difference are social assistence benefits in kind (D623), which cannot be separated in the data. Furthermore, we do not include private pensions (D442), which are not available from the Eurostat database.

useful to compare income in countries with notable differences in their pension systems. Private pension plans are already a part of pre-tax factor income, so that countries with substantial private pensions usually exhibit a significantly lower inequality of pre-tax factor income than those which rely more on the public pension system. Pre-tax national income is considered as the primary variable for income before taxes by the DINA literature.

Post-tax disposable income (POTDI) deducts all direct and indirect taxes from pre-tax national income. It is thus substantially smaller in value than the other three income concepts. The size of the public sector differs between countries, so that post-tax disposable income is challenging to compare. Nevertheless, it conceptually comes closest to household disposable income. The DINA literature, however, focuses more on post-tax national income (POTNI), which is obtained by adding social transfers in kind and collective consumption. Post-tax national income is the key indicator used to describe the distribution once we account for the effects of the public sector.

In addition to the abovementioned DINA income concepts, we use two more income concepts from the EG-DNA initiative: household disposable income (DINC) and adjusted household disposable income (ADINC). Both account only for the household sector and leave the other sectors aside. Disposable income includes all forms of employment and self-employment income as well as imputed rents, capital income and social transfers, net of taxes and contributions. Adjusted disposable income furthermore includes social transfers in kind. By calculating these two additional types of income, we can compare the results of the DINA methodology to those of the EG-DNA approach. Table 1 provides an overview of the income concepts, its calculation from national accounts, and the corresponding variables in the HFCS and SILC datasets (see also the following subsection).

2.2 Matching survey data with national accounts

All income concepts are calculated at the micro-level. Following Piketty et al. (2018), our unit of analysis is the individual adult, aged 20 years or above. We thus exclude children and teenagers, as well as people in institutions.[2] Furthermore, we construct DINA based on two different ways of income splitting, the 'equal' and 'individual split' (Piketty et al., 2018). The first approach divides the pooled household income equally between individuals, thus eliminating inequalities within households. While this concept is similar in idea to the household concept used by the EG-DNA initiative, there are significant differences: In contrast to the latter, the equal split weighs all adult individuals equally and assumes away any economies of scale of living in a shared household. The individual split, in contrast, attributes personal income to individuals and divides income which is only available in the data at the household level, such as capital income, equally between persons. The individual split generally yields higher inequality across individuals than the equal split.

Our methodology comprises two major steps. First, we assign to each income component in the micro dataset a corresponding variable from the national accounts. Since most types of income are defined differently in micro and macro data, the assignment is in some cases only the closest possible approx-

[2]The share of income of these persons is small in European countries. See Zwijnenburg (2019) for a detailed discussion.

Table 1 **Income concepts**

1) *EG-DNA*		**National Accounts**	**HFCS**	**SILC**
Gross wages and salaries		D11	PG0110	PY010G + PY020G + HY110G
Employer social contributions	+	D12	EUROMOD	EUROMOD
Gross operating surplus & rents received	+	B2G + D45	HG0310 + IMP	HY040G + HY030G
Consumption of fixed capital	−	P51C [part]	IMP [B2G]	IMP [B2G]
Gross mixed income	+	B3G	PG0210	PY050G + HY170G
Consumption of fixed capital	−	P51C [part]	IMP [B3G]	IMP [B3G]
Interest w/o FISIM & distributed income of corporations	+	D41G + D42	HG0410 + HG0510	HY090G
FISIM for interest received	+	D41G – D41	IMP [D41G + D42]	IMP [D41G + D42]
Other property income received	+	D43 + D44	IMP [D41G + D42]	IMP [D41G + D42]
Interest paid	−	D41G	DI1412	HY100G
FISIM for interest paid	+	D41 – D41G	IMP [D41G]	IMP [D41G]
Other property income paid	−	D4 (exc. D41G)	neutral	neutral
Primary income (net)	=	PRINC (B5N)		
Current taxes on income: employment income	−	D51 [part]	EUROMOD	EUROMOD
Current taxes on income: simulated property income	−	D51 [part]	EUROMOD	EUROMOD
Current taxes on income: non-simulated property income	−	D51 [part]	IMP [D51]	IMP [D51]
Current taxes on wealth	−	D59	IMP [D51]	IMP [D51]
Employee social contributions	−	D61 – D12	EUROMOD	EUROMOD
Monetary transfers	+	D62	PG0310 + PG0510 + HG0110	PY100G + PY110G + PY090G + PY120G + PY130G + PY140G + HY050G + HY060G + HY070G
Other current transfers net	+	D7R – D7P	neutral	neutral
Disposable income (net)	=	DIINC (B6N)		
Social transfers in kind	+	D63	equal	equal
Adjusted disposable income	=	ADINC (B7N)		
2) *DINA*				
Primary income S14S15		B5N (S14S15)		
Primary income S11	+	B5N (S11)	IMP [D41G + D42]	IMP [D41G + D42]
Primary income S12	+	B5N (S12)	IMP [D41G + D42]	IMP [D41G + D42]
Net operating surplus & mixed income S13	+	B2A3N (S13)	IMP [D41G + D42]	IMP [D41G + D42]
Net property income S13	+	D4 (S13)	IMP [D41G + D42]	IMP [D41G + D42]
Net indirect taxes	+	D2 – D3 (S13)	neutral	neutral
Pre-tax factor income	=	PRTFI		
Social contributions	−	D61 (S14S15)	EUROMOD	EUROMOD
Monetary transfers	+	D62 (S14S15)	PG0310 + PG0510 + HG0110	PY100G + PY110G + PY090G + PY120G + PY130G + PY140G + HY050G + HY060G + HY070G
Difference D61 & D62	+	D61 – D62 (S14S15)	neutral	neutral
Pre-tax national income	=	PRTNI		
Net indirect taxes	−	D2 – D3 (S13)	neutral	neutral
Current taxes on income and wealth S14S15	−	D51 + D59 (S14S15)	EUROMOD + IMP	EUROMOD + IMP
Current taxes on income and wealth (other sectors)	−	D51 + D59 (other)	IMP [D51 + D59 (S14S15)]	IMP [D51 + D59 (S14S15)]
Post-tax disposable income	=	POTDI		
Social transfers in kind	+	D63 (S14S15)	equal	equal
Collective consumption	+	P32 (S13)	equal	equal
Primary surplus S13	+	D2 – D3 + D51 + D59 – D63 – P32 (S13)	neutral	neutral
Post-tax national income	=	POTNI		

Notes: This table shows the composition of income concepts, their variable codes in the raw data or if the data had to be imputed.
Source: SNA, HFCS, SILC.

imation. Our two sources, SILC and HFCS, usually have different variable definitions, so that neither of the two datasets is ideal to be aligned to national accounts. This methodical step corresponds to the DINA methodology with the difference that we use incomes from survey data and not tax data. Since the income definitions in tax data also deviate from national accounts, the abovementioned methodical imprecision is inherent in the DINA methodology.

Table 2 provides a summary of the correspondence of macro and micro data of private households. Gross wages and salaries in the national accounts (D11) include cash and non-cash employee income. The latter is only available in SILC data, where employee income is split in cash or near-cash income (PY010G) and non-cash employee income (PY020G). As a third component, gross income of people younger than 16 years (HY110G) is provided at the household level and is split between the individual adults. For the HFCS, gross cash employee income (PG0110) is the only variable assigned to the corresponding macro variable.

Gross operating surplus (B2G) in the national accounts includes own-account production of accommodation services by owner-occupier households (imputed rents) as well as income from rent. These correspond to the variables HY030G and HY040G in SILC. The HFCS only includes income from real estate property (HG0310). We thus calculate imputed rents following a capital market approach (see next subsection). Furthermore, both SILC and HFCS cover rents from land. We thus add the variable D45G from national accounts.

Gross mixed income (B3G) represents the surplus or deficit of unincorporated enterprises recorded in the household sector. The variable also includes the surplus from underground and own-account production. The corresponding variables of SILC are gross cash benefits or losses from self-employment (PY050G), and the gross value of goods produced for own consumption (HY170G). The HFCS does not cover the value of own consumption, its corresponding variable is thus only the gross self-employment income (PG0210). Furthermore, both surveys do not cover underground activities, which is one of the reasons for a low coverage rate for this income source (see section 3).

For capital income, only received interest not adjusted for FISIM (D41G, received) and distributed income of corporations (D42) have corresponding variables in the micro datasets. SILC provides a summary variable for gross interests, dividends, and profits from capital investment in unincorporated businesses (HY0900G), whereas in the HFCS capital income is divided into gross income from financial investments (HG0410) and gross income from a private business other than self-employment (HG0510). Interest payments in national accounts (D41G, paid) are assigned interest payments for outstanding loans from the HFCS (DI1412) and gross interest payments on mortgages in the SILC (HY100G). Both variables only partly correspond to the national account's definition.

For social benefits (D62), national account data are split into pensions, unemployment benefits, and other social transfers, each of which has corresponding variables in the HFCS (PG0310, PG0510, HG0110). For the SILC, we aggregate old-age benefits (PY100G) and survivor benefits (PY110G) for pensions and assign all other social benefits except unemployment benefits (PY090G) to other social transfers in national accounts data.

Table 2 **Matching macro and micro variables**

National Accounts		HFCS		SILC	
Gross wages and salaries	D11	Gross cash employee income	PG0110	Gross cash or near cash employee income	PY010G
				Gross non-cash employee income	PY020G
				Gross income received by people aged under 16	HY110G
Gross operating surplus	B2G	Gross rental income from real estate property	HG0310	Gross income from rental of a property or land	HY040G
Imputed rents	B2G			Gross imputed rent	HY030G
Rents (rec.)	D45				
Gross mixed income	B3G	Gross self-employment income (profit/losses of unincorporated enterprises)	PG0210	Gross cash benefits or losses from self-employment	PY050G
				Gross value of goods produced for own-consumption	HY170G
Interest (rec., w/o FISIM)	D41G	Gross income from financial investments	HG0410	Gross interests, dividends, profit from capital investment in uncorporated business	HY090G
Distributed income of corporations	D42	Gross income from private business other than self-employment	HG0510		
Interest (paid, w/o FISIM)	D41G	Interest payments	DI1412	Gross interest repayments on mortgage	HY100G
Social benefits other than social transfers in kind (pensions)	D62	Gross income from public pensions	PG0310	Gross old-age benefits	PY100G
				Gross survivor benefits	PY110G
Social benefits other than social transfers in kind (unemployment)	D62	Gross income from unemployment benefits	PG0510	Gross unemployment benefits	PY090G
Social benefits other than social transfers in kind (other)	D62	Gross income from regular social transfers	HG0110	Gross sickness benefits	PY120G
				Gross disability benefits	PY130G
				Gross education-related allowances	PY140G
				Gross family/children-related allowances	HY050G
				Gross social exclusion not elsewhere classified	HY060G
				Gross housing allowances	HY070G

Notes: This table shows the detailled correspondence of macro and micro data for private households.
Source: SNA, HFCS, SILC.

2.3 Imputation of additional variables

The second major step of our methodology is to impute from other sources those variables from national accounts that have no direct correspondence in the survey data (Table 1). We generally follow the guidelines of the DINA methodology (Piketty et al., 2018) but use only information from common sources for all countries for the imputations.

The first and foremost part of imputations concerns taxes and social contributions. Both are usually not separately available in survey data.[3] We thus use EUROMOD, a widely available tax-benefit microsimulation model for the EU to calculate tax and contribution rates along the income distribution for each country. In a first step, we separate direct taxes in national account data (D51) between employment and capital taxes, using data on taxation by the European Commission (2016).[4] Taxes on employment income and capital income are then simulated with EUROMOD at the individual level and are aggregated to average tax rates on these two income sources for each percentile of the income distribution. Finally, they are levied on all individuals within that income segment. A part of capital taxes is not simulated but directly imputed according to the distribution of the simulated part. Furthermore, we impute taxes on wealth (D59) similar to the sum of income taxes (D51). Social contributions are separated into employer (D12) and employee contributions. Both are simulated at the micro level using EUROMOD.

The second part of imputations concerns imputed rents, which are included in SILC but are not covered by the HFCS. To do so, we follow the capital market approach (Balcazar, Ceriani, Olivieri, & Ranzani, 2014). The HFCS provides information on the current value of an owned dwelling[5] and of outstanding mortgages, from which the net value of the residence is derived. The value is then multiplied by an exogenous rate to calculate the imputed rent.[6] For the SILC, we deduct interest paid on mortgages from gross imputed rents to derive their net value.

Further imputations for other items of primary income of the household sector include consumption of fixed capital as well as FISIM and other property income (D43 and D44). The first is split between gross operating surplus and gross mixed income according to their relative macroeconomic proportions and distributed to individuals in the same way as these two variables. The FISIM and other capital income are imputed according to the distribution of interest income and distributed income of corporations (D41 and D42). The FISIM for paid taxes is distributed accordingly. Finally, other capital income paid is distributed neutrally (Table 1).

Moving from the primary income of the household sector to pre-tax factor income, we impute the primary income of the other sectors, which are not included in survey data. Since the primary income of non-financial and financial corporations mostly consist of retained earnings, whose beneficiaries are the owners of these corporations, we generally assume that they are distributed like capital income. We proceed in

[3] The SILC provides both gross and net incomes, however without accounting for different taxes and social contributions separately. The HFCS does not even include any information on net incomes.

[4] See Table A.3 in Appendix A.

[5] For free users, i.e. people who do neither own the residence nor pay rent, the value of the dwelling is not provided by the HFCS. We derive it from the square meter price for similar types of dwelling in the group of homeowners.

[6] Following the literature (e.g. Fessler, Rehm, & Tockner, 2016), we chose a value of 3% for the exogenous rate, which approximately represents the average interest rate of a risk-free asset such as a 10-year treasury bond.

the same way with the operating surplus and the property income of the state. Net indirect taxes (D2 minus D3) are distributed neutrally since they are related to production and not income. Two other sizable and important items are social transfers in kind and collective consumption, both of which we add to calculate post-tax national income. Following Zwijnenburg et al. (2017) we distribute both variables equally between individuals.

3 Results

3.1 Distributions for DINA and EG-DNA income concepts

We adopt the methodology of the previous section to construct Distributional National Accounts for twelve European countries. Since we work with two different surveys (SILC and HFCS) for two specific years (2010 and 2014) and use two different ways to split the household income (equal and individual split, see section 2), we build eight synthetic micro datasets per country. Each DINA dataset consists of adult individuals that represent the whole population. We account for different income components separately so that we can calculate their specific impact on the distribution. With such a dataset, it is generally possible to analyze the distributional effects of specific policy measures, such as a tax reform.

Figure 1 and 2 summarize the distribution of the different income concepts for all twelve countries.[7] Both surveys generally yield similar results, especially for the patterns of inequality in the income concepts.[8]

We start with the primary income of the household sector (PRINCO), which includes mostly variables which are directly covered in the surveys (see Table 1). Primary income is in general, highly unequally distributed in all countries, with a Gini coefficient between 40 and 55. It includes self-employment and capital income, both of which are usually concentrated on the top of the distribution. The top 5% share of the primary income of the household sector ranges from 15 to 25%.

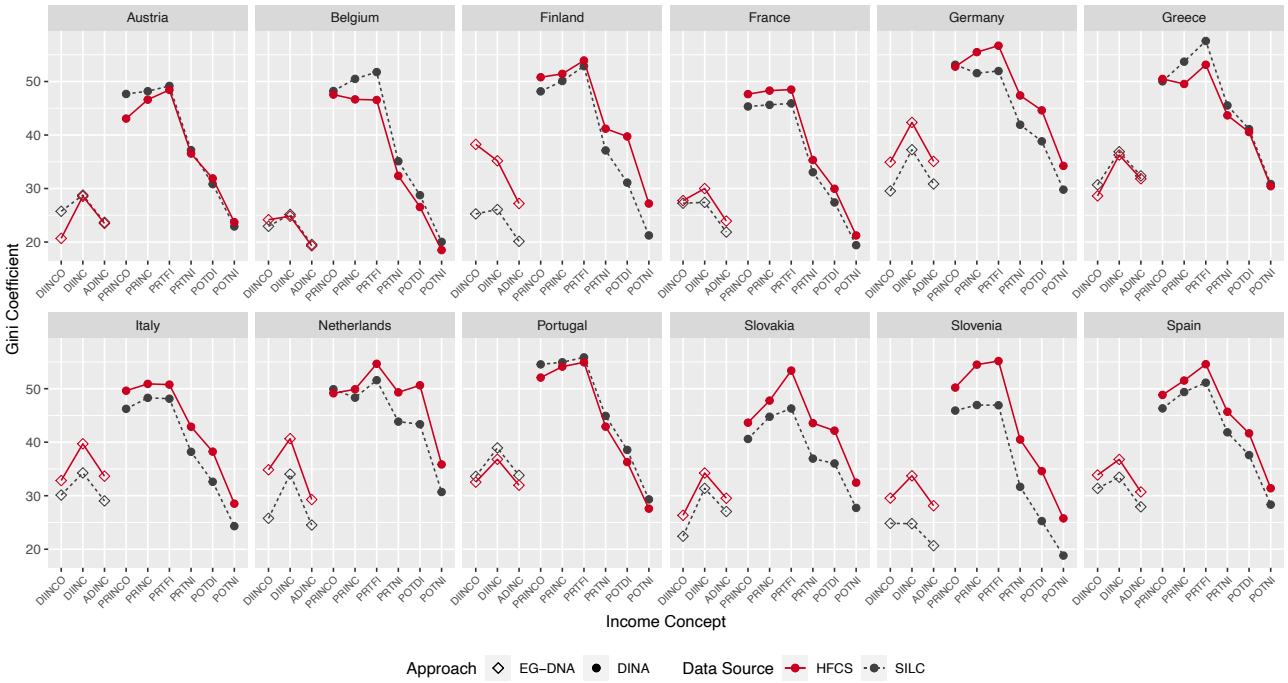
Rescaling the primary income to national accounts data (PRINC) generally increases the Gini coefficient in most of the countries by up to five points. In some cases, however, inequality remains unchanged or even decreases slightly. This is surprising because self-employment and capital income are usually not very well covered in survey data and have thus significant rescaling factors. One reason is that the Gini coefficient generally does not map the inequalities at the top and the bottom well. The top 5% share, which is more appropriate for this purpose, increases considerably in almost all countries due to the realignment of primary income to national accounts.

Even for this variable, however, there are some differences in these patterns with regard to the data source. In Belgium, the top 5% share does not increase for HFCS data. In Slovenia, the increase due to the realignment is rather small for SILC data. These differences can primarily be attributed to variations

[7] Throughout this section, we present the results for 2014. Following Piketty et al. (2018), our standard concept is the equal split. Results for 2010 and the individual split can be found in Appendix B and C.

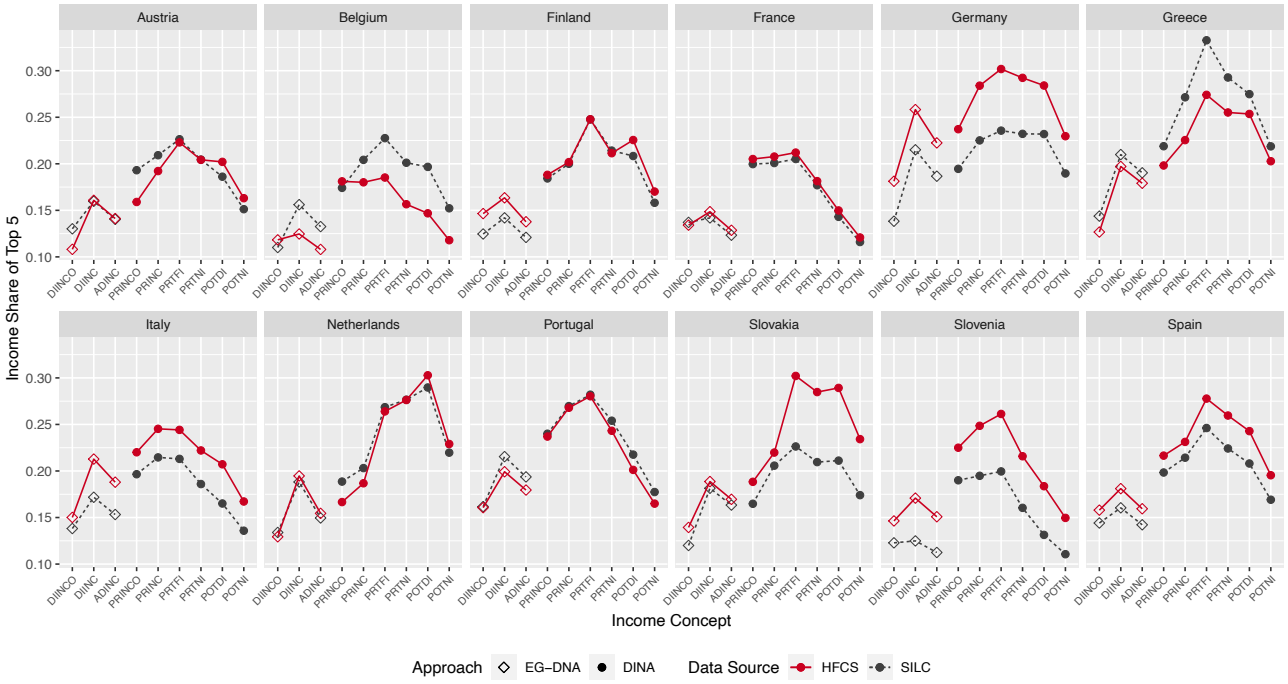
[8] In contrast to the patterns in the income concepts, the *level* of inequality differs considerably between SILC and HFCS in some countries. This is generally due to variations in the coverage rate and the distribution of gross mixed income and/or capital income between these two datasets. See discussion further below.

Figure 1 **Inequality for income concepts: Gini coefficient (SILC & HFCS, Wave 2014, Equal split)**



Notes: This graph shows the Gini coefficient for different DINA income concepts by country.
Source: SNA, SILC, HFCS, own calculations and illustration.

Figure 2 **Inequality for income concepts: Top 5% share (SILC & HFCS, Wave 2014, Equal split)**



Notes: This graph shows the income share of the Top 5% for different DINA income concepts by country.
Source: SNA, SILC, HFCS, own calculations and illustration.

in the coverage ratio or the inequality for gross mixed income and/or capital income.[9] Particularly interesting is the case of France, for which we find no substantial increase in inequality both for HFCS and SILC data. Capital income is very well covered there in both data sets so that realigning it to national accounts does not affect whatsoever on inequality.

The next step, from primary income to pre-tax factor income (PRTFI), usually leads to a small increase in inequality, albeit not in all countries. This effect stems from the inclusion of the primary income of non-financial and financial firms as well as the state. The increase in the top 5% share is usually considerably higher, which confirms once again that the latter maps the inequality at the top better than the former. We find similar peculiarities as for the previous methodological step, which is not surprising because we distribute the primary income of these sectors similarly to capital income. Indeed, the top 5% share only rises slightly in France. Likewise, we find that in Belgium, the increase for the HFCS data is much smaller than for SILC, and vice versa in Slovenia. Furthermore, there is an unusually large variation in the increase of the top 5% share between the two data sources in Slovakia. The reason is that capital income (and consequently also the primary income of the corporate sector in our methodology) is more concentrated at the top in the HFCS data.

Interestingly, differences in national accounts also play a role in these particularities. For example, the ratio of the primary income of non-financial and financial firms to national income is rather low for France, which contributes to the above-mentioned small increase in inequality. Similarly, this effect is almost nil in Italy, because the primary income of this sector is minuscule compared to national income.[10] In contrast, the increase in inequality from PRINC to PRTFI is high in the Netherlands due to the substantial contribution of the primary income of corporations to national income. Likewise, the differences as mentioned earlier in the increase in the top 5% share between the two data sources in Slovakia is enlarged because of a similar circumstance.

The decline in inequality from pre-tax factor income to pre-tax national income (PRTNI) is generally very large in all countries. This reflects the redistributive impact of the welfare state, i.e. social contributions and monetary transfers, including pensions. The movement to post-tax disposable income (POTDI) in contrast results in a much smaller decline in inequality. This decline reflects the effects of the tax system (i.e. income and wealth taxes) and is consistent with the fact that the tax system is not very progressive in many countries. Finally, going from post-tax disposable income to post-tax national income (POTNI) again entails a large decrease in inequality in all countries, which is due to the inclusion of social transfers in kind and collective consumption. We distribute both income components equally over the population, which obviously reduces inequality. This effect is generally strong and as large as the effect of the welfare system in many countries.

To summarize, rescaling household income and imputing income from other sectors increases inequality in most countries. There are considerable differences with respect to the data source. These differences

[9] The coverage ratio for gross mixed income in Belgium is higher for the HFCS than for the SILC. In Slovenia, the top 5% share in gross mixed income is higher for the HFCS. Similar differences can be found for capital income. Such variations are partly also responsible for the differences in outcomes between the two inequality measures. See below for a detailed discussion.

[10] Variations in national accounts are thus another source contributing to the patterns in inequality, which is a characteristic inherent in the DINA methodology. See Table A.2 in Appendix A for a summary of the variables in national account data.

primarily result from the coverage rate and/or distribution of mixed income and capital income, as well as the importance of these variables and the primary income of the corporate sector in national accounts. The increase in inequality due to the first two methodological steps is, however, much smaller than its decrease because of the redistribution by the state. All three steps in the redistribution process, i.e. from pre-tax factor income to pre-tax national income and then to post-tax disposable and national income, contribute significantly to this reduction in inequality. This decline reflects the redistributive effects of the welfare state, the tax system, collective consumption and social transfers in kind. The highest decreases in inequality are due to the first and the last two, whereas the tax system is usually not very progressive in most countries. This generally reproduces the findings of Piketty et al. (2018) and other contributions to the World Income Database.

In general, the results are highly similar for both years for which we construct DINA.[11] There are minor variations as to the coverage rates and distributions of the variables, in particular of mixed income and capital income, so that the patterns of inequality between income concepts change slightly. The distributions of the individual split are more unequal for all income concepts, countries, and years, given the fact that for this methodical approach, personal income remains attributed to individuals (see section 2 and Piketty et al., 2018).

Finally, we calculate the disposable income and adjusted disposable income of the household sector in order to compare the DINA income concepts with those of the EG-DNA approach. Moving from primary income (PRINC) to disposable income (DINC), inequality declines substantially. This step includes the deduction of taxes and contributions from income as well as the addition of monetary transfers. It thus reflects the effect of the state on the income distribution of the household sector. Interestingly, the impact of rescaling of the variables on the distribution of disposable income (i.e., moving from DINCO to DINC) is larger than on that of primary income (from PRINCO to PRINC). This is primarily due to a considerable upscaling of taxes and monetary transfers, variables which are not considered in primary income. The difference between unscaled primary and disposable income (PRINCO vs. DINCO) is therefore also larger than between rescaled incomes (PRINC vs. DINC). Furthermore, the distributions of disposable income of the household sector and post-tax disposable income show considerable differences in most countries. This is due to the inclusion of the primary income of the corporate sector and the state.

Adjusting disposable income for social transfers in kind (ADINC) expectedly decreases inequality, since we distribute the latter equally across individuals. Interestingly, the Gini coefficient of adjusted disposable income is similar to post-tax national income in many countries. Apparently, the distributive effect of the primary income of the other sectors and that of collective consumption, which are the main differences between these two income concepts, partly neutralize each other. The step from disposable (DINC) to adjusted disposable income (ADINC) is nevertheless smaller than the one from post-tax disposable (POTDI) to post-tax national income (POTNI), because the latter also includes collective consumption.

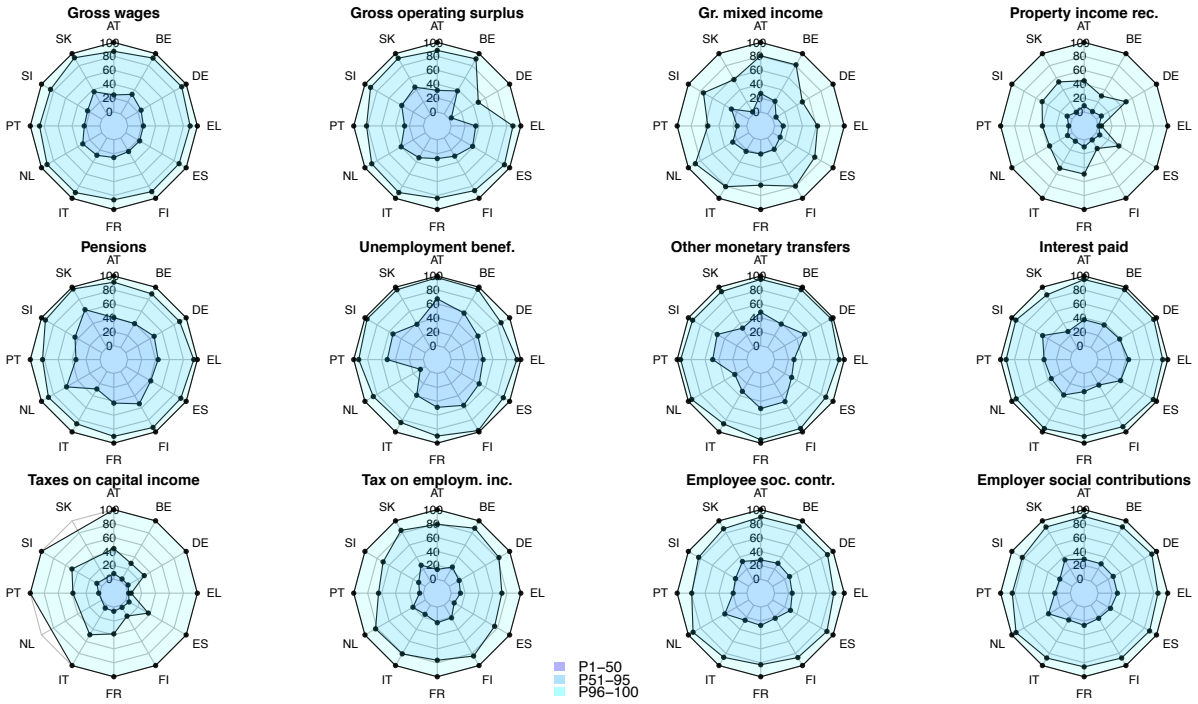
[11] Results for 2010 as well as for the individual split are reported in Appendix B and C.

3.2 Contributions to inequality

In the previous subsection, we saw that inequality varies substantially between different income concepts. Furthermore, we have found differences in the patters with regard to the data sources because of variations in the coverage rates and the distributions of specific variables. In this subsection, we discuss the factors which contribute to inequality in a more systematic manner.

For the variables which are included in the surveys, its impact on the distribution depends on three factors: First, its distribution in the original micro dataset. Second, the ratio between the sum of that variable over all individuals in the dataset and the value of the corresponding aggregated variable in national accounts, i.e. the coverage rate. The inverse of the coverage ratio is the factor by which the income component is multiplied when we rescale that variable (see section 2). Third, the contribution of the specific variable to total national income in the national accounts.

Figure 3 **Original distributions (SILC, Wave 2014, Equal split)**

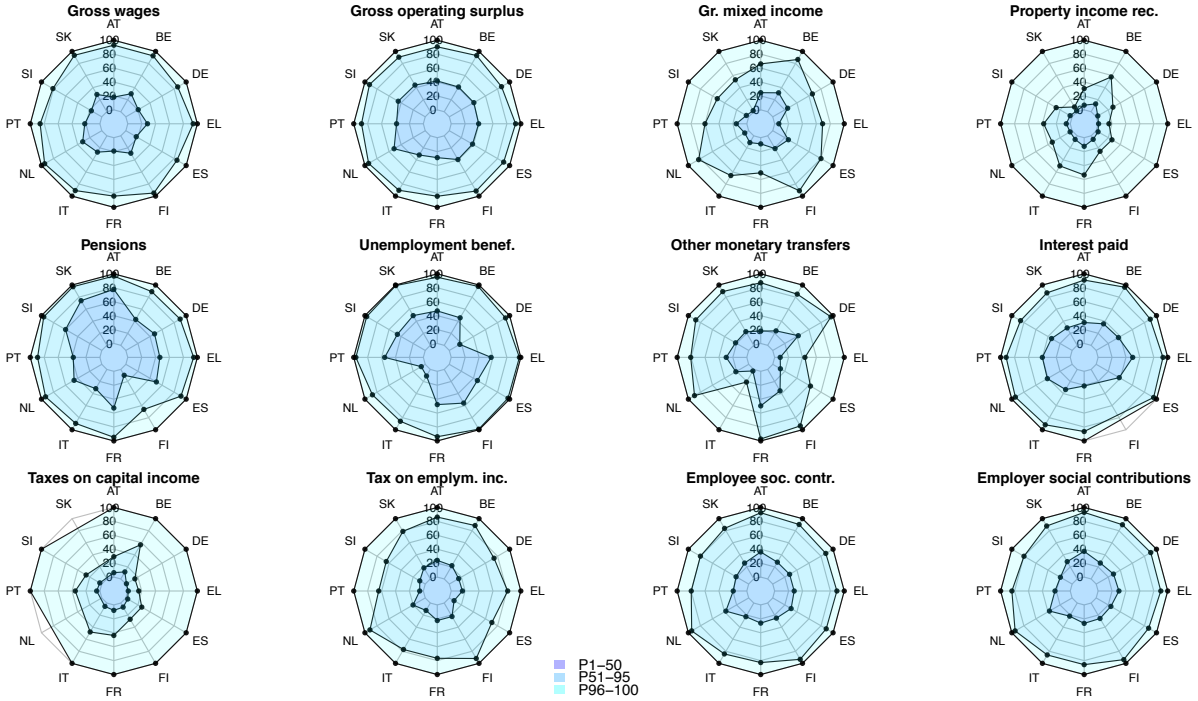


Notes: This graph shows the distribution of income components in the raw data. The income groups are defined on the basis of post-tax national income.
Source: SILC, own calculations and illustration.

Figures 3 and 4 show the distribution of the income components in the original micro datasets.[12] The innermost (dark) areas represent the share of the bottom 50% of the population, whereas the outermost (light) areas account for the share of the top 5%. Gross wages are generally concentrated in the middle of the income distribution; both the bottom 50% share and the top 5% share are low compared to other variables. The share of the bottom 50% in gross operating surplus, which mainly consists of actual and imputed rents, is generally a bit higher than in wages, whereas the top 5% share is quite similar. Mixed

[12]The income groups are defined on the basis of post-tax national income. The results for pre-tax national income are highly similar.

Figure 4 **Original distributions (HFCS, Wave 2014, Equal split)**



Notes: This graph shows the distribution of income components in the raw data. The income groups are defined on the basis of post-tax national income.
Source: HFCS, own calculations and illustration.

income and property income, on the other hand, are much more concentrated at the top of the income distribution.

Pensions, unemployment benefits, and other monetary transfers, in contrast, are more equally distributed. The share of the bottom 50% of these income components is generally high, whereas that of the top 5% is rather negligible. The distribution of social contributions is usually similar to that of wages, whereas taxes on both employment and capital income are more unequally distributed. The tax on the latter is primarily paid by the upper part of the income distribution, the former by the middle. The patterns quantitatively differ somewhat between SILC and HFCS but are qualitatively similar.

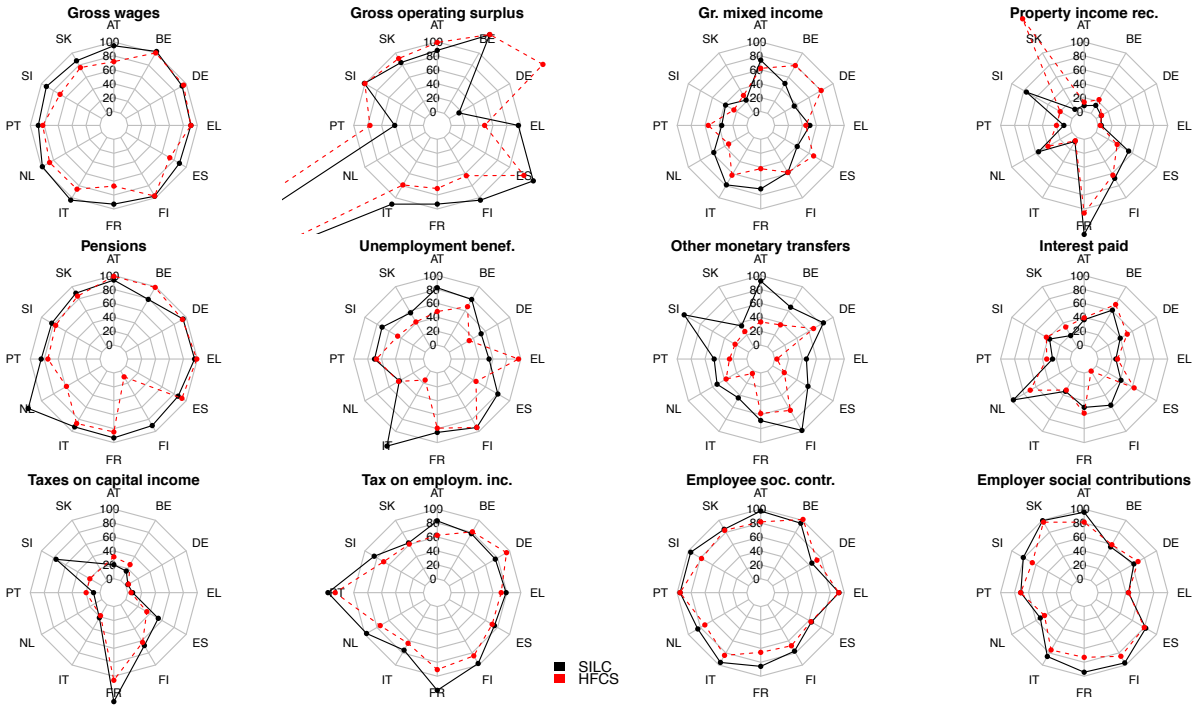
The impact of the income components on the distribution changes considerably when we rescale them to national accounts aggregates, because each variable is distributed differently across the population. Figure 5 shows how well the survey data cover the different income components compared to national accounts. Gross wages and salaries are generally well captured, the coverage rate for SILC data ranges between 90 and 105%. Some countries match their survey data for wages and salaries with administrative data, so that the coverage rate is close to 100%. Gross wages in the HFCS are usually a bit lower, but the coverage rate is still respectable for many countries. The rescaling factor for wages is consequently low for both surveys.

Gross operating surplus, as a ratio of national accounts totals, in general varies a lot across countries. In many cases, the SILC, which reports imputed rents in the data, covers almost 100% of national aggregates.

In some countries, however, coverage is not very good. For the HFCS, we impute rents based on data on homeownership and the capital market approach (see section 2). The deviation between SILC and HFCS can thus be considerable due to these differences in the data sources and the methodology.

The coverage rates of mixed and capital income are generally much lower and vary considerably between countries. With the very exception of the latter in France, their rescaling factors are usually quite large. There is also substantial variation in the coverage rates for pensions, unemployment benefits and other monetary transfers across countries. They are however much higher than for capital income, in particular for pensions for which they amount to almost 100% in some cases. Taxes on employment and social contributions are in general also well covered, given by the fact that they are simulated by EUROMOD and that wages are the primary source for employment income.

Figure 5 **Coverage rates (Wave 2014)**



Notes: This graph shows the coverage of income components in comparison to the aggregates of the national accounts.
Source: SNA, SILC, HFCS, own calculations and illustration.

Figure 6 summarizes the impact of the variables on the distribution of post-tax national income.[13] Naturally, those income components which are unequally distributed in the survey data, whose coverage rate is low (so they consequently have a high rescaling factor) and/or their contribution to national income is high, have the highest impact on the distribution.

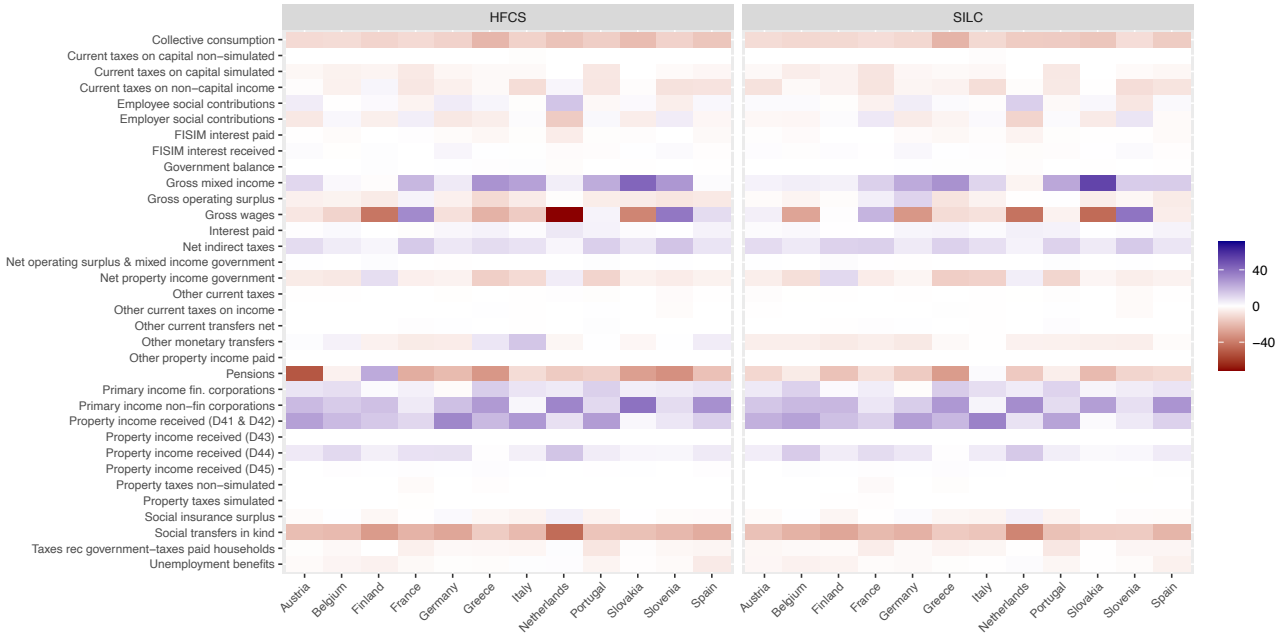
Interestingly, gross wages have mixed effects on the distribution of post-tax national income. In some countries, they reduce inequality when added and in others they increase it. Wages are generally concentrated in the middle, so that their impact depends on how the other variables are distributed.[14] In the case of already highly unequal income distribution, the addition of wages is likely to entail a decrease

[13] The impact of the variables on the distribution of pre-tax national income is highly similar.
[14] This is obviously true for all variables. However, wages stand out in that the sign of their effect varies, not just the size.

in inequality, and vice versa. Furthermore, wages make up a large part of national income, so that their impact (in whatever direction) is higher than for other income components. A similar albeit less marked ambiguous effect can be observed for the operating surplus, and, in some countries, for pensions.

Mixed income and property income on the other hand increase inequality of post-tax national income in all countries. Both are very unequally distributed and have low coverage in the surveys, so that their rescaling factors are high. Even though mixed income is less than 25% and property income less than 10% of national income, their impact on the distribution is substantial.[15]

Figure 6 **Contribution to Inequality of post-tax national income (SILC & HFCS, Wave 2014, Equal split)**



Notes: The contributions of the variables to post-tax national income distribution are based on the Gini coefficient. They are calculated by comparing the income distribution with and without the respective variable. Contributions marked in blue make the distribution more equal; those in red more unequal.
Source: SNA, SILC, HFCS, own calculations and illustration.

Besides the variables discussed above, which directly come from survey data, the imputed variables also have a substantial impact on the distribution. Their effect on the one hand depends on the original distribution of the income component which we use as the basis for their imputation, and on the other hand on its contribution to national income. The highest unequal effect has the primary income of financial and non-financial corporations, due to the fact that they are distributed similarly to property income. Since their relative size in national income is comparable, these two variables have a similar impact on the distribution. Social transfers in kind and collective consumption in contrast have the highest redistributive effects. Both amount to around 10 to 20% of national income and are distributed equally across the population.

The contributions of the different methodological steps to the income distribution are summarized in Figures 7 and 8. Naturally, the scaling effects are particularly high for the top 5% of the income distribution,

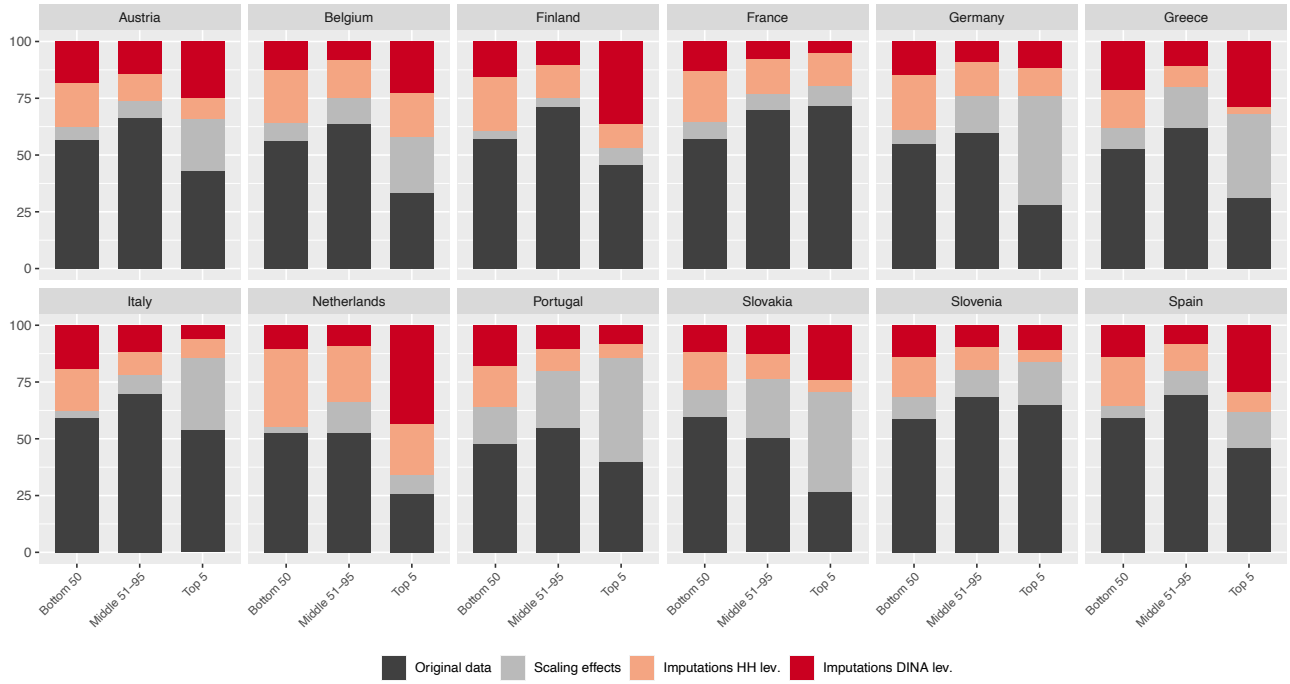
[15] See Table A.2 in Appendix A for the contributions of all variables to national income.

because property income and mixed income, which are concentrated at the top, are usually not covered well by surveys. In some countries, the income of the top 5% is more than doubled by re-aligning micro data to national accounts. In Finland, France, the Netherlands, Slovenia and Spain, where the coverage rates of both capital income and mixed income are rather high, in particular of the SILC, the effect of rescaling is in contrast rather low. For the HFCS, which generally covers these income components less well, the scaling effect on the income of the top 5% is high even for those countries.

The imputations at the level of the household sector increase the income of the bottom 50% relatively more than those of the other income groups. These variables include social transfers in kind, which we distribute equally, so that low income groups generally benefit more from them. In Belgium and the Netherlands however the imputations at the level of the household sector also increase the income of the top 5% substantially. This is due to the inclusion of other capital income, which primarily consists of private pensions and are concentrated at the top of the distribution.

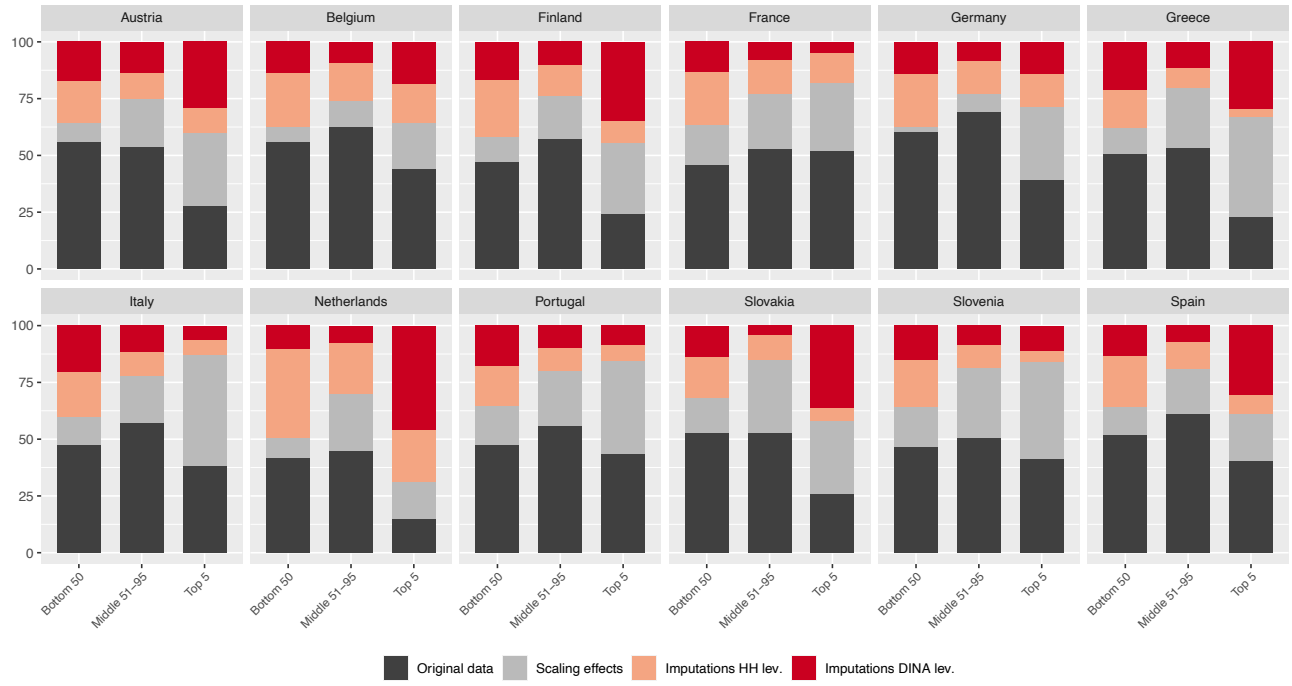
The imputations at the level of the total economy generally benefit the bottom and the top of the distribution most. On the one hand, these variables include collective consumption, which has similar effects as social transfers in kind. On the other, these imputations also include the primary income of financial and non-financial corporations, which we distribute similarly to capital income, and which thus mostly accrues to the top 5%. The effect of the latter is particularly high where the primary income of corporations contributes a substantial amount to national income, such as in the Netherlands.

Figure 7 **Effects of scaling and imputations (SILC, Wave 2014, Equal split)**



Notes: The income of each group is normalized to 100 %. The variables which are rescaled and imputed on the household and national level are summarized in Table 1.
Source: SNA, SILC, own calculations and illustration.

Figure 8 **Effects of scaling and imputations (HFCS, Wave 2014, Equal split)**



Notes: The income of each group is normalized to 100 %. The variables which are rescaled and imputed on the household and national level are summarized in Table 1.
Source: SNA, HFCS, own calculations and illustration.

4 Conclusion

The aim of this paper is to construct Distributional National Accounts (DINA) from household survey data (SILC, HFCS). We build on the insights of both the DINA and the EG-DNA literature to develop a transparent and easily reproducible methodology, which can be used to construct DINA whenever administrative tax data are not available, and apply this methodology to various European countries. By doing so, we build synthetic micro datasets which cover the entire distribution, include all income components separately, and are consistent with national accounts. These datasets are highly comparable across countries.

We find that realigning the primary income of the household sector to national accounts totals and imputing the primary income of the other institutional sectors increase inequality significantly in most countries. Our findings thus generally confirm the need to construct DINA to get a more adequate picture of the income distribution. The distributional effects of these methodological steps however vary between countries and data sources. These differences can primarily be attributed to variations in the coverage rate of mixed and capital income as well as in their distributions in the original micro dataset. Differences in the importance of the variables for national income in national accounts also play a non-negligible role. Furthermore, we find that income inequality declines substantially when we move from pre-tax factor income to post-tax national income due to the redistributive effects of the tax system, the welfare state, and collective consumption, with the first one usually not being very progressive in most countries. Our

findings are similar for the two periods for which we build DINA. Not surprisingly, using the individual split instead of an equal split in general increases inequality. This is all very well in line with the DINA literature.

Comparing the DINA approach to that of the EG-DNA, we find that post-tax disposable income and disposable income of the household sector show considerable differences in distribution in most countries. Including the primary income of the corporate sector and the state usually increases inequality significantly. Interestingly, the distribution of adjusted disposable income is similar to that of post-tax national income in many countries. Apparently, the distributive effects of including the primary income of the other sectors and collective consumption partly neutralize each other.

In general, the impact of a variable in the income distribution depends on its own distribution in the original survey data, the coverage rate, and the contribution of the variable to national income. Mixed income and capital income as well as the primary income of the corporate sector usually have the largest unequal effect. The first two are highly unequally distributed already in the micro data and generally have a low coverage rate. The last one is distributed similarly to capital income and amounts considerably to national income in most countries.

Social transfers in kind and collective consumption in contrast have the highest redistributive effects. Both contribute a significant amount to national income and are distributed equally by means of a per-capita allocation across the population. Wages, pensions, unemployment benefits and other monetary transfers usually reduce inequality, albeit in some cases only slightly. In some countries, wages even increase the Gini coefficient for post-tax national income because they are concentrated in the middle of the income distribution. Our methodological steps benefit income groups differently: The top 5% of the income distribution profit most from the rescaling effect. The imputation of variables, which include the primary income of the corporate sector as well as social transfers in kind and collective consumption, usually increases the incomes at the top and the bottom most.

All the three above-mentioned factors which affect inequality vary considerably across countries, data sources and time, which leads to variations in the income distributions. Our methodology (as well as other approaches in the DINA literature which use survey data) must therefore be applied with some caution, in particular when constructing time series of distributional indicators such as the Gini coefficient or the Top 5% share. The patterns in inequality which we find in the different income definitions however are highly similar for the two periods for which we construct DINA and in general resemble each other for the different data sources, which hints at a certain consistency in the data and confirm the robustness of our methodology. Even though there could be variations in the *levels* of inequality over time due to variations in the data sources, our methodology is generally well suited to construct synthetic micro datasets for policy analysis, given the fact that the impact of, say, a tax reform only needs DINA for a specific year. Our datasets are thus a major improvement over traditional distributional analyses, which ususally use the original survey data.

Comparing our results to those in the literature however reveals that our methodolgy is still likely to underestimate the income inequality due to the absence of tax data. In the case of France, where the coverage of mixed and capital income in the surveys is nearly 100%, our results for the Gini coefficient and

the top 5% income share are significantly lower than those of Garbinti et al. (2018).This shows that surveys, even if they cover these forms of income well, still underrepresent the income concentration at the top of the distribution, given that they only include a sample of the population. For many countries however, increasing the quality of data on mixed and capital income by matching surveys with administrative data, would most probably improve our results. Furthermore, note that it is generally not straightforward to compare our results to those based on more aggregated approaches, given that the latter usually realign *total* income to national accounts.[16] All the above-mentioned three factors which affect inequality in our methodology can contribute considerably to differences in the results, given that we realign and impute each variable separately to national account totals.[17] Nevertheless, the decline in inequality from pre-tax to post-tax national income in our approach is similar to that in the literature for most countries.

[16]The results in the WID database for most European countries are based on Blanchet et al. (2019), who use such an aggregated approach. See section 1.

[17]To some extent, the deviation of our results from WID data for post-tax national income is due to the fact that we choose to distribute collective consumption equally between individuals. See section 2.

References

Alvaredo, F., Atkinson, A. B., Chancel, L., Piketty, T., Saez, E., & Zucman, G. (2016). *Distributional National Accounts (DINA) Guidelines: Concepts and Methods used in WID.world* (Working Paper). World Income and Wealth Database.

Alvaredo, F., Chancel, L., Piketty, T., Saez, E., & Zucman, G. (2017). Global Inequality Dynamics: New Findings from WID.world. *American Economic Review*, *107*(5), 404–409. doi: 10.1257/aer.p20171095

Atkinson, A. B. (1971). The distribution of wealth and the individual life-cycle. *Oxford Economic Papers*, *23*(2), 239–254. doi: 10.2307/2662236

Atkinson, A. B. (1997). Bringing Income Distribution in From the Cold. *The Economic Journal*, *107*(441), 297–321. doi: 10.1111/j.0013-0133.1997.159.x

Balcazar, C. F., Ceriani, L., Olivieri, S., & Ranzani, M. (2014). *Rent imputation for welfare measurement: A review of methodologies and empirical findings.* The World Bank.

Blanchet, T., Chancel, L., & Gethin, A. (2019). How Unequal Is Europe? Evidence from Distributional National Accounts, 1980-2017. *WID.world Working Paper*(2019/06).

European Commission. (2016). *Data on Taxation.* Retrieved 2020-02-21, from https://ec.europa.eu/taxation_customs/business/economic-analysis-taxation/data-taxation_en

Eurostat. (2013). *European system of accounts.* Retrieved 2020-02-21, from https://ec.europa.eu/eurostat/web/esa-2010/overview

Eurostat. (2019). *EU-SILC Operational Guidelines.* Retrieved 2020-02-21, from https://circabc.europa.eu/faces/jsp/extension/wai/navigation/container.jsp

Fesseau, M., & Mattonetti, M. L. (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 Statistics Working Papers No. 2013/04). Paris: Organization for Economic Co-operation and Development.

Fesseau, M., Wolff, F., & Mattonetti, M. L. (2013). *A Cross-Country Comparison of Household Income, Consumption and Wealth Between Micro Sources and National Accounts Aggregates* (OECD Statistics Working Papers No. 2013/03). Paris: Organization for Economic Co-operation and Development.

Fessler, P., Rehm, M., & Tockner, L. (2016). The impact of housing non-cash income on the household income distribution in Austria. *Urban Studies*, *53*(13), 2849–2866.

Garbinti, B., Goupille-Lebret, J., & Piketty, T. (2018). Income inequality in France, 1900–2014: Evidence from Distributional National Accounts (DINA). *Journal of Public Economics*, *162*, 63–77. doi: 10.1016/j.jpubeco.2018.01.012

Household Finance and Consumption Network. (2016). The household finance and consumption survey: methodological report for the second wave. *ECB Statistics Paper Series*, *17*.

Kuznets, S. (1955). Economic Growth and Income Inequality. *American Economic Review*, *45*(1), 1–28. doi: 10.2307/1811581

Piketty, T. (2003). Income inequality in France, 1901–1998. *Journal of Political Economy*, *111*(5), 1004-1042. doi: 10.1086/376955

Piketty, T. (2015). Putting Distribution Back at the Center of Economics: Reflections on *Capital in the Twenty-First Century*. *Journal of Economic Perspectives*, *29*(1), 67–88. doi: 10.1257/jep.29.1.67

Piketty, T., & Saez, E. (2003). Income Inequality in the United States, 19131998*. *The Quarterly Journal of Economics*, *118*(1), 1-41. doi: 10.1162/00335530360535135

Piketty, T., Saez, E., & Zucman, G. (2018). Distributional National Accounts: Methods and Estimates for the United States. *The Quarterly Journal of Economics*, *133*(2), 553–609. doi: 10.1093/qje/qjx043

Piketty, T., Saez, E., & Zucman, G. (2019). Simplified Distributional National Accounts. *AEA Papers and Proceedings*, *109*, 289–295. doi: 10.1257/pandp.20191035

Seneviratne, A. (2016). *Australian National Accounts: Distribution of Household Income, Consumption and Wealth* (Paper presented at the DGINS 2016 Conference).

Stiglitz, J., Sen, A., & Fitoussi, J.-P. (2009). *Report by the Commission on the Measurement of Economic Performance and Social Progress* (Tech. Rep.).

Zwijnenburg, J. (2019). Unequal Distributions: EG DNA versus DINA Approach. *AEA Papers and Proceedings*, *109*, 296–301. doi: 10.1257/pandp.20191036

Zwijnenburg, J., Bournot, S., & Giovannelli, F. (2017). OECD Expert Group on Disparities in a National Accounts Frame-work Results from the 2015 Exercise. *OECD Statistics Working Paper*(76).

Appendices

A Additional Tables

Table A.1 **HFCS reference periods**

	Wave 1: Income	Wave 2: Income
AT	2009	2013
BE	2009	2013
DE	2009	2013
EL	2009	2014
ES	2007	2010
FI	2009	2013
FR	2009	2014
IT	2010	2014
NL	2009	2013
PT	2009	2012
SI	2009	2013
SK	2010	2013

Source: HFCS.

Table A.2 **National Income**

	AT	BE	DE	EL	ES	FI	FR	IT	NL	PT	SI	SK
Incomes of Household Sector												
Gross wages	47.8	46.8	48.9	29.9	48.5	49.6	46.2	35.4	45.2	43.9	55.4	38.4
Employer social contributions	10.1	17.7	11.0	9.7	12.7	11.6	17.2	13.0	12.7	12.3	9.3	11.1
Gross operating surplus and rents	6.6	7.0	5.0	11.4	7.8	8.7	9.9	12.1	0.5	8.9	8.0	5.4
Gross mixed income	9.1	7.9	10.3	26.1	10.9	6.5	6.5	17.2	9.8	14.6	11.8	24.8
Consumption of fixed capital	−4.5	−5.6	−6.3	−7.3	−5.2	−5.3	−4.5	−6.4	−4.5	−7.3	−6.9	−4.4
Interest with FISIM and distributed income	7.9	6.3	13.3	5.6	3.8	4.3	4.0	11.8	2.8	9.4	2.1	0.7
Property income	1.7	3.5	3.0	0.1	1.7	1.2	2.9	1.8	6.1	1.8	0.6	0.7
Interest paid	−0.6	−0.8	−1.5	−1.1	−1.9	−0.5	−0.8	−0.7	−1.3	−3.9	−1.1	−0.5
Other property income paid	0.0	0.1	0.1	0.3	0.1	0.2	0.1	0.2	0.0	0.0	0.0	0.1
Primary income (S1415)	78.1	82.8	83.9	74.7	78.5	76.2	81.6	84.3	71.3	79.7	79.5	76.4
Pre-tax Factor Income												
Primary income (S1112)	7.6	7.9	5.9	12.1	11.7	5.0	3.6	4.0	15.3	8.2	4.0	12.8
Net operating surplus + mixed income	0.1	0.0	−0.1	0.0	0.0	0.3	0.0	0.0	0.0	0.2	0.1	0.0
Net capital income	1.8	2.5	1.4	4.2	1.2	−2.4	1.8	4.7	−1.9	4.7	1.5	1.2
Net indirect taxes	16.0	11.7	11.7	17.5	11.1	16.1	16.1	16.4	11.4	16.7	18.0	12.1
PRTFI	100.0	100.0	100.0	100.0	100.0	100.0	99.5	100.0	100.0	100.0	100.0	100.0
Pre-tax National Income												
Social contributions	−19.8	−25.3	−24.3	−17.2	−17.6	−16.4	−25.5	−18.4	−31.1	−17.5	−21.2	−19.0
Monetary transfers	24.0	24.6	20.6	24.1	20.5	24.1	26.0	26.7	22.0	24.8	24.0	18.6
Difference	−4.2	0.7	3.7	−6.9	−2.8	−7.7	−0.5	−8.3	9.1	−7.4	−2.8	0.5
PRTNI	100.0	100.0	100.0	100.0	100.0	100.0	99.5	100.0	100.0	100.0	100.0	100.0
Post-tax Disposable Income												
Net indirect taxes	−16.0	−11.7	−11.7	−17.5	−11.1	−16.1	−16.1	−16.4	−11.4	−16.7	−18.0	−12.1
Taxes on income and wealth	−12.5	−14.8	−10.3	−6.4	−8.1	−15.3	−8.0	−13.7	−9.2	−5.1	−6.3	−3.8
POTDI	67.7	67.9	74.0	70.9	77.6	64.0	68.9	65.7	76.6	72.0	73.0	79.8
Post-tax National Income												
Social transfers in kind	15.0	19.0	14.6	10.4	14.4	20.3	18.5	13.9	20.5	12.6	15.0	11.4
Collective consumption	9.2	10.8	8.2	13.9	10.6	10.1	10.1	9.7	10.3	10.7	10.2	12.1
Difference	8.1	2.3	3.3	4.8	−2.6	5.6	2.5	10.8	−7.4	4.7	1.8	−3.2
POTNI	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

Table A.3 **Taxes**

	Code	Tax
Austria		
Employment income	D51A_C01	Income tax
	D51A_C03	Tax on industry and trade
	D51A_C04	Contribution to chambers
	D51A_C05	Contribution to chambers
	D51A_C08	Wage tax
Property income (sim.)	D51A_C06	Tax on capital yields
	D51A_C07	Tax on interest
Property income (other)	D51A_C02	EU withholding tax on interest income of non-residents
Belgium		
Employment income	D51A_C02	Business' advance tax payment (PP)
	D51A_C03	Advance payments (PP)
	D51A_C04	Assessments (PP)
	D51A_C06	Special social contributions
	D51A_C07	Contribution large incomes
	D51A_C09	Other taxes on income
Property income (sim.)	D51A_C01	Advance tax payment on movable property (PP)
Property income (other)	D51A_C05	Annual tax on profit sharing
	D51A_C08	Tax on the worker's particip. in the benefit/capital of the comp.
Germany		
Employment income	D51M_C01	Tax on wages
	D51M_C02	Income tax
Property income (sim.)	D51M_C03	Capital gains tax and interest income deduction
	D51M_C04	Trade tax
Greece		
Employment income	D51A_C01	Income taxes on individuals
	D51A_C03	Taxes on Pension (AGAGE)
Property income (sim.)	D51A_C02	Taxes on interest and other taxes on individuals
Italy		
Employment income	D51A_C01	Personal income tax
	D51A_C02	Regional income-tax surcharge
	D51A_C03	Municipal income-tax surcharge
	D51A_C08	Contributions to GESCAL - employees' contribution
Property income (sim.)	D51A_C04	Withholding tax on income from deposits paid by households
	D51A_C10	Withholding tax on company dividends paid by households
Property income (other)	D51A_C05	Local income tax paid by households
	D51A_C06	Substitutive tax on financial instruments and assets
	D51A_C07	10% Surcharge on income
	D51A_C14	Substitute tax on income from rental property
	D51A_C09	Municipal capital gains tax on buildings paid by households
	D51A_C11	Municipal tax on industry, crafts and professions
	D51A_C12	Tax on life insurance and supplementary pension schemes
	D51A_C13	Tax on actuarial reserve
Netherlands		
Employment income	D51A_C01	Wage tax
	D51A_C02	Income tax
Property income (other)	D51A_C03	Dividend tax (households)

Notes: For countries not mentioned in the table, we use the value of the simulated property income taxes from EUROMOD and realign it with the scaling factor of property income. The difference of the sum of property taxes calculated in this manner and D51 in national accounts is then used as the macro aggregate for taxes on employment income.
Source: Own illustration.

B Additional Results for Wave 2014: Individual split

Figure B.1 **Inequality for income concepts: Gini coefficient (SILC & HFCS, Wave 2014, Individual split)**

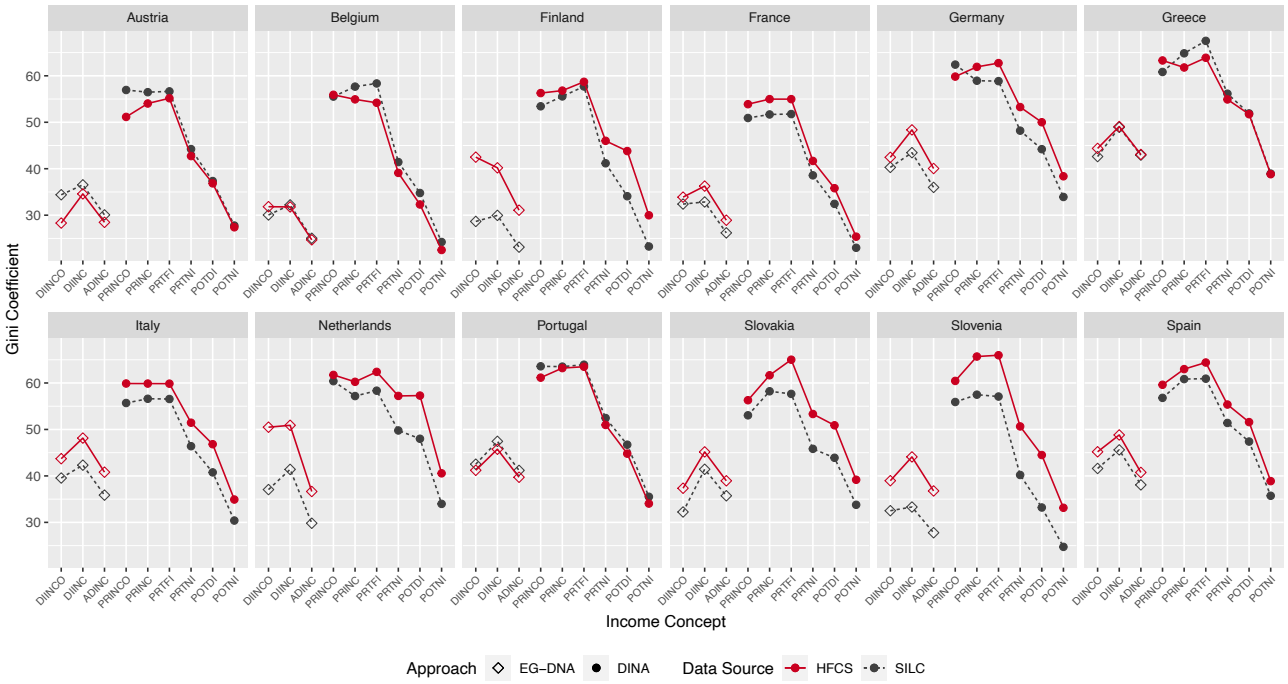


Figure B.2 **Inequality for income concepts: Top 5% share (SILC & HFCS, Wave 2014, Individual split)**

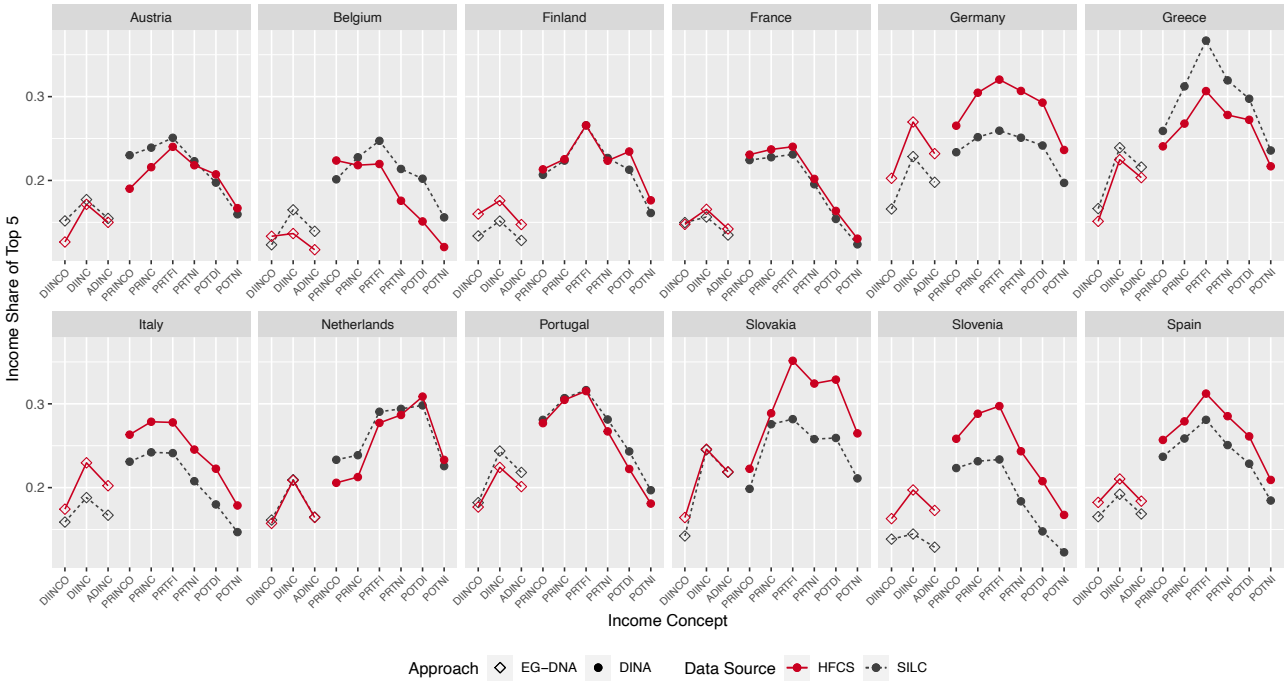


Figure B.3 **Original distributions** (SILC, Wave 2014, Individual split)

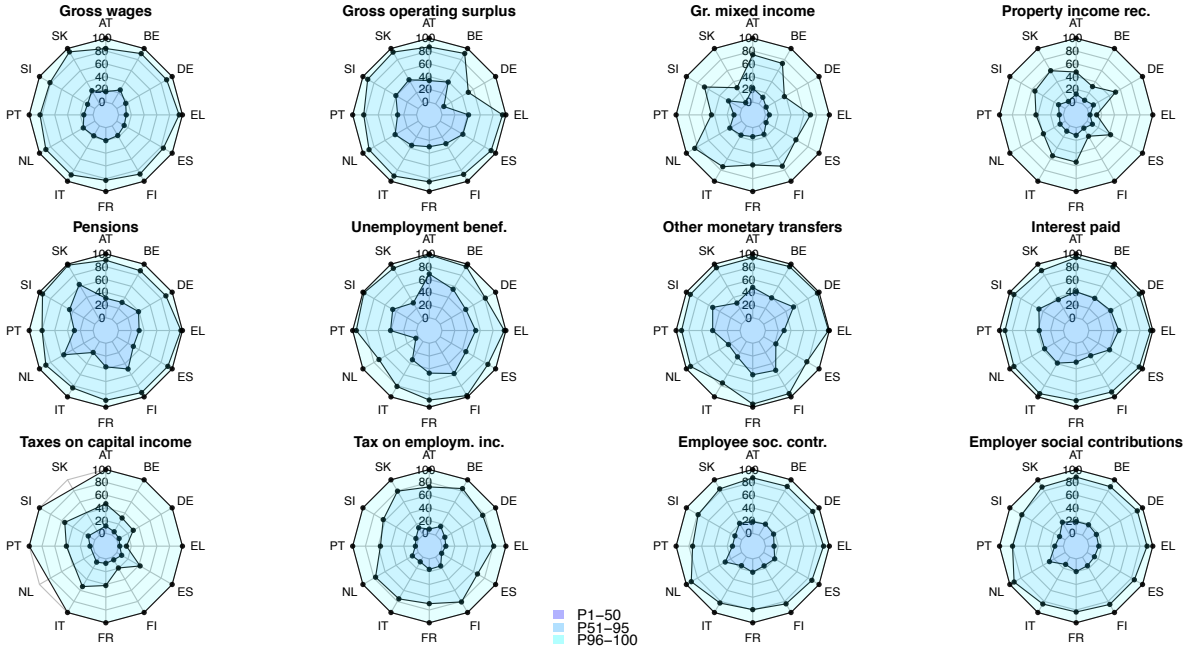


Figure B.4 **Original distributions** (HFCS, Wave 2014, Individual split)

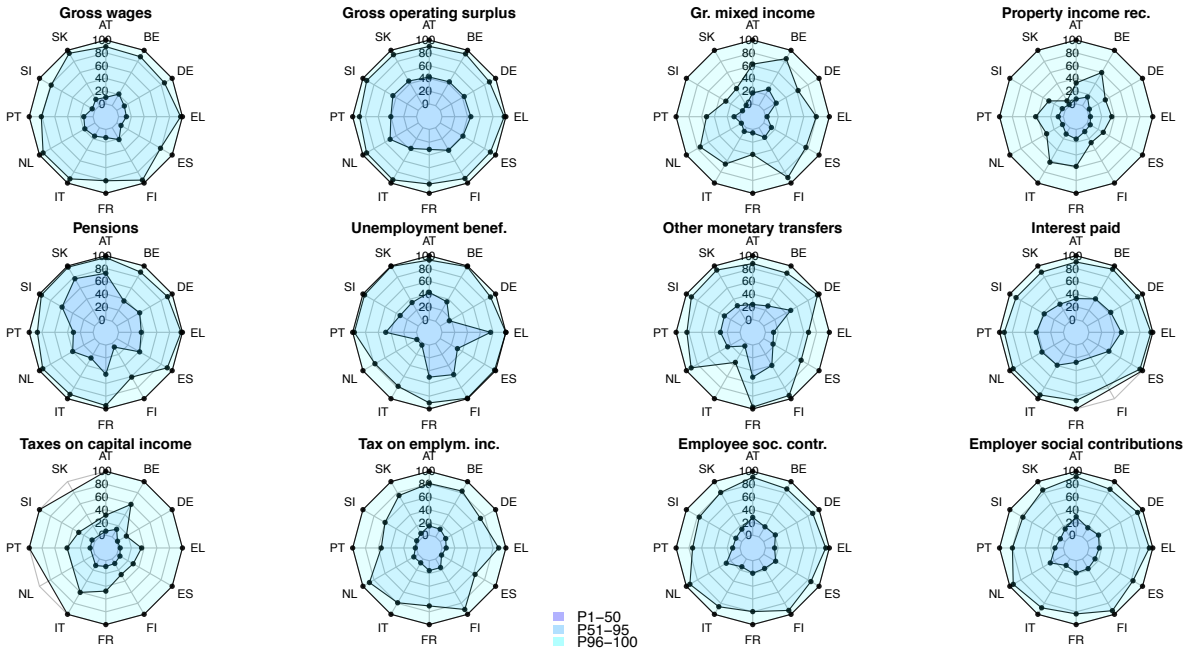


Figure B.5 **Contributions to Inequality** (SILC & HFCS, Wave 2014, Individual split)

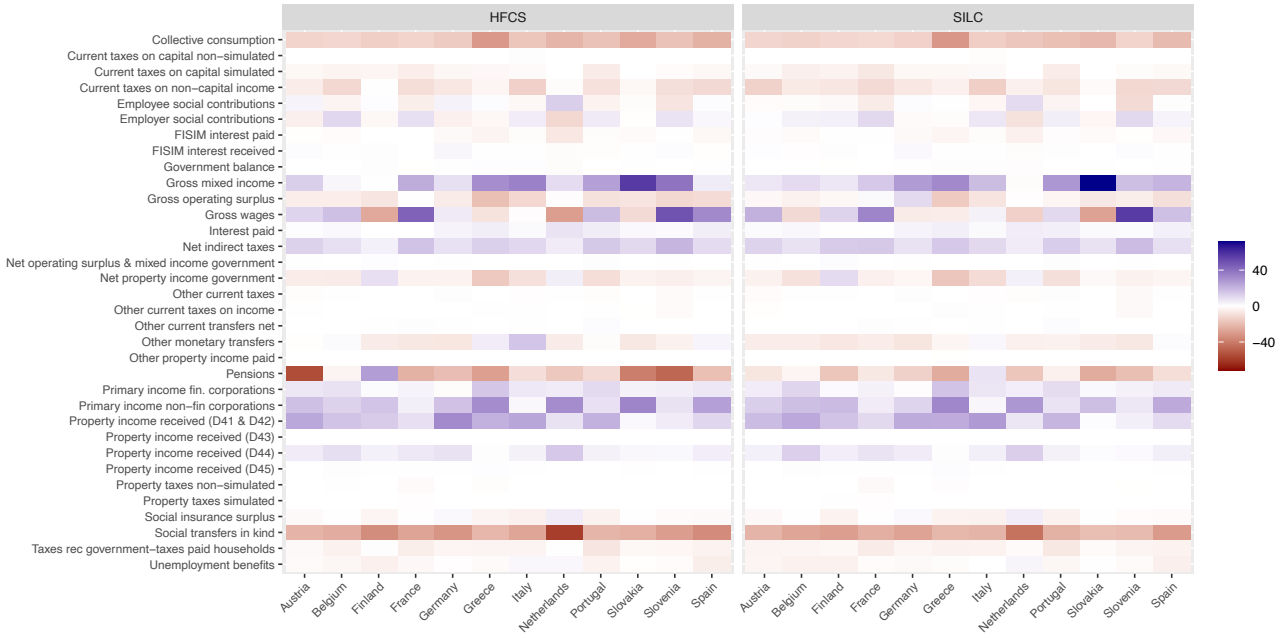


Figure B.6 **Effects of scaling and imputations** (SILC, Wave 2014, Individual split)

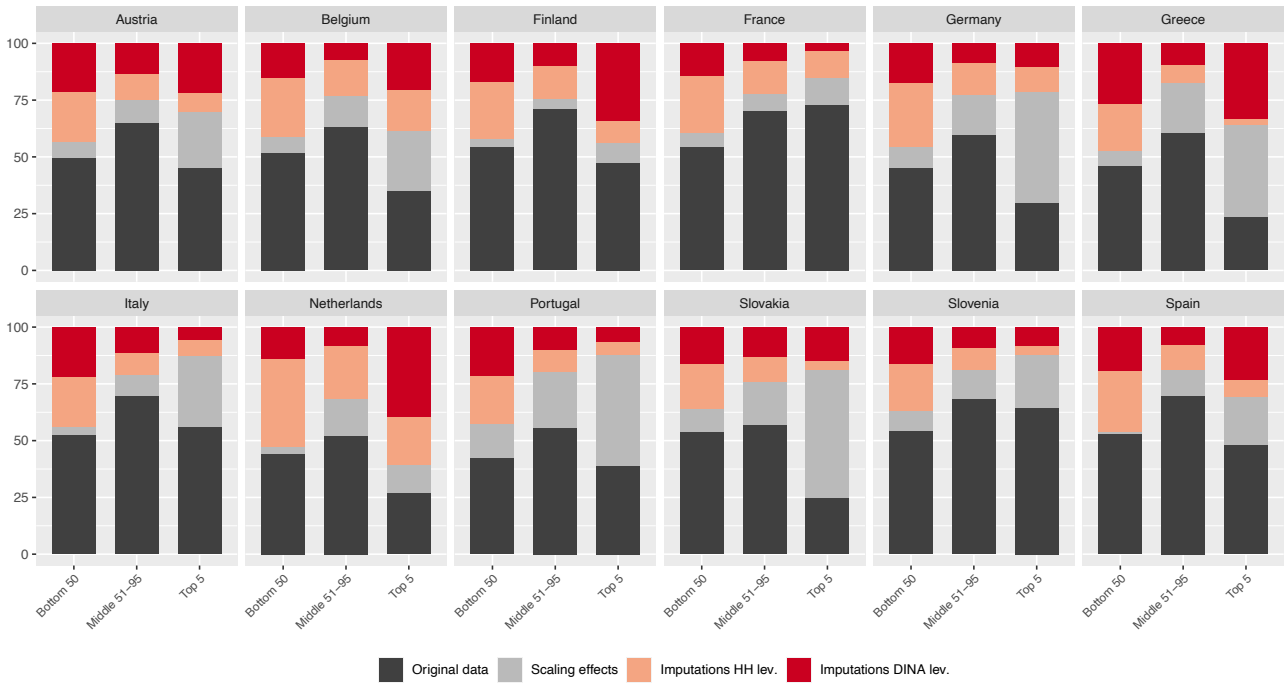
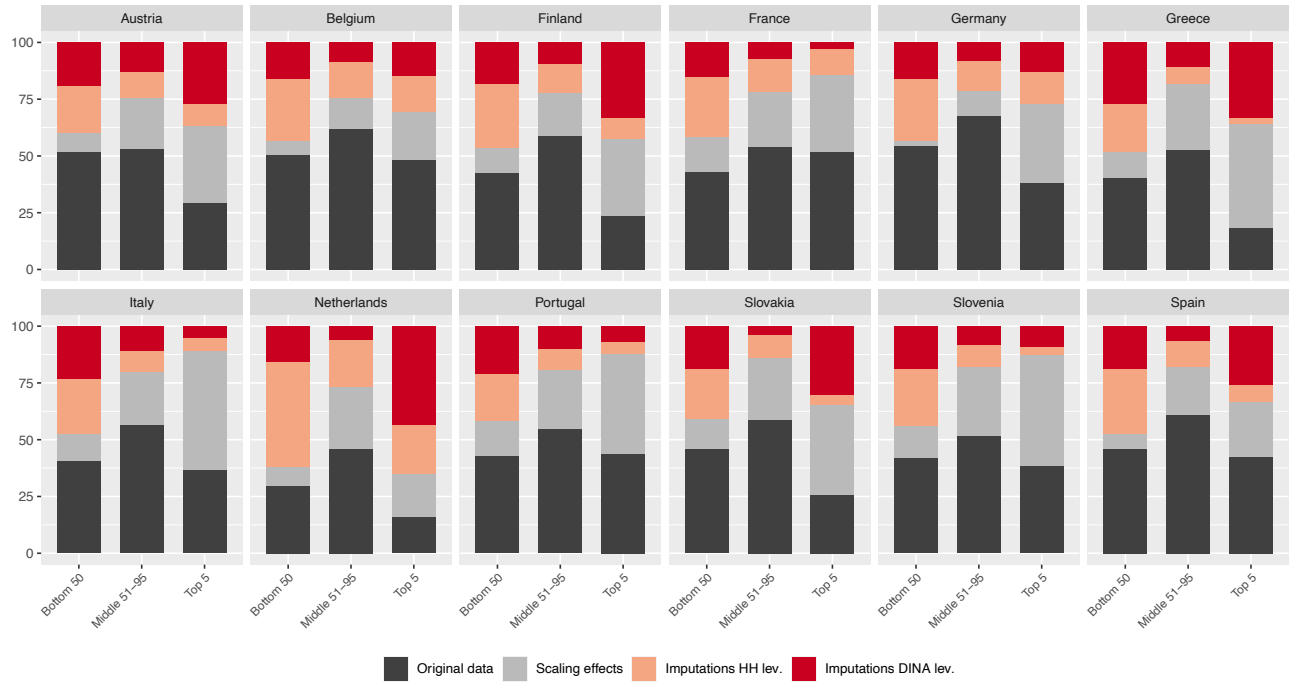


Figure B.7 **Effects of scaling and imputations (HFCS, Wave 2014, Individual split)**



C Additional Results for Wave 2010

C.1 Equal split

Figure C.1 **Inequality for income concepts: Gini coefficient (SILC & HFCS, Wave 2010, Equal split)**

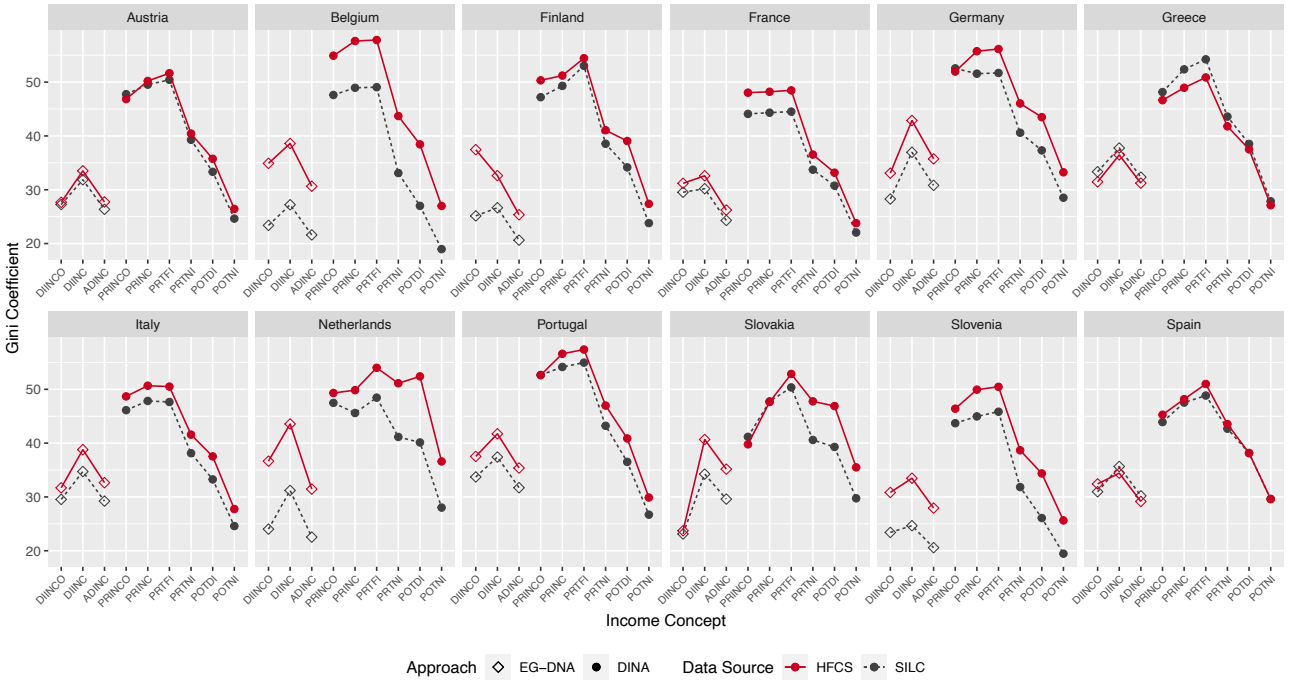


Figure C.2 **Inequality for income concepts: Top 5% share (SILC & HFCS, Wave 2010, Equal split)**

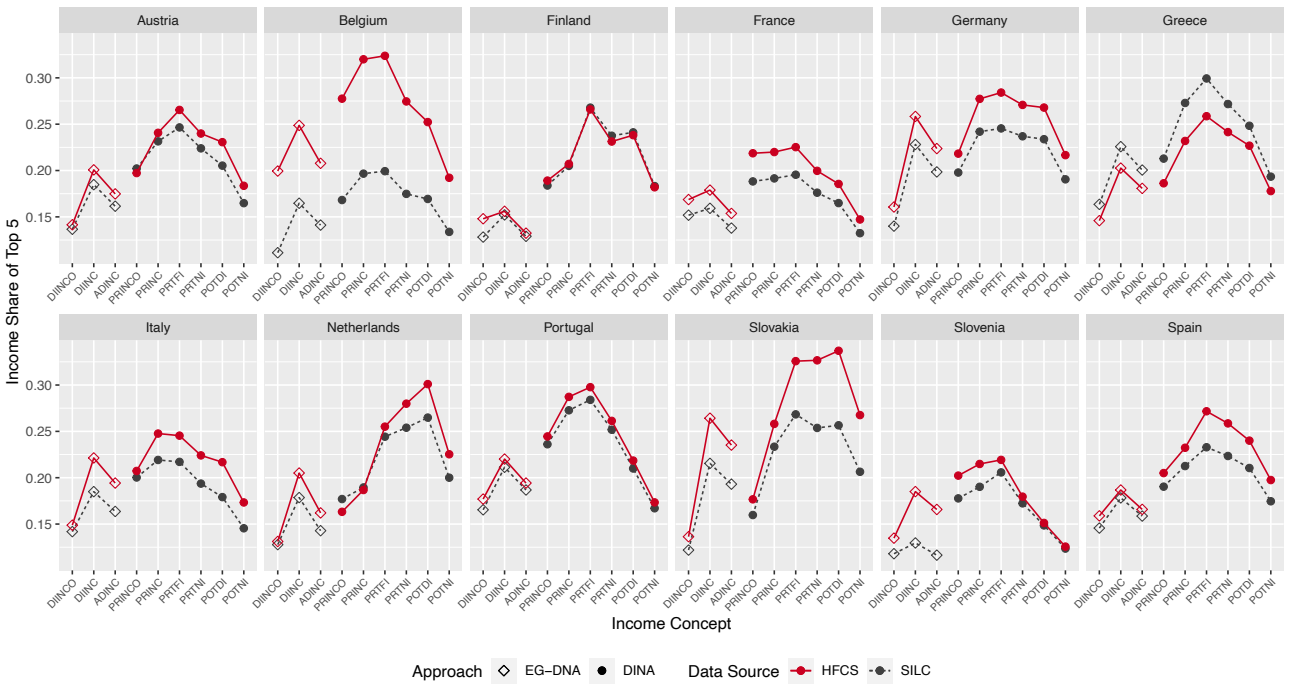


Figure C.3 **Original distributions** (SILC, Wave 2010, Equal split)

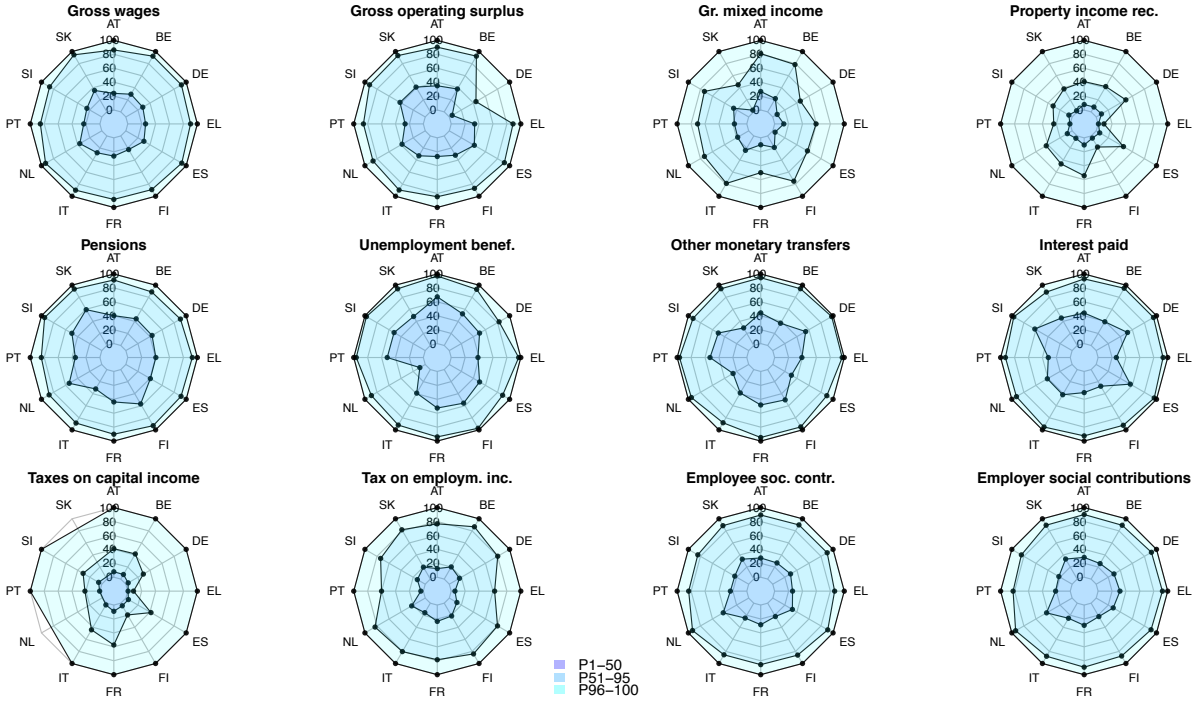


Figure C.4 **Original distributions** (HFCS, Wave 2010, Equal split)

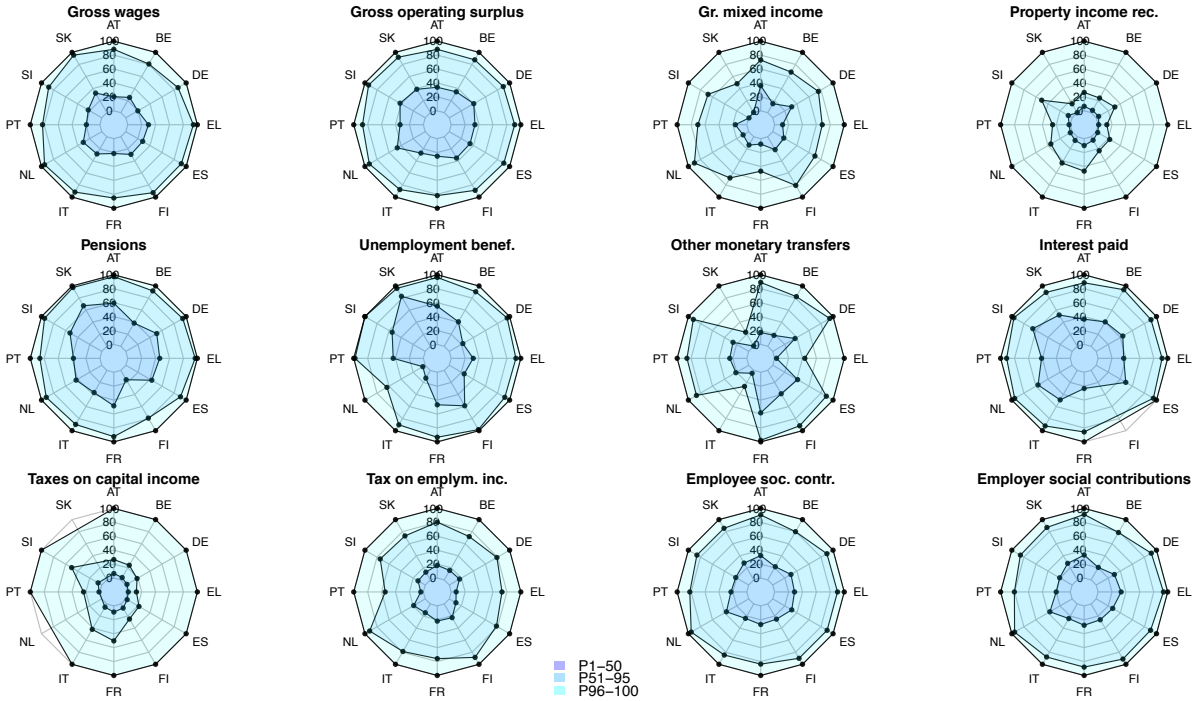


Figure C.5 **Coverage rates (Wave 2010)**

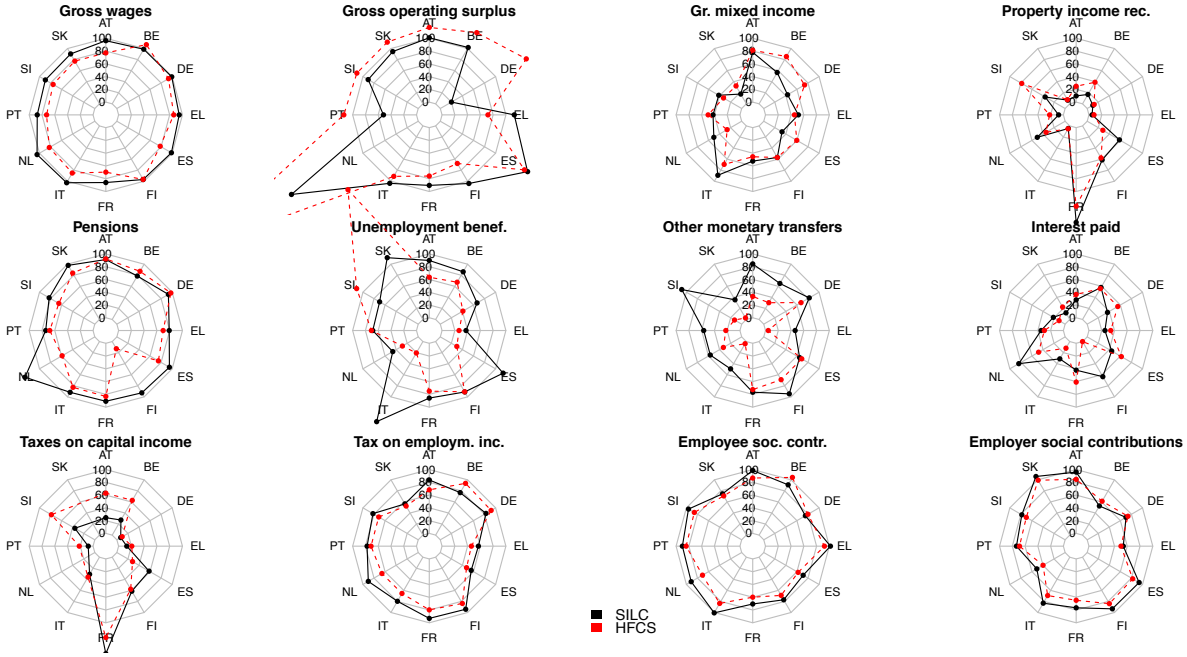


Figure C.6 **Contributions to inequality (SILC & HFCS, Wave 2010, Equal split)**

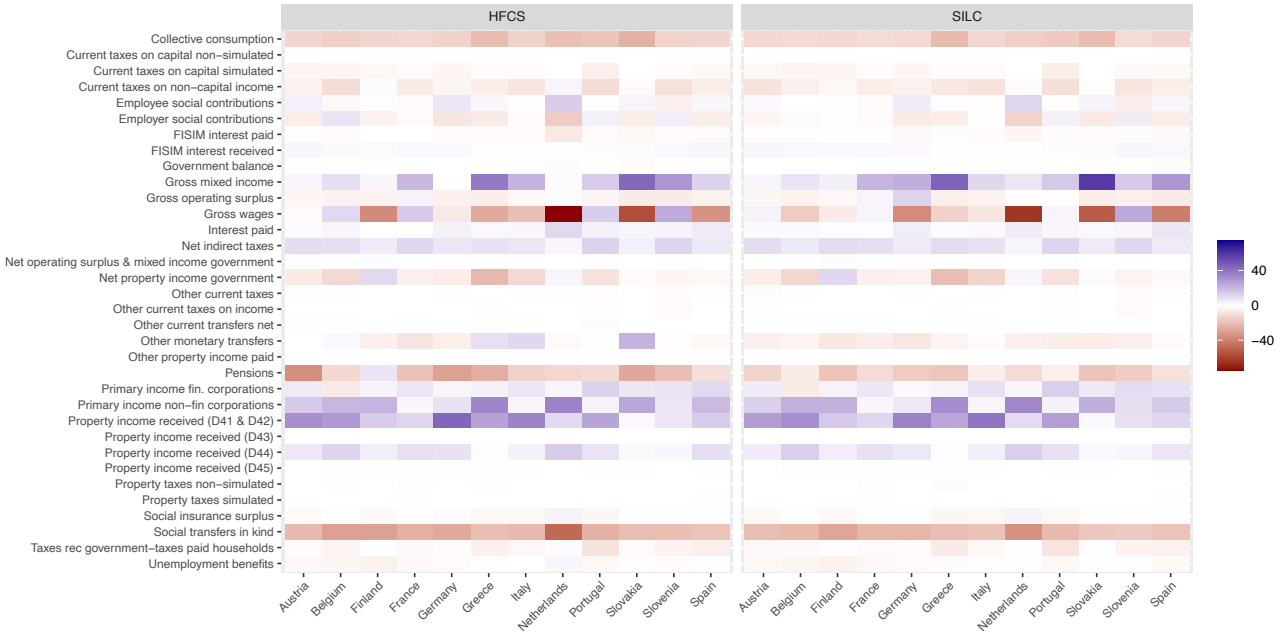


Figure C.7 **Effects of scaling and imputations (HFCS, Wave 2010, Equal split)**

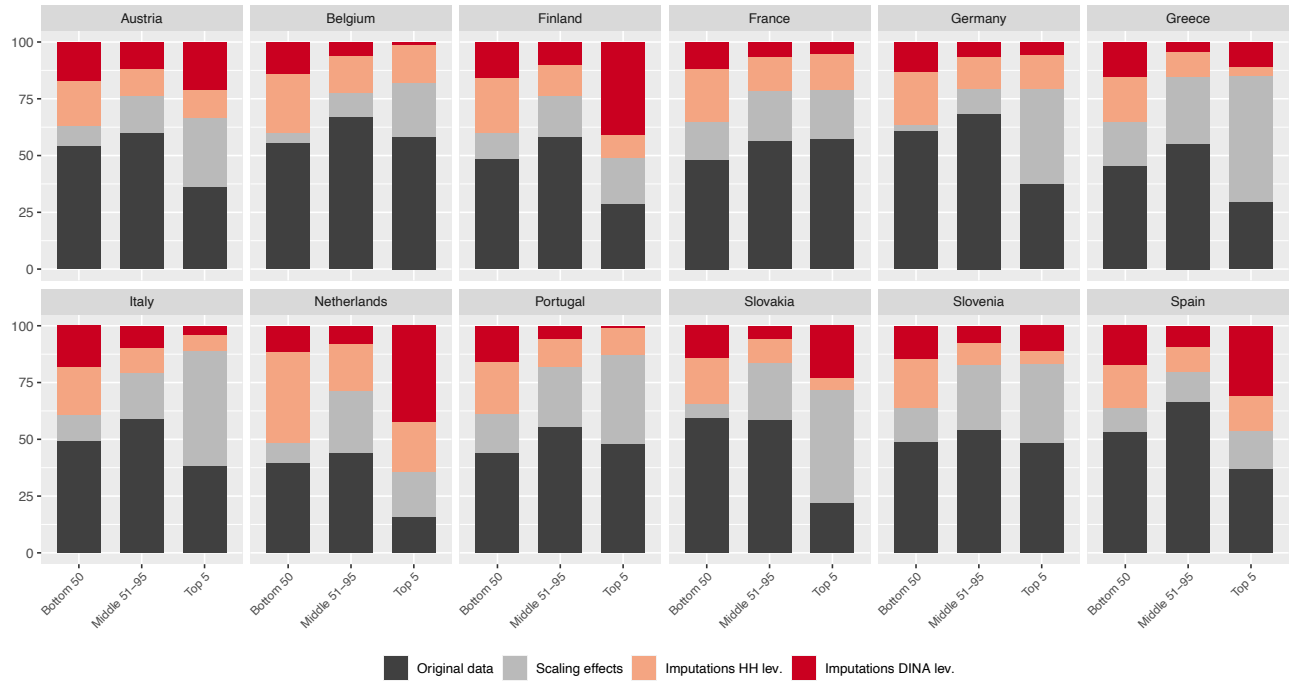
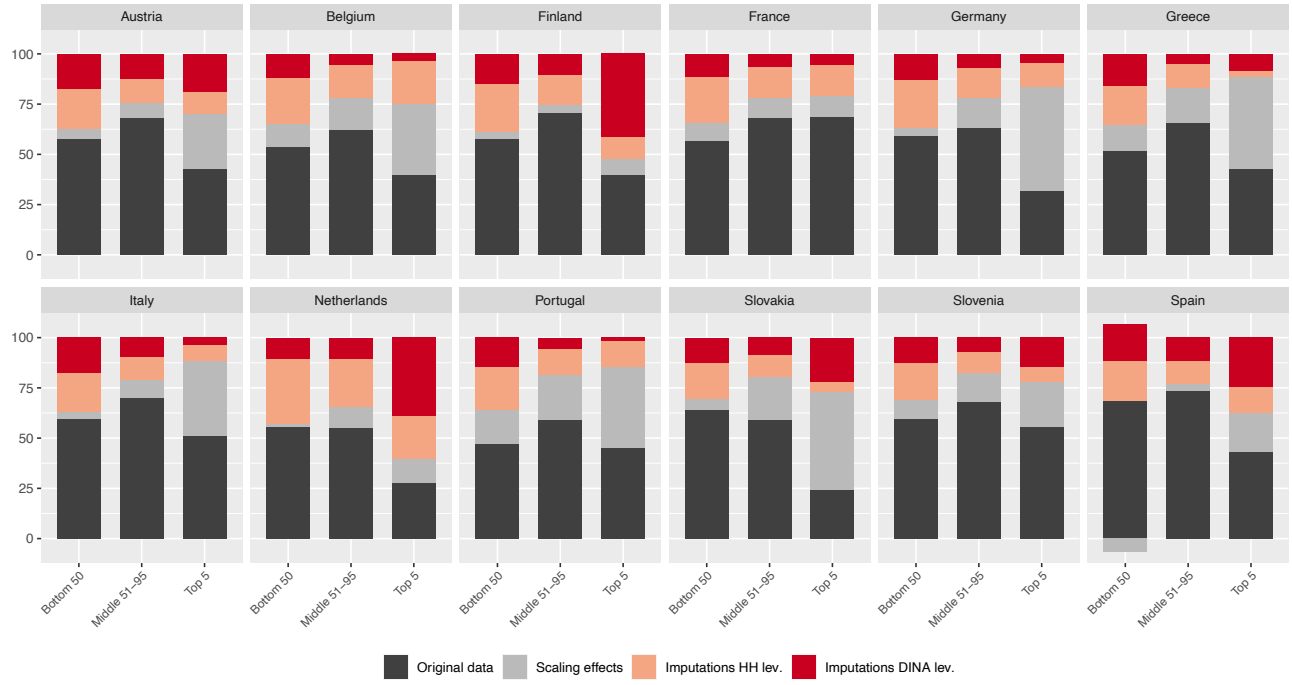


Figure C.8 **Effects of scaling and imputations (SILC, Wave 2010, Equal split)**



C.2 Individual Split

Figure C.9 **Inequality for income concepts: Gini coefficient (SILC & HFCS, Wave 2010, Individual split)**

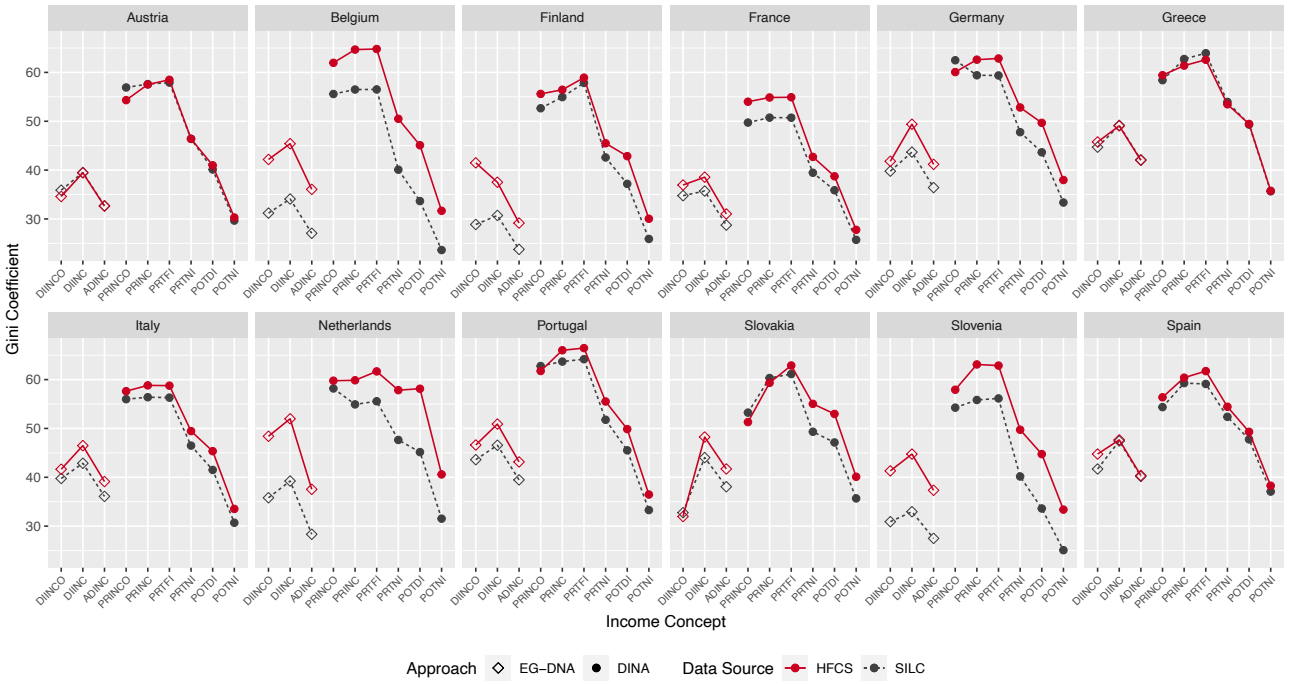


Figure C.10 **Inequality for income concepts: Top 5% share (SILC & HFCS, Wave 2010, Individual split)**

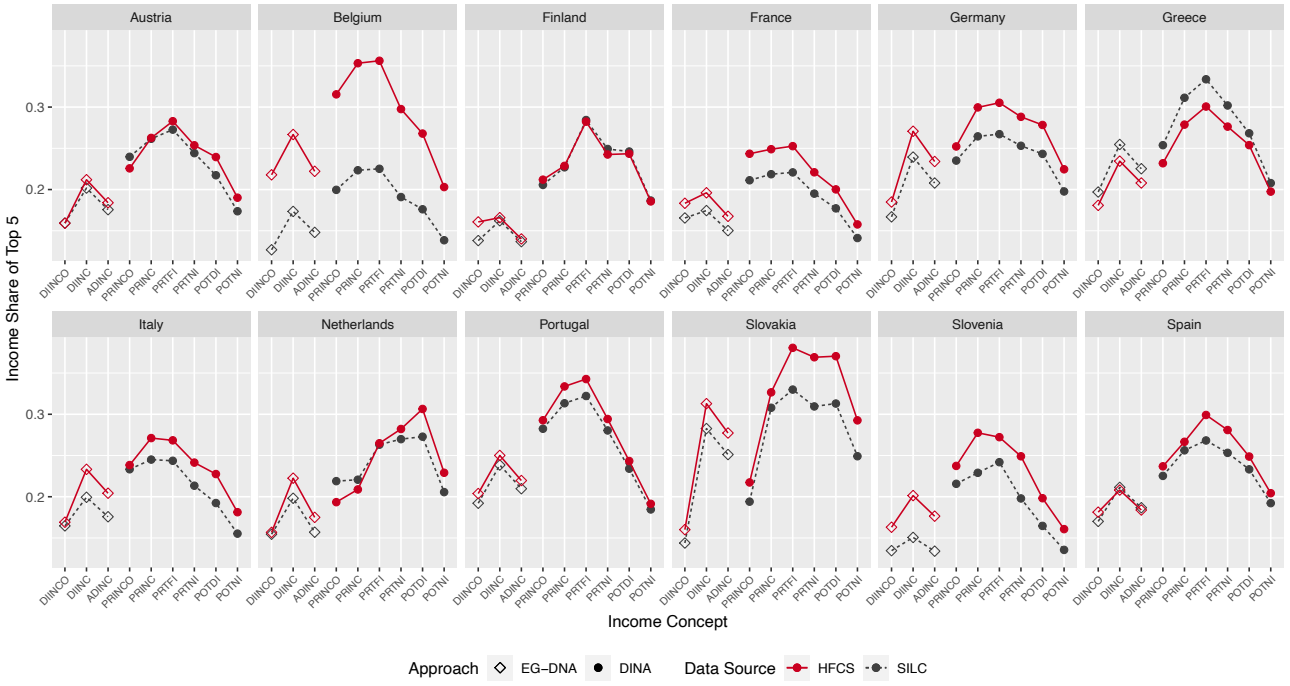


Figure C.11 **Original distributions** (SILC, Wave 2010, Individual split)

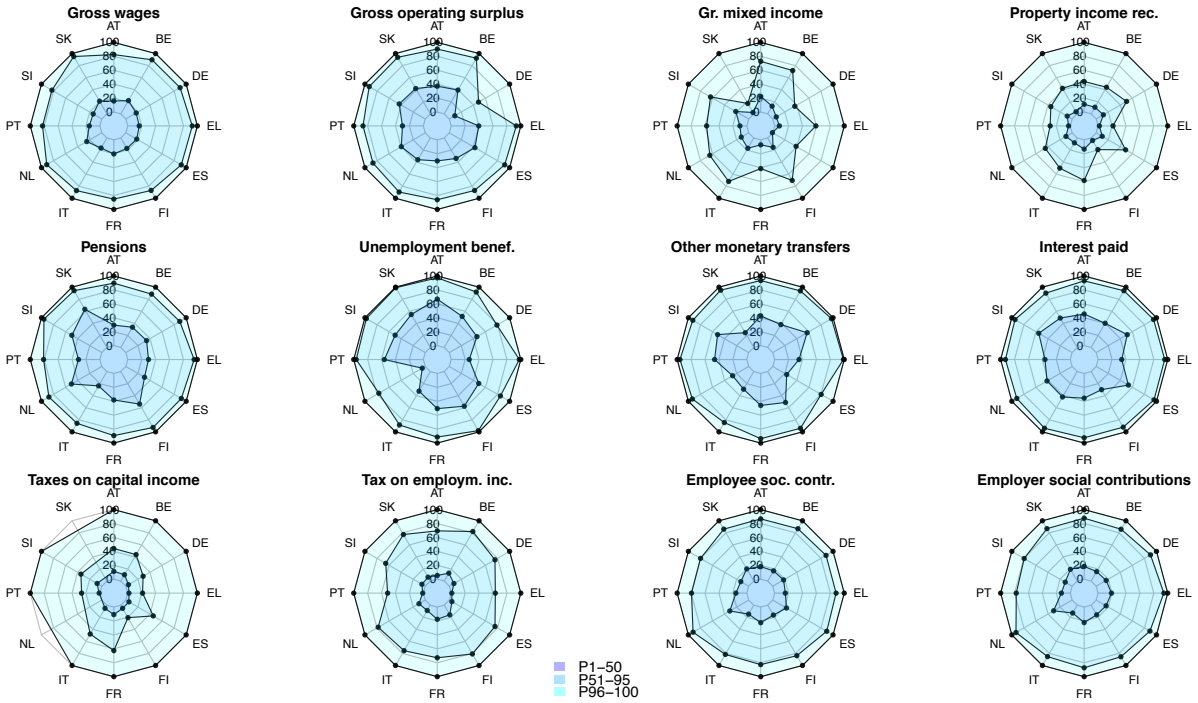


Figure C.12 **Original distributions** (HFCS, Wave 2010, Individual split)

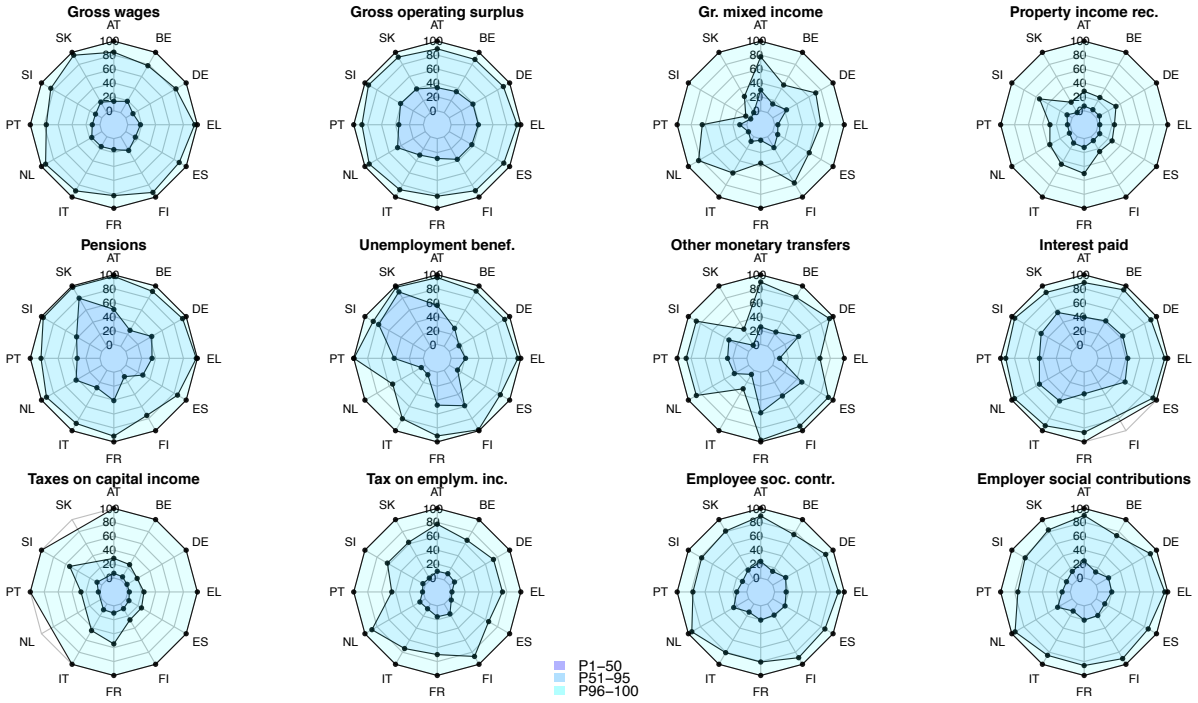


Figure C.13 **Contributions to Inequality (SILC & HFCS, Wave 2010, Individual split)**

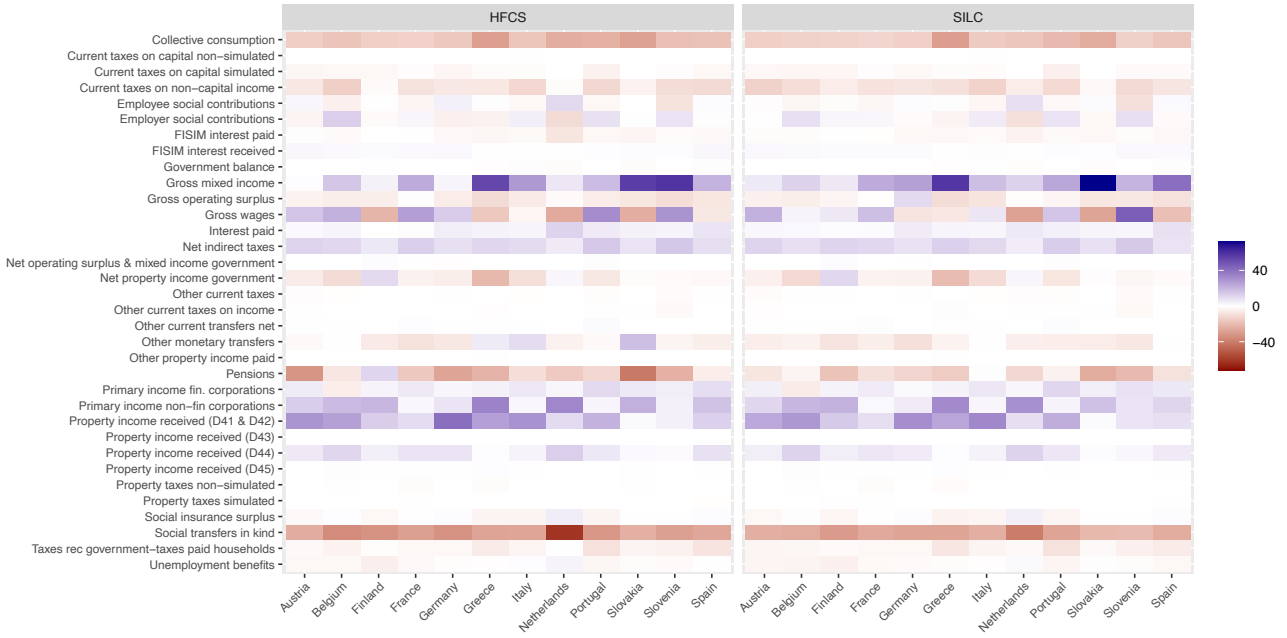


Figure C.14 **Effects of scaling and imputations (SILC, Wave 2010, Individual split)**

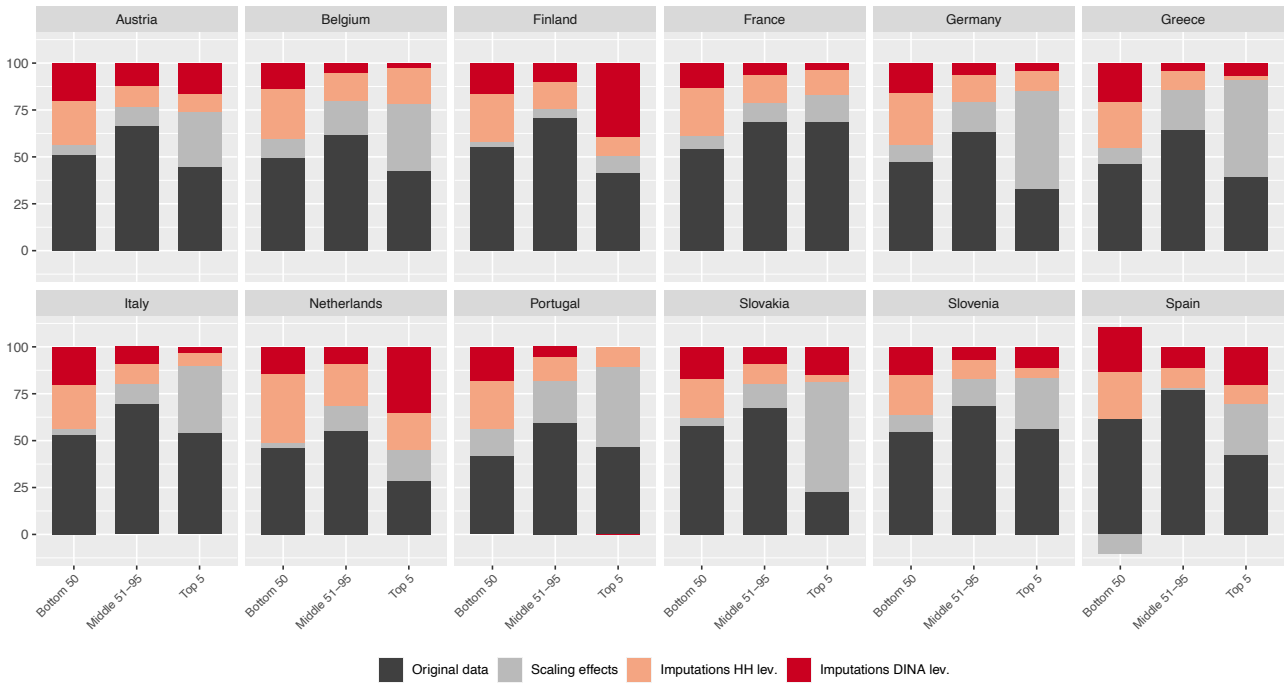
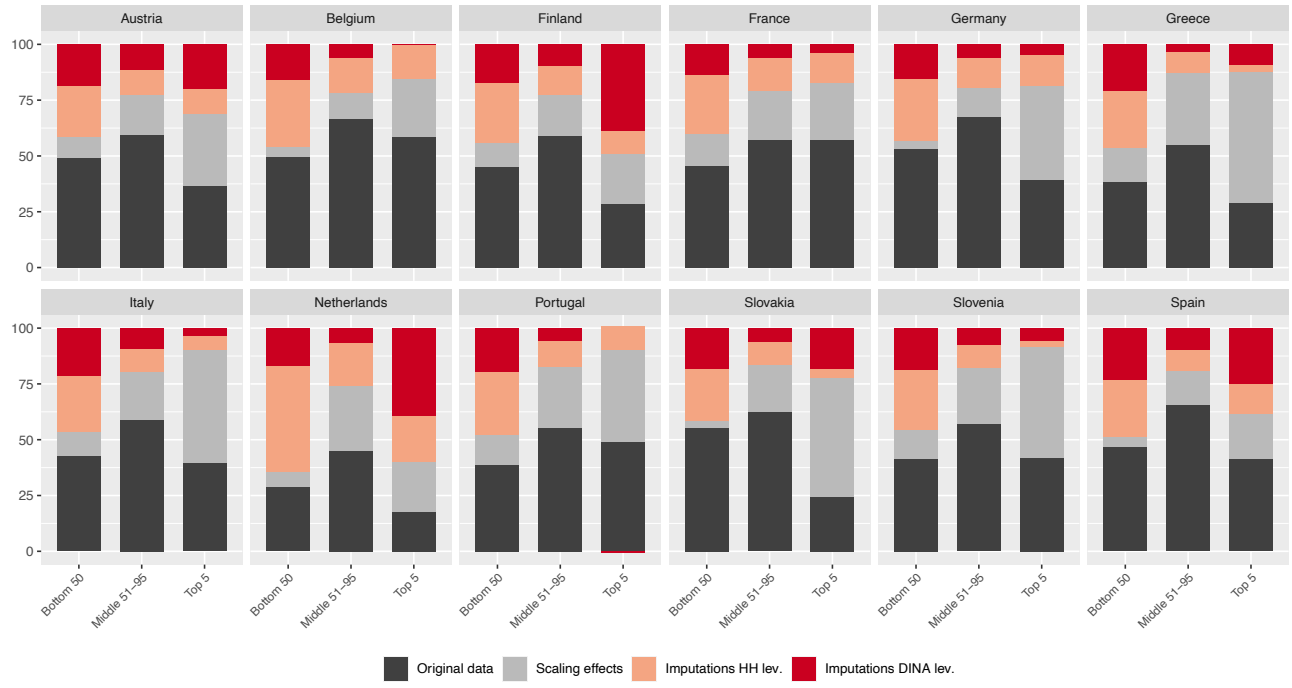


Figure C.15 **Effects of scaling and imputations (HFCS, Wave 2010, Individual split)**



D Components of POTNI by Income Vingtile

Figure D.16 **Components of POTNI by Income Vingtile (HFCS, Wave 2010, Equal Split)**

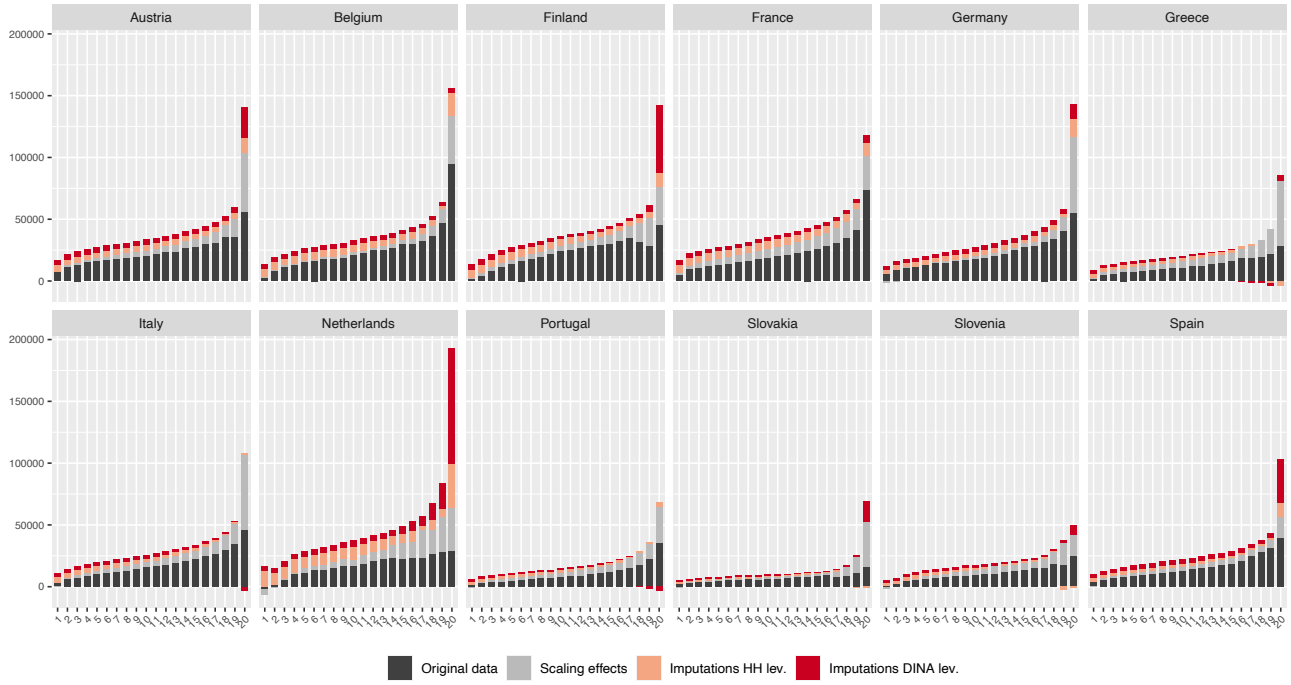


Figure D.17 **Components of POTNI by Income Vingtile (HFCS, Wave 2014, Equal Split)**

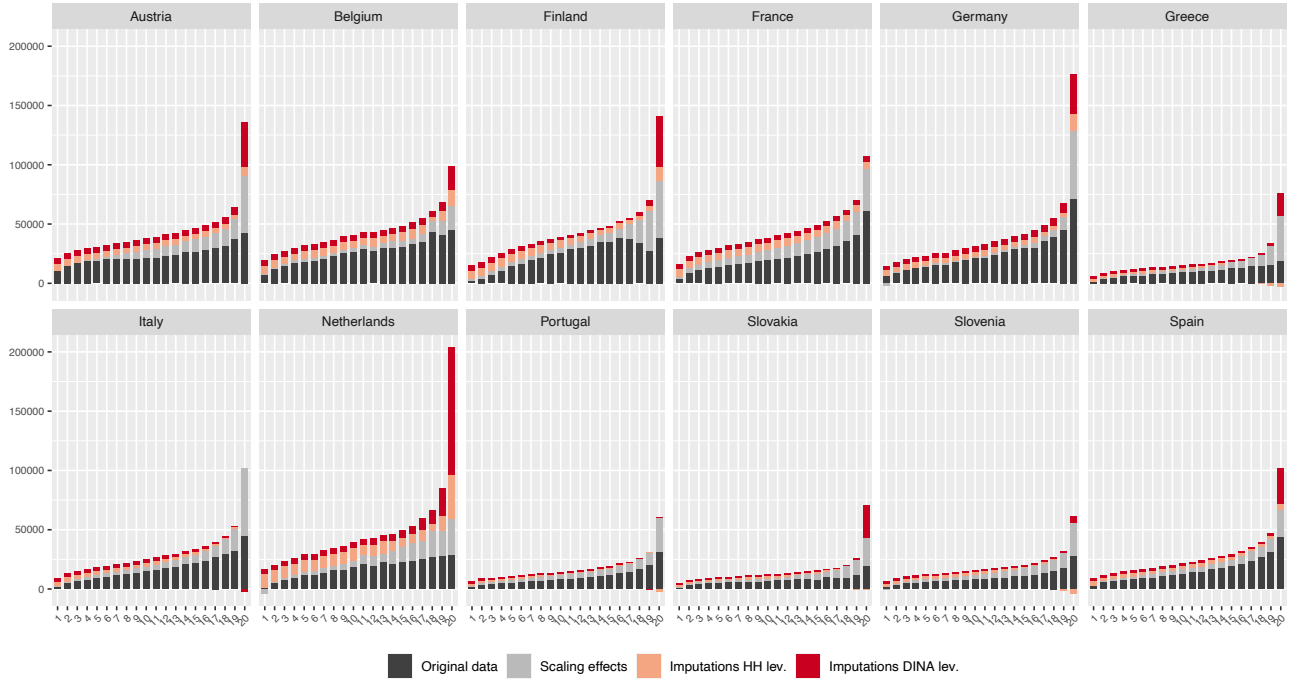


Figure D.18 **Components of POTNI by Income Vingtile (SILC, Wave 2010, Equal Split)**

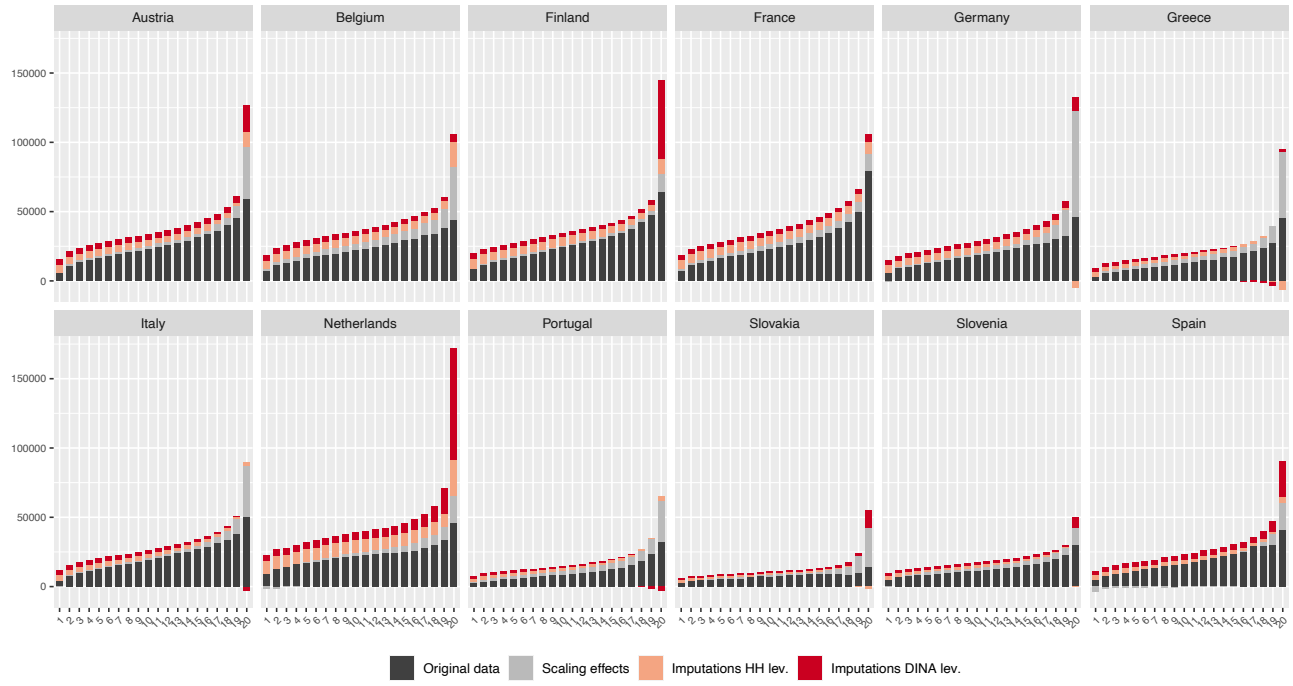


Figure D.19 **Components of POTNI by Income Vingtile (SILC, Wave 2014, Equal Split)**

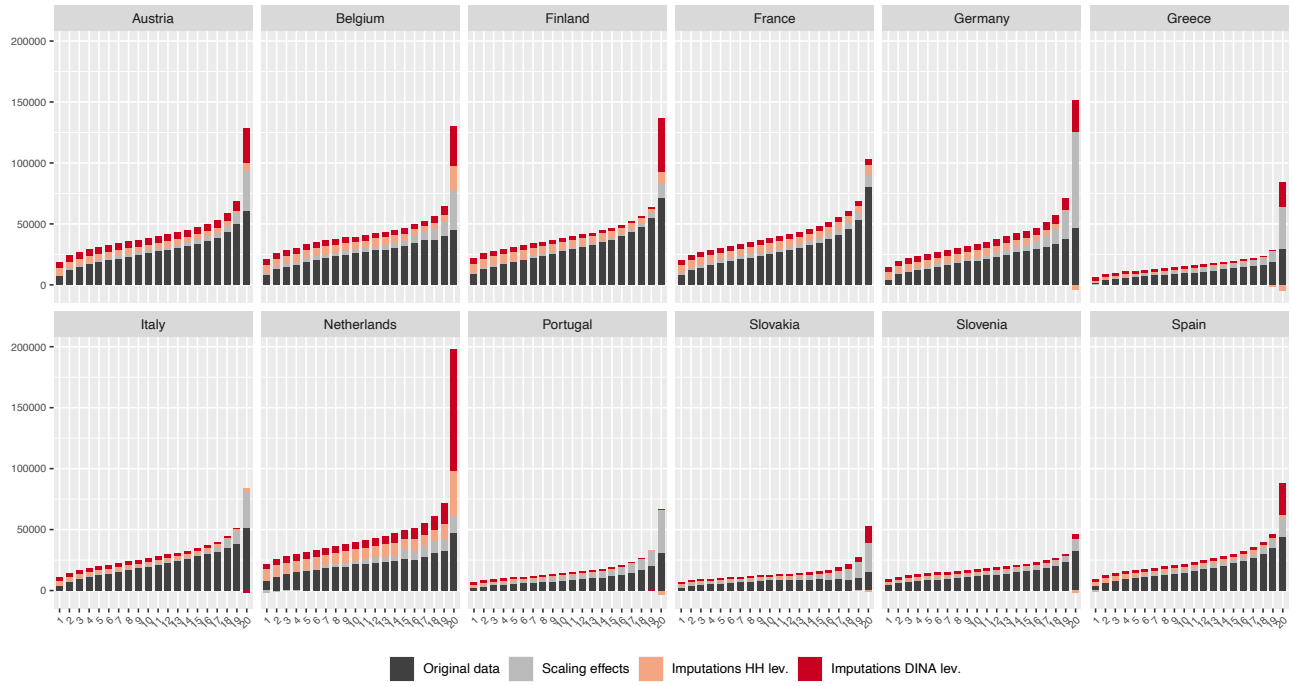


Figure D.20 **Components of POTNI by Income Vingtile** (HFCS, Wave 2010, Individual Split)

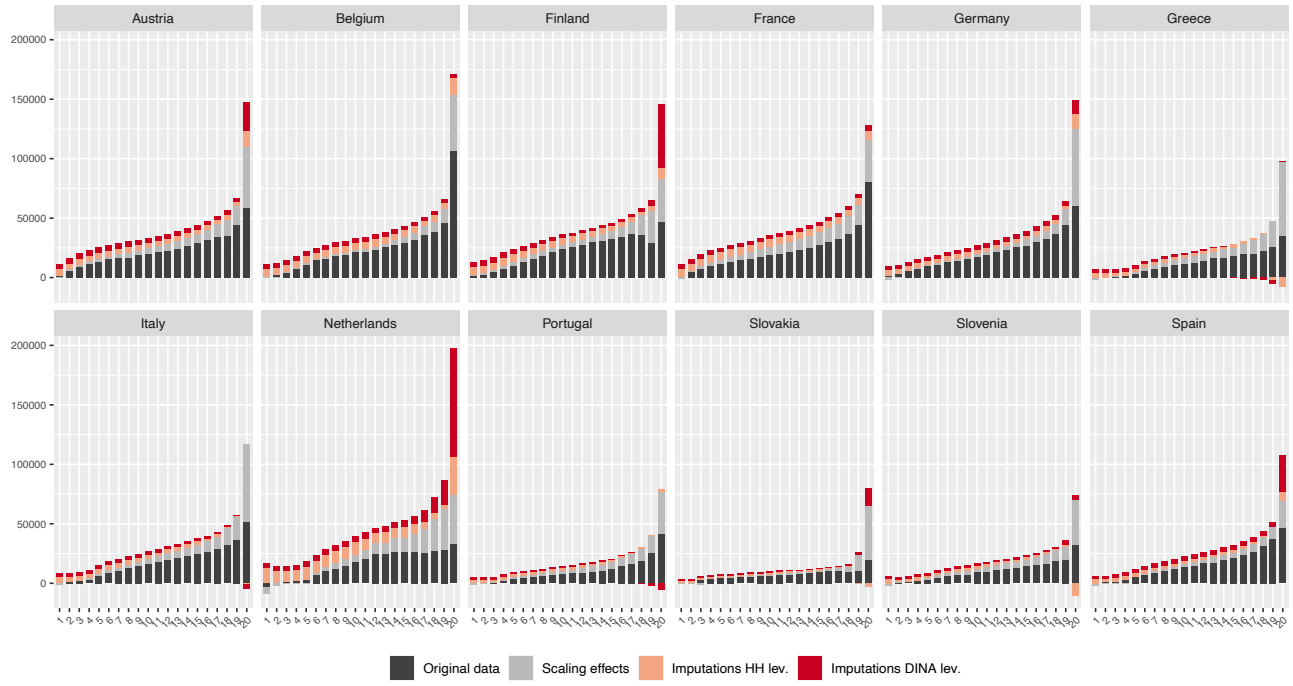


Figure D.21 **Components of POTNI by Income Vingtile** (HFCS, Wave 2014, Individual Split)

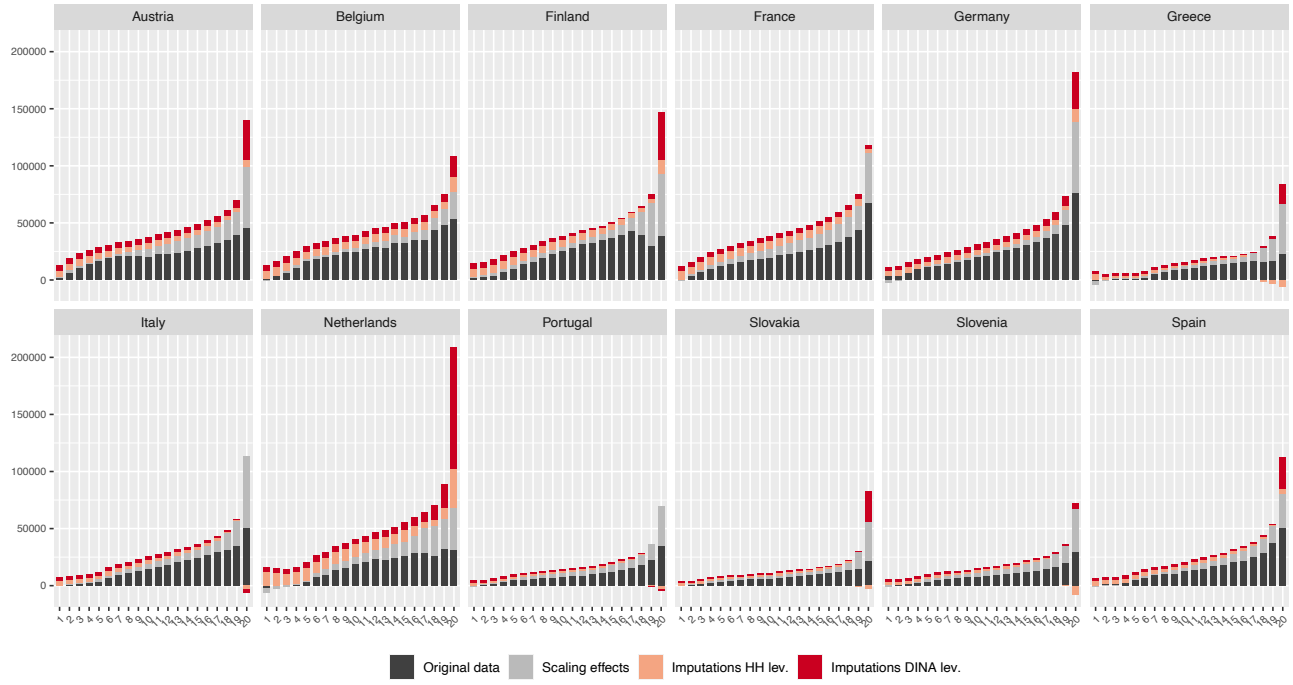


Figure D.22 **Components of POTNI by Income Vingtile** (SILC, Wave 2010, Individual Split)

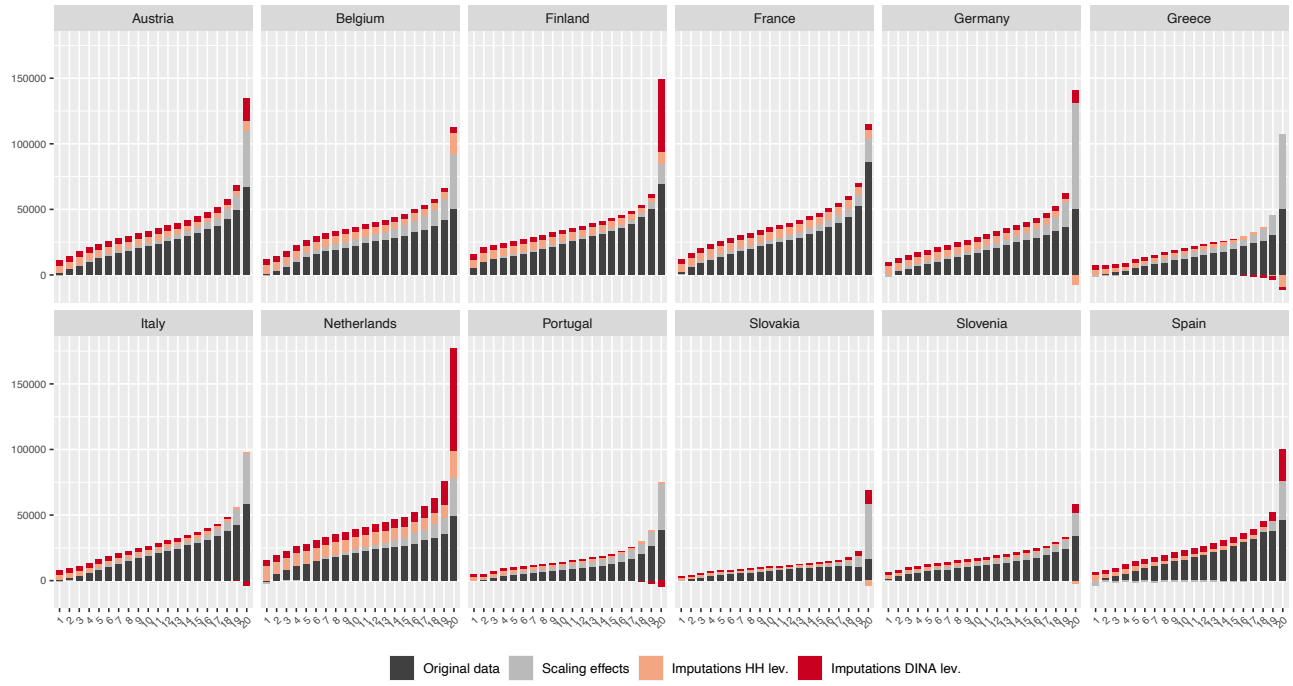


Figure D.23 **Components of POTNI by Income Vingtile** (SILC, Wave 2014, Individual Split)