Adding Inequalities to the SNA Framework: How Macro Disposable Income Benefits and Differs from Micro Disposable Income

Arjan Bruil
(Statistics Netherlands)

Paper prepared for the 35th IARIW General Conference

Copenhagen, Denmark, August 20-25, 2018

Session 4D: Distributional Diversity in the National Accounts

Time: Wednesday, August 22, 2018 [14:00-17:00]

ADDING INEQUALITIES TO THE SNA FRAMEWORK

HOW MACRO DISPOSABLE INCOME BENEFITS AND DIFFERS FROM MICRO DISPOSABLE INCOME

Author: Arjan Bruil

Abstract: Driven by influential reports from Stiglitz *et al*, and the IMF and FSB, the demand for distributional measures in the National Accounts increased. In almost all of the research by National Statistical Institutes, and many of the publications in this field, distributions are added to the national accounts data which serve as the benchmark totals. However, in construction of the household sector accounts micro data sources are often not used, resulting in unaccounted data gaps.

In this article we work the other way around, we create the national accounts totals from micro data, thus including distributions from the very first moment of constructing the national accounts. This approach improves the national accounts in two ways, first the household sector accounts are less dependent on counterpart information or the residual approach, and second, the distributions within the sector are consistent with the macro totals.

As a result of this increased attention for distributions in the national accounts, the understanding of micro macro gaps becomes more prominent as well. The levels of micro and macro disposable income already differed, but now inequality measures as well. In this paper we create the household sector in the national accounts from micro data. We combine many data sources which allow us to present a detailed analysis of distributions within the national accounts. We feel that the present approach adds an important perspective in analysing inequality, because we include otherwise unaccounted income components. We will show this by explicitly linking the micro and macro disposable income.

Keywords: Distributions, Households, Inequality, National Accounts, Sector Accounts

Content

Introduction

In recent years income and wealth distributions in the SNA framework gained more attention. The Report on the Measurement of Economic Performance and Social Progress by Stiglitz, Fitoussi and Sen (2009), and the IMF/FSB report to the G-20 Finance Ministers and Central Bank Governors (2009) made several recommendations in this area. These reports have been followed up by an expert group by the OECD and Eurostat (Expert Group on Disparities in a National Accounts framework), and in a later stage the ECB launched the Expert Group on Linking Micro and Macro statistics. Both Expert Groups focus on the breakdown of the household sequence of accounts by household type. Where the OECD focuses on income, consumption and savings, the ECB addresses the financial balance sheets. The work done internationally inspired NSI's to publish results of these breakdowns, among which the Netherlands.

The SNA framework was developed to depict developments of macro aggregates, where a micro view is more and more desired. In itself these distributional measures were not new. Statistics Netherlands has a history of research on and publications of distributional measures linked to the national accounts (Huigen, Van der Stadt, & Zeelenberg, 1989) (Timmerman & Van de Ven, 1994). In the studies concerning these breakdowns, micro data, which contain the distributional information, is linked to national accounts macro data. However the data sources used to construct the national accounts, and the data used for distributional measures often don't come from the same source. This leads to data gaps, which are addressed excessively in the EGDNA (Fesseau & Mattonetti, 2013a) (Fesseau & Mattonetti, 2013b). Moreover, current practice in the construction of national accounts is that the household sector often depends on counterpart sector or the residual approach due to a lack of data sources.

In this paper we propose a bottom up approach, constructing the household sector from multiple micro data sources, thus including distributions from the very first start of the construction process. We feel that very useful data sources are available for the construction of the sector accounts, which can directly address distributions as well. Our aim is to construct a micro database covering all individuals in the Netherlands, for which we set up the household sector accounts. This database

sums up to macro totals, thus allowing for detailed analysis of the household sector that is always consistent with SNA concepts and totals.

One benefit of this approach is that the household sector is less dependent on counterpart sectors for the macro results, and that the SNA framework is strengthened as well by adding additional checks and balances. Furthermore distributional measures are consistent with SNA totals. To achieve this, every integration decision has to be allocated to individual households. Because of this detailed link, it is also possible to address the differences between the micro disposable income (as taken from the integral income and wealth studies, one of the used data sources), and the macro disposable income.

In the next paragraph the household database is constructed. The compilation process of sector accounts is briefly explained, but a large part of this paper is dedicated to the methodology and data sources. We will focus on the adjustments needed to arrive from micro estimates at SNA values. In the results section we will show the resulting income distributions of the SNA household sector for 2015 and 2016. Also we will go into detail for the specific differences with micro estimates of disposable income and inequality. We end with a summary and conclusion.

Methodology

Process of sector accounts compilation

The foremost reason of our work is that we need to construct national accounts. For this purpose we need data sources covering the economic behaviour of the agents in the framework. In this paper the explicit focus is on the household sector, however, other sectors would benefit from these efforts as well. Second, we want to cover the distributions within the household sector. This approach is more or less consistent with research done in international expert groups (i.e. the OECD Expert Group on disparities in a national accounts framework, and the ECB expert group on linking micro and macro statistics). In these expert groups the national accounts totals serve as the benchmark for the distributions, which are combined afterwards. We feel that, in order to truly benefit from these micro data sources, they should better be included in the construction process. In this respect we differ, however in many other aspects we follow a similar approach. The scope of the household sector is in accordance with the expert groups, but differs from the Distributions in the National Accounts (DINA) project (Alvaredo, et al., 2016) who also allocate the income and net worth of the other economic sectors to households. An overview of the differences between the EGDNA and DINA approaches are set out by Zwijnenburg (2017).

The compilation process of the sector accounts consists of roughly two phases. In the first phase, each sector constructs a full sequence of accounts, based upon available data sources. In the second phase, the estimates of all economic sectors are integrated until a consistent set of accounts is achieved. The resources of one sector are the uses of another, and conflicting estimates that occur in the first phase are balanced out in the second. In many countries the household sector is largely determined through counterpart data. This means that in de integration phase the counterpart information is chosen over the household data. And also, in case there is no data for the household sector at all, the residual approach is used. Also in the Netherlands for most transactions the counterpart sector was considered to have the better, more comprehensive and reliable data sources. The studies presented in this paper strengthens the household sector by being less dependent on these counterpart sectors. The benefit of including the micro data in the construction phase, is that there is a feedback loop, and that also the counterpart sectors should be able to

explain why they should be preferred. In this paper we will focus on the construction of adjusted disposable income, including the income transactions that are needed. This means that consumption expenditures, savings, and wealth are out of scope.

Data sources

There is not one data source that can cover the wide span of transactions on the SNA sequence of accounts. We follow a multi sources approach where we combine census data on the population with integral administrative data, but also surveys, where needed. These data sources are elaborated upon below. In this paper we focus on the current accounts of the SNA, up to and including the *redistribution of income in kind account*, which leads to the adjusted disposable income as the balancing item. Data sources on the net worth of the households are only mentioned in case they were used in the construction or integration of these current accounts as well. We consider two years of national accounts publications, 2015 and 2016. Both years have a final status in national accounts, however the micro data source do not have the same status as the national accounts totals. Due to timeliness of the data not all sources that were available for 2015 could be used in 2016 as well. In that case we found a second best solution for 2016. We will update the distributions in case new data sources arrive, however, only when these are used as a proxy or merely for distributions. Data sources that alter the national account totals cannot be updated, unless macro totals are revised as well.

We use **census data** to determine the population of the household sector. We combine the population on the first and the last day of the year. From each individual we have background characteristics such as the gender, year and month of birth, country of origin, the position in the household, and the (encrypted) social security number. The latter allows us to link this data on the individual level to other data sources. Also we know from each individual to which household she/he belongs, allowing us to aggregate over individuals to households, and break down households to individuals.

The population that is registered on the first of January does not necessarily be present the entire year. Also census data on the deaths within the year are available, and using this we can derive the emigrant flows. Individuals that were present on the first of the year and not on the last day of the year, but which did not pass away, are considered to be emigrated. Also the other way around we can estimate the immigrant flows. Individuals that are present on the last day of the year, but not on the first, while they have of date of birth before the research year, are considered to be immigrated. The resulting population overview is given in Table 1 below. Flows within the year that do not come forward in the population on the used reference dates (for instance immigrants who also leave the country, or pass away), are disregarded in this approach[1].

Table 1: Population overview

	1-jan	immigration	newborns	emigration	deceased	31-dec
2015	16.900.726	198.097	170.341	143.696	146.348	16.979.120
2016	16.979.120	220.861	172.288	142.587	148.175	17.081.507

[1] Our choice to include immigrants and newborns seems straightforward, however it does raise some issues concerning derived household characteristics. For example the household composition is derived for the 1st of January, however these flows change the composition. Therefore the characteristic does not match with the composition later in the year. Because this holds for every reference date we could use, we decide to keep using the 1st of January as the reference date for deriving these kind of characteristics. This does mean though that in a single person household a child might be added, or a partner. Or a couple could be a single person household for most of the year.

The **Registers for Addresses and Buildings** (BAG) is an administrative data source, which describes all buildings, accommodation objects, pitches, berths, residences, public spaces and numbering marks on the territory of the municipalities of the Netherlands. We select only the dwellings (including those with practice), farms, and non-dwellings that are partly used as dwellings. This leaves us with 7,336,995 units of observation. We further select only those dwellings that are owner-occupied, resulting in 4,097,789 units. Of these addresses the building year, surface and region of the address / building is known and a unique key to link the owner to the household population. This data sources is available for 2015 and 2016.

The **Integral Income and Wealth Studies** (IIWS) has been designed to give a picture of the composition and distribution of the income of individuals and households in the Netherlands. It is an integral register of all individuals present in the Netherlands on the first day of the year, which is consistent with census data. The underlying data sources are tax records, but also data from the education administration (in Dutch: DUO), and the organization that implements national insurance schemes in the Netherlands (in Dutch: SVB), are used. The IIWS records detailed income and wealth information by individual and households. Income components are mostly available on the individual level, except for some specific imputed items (income from owner occupied dwellings) which are imputed to a single individual while they might concern others as well. Wealth components are available only on the household level. The IIWS covers both private and institutional households, however on the publication level the latter are left out of scope. This data source is used by Statistics Netherlands to construct, among many other publications, micro disposable income and inequality measures (Table 2). This data source becomes available in a preliminary version and a final version. For the construction of the household sector, only the preliminary version can be used.

Table 2: Micro statistics

	Number of households	Average micro disposable income	Total micro disposable income[2]	Gini coefficient
	x 1,000	*1,000 euros*	*1,000,000 euros*	
2011	7347,6	36,6	268.922	0,288
2012	7412,1	36,8	272.765	0,289
2013	7467,8	37	276.309	0,289
2014	7496,4	39	292.360	0,303
2015	7568,5	38,6	292.144	0,290
2016[3]	7619,8	39,7	302.506	0,288

Source: http://statline.cbs.nl/Statweb/selection/?VW=T&DM=SLNL&PA=83739NED&D1=0-1,5&D2=0&D3=a&HDR=T&STB=G1,G2

The data source used for the construction and distribution of the consumption components of Dutch households is the **Household Budget Survey** (HBS). The HBS is held once every 5 years, the last time for the survey year 2015. Statistics Netherlands asks households to fill out a survey with general questions on the household, fill out a survey on recurring periodic expenses, keep track of all non-

[2] The total micro disposable income is not published, but is derived from the average and number of households.
[3] Preliminary figures

periodic purchases in a diary during 4 weeks, and fill out a survey on very large expenses and holiday activities in 2015. The total sample size is 90 thousand households, in which low and high income groups are oversampled. The response rate is approximately 17%, which leads to a survey population of 15 thousand households. This covers only private households, persons living in institutions (i.e. old age homes, prisons, medical facilities) are excluded. The HBS measures the consumption expenditures on the household level, therefore individual estimates are unavailable.

The **Pension Claims Statistics** (PCS) aims to give an overview of the mandatory (employment related) pension entitlements of the population. The most recent year for this data source is 2014 currently. Timeliness of this data source is an issue, as is coverage. It is a partial register received from pension funds. Not all pension funds are asked to respond, and, moreover, it only covers individuals who are in the contribution phase of their pension, i.e. who are between 21 and 65 years old. Of these individuals the encrypted social security number is known. Statistics Netherlands works on a redesign, which should improve both the coverage and the timeliness of this data source.

The main aim of the **Household Finance and Consumption Survey** (HFCS) is to gather micro-level structural information on euro area households' assets and liabilities. The survey also collects other information in order to analyse the economic decisions taken by households. In the Netherland this survey is performed by the Central Bank and CentERdata. The target reference population for national surveys is all private households and their current members residing in the Netherlands at the time of data collection. This survey is held once every three years, the most recent survey was for 2014. From this survey we use the variable whether or not the respondent has a voluntary pension/whole life insurance, and the average value of this pension/insurance.

Money Transfer Operators (MTO) serve as the channel for remittances to and from the Netherlands. The Dutch Central Bank gathers information of these MTO's for the total flows per receiving and sending country. This an annual data source, which is used only for flows to and from abroad. Flows within the Netherlands are included in the data source, but neglected for our purpose.

Statistics Netherlands compiles a **Satellite un-incorporate enterprises** (SZO), that allows multiple users within the statistical institute to analyse and publish data on self-employed. It contains information about all the self-employed in all industries. This satellite contains individual tax declarations for the self-employed, including fiscal income and profits. For the compilation of the labour accounts this information is used to estimate variables concerning self-employed. The data are also used for estimating the dual classification of production, intermediate consumption and value added and its components in order to link the supply and use tables to the institutional sector accounts. For the purpose of the household breakdown this register contains mixed income from self-employment, interest paid and received, consumption of fixed capital, paid compensation of employees, and the encrypted social security number.

For the benchmark revision year (2015), this data source was available in time. For a final year it is not, because the underling tax records are not sufficiently available yet. In this case, for the macro total an estimate is made based upon the results of previous year, and known developments in self-employment income. For our final year in this paper (2016) we use this macro total, and make use of the IIWS for the distribution over households and individuals.

Also for the **compensation of employees**, we make use of a register data set provided by the labour accounts. Underlying this wage register that is used in the household distribution project, multiple sources are combined. These are confronted and used to publish statistics on jobs, hours worked, wages etc. This data source covers all individuals working in the Netherlands, and distinguishes compensation and social contributions, and again the encrypted social security number. Wages

earned by foreign residents are included because the labour accounts cover the earnings in the national territory. These will not be linked to our household distribution dataset in case one is nor registered in the Netherlands at the first or last day of the year.

In this paper we make use of data of the LISS (**Longitudinal Internet Studies for the Social sciences**) panel administered by CentERdata (Tilburg University, The Netherlands). The LISS panel is a representative sample of Dutch individuals who participate in monthly Internet surveys. The panel is based on a true probability sample of households drawn from the population register. Households that could not otherwise participate are provided with a computer and Internet connection. A longitudinal survey is fielded in the panel every year, covering a large variety of domains including work, education, income, housing, time use, political views, values and personality. The core, longitudinal, study is repeated yearly and is designed to follow changes in the life course and living conditions of the panel members. In addition to the LISS Core Study there is ample room to collect data for different research purposes. This data received covers the year 2014, but in our construction it is used for 2015 and 2016 as well.

This research also uses data from the **Giving in the Netherlands Panel Survey** (GINPS) collected by the Center for Philanthropic Studies at VU University Amsterdam. GINPS is largely supported by a grant from the Netherlands Ministry of Justice (Bekkers & Boonstoppel, 2010) (Bekkers, Schuyt, & Gouwenberg, 2002-2008). It focuses on giving behaviour to NPISH mainly. Giving in the Netherlands aims to chart the giving behavior of households, individuals, funds, companies and charity lotteries. This not only involves monetary contributions, but also the time and effort in the form of volunteer work in many social fields. This survey is held every two years, the most recent data covers 2015. The studies showed that the 20% wealthiest households made 80% of total gifts to NPISH such as charities, and religious institutes. Wealth however is self-recorded, so no clear definition is available.

The social transfers in kind are an important transaction in the redistribution of income. Including these income components leads to an alternative income measure in the national accounts, i.e. the adjusted disposable income. In the welfare statistics these flows are hardly covered, only the received rental subsidy is. A large part of the social transfers in kind is the compensation through the **healthcare insurance act**. This source contains per Dutch resident who is insured via the standard insurance, his costs per year for care insured via the standard insurance. The standard insurance is legally required via the Health Insurance Act for almost all Dutch residents. The costs are those costs actually reimbursed by health insurers. The costs are subdivided into care forms, for instance birth care, hospital care, etc. This data source is not yet available for 2016, but was used for 2015. For the **long term care act** we use average amounts per 5-year-age-class and gender. These data are taken from the Cost of Illness studies[4] by the National Institute for Public Health and the Environment (RIVM) in cooperation with Statistics Netherlands. These are available for 2015 only.

Another large part of social transfers in kind is education. For this we use data from the **Education Administration** (DUO) which is an administrative data source covering education enrolment. The available information includes a unique student number, type of education enrolled in, institution where education is followed, and also here the encrypted social security number.

Finally, a small part of social transfers in kind is the paid for use of legal counsel under the **Law on Legal Aid**. Within Statistics Netherlands the individuals that are registered in the municipal administration, and that have received legal counsel under the law on legal aid are included in a

[4] https://costofillnesstool.volksgezondheidenzorg.info/tool/english/

register. For these individuals the number of times one has received legal aid, distinguished by jurisdiction, is registered.

Creating the household database

In this section we'll discuss the methods to derive the macro estimates - including the distributions – from the micro data sources. We create a household database wherein each individual in the Netherlands is represented, and for whom the sector accounts are created. We do this by linking micro data sources to the population register that is set up from the census data. Ideally we do this through a record linking technique, using the (encrypted) social security number as a unique key, which is available in many of the mentioned data sources. For those sources that do not include this unique key we synthetically impute the database, by allocating an average amount to each individual, a household, or a group of individuals based upon the characteristics that are available in both the household database and the respective data sources. The total number of records in our database is 17.372.269 in 2016. This is higher than the population at a given date due to the fact that we include the flows as well (Table 1). It equals the population at the first of January (16.979.120) + the immigrants (220.861) + the new-borns (172.288).

The importance of this exhaustive population register comes forward when data sources are matched with our census data. Often these data sources focus on the population on a reference date, hence not representing the entire national accounts population. For almost all data sources we need to correct for part of the missing population. Without estimates for these groups, the sum of the micro distributions can never equal the SNA macro totals.

Even though we use a multi-source approach, not all transactions have a micro counterpart. This is mostly the case for specific SNA constructs, such as fisim or the income attributed to insurance policy holders. For the transactions for which no micro data is available we take the SNA total from a previous national accounts status[5] in the construction phase, and use a proxy variable for the distribution over the population. In the integration phase we take over the estimates of the counterpart sector(s) as the macro total and use the same proxy for the distributions. A summary of the data sources and mentioned values is given in Table 6.

Operating surplus

In first instance we record link the Integral Income and Wealth Statistics (IIWS) to our database and estimate the operating surplus. For households this equals the impute rent of owner-occupied dwellings. In the IIWS this amounts to 31.923 million euros. This excludes the intermediate use of fisim. In the construction phase we include the intermediate use of 2015 (-23.772). The intermediate use of fisim is distributed using the interest payments of mortgages as a proxy.

In the integration phase we include the Registers for Addresses and Buildings (BAG), also this data source is record linked to individuals, in this case the owner-occupant of the dwelling. Using this data source we estimate the imputed income based upon a regression analysis, following the practice of the supply and use tables for the macro estimate in the national accounts. The necessary parameters are estimated from the rental survey of Statistics Netherlands. The regression analysis is based upon the characteristics of the dwelling. It includes the region where the dwelling is located (North, Middle, South, or one of the four largest cities), the surface of the dwelling (less than 50 square

[5] National accounts in the Netherlands are made twice. The first estimate has a preliminary status, the second a final one. The benchmark revision of the national accounts creates a third status, revised. In the construction phase of the benchmark revision year (2015), the final version is considered the previous one. In the regular cycle the preliminary version is the previous. For a preliminary year this exercise is not possible due to the timeliness of most data sources.

meters, between 50 and 100, between 100 and 150, between 150 and 200, and 200 square meters or more), and the building year of the dwelling (before 1950, between 1950 and 1970, between 1970 and 1990, or after 1990).

The regression used is:

*Monthly Imputed net rent = 1151.625 + (-73.372) * Northern region + (-45.872) * Southern region + (-33.648) * Middle region + (-507.118) * Surface<50 + (-338.706) * Surface50-100 + (-194.790) * Surface100-150 + (67.390) * Surface150-200 + (-208.807) * BY<1950 + (-215.564) * BY1950-1970 + (-158.273) * BY1970-1990.*

A correction is made for the share of vacancies and non-use of dwellings. These shares are respectively 0.969 for dwellings, 0.9 for dwellings and practices, 0.75 for farms and 0.3 for non-dwellings that are nonetheless used as dwellings. The result of this calculation is the imputed monthly income. This is multiplied by 12 months, and balanced to the macro total of the SUT. Rents are usually raised in the middle of the year, and the rent reflected in the rental survey is used for the second half of the year. On the macro level the rent of the first months is deflated, in our approach we take this step in the balancing process. On the macro level also additional estimates are made for houseboats, garages, and caravans. The imputed income equals output (38,638 mn), and intermediate use has to be deducted to get the operating surplus. For intermediate use we have no micro data available (except for intermediate use of fisim), thus we assume that the distribution of the estimated output equals the distribution of the balancing item.

The reason that output according to the BAG differs from the output using the IIWS is in the approach that is used. In national accounts the characteristics of the dwelling are used, in the IIWS the value of the dwelling. This leads to differences on the macro level, but also in the distributions within the household sector.

Mixed income

Mixed income occurs only within the sector households, it consists of income from self-employed, from the non-observed economy, from rentals of property, and from recuperation. Macro calculations are important for most of these components, especially in 2016.

The largest component of mixed income in the national accounts is the income from self-employed. In 2016 we take the distributions from the IIWS, because the satellite un-incorporate enterprises (SZO) is not available in time[6]. The macro estimate for this income is our benchmark total. From the IIWS we taken the self-employment income (33,554 mn) and add the variables that are considered income from employees in the IIWS, but self-employment in SNA (3,020 mn). This balanced to the macro total (46,856 mn) which is estimated on preliminary estimates by the labour accounts. This gap is considerable, but from the 2015 exercise – when the SZO was available – we know that a large part of this gap is the results from a different recording of interest paid and received by self-employed. This is part of the income concept in the IIWS and part of property income in the national accounts. Also the IIWS is a net figure, whereas the national accounts is gross. In 2016 we allocate the interest proportionally over the self-employed and the consumption of fixed capital according to the distribution of business wealth.

Another part of mixed income is the income from renting out dwellings or company buildings. In the construction phase we link this to IIWS (3,694 mn), in the integration phase we benchmark it

[6] In 2015 we could also use the micro data of the SZO. Our distributional results of 2016 will be updated when a full picture of micro data sources is available.

(proportionally) to SUT macro totals (5,284 mn). We do something similar for the non-observed economy. This is not recorded in tax data of course, but in the supply and use table there is a macro estimate available for legal activities that are not declared to the tax authorities, and illegal activities such as the production and trade of illicit drugs. For the distribution of the macro estimate we use a studies by Kazemier (2014), who did research on the grey economy, based upon sample surveys. His logistic regression of the participation on the hidden labour market included the age of the individual, household size, monthly income, hours worked, education, job specifications, the expected probability of detection, and whether or not the respondent knew someone else who worked on the hidden labour market. In our database we can only link the first three of these determinants, but it is the best we have so far. We use the resulting probability that one participates in the hidden economy both for the grey market work, and for the illegal activities. Specific macro corrections for cost fraud by self-employed are distributed evenly over self-employed, and corrections for exhaustiveness evenly over everyone in the population.

Compensation of employees

For wages and salaries there are two data sources available. The IIWS, and the Labour Accounts Wage register (LA). Both are record linked to our population register. We use the former in the construction, and the latter in the integration phase. We prefer the latter because it is already brought to national accounting concepts. These distributional results are, by construction, consistent with the macro totals. The corrections we still need to make (-5,846 mn on wages and salaries) and (-1.199 mn on social contributions) are because we focus on the population in the Netherlands. Part of the LA total flows abroad, which we identify by record linking this data source to our household population. The LA links their underlying data source to the GBA on a monthly basis to accomplish the same. In case there are still differences, we distribute this proportionally. The income flow coming from abroad are not captured by the LA, but these are covered by the IIWS.

Property income

In constructing the household sector accounts the IIWS is used for the macro estimate and distribution of property income, both paid and received. Interest received in the IIWS amounts to 2,967 mn euros. This consists of interest on savings, and interest on bonds. A correction is made for interest received (61 mn), because the IIWS has a different recording of interest paid for business purposes by self-employed. For 2015 this amount is known from the SZO, and this is projected to 2016. In the integration phase macro corrections are made because mortgage related savings are not included in IIWS, hence the income flows are not either. We allocate this macro correction (1,775 mn) proportionally over households with a mortgage. This way we assume that households with a higher mortgage also have high savings, and high interest flows. In the future, loan level data on mortgages will become available, which can supplement this data source on this subject.

Paid interest consists of interest paid on mortgages, on student loans and other loans, in total 31,177 mn euros in the IIWS. The corrections (2,474 mn) reflect here again the interest paid by self-employed, similar to the received interest but at a higher level. In the integration phase we adjust the interest paid downwards, in relation to the integration of total (mortgage) debt which is lower according to the banks than it is according to the households. The relative interest rates remained the same in construction and integration phase.

In national accounts the interest received and paid is corrected for the financial intermediation services indirectly measured (fisim). There is no micro data available for this interest margin, therefore we use the distribution of the related interest flow as the proxy for the distribution of fisim. On the macro level fisim does not influence disposable income much, however for individual household it can be influential. A household that pays much interest has a relatively low micro

disposable income. Because national accounts reassigns part of this interest payment to consumption, the households' macro disposable income is higher.

Dividends (including those attributable to collective investment fund shareholders) are taken from the IIWS. The share for collective funds is in the integration phase decided by the value of the previous year (675 mn). In the IIWS dividends include the dividends for owners / major shareholders. This is a difficult variable to interpret, because these dividends do not necessarily had to be paid out in the fiscal year. The value responds strongly to fiscal regulation, in years where the tax percentage was slightly favourable, much more dividends were reported. Therefore in national accounts we decided to set a maximum level, on 4.2 percent of the outstanding shares of these companies. In 2016 this maximum was not reached, but in 2015 it was (5,500 mn). Because these dividends are mostly earned by the top 1% of the income distribution we influence inequality measures as well with an adjustment on this variable. In 2016 an adjustment for excessive dividend flows was not needed though. In fact we found the total reported dividend flows by the non-financial corporations was far higher than reported by the receiving sectors (as was the total value of shares). Therefore in the integration phase total dividends more than doubled (6.989 mn). Also this is allocated to those who have reported dividends. The retained earnings attributable to collective investment fund shareholders are distributed using the dividends (other than for owners / major-shareholders) as a proxy variable.

Table 3: Dividends (mn euros)

	2015	2016
IIWS dividends	14,519	6,440
Shift to D.4431	-285	-675
Maximum 4.2% of share value	-5,500	-
Integration	3,439	6,989
Publication	12,173	12,754

Quasi-corporations do not exist in the Netherlands. All self-employed are included in the SZO and thus in mixed income. The reported value of income from quasi corporations represents the income received from holiday homes abroad. This is estimated on the macro level, and distributed over household using a proxy variable from the IIWS (value of real estate owned, other than the dwelling for main residence). Also other property income flows have no micro counterpart. In the construction phase these are taken from the previous year, and in the integration phase determined by the counterpart sectors. For the investment income attributable to insurance policy holders, we use the HFCS for the distribution over individuals. From the HFCS we know for a combination of age and gender the average value of voluntary pension benefits. We impute this average value into our register for each individual. The sum is proportionally balanced to the total life insurance and annuity entitlements (F.62), and used as a proxy for the income flows related to these assets.

Similar, we use the Pension Claims Statistics (PCS) to distribute investment income payable on pension entitlements over individuals. The PCS of 2014 is an incomplete survey, so imputation is needed. The PCS does contain the unique key which allows us to record link the data source to our population register. Second, we impute the PCS by aggregating over age, gender, and income group. The resulting amount reflects the yearly entitlement for the population aged 21-65. For the population over 65 years old, we use the actual pension benefit from the IIWS. For the entire population we recalculate this amount into the total entitlements using the prescribed discount rate, the life expectancy by age and gender, and the retirement age forecast. The resulting entitlements

reflect our estimates for the pension entitlements (AF.63) in the opening balance sheets, and are also used as a proxy for the income attributed to insurance policy holders.

Taxes

The current tax on income is initially taken from the IIWS (56,179 mn). We correct these for the missing population, i.e. immigrants. For these individuals we know their labour income through wages and salaries (from the LA), but only in the integration phase[7]. We use a long-term tax rate on wages and salaries, as derived from the national accounts historic results (16.7%). Moreover, we correct for implausible results, for example when taxes are paid by young children. However, there remains a large micro-macro gap. There are two reasons for this. First we combine different data sources; we use the LA for wages and the IIWS for taxes. A cross-check of the wage data in both sources shows for example that some households have high wages according to the IIWS and high taxes. If, in the LA, these wages turn out to be much lower, or even 0, high taxes remain. Because the tax in the IIWS is a calculated tax, we re-evaluate this in case we adjust the wages as well[8]. Second, there is a difference in recording of these taxes in national accounts government sector, and the IIWS. As just mentioned the IIWS calculates the tax, following the tax forms. This leads to the amount that should be paid over the income earned in that year. The government sector in the national accounts records these taxes differently, they use the date of receipt by the tax authority (with a one-month delay). However, a receipt in 2016 can just as well be related to income earned a year ago, or even further back. Therefore the micro distribution does not necessarily relate to the macro totals. This is a subject for further study. Currently we make the abovementioned corrections, and proportionally allocate the remaining difference over the population.

Some transactions which are in the SNA considered income components are in the welfare statistics considered consumption. Other current taxes for instance, which include motor vehicle tax or environmental taxes. We use the HBS to impute data to the population register, using the household composition as the linking characteristic. We choose this characteristic because it is both distinctive for many classes, and the related taxes often depend on the household size. This amounts to 8,144 million euros. We assume that the missing population (in institutions) behaves similar as the private households, but we assume they don't pay taxes regarding the use of a vehicle as these individuals are mostly very old, or living in a prison or hospital for a long time. For immigrants we assumes that they only lived in the Netherlands for half a year thus we allocate half of the amount according to the HBS to them. The HBS covers the expenses of 2015, for 2016 we assume growth comparable to the consumer price index of the related taxes. Total corrections add up to 174 million euros. In the integration phase, we use the proportional allocation of the difference.

Social contributions and benefits

Part of the social contributions and benefits are by definition equal to the employers' social contributions. Also the supplementary contributions equal the imputed income. The households' pensions (8,838 mn euros) and non-pension contributions (64,629 mn) are taken from the IIWS. The latter consist of multiple schemes, such as unemployment or disability. Again corrections are made for implausible results and estimate the contributions immigrants would have made given a long-term average ratio between wages and contributions. For pensions this is 4.7% and non-pension contributions 24.4%. The social contributions have the same recording issue as the taxes on income.

[7] In the construction phase we allocate an average value to the immigrants, based upon the value of previous year.

[8] This example would indicate that the IIWS could be preferred over the LA because of the consistency between transactions. However this consistency is not always present, it also happens that a household has high wages in the IIWS, but does not pay taxes. In this case lower wages in LA seem more plausible.

Moreover the split between taxes and social contributions is not exact and this can only be determined after several years. Because the split between the pension and non-pension part might not always be exactly clear we decide to integrate the two components of households' contributions jointly so as not to lose distributional information. This is done proportionally over the households.

The social security benefits are done similarly. First they are matched with the IIWS, and records are corrected for results that cannot be true according to the regulations. For instance, if individuals young than 65 receive a public pension benefit, or in case children receive an unemployment benefit. In integration phase we further analyse the differences per social scheme (Table 4). The macro totals per scheme are available in the government sector and combined with the micro totals. Differences are allocated to household that take part in the relevant schemes. The integration correction we make for the disability schemes is expected, and shifted to the other current transfers. In the IIWS it is known that also private income insurance is included in the data, but this cannot be allocated exactly. The integration amount for this scheme (-1,359 mn euros) is therefore considered to reflect at least part of the integration amount of the other current transfers as well.

Table 4: Social security non-pension benefits in cash (mn euros)

	IIWS	corrections	GOV[9]	integration
Unemployment	6,778	-0	6,729	-49
Survivor	383	-	416	33
Old age	35,669	-262	35,407	0
Disability	11,705	-	10,346	-1,359
Other	1,127	-	401	-725
Total	55,661	-262	53,299	-2,100

The pension benefits amount to 43,225 mn euros in the IIWS. These also include the voluntary pensions. In national accounts these should be recorded as financial transactions. We cannot identify these flows in the construction phase. But the confrontation with counterpart sector information gives us a data gap that should match these flows. We assume that the highest pension benefits are not work-related. The maximum benefit received by an individual in 2016 was 4.5 million euros. This is likely to be a voluntary scheme or a non-life insurance, and not a work-related scheme. We top off the highest pension benefits until the data gap is closed. This is not a perfect approach, because we cut off the entire benefit, while there might be both a voluntary and employment related component.

Social service add up to 20,907 mn euros in the IIWS. For new-borns we impute child benefits[10] by modelling the social system for these schemes. The integration is done per scheme. An additional correction is made for immigrant that can receive welfare as well. We allocate the average of the entire population for this group of individuals as well. The remainder is still done proportionally.

[9] The details are known only from the government point of view. These include flows abroad. We assume that only old age benefits and other schemes flow abroad. The integration total for old age is set to zero and the remainder is allocated to 'other'.

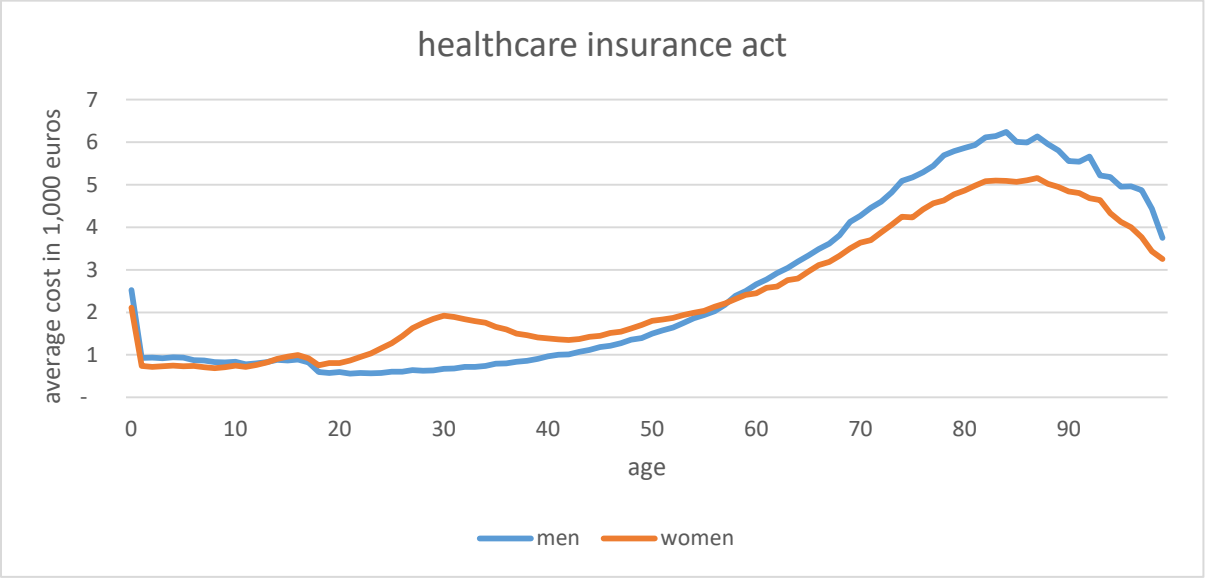
[10] In Dutch: Kinderbijslag and kindgebonden budget.

Table 5: Social service benefits in cash (mn euros)

	IIWS	corrections	GOV[11]	integration
Welfare	5,227		5,685	458
Sickness	608		788	180
Child care	5,080	59	5,178	39
Student loans	1,770		550	-1,220
Healthcare allowance	4,316		4,394	78
Other	3,906		4,298	392
	20,907	59	20,893	-73

Finally, social transfers in kind are included only in the integration process. In 2015 we could record link the healthcare act to individuals, but this data source is not yet available for 2016. Instead we use the average cost by age and gender to impute the 2016 register (Figure 1). For the long-term care act we do something similar. This data is (currently) only available for 2015 on the aggregate level of age-class and gender[12]. The coverage of the healthcare insurance act is 85% and for the long-term care act 100 %. For education, we record link the enrolled individuals from the education registration and allocate the total amount evenly over the participants. We do this for primary education, secondary (vocational) education, and tertiary education. Similarly, we record link the individual data of the Law on Legal Aid. These reflect the number of times one has had legal aid under this law. We use this to proportionally allocate the total amount. Rent subsidies are taken from the IIWS (3,409), these are part of micro disposable income, but macro adjusted disposable income. Other components of social transfers in kind are either allocated through a proxy, or by evenly distributing them over an age group of eligible individuals.

Figure 1: Average cost for the healthcare insurance act (2015)



[11] The details are known only from the government point of view. These include flows abroad. We assume that only child care benefits flow abroad. The difference between the government total received and household total paid is in this table allocated to child care.

[12] https://statline.rivm.nl/#/RIVM/nl/dataset/50040NED/table?graphtype=Table&ts=1512975518824

The costs of social schemes are made explicit in the national accounts, but not in the IIWS nor another micro data source. We decide to evenly distribute these over all individuals that participate. This means that an insurer or pension fund makes just as much costs for someone with low pension wealth as they do for someone with high pension wealth.

Other current transfers

The non-life insurance premiums and claims consist of insured injuries or damage suffered by persons or goods. The private insurance against injuries of persons is also part of the IIWS. However the premiums are identifiable (1,400 mn euros), but the claims are not. The claims are included in the social security schemes mentioned above. In the integration we include the remainder of the disability benefit (Table 4) first. The remaining difference is allocated proportionally.

The contributions for insurance against damage are taken from the HBS, and linked similarly as the taxes. The claims are unknown and set at the same level as the contributions. This does mean that every household is insured perfectly in our approach, an assumption that is fairly accurate on the macro level, but should be different on the household level. However, even though we combine many data sources, we are limited here. The amount in the HBS is partly a service and partly an income transfer. We assign 0.28 of the HBS amount to the service, based upon the share of services in the paid premiums for 2015, taken from an overview of the supervisor (Dutch Central Bank).

The miscellaneous current transfers consist of many different flows. Here we combine the HBS, the LISS panel, the Giving in the Netherlands Panel Survey, the IIWS, and MTO data. The alimony payments from the IIWS are record linked. The other data sources are imputed.

From the HBS we take the contributions to NPISH, and link it in the same way as taxes. We exclude the contribution to gyms, these are considered to be non-financial corporations. The macro total for these contributions is taken from the European Health and Fitness Market report from 2017. This correction is distributed evenly over the number of individuals in the age category 15-75. Also from the HBS we take the expenditures on lotteries. This is imputed in the same way as the private insurance premiums. In this case we arbitrarily set the share of services and income transfers to 50%. The income transfers is between households, so on the macro level there is no effect on disposable income, but on the household level it should. However we don't have any data on lottery winnings, thus set the resources equal to the uses.

For the gifts to non-profit institutions serving households (NPISH) we use the Giving in the Netherlands Panel Survey (GiN). The data are imputed using the combination of income quintile and wealth group, and if applicable, country of origin. The income quintile and wealth group are mimicked as closely as possible, but an exact match cannot be made. The reason is that wealth is self-reported and it is not exactly sure what a household will include. Also the income concept does not have an exact match. Still these characteristics were preferred over other household characteristics because the studies explicitly claims that the 20% wealthiest households made 80% of total gifts to NPISH such as charities, and religious institutes. Our approach does alter the weights that are found per household, therefore we cut off our total in 2015 to the amount in the survey (2,611 million euro). For 2016 we assume the total amount increases with the same percentage as the deductible item for gifts in the IIWS.

Also included in the SNA are the payments from parents to their children, which we impute using the LISS panel. We allocate an average value to households by background characteristics. The growth we assume is taken from the consumer price index of 2016. These data are imputed in the register. We link the amounts children receive by the age of the individual, and the amounts parents pay by

the income class. These flows explicitly mention that they flow from parents to (studying) children living on their own. However we cannot identify an exact match of parents and children if they live on their own (if it would be a flow within the household we could have). We specify the links a bit by assuming an age group for the paying individuals, where they are most likely to have studying children living on their own. And for the receiving end we further determine that their socio-economic class is student.

We estimate the contributions to home owners associations by imputing an average amount to each household with an own home. This amount (641 euros in 2016) is based on selected services in the HBS (insurance, maintenance and repair). This amount is multiplied by the chance that the household is a member of a home-owners association. This is determined by the total number of households in such an association as a share of the total household with an own home (known for 2015).

Finally, the paid and received remittances are estimated using data from Money Transfer Operators (MTO) received from the Central Bank. This data is linked to the population register by country of origin for individuals of 18 years and older. Because MTO's are likely not to be the only channel, these totals are grossed up with 50%, following the findings of Plaza *et al.* (2011), that for 5 investigated African countries MTO's were in approximately 66% of the times used by migrants to OECD countries. We do cut off the household total in order not end up with extreme values.

These flows are hardly integrated, because for many of these flows the household sector is its own counterpart sector. There are some additional flows added in the integration phase. These are from the government sectors, mainly fines. These are allocated according to the paid fines in the HBS.

Table 6: Micro-macro links

Column		A	B	D	E	F	G	H	I	K	L				M	J	
			construction phase							integration phase							
			micro data sources							micro data sources							
Transaction		NA total	IIWS	HBS	Giving in the NLD	LISS panel	MTO	corrections to the data sources	Construction total	Labour Accounts	BAG	IIWS	Healthcare Act	Cost of Illness	integration amount	Coverage rate construction versus NA total	
Balancing items																	
B2G3G	*Operating surplus / mixed income, gross*	76.881	72.191					-5.440	66.751						10.130	87%	
B2G	*Operating surplus, gross*	12.125	31.923					-23.772	8.151		38.638				26.513	67%	123%
B3G	*Mixed income, gross*	64.756	40.268					18.332	58.600						6.156	90%	
Uses																	
D.41A	*Interest*	31.570	31.177					2.474	33.651						-2.081	107%	
P.119C	*Fisim*	-25.448						-26.239	-26.239						791	103%	
D.45	*Rent*	273	74						74						199	27%	
D.51	*Taxes on income*	49.327	56.179					252	56.431						-7.104	114%	
D.59	*Other current taxes*	7.428		8.144				174	8.318						-890	112%	
D.6111	*Employers' actual pension contributions*	22.213						20.505	20.505						1.708	92%	
D.6112	*Employers' actual non-pension contributions*	36.543						37.743	37.743						-1.200	103%	
D.6121	*Employers' imputed pension contributions*	757						714	714						43	94%	
D.6122	*Employers' imputed non-pension contributions*	11.069						10.683	10.683						386	97%	
D.6131	*Households' actual pension contributions*	12.931	8.838					71	8.909						4.022	69%	
D.6132	*Households' actual non-pension contributions*	66.187	64.629					368	64.997						1.190	98%	
D.6141	*Households' pension contributions supplements*	33.716						33.750	33.750						-34	100%	
D.6142	*Households' non-pension contributions supplements*	501						337	337						164	67%	
D.61SC	*Social insurance scheme service charges*	-9.258						-8.715	-8.715						-543	94%	
D.6222	*Other social insurance non-pension benefits*	256						204	204						52	80%	
D.711	*Net non-life direct insurance pemiums*	9.289	1.400	6.986				-1.749	6.637						2.652	71%	
D.75	*Miscellaneous current transfers*	11.259	674	3.901	2.611	3.097	598	-1.228	9.653						1.606	86%	
Resources																	
D.11	*Wages and salaries*	262.959	264.641					-9.823	254.818	268.805					-5.846	97%	102%
D.12	*Employers' social contributions*	70.582	57.959					11.686	69.645	71.781					-1.199	99%	102%
D.1211	*Employers' actual pension contributions*	22.213	20.401					103	20.504						1.709	92%	
D.1212	*Employers' actual non-pension contributions*	36.543	37.558					186	37.744						-1.201	103%	
D.1221	*Employers' imputed pension contributions*	757						714	714						43	94%	
D.1222	*Employers' imputed non-pension contributions*	11.069						10.683	10.683						386	97%	
D.41A	*Interest*	4.803	2.967					61	3.028						1.775	63%	
P.119C	*Fisim*	-2.253						-2.779	-2.779						526	123%	
D.421	*Dividends*	12.754	6.440					-675	5.765						6.989	45%	
D.422	*Withdrawals from income of quasi-corporations*	1.404							-						1.404	0%	
D.441	*Investment income attributable to insurance policy holders*	4.531						5.130	5.130						-599	113%	
D.442	*Investment income payable on pension entitlements*	34.217						34.087	34.087						130	100%	
D.4431	*Dividends attributable to collective investment fund shareholders*	372						675	675						-303	181%	
D.4432	*attributable to collective investment fund shareholders*	1.163						1.278	1.278						-115	110%	
D.45	*Rent*	4							-						4	0%	
D.6122	*Employers' imputed non-pension contributions*	256						204	204						52	80%	
D.6212	*Social security non-pension benefits in cash*	53.299	55.661					-262	55.399						-2.100	104%	
D.6221	*Other social insurance pension benefits*	39.772	43.225						43.225						-3.453	109%	
D.6222	*Other social insurance non-pension benefits*	11.069						10.683	10.683						386	97%	
D.623	*Social assistance benefits in cash*	20.893	20.907					59	20.966						-73	100%	
D.631	*Social transfers in kind*	122.829															
	... healthcare insurance act	38.777											32.903		5.874		85%
	... longterm care act	17.972												17.962	10		100%
	... rent subsidies	3.339										3.409			-70		102%
	... other	62.741													62.741		
D.721	*Non-life direct insurance claims*	9.386	-					5.236	5.236						4.150	56%	
D.75	*Miscellaneous current transfers*	6.022	559			2.040	118	3.025	5.742						280	95%	

Results

SNA income distribution

Household gross (macro) disposable income adds up to 349.5 bn euros in 2016. The Gini-coefficient for primary income (before redistribution through taxes, social schemes, or private household decisions) comes to 0.526. The effect of these redistributive efforts is that inequality drops to 0.331. The Gini-coefficients are calculated on equivalised incomes, using the oxford modified equivalence scale[13].

Table 7: SNA results

		2015		**2016**	
		mn euros	*Gini*	*mn euros*	*Gini*
B5G	primary income	448,270	(0.528)	461,022	(0.526)
B6G	disposable income	340,419	(0.336)	349,501	(0.331)
B7G	adjusted disposable income	462,398		472,330	

If we cluster households in groups of equivalised macro disposable income and depict the income distribution of primary, disposable, and adjusted disposable income for those groups, the effect of redistribution for each income group is very clear (Figure 2). Primary income and disposable income cross at the 38th percentile, both in 2015 and 2016. The households below (except for the very first percentile) have a higher disposable income than primary income meaning they profit from the redistribution through taxes and social and current transfers. Including the social transfers in kind we find that the households until the 62nd percentile benefit from the redistributions. Both Figure 2 and Figure 3 show that the top percentile receives a large portion of the income distribution. In 2016 the equivalised disposable income of the households in the highest percentile was 2.5 times higher than that of the household in the 99th percentile. Compared to the households in the 50th percentile the factor was 8.1.

[13] The oxford modified equivalence scale assigns a value of 1 to the household head, of 0.5 to each additional adult member and of 0.3 to each child.

Figure 2: Income distribution (2015)

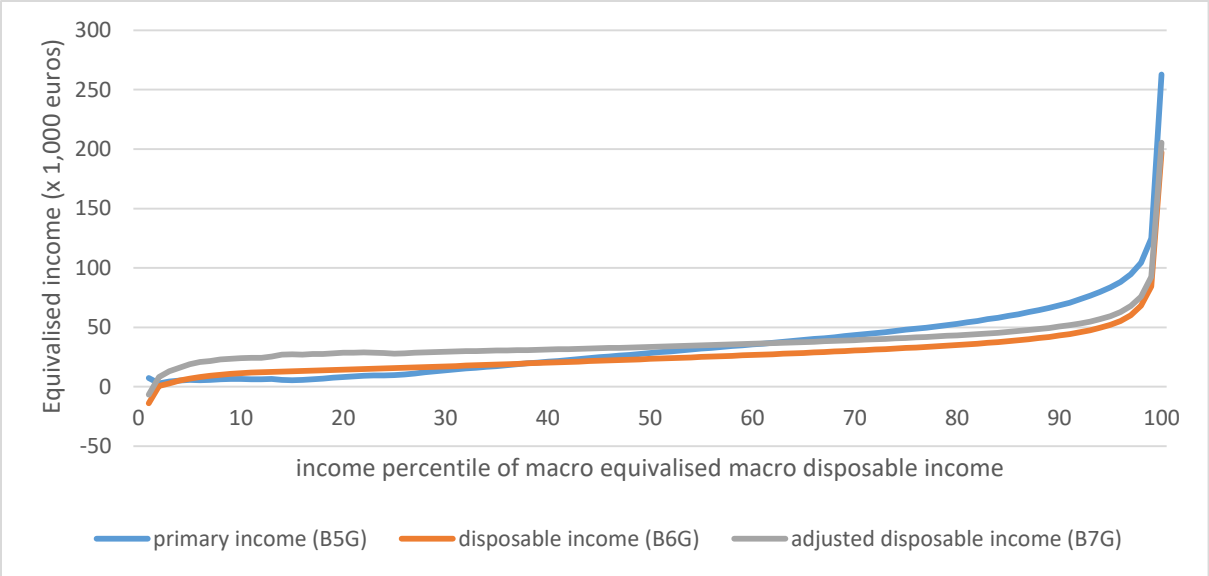
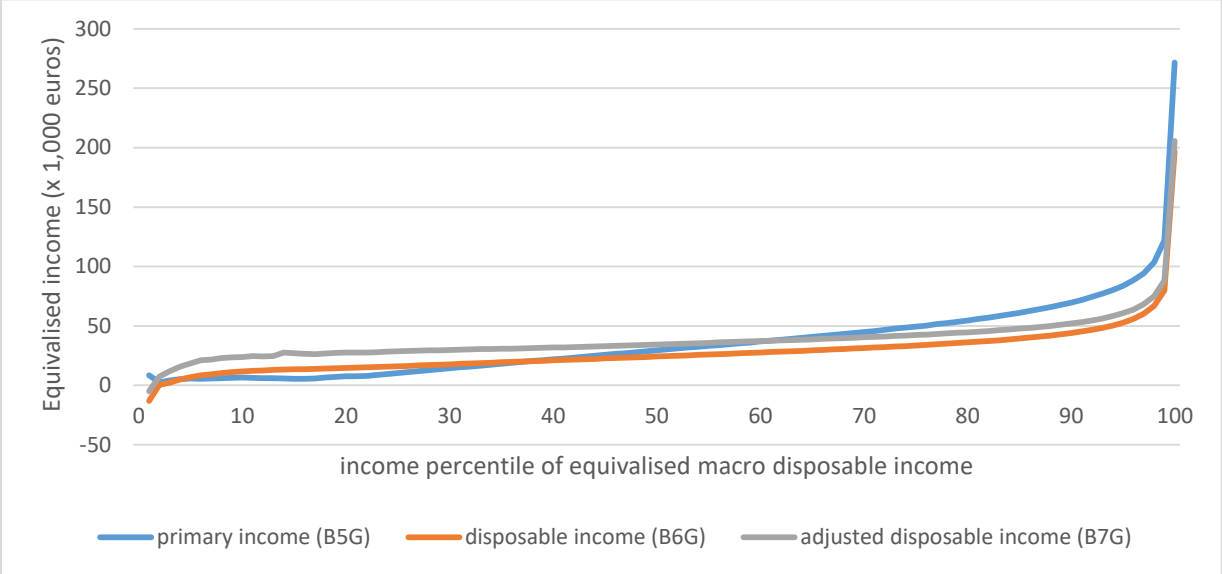


Figure 3: Income distribution (2016)

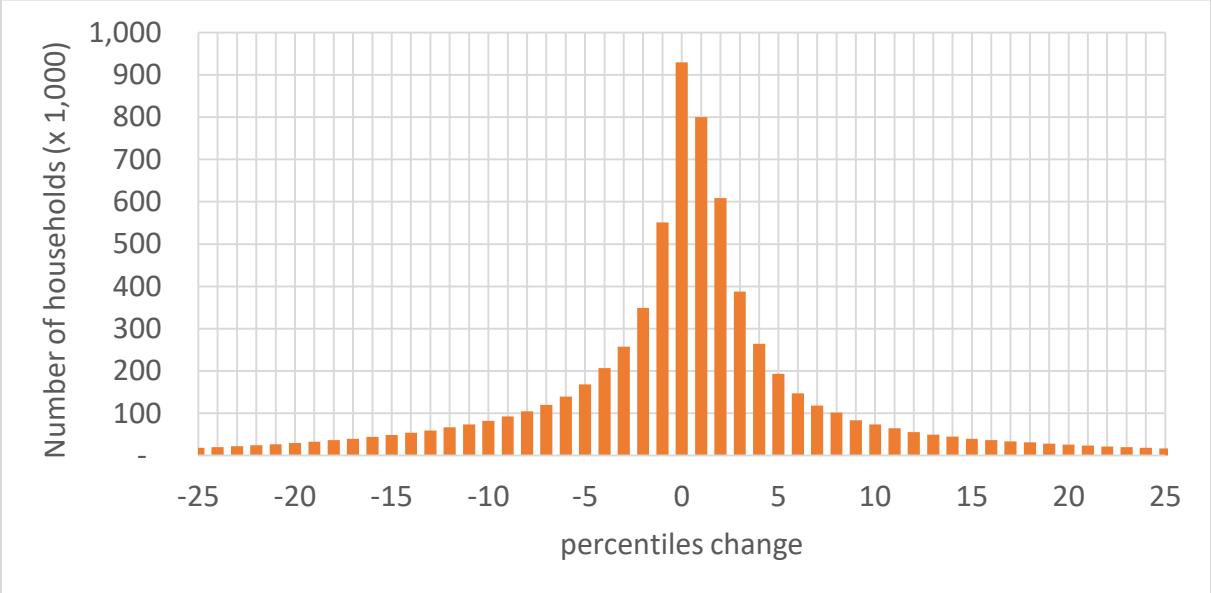


From Figure 2 and Figure 3 it becomes clear that income distributions are quite similar in both years. For the social transfers in kind this is for a large part by construction, because we used the same average values by age for the healthcare insurance act, and the long-term care act. Even though the income distribution is quite similar in 2015 as it is in 2016, the Gini-coefficient drops, more than in the publication of the welfare statistics (Table 2). There is a lot of dynamics within the percentiles, as a result of economic behavior. People losing jobs, self-employed performing better or worse than they did a year ago, students entering the labour market etc. To check the dynamics of households we link our data set of 2016 with the dataset in 2015. An advantage of our approach to create a database of all individuals who can be identified using a unique key, is that we can follow them over time as well. Not all households can be found in both years though, in case households leave the population, or the reference person changes, a match cannot be made. Also it won't be possible to

link new entrants in 2016 to the dataset of 2015. For 7,344,567 households we are still able to link both years together, allowing for an analysis of the income dynamics in the population.

Figure 4 shows these dynamics by presenting the number of households per percentile change in these years. 929.5 thousand households are in the same income percentile in both years, this equals 12.7% of the total number of households that can be identified in both years. Almost two-thirds (64,2%) of this population moves at most five percentiles up or down. The extreme case where households move from the bottom to the top percentile (1,219 households) or the other way around (996 households) occur as well. These dynamics are important to understand economic behavior, such as savings decisions.

Figure 4: Income dynamics in macro disposable income (2015-2016)



Differences between micro and macro disposable income

Disposable income according to the IIWS micro data equals 305,825 mn euros, compared to 349,501 for macro disposable income. The publication of the welfare statistics does not include the total micro disposable income, but this can be derived from the average and the number of households (Table 2). This is lower than the IIWS data source because of the scope of the publication. Moreover for 2015 we used the preliminary version of the IIWS data source, and the publication in Table 2 is final.

Both income concepts are equal in name[14], and focus on the household population. From previous elaborations it is already clear that SNA covers more than the IIWS alone (Table 6). Only a few components of micro disposable income are not part of macro disposable income. These are conceptually excluded, because these components are mostly considered financial transactions in the national accounts. Also rent subsidies is not part of disposable income, but is included in social transfers in kind and thus in the adjusted disposable income concept. Social transfers in kind are neglected in the micro concept, apart from these rent subsidies.

[14] In English both income concepts are disposable, in Dutch a distinction is made.

Table 8: From micro to macro disposable income (2016)

	2016		
Micro disposable income (IIWS)	**305,825**		
Conceptually excluded from macro disposable income	-2,208	+	
Excluded because other data sources is preferred	31,923	-	Operating surplus
	38,638	+	
	322,600	-	Compensation of employees
	340,586	+	
Included in macro, no microdata source	23,178	-	Intermediate use of fisim
	-25,448	-	Fisim (uses)
	-2,253	+	Fisim (resources)
	41,315	+	Other property income
	34,217	-	Supplementary pension contributions
	-9,258	-	Service charges
	9,212	+	Non-observed economy
Included in macro, but in other microdata source	8,144	-	Other current taxes
	2,158	+	Other current transfers
	17,193	-	Other current transfers
Corrections, including integration	3,335	-	Operating surplus
	15,276	+	Mixed income
	-7,568	-	Taxes
	-7,045	+	Compensation of employees
	7,934	+	Net property income
	1,281	-	Other current transfers
	12,691	+	Other current transfers
	7,461	-	Social contributions
	-5,573	+	Social benefits
	3	+	Statistical discrepancy
Macro disposable income	**349,501**		

Furthermore we have seen that for compensation of employees, and the operating surplus a different data source is preferred. For the former this is necessary to comply with the SNA practices for the macro estimate of operating surplus. Also the latter is consistent with SNA concepts already,

but we still need to make integration decisions for the compensation that flows abroad. In 2015 also mixed income would for a large part be overwritten by the SZO, but this is not used in 2016.

These other data source can affect inequality measures as well. For individual records in the LA and the IIWS outcomes can be different as explained, and taxes are adjusted in accordance with these differences. With regard the operating surplus though, the income we attribute to households depends on the approach we use, whether the characteristics of the dwelling are used as the determinants, or the value of the dwelling. As an alternative the operating surplus is estimated using the following regression based upon the value of the dwelling: Monthly Imputed rent = 365.317 + (1.836 * (Value / 1000)). These parameters are consistent with the rental survey - that is also used to estimate the parameters for the characteristics – but not with the IIWS. If households are ranked following their operating surplus only 19% of the 4,1 million households with an owner-occupied dwelling is allocated to the same income decile by both approaches, and there is a rather wide dispersion over the other deciles.

Figure 5: Cross check of operating surplus by two approaches (2015)

		Decile group of operating surplus (characteristics of the dwelling)									
		1	2	3	4	5	6	7	8	9	10
Decile group of operating surplus (value of the dwelling)	1	36%	30%	10%	11%	5%	4%	1%	1%	2%	1%
	2	21%	20%	14%	13%	13%	11%	3%	2%	2%	0%
	3	13%	13%	14%	11%	17%	15%	7%	6%	3%	0%
	4	9%	10%	13%	10%	17%	17%	10%	9%	4%	1%
	5	6%	8%	11%	10%	15%	16%	12%	11%	6%	3%
	6	4%	6%	10%	9%	13%	13%	13%	15%	9%	6%
	7	3%	5%	9%	9%	10%	9%	13%	17%	12%	12%
	8	3%	4%	8%	10%	6%	7%	12%	15%	17%	20%
	9	2%	3%	6%	10%	3%	5%	12%	12%	22%	27%
	10	2%	2%	3%	8%	1%	4%	17%	12%	23%	29%
		100%	100%	100%	100%	100%	100%	100%	100%	100%	100%

There are many transactions in the household sector accounts that have no counterpart in the IIWS. Some of them can be found in another micro data source (HBS, LISS, Giving in the Netherlands), but for others there is no micro data at all. The macro effects on disposable income are limited, because in some cases these flows cancel each other out. The adjustments of other property income consist of the largest part of the income attributed to insurance policy holders, which is also considered to be paid back to the insurer or pension fund in the form of supplementary pension contributions. This is an identity that holds on the macro level, but also on the individual level, therefore inequality will not be affected either.

The sum of the three fisim components is not very high either, however for an individual household this can make a difference. In case a household pays a lot of interest, it also has large fisim expenditures in our approach. These expenditures are shifted towards consumption, thus increasing the disposable income compared to the micro statistics. In case consumption would be included in the comparison, this household would have higher consumption as well, and savings would be equal in both statistics. The same reasoning goes for the service charges.

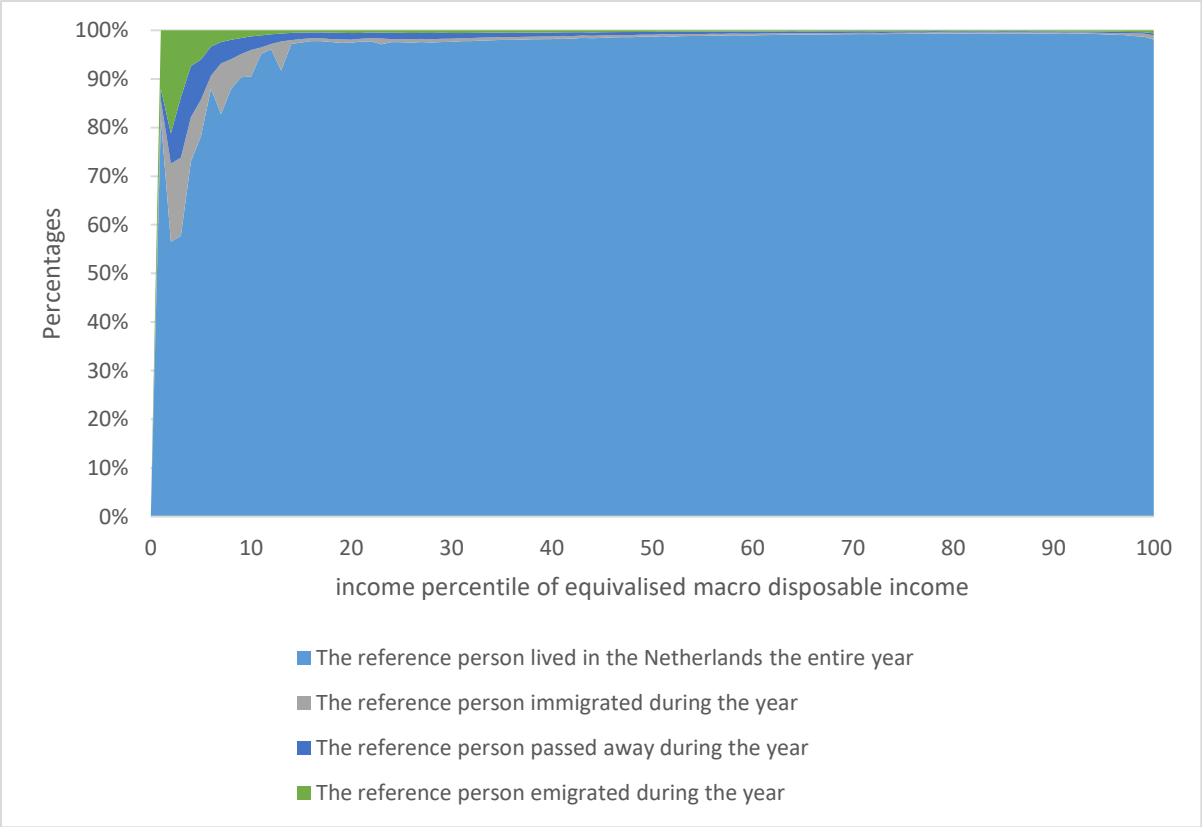
The non-observed economy is not covered by any data source. Our method to distribute the macro estimate over households depends on income, household size and age. As a result we do allocate more income to the lower percentiles (also there the household size is smaller), but apart from the first decile, the distribution is rather flat. The first 10 percentiles receive together around 16% of total non-observed income, the other deciles between 8% and 10%.

The corrections made to the data, either in the construction, or the integration phase, are still substantial. Some of them are the result of the limited data availability in 2016 compared to 2015. Integration of mixed income will be much lower when the SZO is used, but for now we need to correct for interest paid and received and consumption of fixed capital. Not all corrections are a sign of weakness though. Other corrections can be identified quite clearly, such as the compensation of employees flowing abroad. However also there are also integration corrections where our understanding of the allocation is limited. This is true for the other current transfers, where we have only partly information on claims, and underestimate the redistribution through these private insurance schemes, or taxes where recording differs between the micro data source and the macro estimate.

The estimates we make for the population create a difference between the two data sources. First the micro disposable income as published includes only the private households, leaving the institutionalized households out of the scope. This equals 255 thousand households in 2016, mainly elderly people. Because we use the IIWS data source and not the published result, we already include these households. Second, the micro disposable income does not include immigrants and new-borns, which are 393 thousand individuals in 2016. The children are economically inactive, but immigrants are not. The labour accounts show that the compensation of employees is on average 8.5 thousand euros, which adds up to 1.9 bn euros in total. Also other transactions, such as taxes and social contributions, are influenced by this. On total disposable income including the population flows has a minor impact though, 0.6% of macro disposable income is earned by immigrants and new-borns. In this respect these studies can be regarded as the first steps, currently we do not include imputations for property income from immigrants, and hardly any for social benefits received, simply because we do not have any data to start from. Using this approach, including the longitudinal information that we can derive from it, hopefully gives us more insights in these imputed parts of our database.

Still, including households that lived only part of the year in the Netherlands influences inequality measures. These households have by definition a relatively low disposable income. Figure 6 shows that households (considered by the status of the reference person), that were only registered in the Netherlands for part of the year, are mostly in the bottom percentiles. Some of these flows were already included in the IIWS (emigrants, and deaths), we added the immigrants. It must be noted that we also included immigrants and new-borns to already existing households, these will end up elsewhere in the income distribution. In the end 88% of the immigrant households end up in the lower half of the income distribution, 12% in the top half. For emigrant households the same percentages apply.

Figure 6: Households by status of the reference person and income percentile (2016)

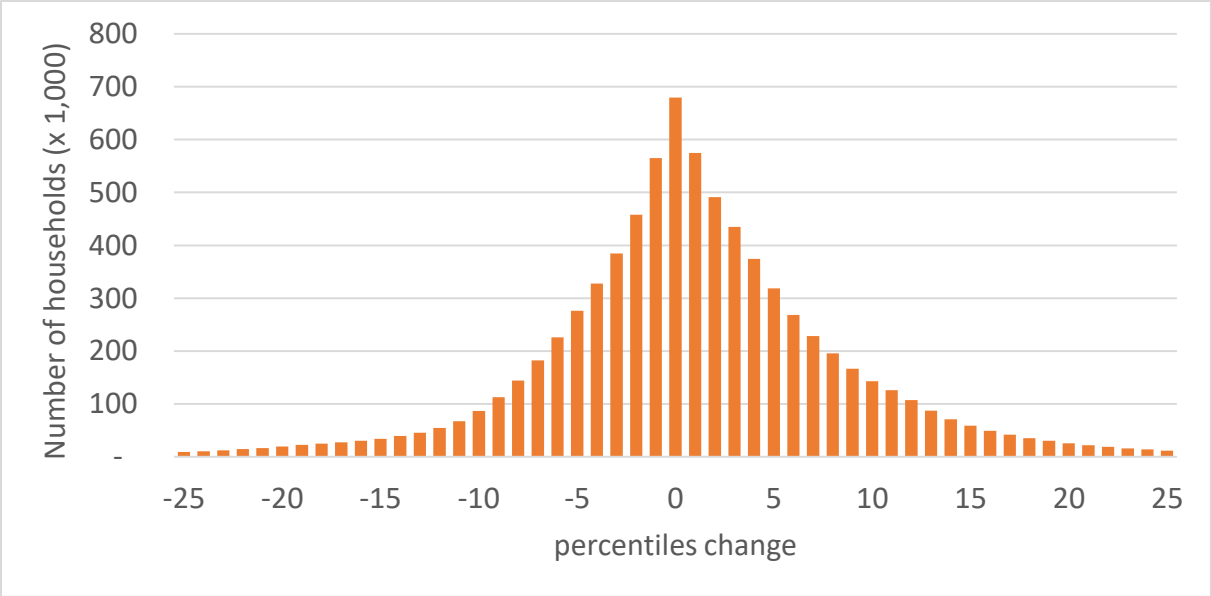


In the end, inequality in macro disposable income (0.331) is higher than in the IIWS micro data (0.304). A last difference that should be noted here is the equivalence scale that is used to account for the economies of scale. The micro data uses the Statistics Netherlands equivalence scale, in this exercise the Oxford Modified Equivalence scale is used. The use of a different equivalence scale has an effect on inequality of 0.007 in both years on the equivalised micro disposable income. This decreases inequality according to the micro data (0.297), but further increases the difference with the macro statistics. If only those household are included that lived in the Netherlands the entire year, inequality is decreases by +0.01 (Table 9).

Table 9: Gini-coefficients

		2015	2016
IIWS	data source	0.320	0.304
	data source (OMS)	0.313	0.297
SNA	B6G	0.336	0.331
	B6G (entire year only)	0.327	0.321

Figure 7: Number of households per percentiles change between micro and macro income percentile



Finally we check the dynamics between the two frameworks, and compare the number of households per income percentile if we rank according to micro equivalised disposable income, and macro equivalised disposable income. 9% of the households are allocated in the same income percentile following both concepts. 61% of the households (2016) are allocated maximum 5 percentiles higher or lower in the macro framework than in the micro data. Remarkable is also that 6,289 households are among the richest 1% according to the IIWS, while they belong to the poorest 1% in macro disposable income. For only 115 households the exact opposite is true.

Summary and conclusion

Using micro data in the national accounts serves foremost the purpose of creating the national accounts. This integrated framework benefits from the combination of data sources from different economic agents, and the work presented in this paper shows there are many data sources to include for the household perspective. Moreover using these large and detailed data sources allows us to answer the increased demand for distributions within the household sector. And with this increased demand for distributions, also the demand for understanding the differences between the micro and macro statistics gained more attention.

The advantages of the use of micro data are not that we have a perfect match between micro and macro disposable income. This is not necessary, from the elaboration on data sources and methodology it is clear that the household sector in the welfare statistics and the macro-economic framework differ in many aspects, even though the key indicators (disposable income) have the same look and feel. Micro macro gaps are not necessarily bad, however limited understanding of the differences makes it hard to present accurate distributions. The benefit of our approach is that - because we included these many micro data sources in the construction and integration process – we have a feedback loop in our process. This improves our understanding of the differences and allows us to make informed decisions on the allocation of these gaps. Moreover it becomes clear where further research is needed, and where additional data are welcomed.

The corrections in Table 8 indicate where there is room for improvement. Apart from income components that lack micro data, also imputations can be improved, for instance those for immigrants. Currently no corrections are made for property income received and paid, because we have no data about this. The experiences with this data set and the possibility to link it with other years might give us more insight in these imputations, and improve our methods for future years.

The resulting income distributions from this exercise shows that the distributions in 2015 and 2016 are rather similar. Inequality, as measured by the Gini-coefficient, decreased in 2016. This is also seen in the welfare statistics. Our methods also allow us to analyse income dynamics. This shows that almost two-thirds of this population moves at most five percentiles up or down between 2015 and 2016. For the other households these movements are larger, which is important to understand, for example when negative savings of low income households are considered.

Finally, if we isolate the link between the micro and macro disposable income, we understand that totals differ because of several reasons. The main reason being that SNA covers more income components than the welfare statistics does. Also because of corrections we make in the construction or integration phase we deviate from micro disposable income. For the income totals this is not a problem, but when the distributions are considered this could be. For parts of these deviations other micro data sources are available, for parts there aren't. In the latter case we are more and more dependent on subjective allocations and as a final resort, proportional allocation. In our methods we combined many data sources, to fill as much of these white spots as possible.

It is clear that including more income components could influence inequality measures as well. But also including more households, and the choice for an equivalence scale does. And even when similar estimates are made (operating surplus), simply by focusing on different determinants might lead to differences on the micro level. This paper focused for a large part on methodology, so as to benefit from the large amount of data that is available, but also to understand how we deviate from it.

References

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

Bekkers, R., & Boonstoppel, E. (2010). *Giving in the Netherlands Panel Survey Codebook.* Amsterdam: VU University, Philantropic Studies.

Bekkers, R., Schuyt, T. N., & Gouwenberg, B. M. (2002-2008). Giving in the Netherlands Panel Survey (GINPS): 2002-2008. Amsterdam: VU University, Philantropic Studies.

Fesseau, M., & Mattonetti, M. (2013a). *A Cross-country Comparison of Household Income, Consumption and Wealth between Micro Sources and National Accounts Aggregates.* Parijs: OECD.

Fesseau, M., & Mattonetti, M. (2013b). *Distributional Measures Across Household Groups in a National Accounts framework - Results from an experimental cross-country exercise on household income, consumption and saving.* Parijs: OECD.

Huigen, R., Van der Stadt, H., & Zeelenberg, K. (1989, September). Socio-Economic Accounts for the Netherlands. *Review of Income and Wealth*(Series 35, Number 3), 317-334.

IMF/FSB. (2009). *The Financial Crisis and Information Gaps. Report to the G-20 Finance Ministers and Central Bank Governors.* http://www.financialstabilityboard.org/wp-content/uploads/r_091029.pdf.

Kazemier, B. (2014). Hidden workers and the hidden worker potential in the Netherlands. *Economic Analysis and Policy*, 39-50.

Plaza, S., Navarrete, M., & Ratha, D. (2011). *Migration and Remittances Household Surveys in Sub-Saharan Africa: Methodological Aspects and Main Findings.*

Stiglitz, J. E., Sen, A., & Fitoussi, J.-P. (2009). *Report by the Commission on the Measurement of Economic Performance and Social Progress.* Retrieved from www.stiglitz-sen-fitoussi.fr.

Timmerman, J. G., & Van de Ven, P. J. (1994). *A Social Accounting Matrix for the Netherlands: concepts and results.* Voorburg: Statistics Netherlands.

Zwijnenburg, J. (2017). Unequal distributions? A study on differences between the compilation of household distributional results according to DINA and EGDNA methodology. *Paper prepared for the First WID.world Conference*, (p. 37). Paris.