EUROPEAN CENTRAL BANK

EUROSYSTEM

Working Paper Series

Janina Engel, Pau Gayà Riera,
Joseph Grilli, Pierre Sola

Developing reconciled quarterly distributional national wealth – insight into inequality and wealth structures

No 2687 / July 2022

Abstract

Distributional accounts for households enable measurement, study developments and identify drivers of inequality. Distributional information on households' wealth is available from the Household Finance and Consumption Survey only for three points in time (2009 – 2018), while aggregates are available quarterly. This paper presents a novel methodology for deriving quarterly distributional national wealth by (i) improving the alignment of survey fieldwork periods with the national accounts' dates; (ii) correcting for differences in several concepts; (iii) estimating missing wealthy households; (iv) developing time series; and (v) computing euro area aggregates. This paper finds an increase in the net wealth Gini of most euro area countries since 2009; that the richest 1% holds 28% of total net wealth, while the bottom 50% holds 4%; and that the net wealth of the top 1% has grown by almost 50%, compared to 28% for the remaining 99%, with a decrease in the bottom 20%.

Keywords: Wealth inequality, HFCS, National Accounts, Pareto distribution

JEL Classification: C46, D31, E27, G51, N34

Non-technical summary

Policymakers have repeatedly emphasised the need for granular household distributional information, as highlighted, for example, by the G-20 Data Gaps Initiative, which was launched in 2009. Such distributional information would facilitate granular analyses to better understand the economic dynamics of the household sector and the effect of policies and regulations on wealth inequality. This paper contributes to this open request by providing a method for the construction of reconciled quarterly distributional national wealth (DNW).

Distributional accounts are commonly derived from micro-economic survey data that are matched with macroeconomic aggregates. DNW is based on the Household Finance and Consumption Survey (HFCS), for which three surveys (or waves) were conducted during the period 2009-18, providing information on the assets, liabilities, income and consumption of households of 20 participating EU Member States. To derive DNW, the HFCS is matched with national accounts (NtlA) statistics, which provide quarterly aggregates of indicators from financial accounts (FA) and non-financial assets. Essentially, NtlA aggregates are assumed to correctly describe the instrument totals and are distributed across household sub-groups based on the HFCS, which captures the heterogeneity within the household sector.

The paper begins by outlining the process of linking the HFCS and NtlA, as derived by the Expert Group on Linking Macro and Micro Data for the Household Sector (EG-LMM). The linked wealth components commonly reveal substantial discrepancies between the HFCS and the NtlA aggregates. Building on that linking process, a novel model is presented to construct micro-economic time series for distributional accounts that match the quarterly NtlA aggregates and address shortcomings in the survey data. First, the linking of the HFCS to quarterly NtlA data is improved by comparing the time frame of the survey to the end quarter dates of the NtlA. Furthermore, non-financial assets in NtlA and in the HFCS are not aligned due to the scarce availability of household data for those assets. This is explained, and an estimation model to align them is presented. Related literature has shown that wealthy households are less likely to be covered by survey data (including the HFCS), resulting in a severe underestimation of the upper tail of the wealth distribution (and likewise inequality). In order to capture the missing wealthy, a Pareto distribution is fitted to the upper tail of the net wealth distribution enhanced with publicly available rich lists, following Vermeulen (2018). The paper presents a novel approach to sampling the missing wealthy households from the fitted Pareto distribution.

The resulting population is then matched with the NtlA population, which also includes persons living in institutions. While these adjustments eliminate some of the discrepancies between the HFCS and NtlA based on identified issues, some differences still remain. A proportional allocation is applied to reconcile the adjusted micro data with NtlA aggregates. This provides three waves of reconciled DNW for the period 2009-18. With a view to increasing the frequency of DNW, an inter- and extrapolation model is presented that enables the computation of quarterly reconciled DNW. The resulting distributional accounts time series for each euro area country can subsequently be aggregated to compile quarterly euro area DNW.

Finally, the derived quarterly reconciled DNW enables an assessment of inequality over time. Our results show that net wealth inequality has increased in the euro area over the last decade, which corroborates findings of related literature. Furthermore, we find evidence that this development is largely attributable to the heterogeneity in the portfolio composition of different net wealth quantiles. While the portfolios of wealthy individuals are composed mainly of equity, which have risen over the last decade, the portfolios of individuals from the middle and bottom of the distribution are dominated by housing wealth and deposits, together with high leverage. In addition, we observe that inequality has not changed homogeneously across euro area countries, which can be explained to a large extent by the different price developments of the different asset classes held by each category of households. Finally, our analysis reveals that the evolution of each decile of the net wealth distribution follows a cyclical pattern, mirroring business cycle movements, with the exception of the bottom 20%, for which a stagnant pattern is observed. Net wealth growth in the household sector seems to be largely concentrated at the top of the distribution.

1 Introduction

Following the 2008 financial crisis and the pursuit of unconventional monetary policy in the years that followed, the distribution of wealth among households and the composition of their portfolios has come under increased scrutiny. This has led to demands from academics and policymakers for timely distributional data on household wealth in order to better assess the heterogeneity in the household sector. In particular, the G20 Data Gaps Initiative (Recommendation II.9) (see FSB and IMF, 2009, 2019) encouraged countries to compile distributional data including wealth distributions.

The research on the impact of wealth inequality has explored topics such as monetary policy (see Lenza and Slačálek, 2018; Dossche et al., 2021), consumption and saving (see Arrondel et al., 2015; Carroll et al., 2014; Kavonius and Honkkila, 2013), real estate prices (see Mathä et al., 2018), price levels (see Adam and Zhu, 2016) and government debt (see Vogel, 2014). The seminal work by the World Inequality Database (WID) (see Piketty and Zucman, 2014; Blanchet et al., 2017, 2019) established the baseline models used to construct distributional data and the standard measures and terminology for modern literature on inequality. Results have since been produced by the Federal Reserve System (FED) (Batty et al., 2020), the Organisation for Economic Co-operation and Development (OECD) (Zwijnenburg et al., 2017), the European Statistical Office (Eurostat) (Törmälehto, Veli-Matti, 2019), and other bodies such as Statistics Netherlands (Bruil and Koymans, 2014) and Credit Suisse (Credit Suisse Research Institute, 2019).

Despite this growing interest, the work still faces limitations owing to the low frequency of distributional data, the lack of detailed information on the very wealthy and differences between national aggregates and survey data. This paper aims to address these short-comings by constructing distributional national wealth (DNW) that captures the missing wealthy and reconciles survey data with quarterly national aggregates to provide a distributional measure of wealth in the economy at a quarterly frequency.

The construction of DNW is based on two main data sets: (i) the Household Finance and Consumption Survey (HFCS), providing micro data; and (ii) national accounts (NtlA), providing macro data.[2] The HFCS, launched in 2007, collects household-level data on households' finances and consumption and is administered by the Household Finance and Consumption Network

[2]Throughout this paper, all figures from the HFCS and NtlA refer to nominal amounts, which means the effects of price changes over time are not removed.

(HFCN), see, for example, Household Finance and Consumption Network (2020a).[3] The survey is based on more than 91,000 interviews conducted in all euro area countries, as well as Croatia, Hungary and Poland, and to date, three waves of the survey have been completed.[4] The fieldwork of wave 1 took place mostly in the period 2010-11, wave 2 in the period 2013-15, and wave 3 in 2017. By contrast, NtlA are derived from a variety of sources, including in particular counterpart information from financial institutions (e.g. banks reporting the deposits collected from resident households) and are published at a quarterly frequency for most countries, starting from the first quarter of 1999.

All national financial and non-financial accounts data are covered by the European System of Accounts (ESA 2010) and the ECB Guideline on financial accounts.[5] Based on that, and other sources, the ECB and Eurostat compile the integrated accounts for the euro area and for the EU, including results for the euro area household sector.[6]

While survey data provide distributional data on the household sector, these data often differ from those reported in NtlA for a variety of reasons.[7] The Expert Group on Linking Macro and Micro Data for the Household Sector (EG-LMM) (see EG-LMM, 2020) outlined conceptual comparability between measures of instruments in the HFCS and the NtlA. Where similarities are found, the EG-LMM developed a linking of HFCS wealth instruments to their respective counterparts in the NtlA. The linking covers the following items that in sum constitute the considered concept of net wealth: (i) financial assets, comprising debt securities, deposits, investment fund shares, life insurance and voluntary pensions[8], listed shares and financial business wealth (i.e. unlisted shares and other equity); (ii) non-financial assets comprising non-financial

[3]The HFCS micro data are publicly not available. However, researchers can request access here: `https://www.ecb.europa.eu/pub/economic-research/research-networks/html/researcher_hfcn.en.html`.

[4]For some countries, only one or two waves of the survey have been conducted, see, for example, Household Finance and Consumption Network (2020b).

[5]Guideline 2014/3/EU of the European Central Bank of 25 July 2013 on the statistical reporting requirements of the European Central Bank in the field of quarterly financial accounts (recast) ECB/2013/24) (OJ L 2, 7.1.2014, p.34).

[6]For further information on national accounts, see, for example, OECD (2017) and the ECB's quarterly household sector report. Regular updates of the quarterly report can be found at `https://sdw.ecb.europa.eu/reports.do?node=1000004962`.

[7]A full discussion on the differences between wealth statistics compiled in macro statistics and NtlA, as well as the reasons for these differences, is presented in OECD (2013).

[8]Other accounts receivables are excluded.

business wealth and housing wealth;[9] and (iii) liabilities, which they suggest to be split into mortgages and other liabilities in future works (EG-LMM, 2020). Batty et al. (2020) follow a similar linking approach, augmented with estimated distributions, to construct DNW for the FED.[10] This paper builds on the linking of HFCS and NtlA derived by the EG-LMM and follows their wealth concept. As in EG-LMM, 2020, items of low conceptual comparability between the HFCS and NtlA, or items that are not covered by the HFCS, are excluded from the DNW approach. This concerns "pension entitlements", "non-life insurance technical reserves", "other accounts receivable" and "banknotes and coins" .[11] Overall, the DNW covers the most significant instruments of households' net wealth, i.e. more than 90% of the value of households' financial and non-financial assets as recorded in the NtlA.

However, even with the linking process, HFCS instrument totals do not equal the respective NtlA aggregates. One of the main causes of this is differential non-response, which refers to the observation that very rich households are less likely to be sampled and to respond to wealth surveys, leading to an incomplete top tail of the wealth distribution (Vermeulen, 2016, 2018; Piketty et al., 2006; Chakraborty and Waltl, 2018; Chakraborty et al., 2019). Vermeulen (2018) and Chakraborty and Waltl (2018) show how survey data can be corrected for the bias stemming from differential non-response using rich list observations (e.g. Forbes) to estimate a Pareto distribution at the top of the wealth distribution, thereby enabling an imputation of the missing wealthy households. Augmenting the survey data with rich lists gives more density to the upper part of the tail, suggesting more rich households than what would be inferred from a regression on the survey data alone. Cantarella et al. (2021) use this estimation approach to adjust HFCS household weights and to add a single observation representing the non-responding rich households, before matching the HFCS wealth instrument totals to the NtlA. This paper builds on this work and the results achieved so far by the EG-LMM.

[9] While luxury durable goods could be considered a form of wealth, they are not included in the wealth concept presented here. As part of the response to the G20 Data Gaps Initiative, the distribution of consumption goods, including durable goods, is investigated in the OECD distributional national accounts. Furthermore, these goods are recorded as consumption in NtlA, meaning that treatments made for assets, such as depreciation adjustments are not applied. We therefore consider these high value goods as part of the distribution of consumption and so omit them from the measurement of wealth assets.

[10] The FED refers to their distributional data as distributional financial accounts. As that approach also covers the concepts of housing and non-financial business, we have taken the name of distributional national wealth instead.

[11] The NtlA concept of "non-life insurance technical reserves" would include some expected reimbursements by insurance corporations to households, as compensation for some damage or illness, for example. "Other account receivables" would cover claims related to, for example, purchases that have already been paid for but not yet delivered. Most households would not regard these instruments as part of their individual wealth anyway.

Among those households that do respond to the wealth surveys, under-reporting behaviour has been proposed as a possible reason for the gap between the survey totals and the national aggregates (Neri and Ranalli, 2012; Meriküll and Rõõm, 2021). Cussen et al. (2018) investigate the role of under-reporting in survey data for deposits, finding that there is no evidence that the level of under-reporting varies across wealth groups or demographic groups, and so could be treated uniformly across the population. Meriküll and Rõõm (2021) use administrative data linked to the HFCS in Estonia to investigate the role of item non-response in wealth surveys. They find that non-response contributes to the underestimation of wealth inequality in the survey data. This results in downward biases at the top of the wealth distribution, affecting the Gini coefficient and also the top wealth shares. They conclude that, in Estonia, these results are driven by item non-response rather than unit non-response, but that imputing these values does not eliminate the bias attributable to the missing data at the top of the distribution.

Finally, wealth surveys are usually conducted to gain insights into household characteristics, behaviour and valuations. By contrast, the NtlA are a measure of the total values of instruments in an economy at regular time intervals. This can lead to differences in the conceptual definition and reference period of the valuation of instruments, resulting in discrepancies.

In addition to the issue of closing the discrepancies between the HFCS and NtlA, it is desirable to provide distributional information at a much higher frequency than the availability of the pure survey data, which is conducted only every 3-4 years, and at a more granular level than the NtlA, which provide only instrument aggregates. Such granular and timely data sets could provide higher frequency measures of wealth inequality and enable a much more timely analysis of the impact of monetary policy on inequality, even in absence of readily available administrative sources of household-level data. In this paper, we propose a method for constructing reconciled quarterly distributional national wealth (DNW), building on existing models and proposing novel approaches to addressing the current limitations of reconciled accounts.

Based on the derived DNW, we find that the net wealth Gini coefficient of 14 euro area countries has increased since 2009. The results also show a significant concentration of wealth at the top of the distribution, with the top 1% of households in the euro area holding 28% of total net wealth. The top 10% is found to hold 58% of total net wealth, while the bottom 50% only hold 4%, with the total net wealth of the bottom 20% decreasing since 2009. This disparity over the distribution is found not only in the levels, but also the rate of growth, with the total net wealth of the top 1% growing by almost 50% since 2009 compared to only 28% for the remaining

99% of the distribution.

In this paper, Section 2 proposes solutions for the main issue: e.g. (i) a novel approach to estimating the missing wealthy households is presented; or (ii) an inter- and extrapolation model is developed to derive quarterly time series of DNW. The derived reconciled time series of DNW enables a detailed analysis of the wealth composition, inequality and development over time. Results of such analyses are provided in Section 3, offering an insight into euro area distributional information and the evolution and drivers of inequality over the past decade. Finally, Section 4 concludes.

2 Methodology for deriving reconciled quarterly DNW

This section provides a detailed explanation of all the approaches that have been developed in order to obtain quarterly DNW time series that are consistent with NtlA figures. First, we consider the matching of the timing of the observations in the HFCS with the corresponding NtlA reference dates (Section 2.1) and a consistent treatment of the instruments of housing wealth and business wealth in the HFCS and NtlA (Section 2.2). Subequently, a solution is presented to impute and represent missing wealthy households in the DNW (Section 2.3). Afterwards, differences in the population coverage of both data sources are addressed (Section 2.4). Lastly, the remaining discrepancies between NtlA and the HFCS are eliminated via a proportional allocation (Section 2.5). Having constructed reconciled DNW for the quarters for which the HFCS was conducted, we apply an inter- and extrapolation method to derive a quarterly time series of reconciled DNW (Section 2.6). Using the time series data, the national results of euro area countries can be aggregated to yield quarterly reconciled euro area DNW (Section 2.7). Finally, the impact of the addition of wealthy households and the proportional allocation method on inequality measures is shown (Section 2.8).

2.1 Linking HFCS fieldwork periods with NtlA reference dates

The first step consists in an improved matching of the timing of each of the three HFCS waves to the corresponding quarter of the NtlA. For some survey observations, this is straight-forward as the values are reported for a given date that is covered by the NtlA. In the majority of cases however, the valuations are computed by households on the date of the survey, which may be any day within the fieldwork period (i.e. the time period over which the survey is conducted).

In such situations, the mid-point of the fieldwork period has been calculated and matched to the closest quarter in the NtlA.

Figure 1 shows the fieldwork periods for the HFCS in each country, which vary in starting date and duration, and the corresponding reference quarter in the NtlA. While some countries' surveys span a short period of time, meaning that the data can be attributed to a single quarter of the NtlA, for other countries, the survey periods are much longer.

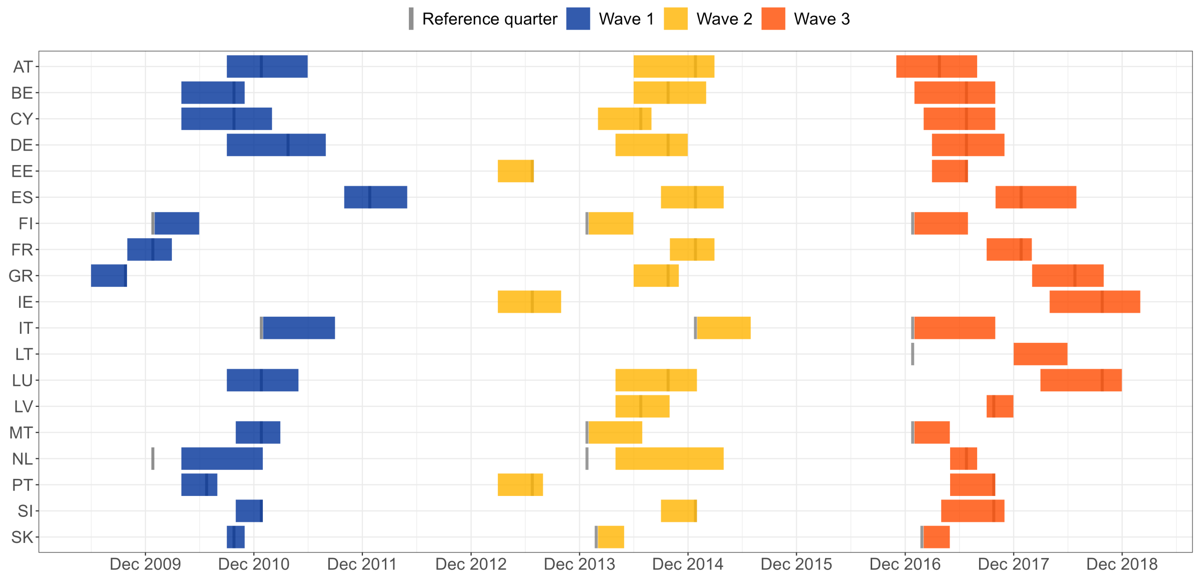

Figure 1: Gantt chart of fieldwork periods of HFCS waves by country.
Notes: The reference quarter refers to the quarter of the NtlA to which the HFCS is matched. Instances where the reference quarter falls before the fieldwork period refer to countries using a valuation date approach, meaning that the survey questions explicitly refer to the specified reference quarter. Annex A explains and analyses the different valuation methods.

2.2 Consistent treatment of business and housing wealth

Conceptual differences between the HFCS and the NtlA contribute to the discrepancies between micro and macro statistics. While the HFCS is conducted only on households, in the NtlA, the non-financial assets are generally recorded as an aggregation of households and non-profit institutions serving households (NPISH). In addition, the value of land owned by this pooled sector of households and NPISH is available only as a total, i.e. the separation of land underlying dwellings and other land is not generally available.[12] The application of the following method

[12]On a voluntary basis, some countries report the split between households and NPISH and the split of land underlying dwellings and other land. However, this is limited to only a few cases.

aims at improving the comparability of the two data sets, more than improving the coverage ratios of housing and business wealth.

2.2.1 Housing wealth

The concept of housing wealth is defined as dwellings owned by households plus the land underlying those dwellings, with Figure 2 showing the coverage ratios for housing wealth using the available information in the NtlA.[13] However, the NtlA cover dwellings and the value of total land owned by the sum of households and NPISH. In addition, housing abroad is not recorded as a non-financial assets; instead, in line with the standard methodology (as described in the ESA 2010) it is recorded as unlisted shares. Thus, housing abroad is part of the NtlA instrument of unlisted shares and other equity.[14] We therefore propose adjustments to be applied to the value of housing wealth in NtlA to more closely align this value with the concept used in the HFCS.

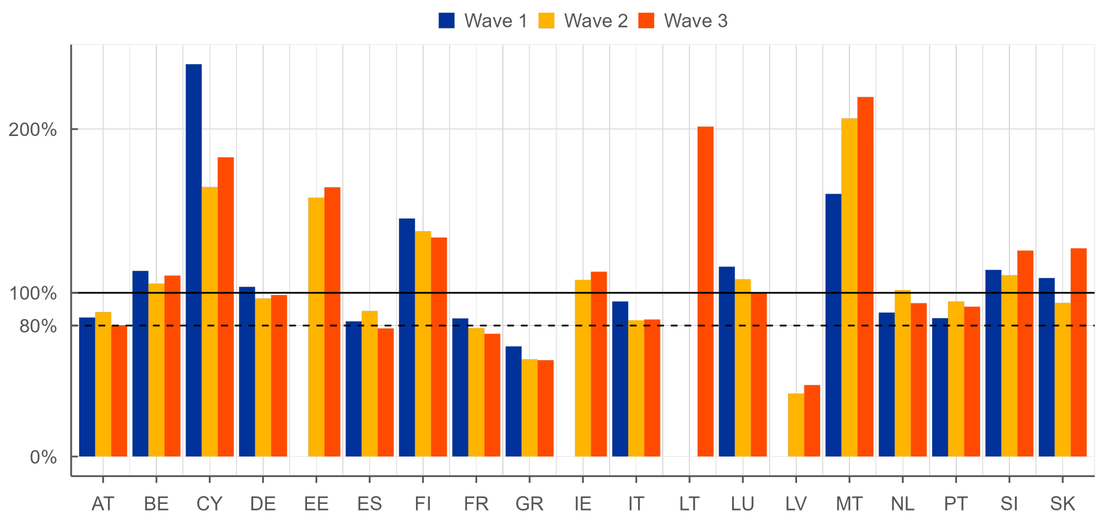

Figure 2: Housing wealth coverage ratio by wave and by country

We can decompose housing wealth in our micro and macro data sources and compare them as follows, where the gap denotes the difference between the implied estimate of total HFCS in the DNW micro data set and the total housing wealth in the NtlA. Let $S.14$ denote households

[13] A coverage ratio is defined as the implied estimate of total HFCS divided by the NtlA total for a specific instrument. In other words, the proportion of NtlA that is reported in the survey data.

[14] Properties abroad are recorded in NtlA as if the household has set up an unlisted company and the company is the owner of the property.

and $S.1M$ denote the aggregation of households and NPISH, then:[15]

$$
\begin{aligned}
&\text{Housing wealth}_{HFCS} + \text{Housing wealth}_{Wealthy\ households} + \textbf{Gap} \\
&= \underbrace{\frac{\text{Dwellings}_{S.14}}{\text{Dwellings}_{S.1M}}}_{a} \text{Dwellings}_{S.1M} \\
&\quad + \underbrace{\frac{\text{Land}_{S.14}}{\text{Land}_{S.1M}}}_{b} \underbrace{\frac{\text{Land underlying dwellings}_{S.14}}{\text{Land}_{S.14}}}_{c} \text{Land}_{S.1M} \\
&\quad + \underbrace{\frac{\text{Housing abroad}}{\text{Real estate abroad}}}_{d} \text{Real estate abroad.}
\end{aligned} \tag{1}
$$

For more details on the wealthy households see Section 2.3. The parameters represent the share of each of the available assets that can be attributed to households in the case of a and b, the share of the land that is underlying dwellings for the case of c, and finally, the share of total real estate abroad that can be attributed to housing wealth in the case of d.

In order to estimate the four parameters, it is necessary to take a set of assumptions and constraints to ensure that 1 is solvable and the results are plausible.

Using data available on a voluntary basis from different countries (i.e. split of households and NPISH, split of land underlying dwellings and other land, and split of housing and other real estate abroad) we can assume the following:

- *Assumption I*: the share of dwellings a and the share of total land b are linearly related. Based on data from several euro area countries, we assume that $b = 0.99a$.

- *Assumption II*: the share of land underlying dwellings over total land c can be estimated as the weighted average of the euro area countries that report this split for the sector households plus NPISH. This share is assumed, on average, to be $c = 0.86$.[16]

- *Assumption III*: the share of housing over total real estate abroad d is assumed to be the average of the euro area countries that report the split. This share is assumed, on average, to be $d = 0.85$.[17]

- *Constraint I*: compared with households, NPISH represent a small fraction of most economies. Based on available data, we can assume that $a \in [0.96, 1]$.

[15] Following the nomenclature defined in the ESA 2010.

With the assumptions and constraint defined above, we can solve equation 1 using a least squares method to minimise the gap for all available waves, country by country. With the assumptions taken, the problem simplifies to solving it for a.[18]

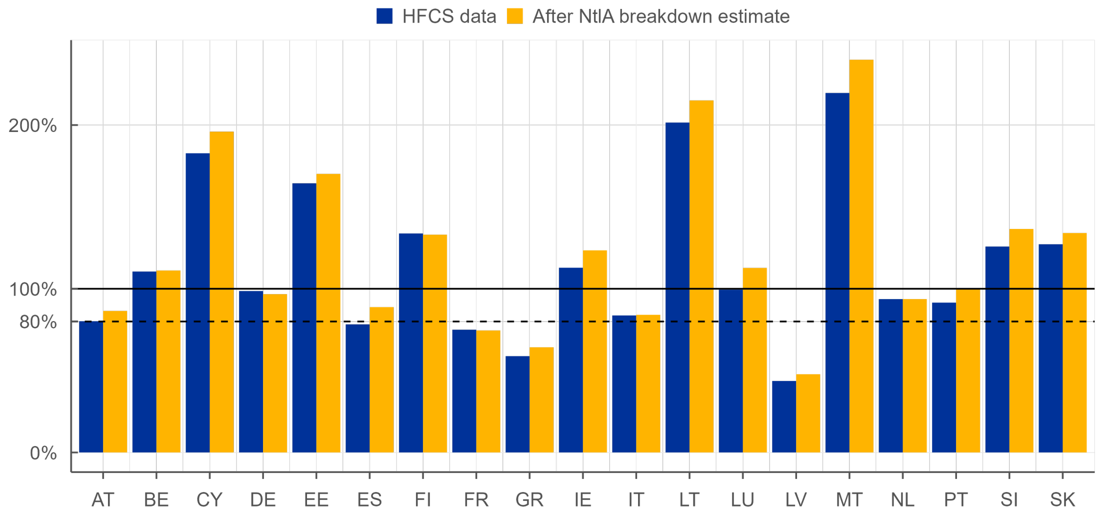

Figure 3: Housing wealth coverage ratio by country.
Note: the data refers to wave 3 (which covers broadly the period 2016 Q4 - 2018 Q3).

Figure 3 shows the effects of the refinements.[19] The blue bars show the ratio between the pure HFCS and the available NtlA values, while the yellow bars include the estimation of the missing breakdowns made to more closely align the definitions. In general, this adjustment noticeably increases the coverage ratio, with only Germany in Figure 3 showing a significant decline. These increases in the coverage ratios are mainly attributable to the estimation of the split of land underlying dwellings and other land, which has a sizeable impact if the value in the NtlA side is reduced. These adjusted values for the NtlA are used going forward as they more accurately align to reporting by households in the micro data underlying the distribution.

2.2.2 Business wealth

The concept of business wealth is defined as the value of financial business wealth (corresponding to unlisted shares and other equity) and non-financial business wealth (comprising all non-

[16] The average share for the reporting countries varies from 0.76 to 0.89.

[17] The average share for the reporting countries varies from 0.67 to 1. However, the total amount of real estate abroad represents, on average, less than 1% of housing wealth.

[18] If the solution given by the resulting linear equation does not satisfy the constraint i.e. $a \notin [0.96, 1]$, the resulting a will be a corner solution i.e. $a = 0.96$ or $a = 1$.

[19] In order to provide a better understanding of the method, a numerical example is available in Annex C.

financial assets covered by NtlA with the exception of housing wealth). As with housing wealth, the lack of a split between households and NPISH, together with little to no information on the split of land underlying dwellings and other land, requires some estimation from our side in order to match the concepts in the HFCS and the NtlA.

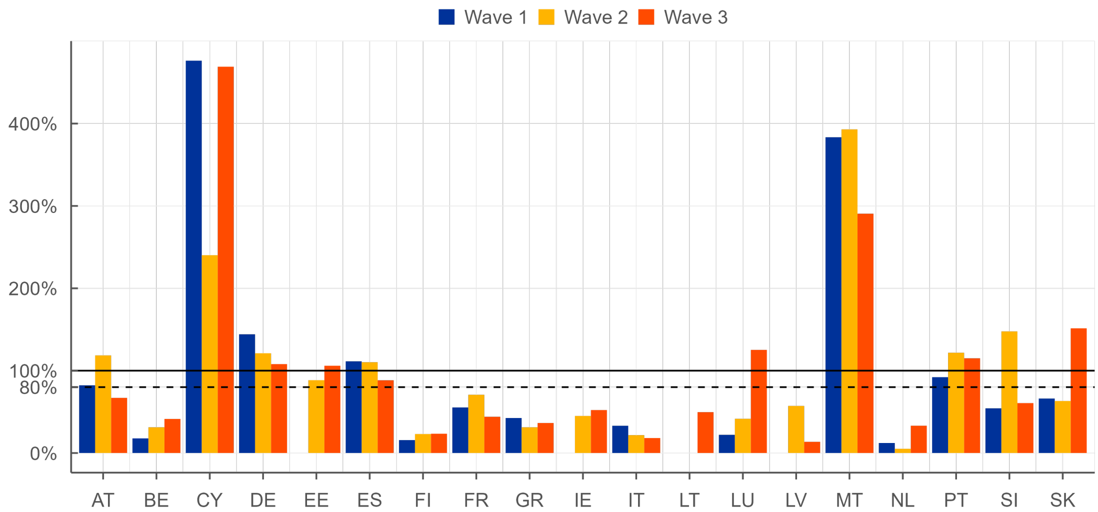

Figure 4: Business wealth coverage ratio by country

Figure 4 shows the coverage ratios for business wealth using only the information available in the NtlA. In the same way we did for housing wealth, we can decompose business wealth in our micro and macro data sources and compare them as follows, where the gap denotes the difference between the implied estimate of total HFCS in the DNW micro data set and the total business wealth in the NtlA:

$$
\begin{aligned}
&\text{Business wealth}_{HFCS} + \text{Business wealth}_{\textit{Wealthy households}} + \textbf{Gap} \\
&= \underbrace{\frac{\text{Non–financial assets}_{S.14}}{\text{Non–financial assets}_{S.1M}}}_{e} \text{ Non–financial assets}_{S.1M} \\
&\quad - \underbrace{\frac{\text{Dwellings}_{S.14}}{\text{Dwellings}_{S.1M}}}_{a} \text{ Dwellings}_{S.1M} \\
&\quad + \underbrace{\frac{\text{Land}_{S.14}}{\text{Land}_{S.1M}}}_{b} \left(1 - \underbrace{\frac{\text{Land underlying dwellings}_{S.14}}{\text{Land}_{S.14}}}_{c}\right) \text{Land}_{S.1M} \\
&\quad + \text{Unlisted shares and other equity}_{S.14} \\
&\quad - \underbrace{\frac{\text{Housing abroad}}{\text{Real estate abroad}}}_{\text{d}} \text{ Real estate abroad.}
\end{aligned}
\tag{2}
$$

Section 2.3 provides more details on the wealthy households. Parameters a, b, c, and d are those estimated above in equation 1. Parameter e represents the share of total non-financial assets that can be attributed to households. In order to ensure plausible results, we take into account the following additional two constraints:

- *Constraint II*: compared with to households, NPISH constitute a small fraction of most economies. Based on available data, we can assume that $e \in [0.95, 1]$.
- *Constraint III*: as a non-financial asset, dwellings are a component of the total. Thus, total non-financial assets have to be at least as large as total dwellings.

With the assumptions and constraints defined above, we can solve equation 2 using a least squares method to minimise the gap for all available waves, country by country. With the assumptions taken, the problem boils down to solving it for e.[20]

[20] If the solution given by the resulting linear equation does not satisfy the constraints i.e. $e \notin [0.95, 1]$, the resulting e will be a corner solution i.e. $e = 0.95$ or $e = 1$.

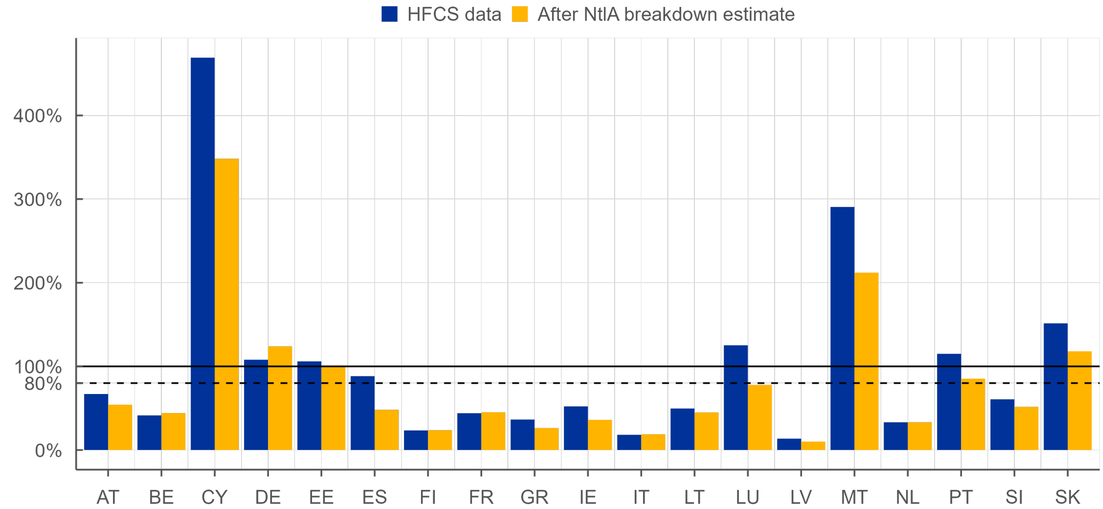

Figure 5: Business wealth coverage ratio by country.
Note: the data refers to wave 3 (which covers broadly the period 2016 Q4 - 2018 Q3).

Figure 5 shows the effects of these refinements. The blue bars show a simple comparison between the pure HFCS and the available NtlA values, while the yellow bar include the estimation of the missing breakdowns. By making these adjustments, we more closely align the concepts in the HFCS and NtlA. The largest effect on the coverage ratios is the result that land not underlying dwellings, which was removed from housing wealth, is now correctly allocated to non-financial business wealth, increasing the NtlA totals.

To sum up, the aforementioned methodology uses the scarce data available to present a set of plausible assumptions and constraints in order to eliminate the small conceptual differences between the HFCS and NtlA and improve the comparability of the two data sets.

2.3 Capturing missing wealthy households

The literature identifies differential non-response as a contributing factor to the discrepancies between the instrument totals of the HFCS and the NtlA: the wealthiest households are generally not covered in the HFCS, leading to an underestimation of the upper tail of the wealth distribution. Vermeulen (2018) demonstrates how the survey data tend to lack observations for wealthy households, such as those that are covered in public lists of wealthy individuals, resulting in missing density in the top tail of the wealth distribution. Schröder et al. (2020) analyse the implications of differential non-response by conducting a special survey that covers, in particular, wealthy German households. The authors conclude that wealth surveys that suffer from differential non-response lead to a significant underestimation of inequality, with the Gini coefficient increasing by four percentage points when additional data on rich households are used.

A wide stream of economic literature advocates that the top tail of wealth distributions can be represented by a Pareto distribution (see Atkinson, 2017; Blanchet et al., 2018; Lakner and Milanovic, 2016; Piketty and Saez, 2003; Piketty, 2003; Ruiz and Woloszko, 2016). On that basis, the missing wealthy households can be estimated based on a Pareto distribution fitted to the available data. Since the HFCS generally does not fully cover the wealthiest households, estimating the Pareto distribution solely on the basis of the HFCS data leads to an underestimation of the upper tail. Vermeulen (2018) shows that there is already a significant improvement in the estimation of the Pareto distribution when a few very wealthy observations are added. Such observations are mostly available from "rich lists", e.g. Forbes World's Billionaires List. Rich lists are an important input factor for the estimation of the Pareto distribution, and hence the imputation of missing wealthy households. Table 1 provides an overview of available rich lists for euro area countries. The length and quality of the rich lists varies substantially across countries and years, with some using media sources and others using national credit and business registers to compute the lists. This can have an impact on the Pareto distribution estimation and hence the imputed wealthy households.

Table 1: Overview of several rich lists for euro area countries and the source.
Note: The number in parentheses is the number of observations provided by the rich list at the time of the third HFCS wave.

	Rich list		Rich list
AT	Trend (100)	GR	Forbes (4)
BE	Forbes (3)	IE	Independent.ie (175)
CY	Forbes (5)	IT	Forbes (43)
DE	Manager Magazin (984)	LV	Public broadcasting of Latvia (10)
ES	El Mundo / Wealth Tax Data (200)	NL	Forbes (10)
FR	Challenges (504)	PT	Forbes (4)

The EG-LMM developed a first approach to capturing the missing wealthy households, which consisted in adjusting the HFCS household weights (see EG-LMM, 2020). While this approach facilitates the construction of an adjusted granular data set, it does so by altering observation weights, which have been precisely calibrated by the HFCN.[21] In addition, this approach assumes

[21] The HFCS weights are calibrated using a complex survey design including features for stratification, clustering and unequal selection properties.

that HFCS portfolios of wealthy households are representative for the missing wealthy households. However, when taking a closer look at the portfolio allocation, this assumption does not appear fully realistic. Figure 6 visualises the portfolio allocation of the HFCS households by net wealth quantiles for Germany, Spain, France and Italy. In all four countries, even the richest 2% of households hold most of their wealth in housing wealth, with financial and non-financial business wealth, as well as listed shares and investment fund shares, playing only a very small role. This seems unlikely to reflect the portfolio allocation of very wealthy households adequately, which are generally identified from their holdings of substantial assets in business wealth and shares (see, for example, The Economist, 2018). A variant on this approach has, therefore, been developed that maintains the HFCS household weights and allows for a flexible adjustment of the portfolio allocation of the missing wealthy households.

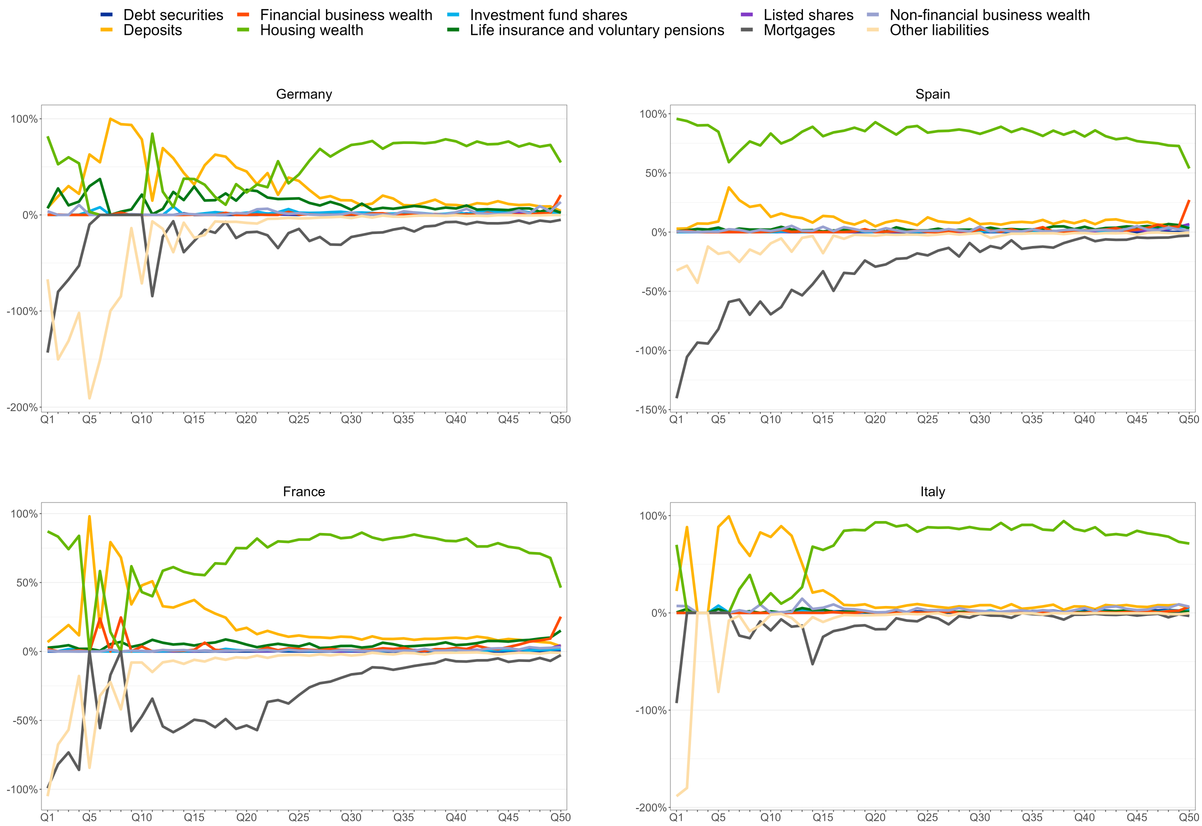

Figure 6: Households' portfolio allocation according to the HFCS.
Note: The data refers to wave 3 (which covers broadly the period 2016 Q4 - 2018 Q3). The portfolio allocation is relative to gross wealth and by fifty net wealth quantiles (i.e. Quantile 1 denotes the 2% poorest households and Quantile 50 the 2% richest households).

Let $x_{ij} \in \mathbb{R}$ denote the holding of household i in instrument j, for $i = 1, 2, \ldots, n$; where n denotes the number of households participating in the HFCS in a certain country and $j \in Instruments$, where the set of considered instruments is given by:

$$\begin{aligned} Instruments = \Big\{ & \text{Deposits, debt securities, listed shares, investment fund shares,} \\ & \text{life insurance and voluntary pensions, financial business wealth,} \\ & \text{non-financial business wealth, housing wealth,} \\ & \text{mortgages, other liabilities} \Big\}. \end{aligned} \tag{3}$$

Moreover, liabilities (i.e. mortgages and other liabilities) are given as negative numbers or zero, while all other instruments are denoted as non-negative numbers. This allows the convenient representation of the net wealth of household i as

$$x_{i \text{ net wealth}} = \sum_{j \in Instruments} x_{ij}. \tag{4}$$

Let $X_{\text{net wealth}}$ denote the random variable of households' net wealth, which above a certain threshold $w_0 \in \mathbb{R}_{>0}$ is assumed to follow a Pareto distribution. Following Vermeulen (2018) and Waltl (2020), the Pareto scale parameter w_0 is fixed to equal one million euros, i.e. $w_0 = 10^6$.[22] The Pareto distribution is given by

$$F(x) = P(X_{\text{net wealth}} \leq x) \begin{cases} 1 - \left(\frac{w_0}{x}\right)^{\alpha}, & \text{for } x \geq w_0, \\ 0, & \text{for } x < w_0, \end{cases} \tag{5}$$

where $\alpha \in \mathbb{R}_{>0}$ denotes the Pareto shape parameter. Hence, fitting a Pareto distribution to a given data set essentially means to find a Pareto shape parameter $\alpha \in \mathbb{R}_{>0}$ that best describes the data. Here, a Pareto distribution is fitted via the methodology suggested by Vermeulen (2018) and implemented by the EG-LMM.

The upper part of the net wealth distribution is shown in Figure 7 for the largest four euro

[22] An alternative to this methodology is to use a non-fixed parameter and define a way to identify it automatically based on other relevant data. This is done in Disslbacher et al. (2020), for example.

area countries. The following three distinct intervals can be identified

$$
\begin{aligned}
I_1 &= [\text{all HFCS households with wealth above } w_0]\,,\\
I_2 &=]\text{gap between the HFCS observations and the rich list}[\,,\\
I_3 &= [\text{from the observations of the rich list until infinity}[\,.
\end{aligned}
\tag{6}
$$

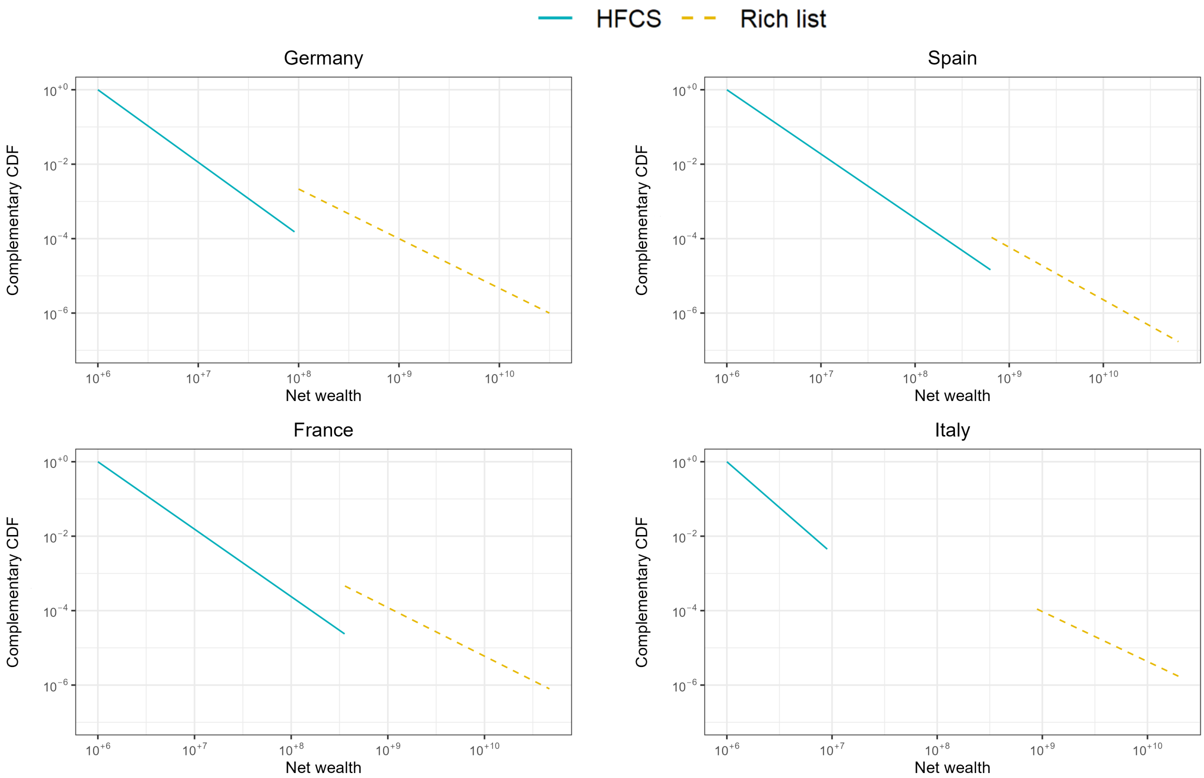

Figure 7: Illustration of the upper tail of the net wealth distribution.
Notes: The data refers to wave 3 (which covers broadly the period 2016 Q4 - 2018 Q3). The solid blue line represents the Pareto distribution fitted solely on the HFCS data, while the dashed yellow line represents the Pareto distribution fitted also on additional observations provided by the rich list. Owing to data confidentiality constraints, the individual observations are not shown. However, the coverage of the HFCS is reflected in the length of the Pareto distribution fitted solely on the HFCS observations (solid turquoise line), where the upper bound shows an order of magnitude.

Depending on the country specific situation, the length of each of the three intervals can differ substantially. While the rich lists of Germany, Spain, France and Italy (see Figure 7) show that there are households with a net wealth of over EUR 10 billion, the sampling of the HFCS varies considerably. The Spanish HFCS shows the highest coverage, including households with a net wealth of up to almost EUR 1 billion. The coverage of the French and German HFCS includes households with a net wealth of around EUR 100 million. By contrast, the coverage of the Italian HFCS stops at a net wealth of slightly below EUR 10 million, leading to a substantial interval of unobserved households between the HFCS and the rich list observations. In particular, if the HFCS uses oversampling techniques to better cover the most wealthy households and the rich list is very long, the interval that is well covered by the HFCS (I_1) and the interval comprising the rich list observations (I_3) might even touch each other or overlap (as, for example, in the case of Spain). In such cases, an interval without any observations (I_2) does not exist. However, this occurs only in a very small number of countries. For most cases, the interval (I_2) that contains no observations and lies between the HFCS and the rich list, is substantial (as, for example, in the case of Italy). Obviously, all observations that are taken from the rich list and fall outside of the interval covered by the HFCS (I_1) constitute wealthy observations that are not represented in the HFCS. Hence, an intuitive way of capturing these missing wealthy households is simply to extend the HFCS by the observations taken from the rich list. Moreover, it seems implausible that there would be no households in the interval without observations (I_2). Therefore, an obvious way of estimating these missing wealthy households is to sample them from the fitted Pareto distribution. Given the available data, the best guess for the Pareto distribution is the one fitted on both the HFCS and the rich list. Synthetic households can subsequently be sampled from the fitted Pareto distribution.

More precisely, based on the fitted Pareto distribution and the interval that is well covered by the HFCS, the number of observations that one would expect in any other interval of the Pareto distribution can be computed. Let

$$\begin{aligned} w_1 &:= \max\left\{\text{HFCS households' net wealth}\right\} = \max\left\{x_{i \text{ net wealth}} : i = 1, 2, \ldots, n\right\}, \\ w_2 &:= \min\left\{\text{net wealth of rich list observations}\right\}, \end{aligned} \tag{7}$$

such that the three intervals defined in 6 are given by

$$I_1 = [w_0, w_1], \qquad I_2 =]w_1, w_2[, \qquad I_3 = [w_2, \infty[. \tag{8}$$

Further, let the net wealth of wealthy households (i.e. households with a net wealth of at least w_0) be denoted by the random independent and identically distributed (i.i.d.) variables

X_i, for $i \in \mathbb{N}_{>0}$, following the fitted Pareto distribution.[23] While the total number of wealthy households is unknown, owing to the gap between the HFCS observations and the (sometimes scarcely available) rich lists, an estimate can be derived as follows. Let the random variables M_1, M_2, and M_3 denote the number of wealthy households in the three intervals I_1, I_2, and I_3 respectively, and $M := M_1 + M_2 + M_3$ the total number of wealthy households.[24] The expected total number of wealthy households $\mathbb{E}[M]$ is given by

$$\begin{aligned}
\mathbb{E}[M_1] &= \mathbb{E}\left[\sum_{i=1}^{M} \mathbb{1}_{\{X_i \in I_1\}}\right] = \mathbb{E}[M]\,\mathbb{E}\left[\mathbb{1}_{\{X_1 \in I_1\}}\right], \quad \text{by Wald's lemma, for } (X_i) \text{ i.i.d.} \\
&= \mathbb{E}[M]\,P\Big(X_1 \in [w_0, w_1]\Big) = \mathbb{E}[M]\Big[P(X_1 \leq w_1)\Big] \\
&= \mathbb{E}[M]\left[1 - \left(\frac{w_0}{w_1}\right)^{\alpha}\right] = \mathbb{E}[M]\,\frac{w_1^{\alpha} - w_0^{\alpha}}{w_1^{\alpha}}, \\
\Leftrightarrow \quad \mathbb{E}[M] &= \mathbb{E}[M_1]\,\frac{w_1^{\alpha}}{w_1^{\alpha} - w_0^{\alpha}}.
\end{aligned} \tag{9}$$

Likewise, the expected number of wealthy households in I_2 is given by

$$\begin{aligned}
\mathbb{E}[M_2] &= \mathbb{E}\left[\sum_{i=1}^{M} \mathbb{1}_{\{X_i \in I_2\}}\right] = \mathbb{E}[M]\,P\Big(X_1 \in]w_1, w_2[\Big) \\
&= \mathbb{E}[M]\left[\left(1 - \left(\frac{w_0}{w_2}\right)^{\alpha}\right) - \left(1 - \left(\frac{w_0}{w_1}\right)^{\alpha}\right)\right] \\
&= \mathbb{E}[M_1]\,\frac{w_1^{\alpha}}{w_1^{\alpha} - w_0^{\alpha}}\left[\left(\frac{w_0}{w_1}\right)^{\alpha} - \left(\frac{w_0}{w_2}\right)^{\alpha}\right], \quad \text{by eq. (9).}
\end{aligned} \tag{10}$$

As the number of wealthy households in I_1 is given by the HFCS, and let this number be denoted by $\hat{m}_1$, we can estimate the number of wealthy households in I_2 from equation 10 by

$$\hat{m}_2 := \hat{m}_1 \frac{w_1^{\alpha}}{w_1^{\alpha} - w_0^{\alpha}}\left[\left(\frac{w_0}{w_1}\right)^{\alpha} - \left(\frac{w_0}{w_2}\right)^{\alpha}\right], \quad \text{by eq. (9).} \tag{11}$$

Once the number of unobserved households in interval I_2 is estimated, $\hat{m}_2$ synthetic households can be randomly sampled from the fitted Pareto distribution via inverse transform sampling. Let $U \sim \text{Unif}\left[F(w_1), F(w_2)\right]$, then $F^{-1}(U)$ constitutes a random variable that follows

[23] Further research may be carried out in the future on the treatment of survey data that would not be truly i.i.d..

[24] Cantarella et al. (2021) propose an estimation of the top-tail population by summing the weights of HFCS households included in the top tail ($\hat{P}_{obs}$) and scaling this by the proportion of the estimated Pareto density tail covered by these households to estimate the total population (P_T) of the tail and therefore the number of missing households $P_T = \hat{P}_{obs}\left[1 - (w_0/w_1)^{\alpha}\right]^{-1}$.

the Pareto distribution in I_2. More precisely, let u denote a realization of U, then

$$F^{-1}(u) = w_0 (1-u)^{-1/\alpha} \tag{12}$$

gives the corresponding sampled net wealth of a synthetic household in I_2.

The approach described above is likewise applicable to any interval of the Pareto distribution. This means that, for those countries for which only a sparse rich list is available, the interval I_3 can also be complemented with synthetic households. More precisely, the number of expected observations in interval I_3 can be computed and the difference vis-à-vis the actual number of observations given by the sparse rich list can be randomly drawn from the Pareto distribution. Even if there is no rich list available, an upper bound of net wealth can be defined and synthetic wealthy households can be sampled according to the Pareto distribution. Figure 8 provides an illustration of the upper tail of the net wealth distribution enhanced with synthetic households where necessary.

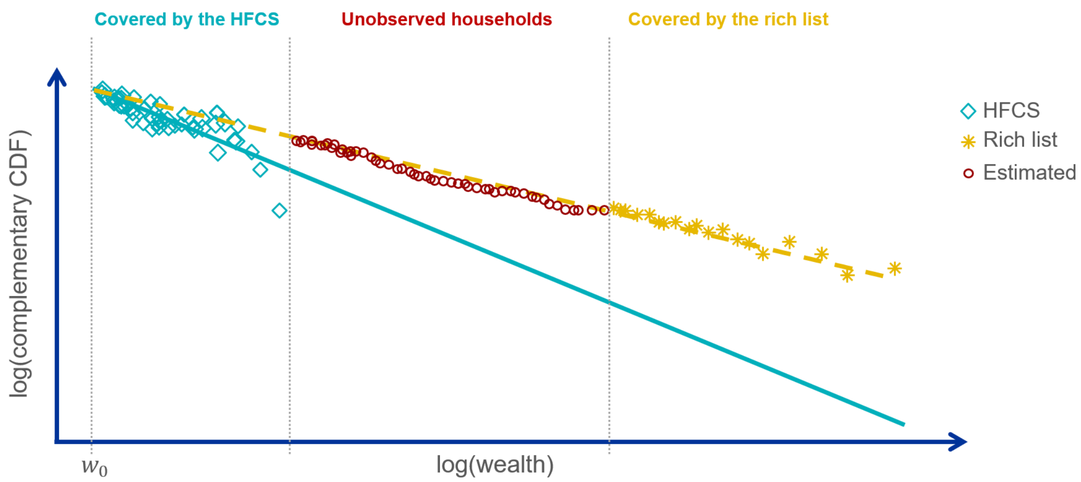

Figure 8: Stylised illustration of the upper tail of the net wealth distribution enhanced with synthetic households where necessary.
Notes: The solid blue line represents the Pareto distribution fitted solely on the HFCS data, while the dashed yellow line represents the Pareto distribution fitted also on additional observations provided by the rich list.

Note that both the observations from the rich list and the synthetic households are given in terms of their net wealth. In order to estimate both assets and liabilities of these added wealthy households, two further steps are required: (i) the size of liabilities which these households are most likely to hold needs to be estimated, and (ii) a realistic portfolio allocation for these wealthy households has to be identified.

Step (i): Estimation of liabilities held by the added wealthy households

Given that the wealthy households are those able to obtain larger credit amounts from financial institutions, and assuming that the quality of the HFCS data on liabilities of poorer households is generally good, most of the gap between the NtlA values of loans to households and the same data collected in the HFCS can be assumed to be attributable to those very wealthy households not covered by the HFCS.[25] More precisely, the results presented in this paper are based on the assumption that 75% of the discrepancy in liabilities is held by the household above a certain threshold (w_0). However, this rate may be adjusted depending on additional information, obtained, for example, from credit registers, where available.

Step (ii) Estimation of the assets held by the added wealthy households

As only net wealth is estimated from the Pareto distribution, assumptions or external data are needed to design a portfolio of instruments for the imputed wealthy households. To construct these portfolios, we use The Global Family Office Report from UBS/Campden Research (2018), as published in The Economist (2018). Table 2 shows the published portfolio structure of more than 300 wealthy households who use "family office" wealth management services.

In this context, the approach developed so far builds on the information provided by The Economist, but other sources of information or additional national data may of course also be used, where available, to construct an appropriate portfolio allocation for the added wealthy households.

The gross wealth of all added wealthy households can then be calculated as the sum of estimated net wealth and liabilities attributed in *Step (i)*. These total assets are attributed to NtlA instruments along the lines measured by UBS/Campden, as outlined in Table 3. The definition of the concepts of the instruments in The Economist publication differs in several cases from the definitions and categories defined in the HFCS and the NtlA. Therefore, when the pure information from UBS/Campden Research (2018) leads to holdings that are even higher than those recorded in the NtlA for a given instrument, the portfolio allocation of the added wealthy households is adjusted to transfer the excess to other plausible instruments that are still under-covered, as described in Table 4: a key assumption in so doing is that holdings of certain instruments, e.g. debt securities or listed shares, may be held directly or via vehicles such as

[25] Coverage of liabilities varies across countries, with mortgage liabilities often having higher coverage rates than non-mortgage liabilities.

investment fund shares or life insurance investment schemes.

Table 2: Portfolio of surveyed family offices(see The Economist, 2018).

Instrument	Portfolio allocation
Bonds	16.2%
Equities	28.0%
Direct investments	14.0%
Funds	7.6%
Hedge Funds	5.7%
Property	18.1%
Commodities	3.4%
Cash or equivalent	7.0%

Table 3: Portfolio of added wealthy households.

Instrument	Portfolio allocation
Debt securities	16.2%
Listed shares	28.0%
Financial business wealth	21.6%
Investment fund shares	5.7%
Housing wealth	18.1%
Non-financial business wealth	3.4%
Deposits	7.0%

Table 4: Transfer of excess in certain instruments to other plausible under-covered instruments in terms of the portfolio allocation of added wealthy households.

excess in	can be transferred to
Debt securities	Investment fund shares and life insurance and voluntary pensions
Listed shares	Business wealth, investment fund shares and life insurance and voluntary pensions
Housing wealth	Listed shares, business wealth, investment fund shares, and life insurance and voluntary pensions

Future research could consider implementing a modelling approach for deriving portfolios of the missing wealthy, incorporating HFCS households and external sources.

2.4 Population matching

Another reason for the discrepancies between the HFCS and the NtlA is that the coverage of households in the NtlA is usually broader than that in the HFCS. In Household Finance and Consumption Network (2020a), a household is defined as "*a person living alone or a group of people who live together in the same private dwelling and share expenditures, including the joint provision of the essentials of living*". Persons living in institutions, e.g. in prisons or retirement homes, as well as homeless people, are excluded from this definition, but they are included in the NtlA definition.

As the aim is to construct DNW time series that are in line with the NtlA, the augmented (with the added wealthy households) HFCS population is adjusted to match the NtlA population. In the absence of any other information, the proportional adjustment is an intuitive solution in that it rescales all household weights with the same factor.[26] We provide the mathematical approach taken together with a summary table with the detailed matching in Annex B.

2.5 Proportional allocation

The methodology presented thus far adjusts and extends the HFCS data by correcting issues identified with the survey in order to adequately cover all classes of household and match the aggregate figures provided by the NtlA. While these adjustments are a step towards reducing the discrepancies between HFCS and NtlA instrument totals, and for a number of countries enable the gaps to be closed for many instruments, in general there are still some discrepancies. To fully reconcile both data sources, a proportional adjustment, which rescales the holdings of all households in a certain instrument by the ratio of the corresponding NtlA total and the adjusted HFCS aggregate, is applied as a final step. An important property of the proportional allocation is that it does not change the Gini coefficient at the level of each instrument. Likewise, the household ranking at the instrument level is preserved, for example, if household A has a higher deposits account than household B, this will also hold true after the final adjustment. Further insights into the proportional allocation and the underlying optimisation problem, as well as the upper and lower bounds of the Gini coefficient that are reached for extreme allocations of the remaining discrepancies between the HFCS and NtlA instrument totals, are provided by Ohlwerter (2021). After the final step of the proportional allocation the adjusted and extended HFCS data match the figures of the household sector instrument holdings given by the NtlA.

2.6 Compilation of quarterly time series

Following the steps developed in previous sections to close the gaps between the HFCS and the NtlA at the end of the quarter that is closest to the relevant HFCS wave, this section focuses on developing time series for the distribution of households' wealth. The demand for annual DNW time series data was also highlighted in a user consultation and it is concluded in Expert Group on Linking macro and micro data for the household sector (2020) that this issue needs to be addressed as soon as possible. The proposed methodology is aligned with the frequency of the NtlA and can thus be applied to generate both annual and quarterly time series. Currently, three waves of the HFCS are available, covering approximately the years 2010, 2014, and 2017.

[26] This assumption is a simplified approach, which is justified by the small share of households affected.

To fill the gaps between the waves and to increase the frequency of the data, a model for inter- and extrapolation has been constructed, which results in a quarterly time series.

A number of approaches for interpolation and extrapolation have been used in the literature, including a linear interpolation of asset holdings at the net wealth quintile level (Kavonius and Honkkila, 2016), the Chow-Lin method (Chow and Lin, 1971; Batty et al., 2019), proxies for growth in wealth from higher frequency data (Honkkila et al., 2018), and simulation models (Ampudia et al., 2014). In the light of the limited availability of data, i.e. the HFCS currently provides only three time points, we opted for the linear interpolation suggested by Kavonius and Honkkila (2016).[27] This method has been applied to not only an interpolation between the three available HFCS waves, but also an extrapolation to the latest available NtlA data, as there are no trends in the distribution of holdings over the three HFCS waves. This is obviously a simplification, but it enables the provision of potentially useful information to analysts on what the data would look like under this basic assumption. As shown in Annex D, the resulting estimates by deciles (classified by level of net wealth) show quite a stable pattern for most countries and instruments since 2009, based on the three available HFCS data sets, suggesting that the assumption made reveals a plausible broad picture of recent developments.

It should be noted that this approach does not leave the wealth distributions constant. Depending on the participation rates of households in various instruments, the portfolio composition, and the different developments of those instruments as measured in the NtlA (i.e. transactions by households and price changes over time), the wealth shares of household groups change. For example, assuming the poorest households do not hold shares, a high increase in share prices will alter the interpolated or extrapolated distribution towards richer households. However, what the proposed method cannot capture is the active changes of portfolios by specific household groups, for example, if poorer households were to increasingly invest in shares, or some richer households were to keep constant or reduce such holdings. Obviously, this caveat may affect the extrapolated results more than those of interpolations between HFCS waves. This means that economic shocks that occur outside the HFCS waves, such as the COVID-19 pandemic, will be captured by their impact on the total values recorded in the NtlA, but not by any adjustments made within a household's portfolio.

[27] The approach has originally been developed on quintile level, but of course any other level of quantiles can likewise be used. The results presented in this paper have been produced by an inter- and extrapolation on decile level. Section 2.6.1 explains how the quintiles are computed and Section 2.6.2 how these quintiles are applied on household level to derive an inter- and extrapolated micro data set.

Extensions and refinements to this basic model can be assessed and added in the future, and when a sufficient number of HFCS waves become available, more sophisticated models can also be explored.[28]

The following sections explain in more detail the linear inter- and extrapolation approach that has been developed. To avoid confusion, the model is described step by step. Section 2.6.1 covers (i) the linear interpolation, and (ii) the extrapolation, following Kavonius and Honkkila (2016), while Section 2.6.2 presents the novel extension to an inter- and extrapolation at the household level, generating the time series for the micro data.

2.6.1 Time-series by quintile

Interpolation between HFCS waves, taking into account quarterly NtlA data

Let t_1 and t_2 denote two consecutive reference dates of the HFCS (e.g. 2010 Q4 and 2014 Q4). Moreover, let Q1 to Q5 denote five net wealth quintiles. This means that Q1 comprises the 20% poorest households in terms of net wealth and Q5 comprises the 20% of households with the highest net wealth. By aggregating the holdings of each instrument $j \in Instruments$ for each quintile $k = 1, 2, \ldots, 5$ and dividing this by the instrument total $T_j \in \mathbb{R}$, we get the ratios r_{jk} that are held for each instrument by each net wealth quintile. More precisely,

$$r_{jk} = \frac{1}{T_j} \sum_{i \in Q_k} d_i x_{ij}, \tag{13}$$

where Q_k denotes the set of households that fall into quintile k, x_{ij} denotes the holdings of household i in instrument j, and d_i the corresponding household weight. In order to conveniently denote the corresponding reference date, the ratios can be considered as a function over time, i.e. we denote $r_{jk}(t_1)$ for the ratio of instrument j and quintile k computed at time point t_1 (e.g. for wave 1, t_1= 2010 Q4). Given these ratios for two consecutive reference dates t_1 and t_2, we can estimate these ratios for all time points between t_1 and t_2 via a linear interpolation. Let

[28] As a comparison, the FED is using a Chow and Lin (1971) approach to construct quarterly distributional financial accounts for the United States; (see Batty et al., 2019). The Chow-Lin method incorporates the time series profile of higher frequency into the lower frequency data and it corresponds in practice to the instrument level interpolation referred to above.

$t_1 \leq t_1 + \ell\delta \leq t_2$, then the interpolated ratio for time point $t_1 + \ell\delta$ is given by

$$\begin{aligned} r_{jk}(t_1 + \ell\delta) &= r_{jk}(t_1) + (r_{jk}(t_2) - r_{jk}(t_1)) \frac{(t_1 + \ell\delta) - t_1}{t_2 - t_1} \\ &= r_{jk}(t_1) + (r_{jk}(t_2) - r_{jk}(t_1)) \frac{\ell\delta}{t_2 - t_1}, \end{aligned} \tag{14}$$

where ℓ is the period to be estimated. In the case of an annual interpolation δ would denote one year, while in the case of a quarterly interpolation δ can be conveniently set to $\delta = 0.25$.

The linear interpolation is illustrated further in Table 5 and an example is provided in Table 6.

Table 5: Illustration of a linear interpolation at the instrument and quintile level. The ratios for the reference dates t_1 and t_2 are given, while the figures for the dates that lie in between, i.e. $t_1 + 1\delta, t_1 + 2\delta, \ldots, t_1 + \left(\frac{t_2 - t_1}{\delta} - 1\right)\delta$ are derived via the linear interpolation.

	Q1	Q2 – Q4	Q5
t_1	$r_{j1}(t_1)$	$\ldots$	$r_{j5}(t_1)$
$t_1 + 1\delta$	$r_{j1}(t_1) + (r_{j1}(t_2) - r_{j1}(t_1)) \frac{1\delta}{t_2 - t_1}$	$\ldots$	$r_{j5}(t_1) + (r_{j5}(t_2) - r_{j5}(t_1)) \frac{1\delta}{t_2 - t_1}$
$t_1 + 2\delta$	$r_{j1}(t_1) + (r_{j1}(t_2) - r_{j1}(t_1)) \frac{2\delta}{t_2 - t_1}$	$\ldots$	$r_{j5}(t_1) + (r_{j5}(t_2) - r_{j5}(t_1)) \frac{2\delta}{t_2 - t_1}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
t_2	$r_{j1}(t_2)$	$\ldots$	$r_{j5}(t_2)$

Table 6: Example of a linear interpolation at the instrument and quintile level. The ratios for the reference dates 2010 Q4 and 2014 Q4 are given, while the figures for the date that lies in between, i.e. 2012 Q1, are derived via the linear interpolation.

	Q1	Q2	Q3	Q4	Q5
2010 Q4 $= 2010.75$	3%	5%	10%	8%	74%
2012 Q1 $= 2012.00$	$r_{j1}(2012.00)$ $= 3\% + (5\% - 3\%) \frac{2012.00 - 2010.75}{2014.75 - 2010.75}$ $= 3.6\%$	5.3%	9.7%	8.6%	$r_{j5}(2012.00)$ $= 74\% + (70\% - 74\%) \frac{2012.00 - 2010.75}{2014.75 - 2010.75}$ $= 72.8\%$
2014 Q4 $= 2014.75$	5%	6%	9%	10%	70%

The estimated interpolated ratios are subsequently multiplied by the instrument total given by the NtlA for the respective time point.

Extrapolation from the most recent HFCS to the latest available NtlA data

Regarding the extrapolation, constant ratios with respect to the closest HFCS reference date are suggested. Unless a clear trend is identified in the back data, it is deemed best to assume a stable distribution since the last survey was conducted. More precisely, if t_2 denotes the latest available HFCS wave, the ratios for any subsequent time point $t_2 + \ell\delta > t_2$, for which the NtlA instrument totals are available, are given by

$$r_{jk}\left(t_2 + \ell\delta\right) = r_{jk}\left(t_2\right). \tag{15}$$

Likewise, a backwards extrapolation is given for any time point $t_1 - \ell\delta < t_1$ (where t_1 denotes the first available HFCS wave), by

$$r_{jk}\left(t_1 - \ell\delta\right) = r_{jk}\left(t_1\right). \tag{16}$$

An example of the constant extrapolation and the multiplication with the NtlA instrument totals is provided in Table 7.

Table 7: Example of a constant extrapolation at the instrument and quintile level. The ratios for the reference dates 2017 Q4 are available, while the figures for 2020 Q1 are derived via the constant extrapolation.

	Unit	Q1	Q2	Q3	Q4	Q5	NtlA total
2017 Q4	totals	7	12	23	19	173	234
2017 Q4	ratio	$\frac{7}{234} = 3\%$	5%	10%	8%	$\frac{173}{234} = 74\%$	234 (100%)
2020 Q1	ratio	$r_{j1}\,(2020.00) = 3\%$	5%	10%	8%	$r_{j5}\,(2020.00) = 74\%$	310 (100%)
2020 Q1	totals	$310\,r_{j1}\,(2020.00) = 9$	17	31	25	$310\,r_{j5}\,(2020.00) = 217$	310

2.6.2 Time-series at the household level

The inter- and extrapolation model presented so far result in aggregated quintile figures for each instrument. However, based on just five numbers for each instrument inequality indicators such as the Gini coefficient cannot be computed. Inter- and extrapolated micro data sets are needed to enable an adequate analysis of inequality for the inter- and extrapolated time line. Moreover, since the reference dates of the HFCS waves differ across countries, euro area aggregates on DNW (discussed in more detail in Section 2.7) cannot be computed by simply aggregating the micro data sets for certain waves. To compile euro area DNW properly, quarterly (or annual) micro data sets for all euro area countries are required. Therefore, in the following, an extension to the inter- and extrapolation approach is presented in order to develop inter- and extrapolated micro data sets. Essentially, the idea is simply to apply the interpolated ratios derived by the methodology presented above to each household in the respective quintiles.

Let $\tilde{x}_{ij}(t)$ denote the ratio of household i's holdings in instrument j at time point t, i.e.

$$\tilde{x}_{ij}(t) = \frac{x_{ij}(t)}{\sum_{i=1}^{n(t)} d_i(t)\, x_{ij}(t)} = \frac{x_{ij}(t)}{T_j(t)}, \tag{17}$$

where $x_{ij}(t)$ as usual denotes the holdings of household i in instrument j at time point t, $d_i(t)$ denotes the corresponding household weight, and $n(t)$ the number of households in the micro data set at time point t.[29] As the adjusted micro data set matches the NtlA instrument total, i.e. $\sum_{i=1}^{n(t)} d_i(t)\, x_{ij}(t) = T_j(t)$, the ratio r_{jk}, introduced in Equation (13) can be rewritten as

$$r_{jk}(t) = \frac{1}{T_j(t)} \sum_{i \in Q_k} d_i(t)\, x_{ij}(t) = \sum_{i \in Q_k} d_i(t)\, \tilde{x}_{ij}(t). \tag{18}$$

Hence, the weighted sum of the households' instrument ratios adds up to the quintile ratios. Via a proportional adjustment of all households' instrument ratios with respect to the interpolated quintile ratios, an interpolated micro data set that matches the interpolated quintile ratios is achieved. More precisely, starting from a reference date t_1, the interpolated instrument ratio of household i, for $i = 1, \ldots, n(t_1)$, belonging to quintile k is given by

$$\tilde{x}_{ij}(t_1 + \ell\delta) = \tilde{x}_{ij}(t_1)\, \frac{r_{jk}(t_1 + \ell\delta)}{r_{jk}(t_1)}. \tag{19}$$

Note that the weighted sum of the interpolated household instrument ratios equals the desired

[29] A word of caution is added here: the introduced notation should not be interpreted as a certain household i appearing in all waves, i.e. i at t_1 and i at t_2 refer to two different households.

interpolated quintile ratios derived above,

$$\begin{aligned}\sum_{i\in Q_k} d_i(t_1)\,\tilde{x}_{ij}(t_1+\ell\delta) &= \frac{r_{jk}(t_1+\ell\delta)}{r_{jk}(t_1)}\sum_{i\in Q_k} d_i(t_1)\,\tilde{x}_{ij}(t_1)\\ &= \frac{r_{jk}(t_1+\ell\delta)}{r_{jk}(t_1)}\sum_{i\in Q_k} d_i(t_1)\,\frac{x_{ij}(t_1)}{T_j(t_1)}\\ &= r_{jk}(t_1+\ell\delta)\,.\end{aligned} \tag{20}$$

Likewise, an interpolated micro data set can be constructed from the following reference date t_2,

$$\tilde{x}_{(n(t_1)+i)j}(t_1+\ell\delta) = \tilde{x}_{ij}(t_2)\,\frac{r_{jk}(t_1+\ell\delta)}{r_{jk}(t_2)}. \tag{21}$$

Since both interpolated micro data sets, i.e. interpolated from t_1 and t_2, are equally likely (i.e. it is not possible to determinate which HFCS wave is more representative of periods in between), the collection of both micro data sets is used for the interpolated micro data set. Since the weighted sum of the interpolated household instrument ratios would then add up to 200%, the household weights of both sets are proportionally adjusted. Let $d_{\text{total}}(t_1+\ell\delta)$ denote the total household population for the interpolated time point, given by a linear interpolation of the population total for both reference dates, i.e.

$$d_{\text{total}}(t_1+\ell\delta) = \left[\sum_{i=1}^{n(t_1)} d_i(t_1)\right] + \ell\delta\left\{\left[\sum_{i=1}^{n(t_2)} d_i(t_2)\right] - \left[\sum_{i=1}^{n(t_1)} d_i(t_1)\right]\right\}. \tag{22}$$

The household weights for the interpolated micro data set are proportionally adjusted, such that each of the two derived interpolated micro data sets contributes 50% of the household population,

$$d_i(t_1+\ell\delta) = \frac{d_i(t_1)}{\sum_{i=1}^{n(t_1)} d_i(t_1)}0.5\,d_{\text{total}}(t_1+\ell\delta)\,, \qquad \text{for } i=1,\ldots,n(t_1) \tag{23}$$

and

$$d_{n(t_1)+i}(t_1+\ell\delta) = \frac{d_i(t_2)}{\sum_{i=1}^{n(t_2)} d_i(t_2)}0.5\,d_{\text{total}}(t_1+\ell\delta)\,, \qquad \text{for } i=1,\ldots,n(t_2)\,. \tag{24}$$

By multiplying the interpolated household ratios with the corresponding NtlA instrument totals, a complete adjusted micro data set is achieved which matches the interpolated quintile ratios and the totals given by the NtlA.

The construction of extrapolated micro data sets is straightforward. As the instrument quintile ratios stay constant for the extrapolation, the extrapolated household ratios are simply given by

$$\tilde{x}_{ij}\left(t_2+\ell\delta\right)=\tilde{x}_{ij}\left(t_2\right). \tag{25}$$

Likewise a backwards extrapolation from the first reference date to the past is given by

$$\tilde{x}_{ij}\left(t_1-\ell\delta\right)=\tilde{x}_{ij}\left(t_1\right). \tag{26}$$

Again, multiplying the extrapolated household ratios by the corresponding NtlA instrument totals generates a complete adjusted micro data set which matches the extrapolated quintile ratios and the totals given by the NtlA.[30]

2.7 Compilation of euro area aggregates

Analysing and understanding the development of distributions of wealth and inequality is of interest and importance at both the country level and the euro area aggregate level. Given the national quarterly reconciled time series derived in the sections above, the compilation of euro area aggregates is straightforward. The quarterly times series contains, for each country and quarter, an adjusted micro data set of households that matches the instrument totals of the countries' NtlA. Hence, combining the micro data sets for all euro area countries creates a micro data set that represents the entire euro area population. Aggregating the micro data sets of all euro area countries into one data frame for each quarter generates a micro data set for the entire euro area times series.

The NtlA for the euro area do not simply consist of the sum of all euro area countries' NtlA figures, but are instead subject to certain harmonization procedures. Therefore, our aggregated euro area micro data set also needs to be adjusted to properly match the euro area NtlA. Since these adjustments are generally rather minor, a simple proportional adjustment for each instrument seems sufficient. More precisely, the holdings of all euro area households $i=1,2,\ldots,n$ (where n denotes the sum of all households given by the interpolation and across all euro area

[30] The example provided here is an illustration of the methodology. This method is illustrated at the quintile level, but it can also be applied at the decile level, percentile level, etc. In our case, we interpolate and extrapolate the distribution at the decile level.

countries) are adjusted for each instrument $j \in \mathit{Instruments}$ by

$$x_{ij}^{(\text{new})} = x_{ij} \frac{T_j^{(\text{EA})}}{\sum_{i=1}^{n} d_i x_{ij}}, \tag{27}$$

where $T_j^{(\text{EA})} \in \mathbb{R}_{>0}$ denotes the euro area total of instrument j given by the NtlA and d_i denotes the household weights (adjusted to the NtlA population total, as explained in Section 2.4).

To provide further insight into the model presented, Figure 9 illustrates some results for the inter- and extrapolation for two instruments for the euro area.[31]

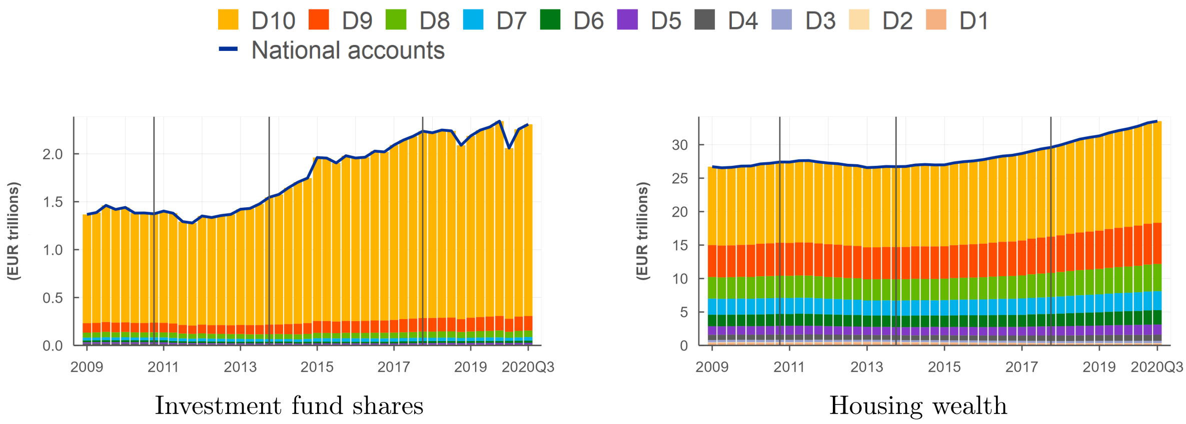

Figure 9: Households' instrument holdings by net wealth deciles, derived via the presented inter- and extrapolation at household level.
Note: D10 denotes the richest net wealth decile, i.e. comprising the 10% of individuals with the highest net wealth, and D1 denotes the poorest net wealth decile, i.e. comprising the 10% of individuals with the lowest net wealth.

2.8 Impact of the model on inequality measures

The aim of the derived microeconomic data set is to enable a detailed analysis of wealth inequality in the household sector across countries and time. The level of inequality measured on the derived data set naturally depends on the inequality captured by the underlying data sets, i.e. the HFCS and the NtlA, as well on the applied modelling approach. It is therefore of interest to analyse the level of inequality captured by the raw HFCS and how this changes with the steps of the proposed model. This section sheds some light on those questions and provides explanations, thereby advocating a better understanding of the impact of the modelling assumptions on measured inequality.

[31] For information, Annex D presents all instruments and a net wealth per capita.

Table 8 presents the most common inequality measures computed for the four largest euro area economies on the raw HFCS and on the main model steps, i.e. after including missing wealthy households and after allocating the remaining discrepancies between the HFCS aggregates and the NtlA figures via the proportional allocation.[32]

Table 8: Inequality measures computed on the HFCS data set and along the main steps of the presented model.

	Number of households with a net wealth above EUR 1 million	Top 1% share of net wealth	Top 5% share of net wealth	Top 10% share of net wealth	Bottom 50% share of net wealth	Net wealth Gini coefficient
Germany						
HFCS	1,562,897	19%	42%	56%	2%	0.742
Wealthy added	1,581,988	26%	47%	60%	2%	0.763
DNW	1,663,633	22%	42%	57%	2%	0.739
Spain						
HFCS	672,442	21%	40%	53%	7%	0.680
Wealthy added	701,137	24%	43%	56%	7%	0.695
DNW	932,755	25%	44%	56%	7%	0.695
France						
HFCS	848,587	17%	37%	51%	4%	0.690
Wealthy added	869,906	26%	43%	56%	4%	0.720
DNW	1,908,899	27%	44%	57%	5%	0.721
Italy						
HFCS	591,456	12%	31%	44%	9%	0.616
Wealthy added	671,539	29%	45%	55%	7%	0.693
DNW	1,088,256	35%	50%	60%	7%	0.721

The inclusion of missing wealthy households leads by construction to a fatter upper tail of the net wealth distribution, increasing inequality. By how much the inclusion of missing wealthy households increases inequality depends on two aspects: the number of added wealthy households, and their net wealth. The number of wealthy households that is estimated to be missing increases with the number of observations provided by a rich list (and that are missing in the HFCS), as well as with the length of the interval of unobserved households (see Figure 7). Additional wealthy observations provided by a country's rich list and the interval of unobserved households between the HFCS and the rich list indicate how well the HFCS covers the wealth spectrum of households and how severely inequality might be underestimated in the raw HFCS. Across the four countries presented in Table 8 inequality measures increase the most for Italy

[32] The other adjustments of the model have a minor impact on inequality and are omitted for the sake of clarity in order to focus on the relevant aspects.

when capturing the missing wealthy. This is in line with the observation in Figure 7, which indicates that the Italian HFCS misses more wealthy households than the German, French, and Spanish HFCS.

The proportional adjustment allocates the remaining discrepancies between the HFCS and NtlA proportionally across all households, maintaining the inequality measured by the Gini coefficient within each wealth instrument (see Ohlwerter (2021)). The impact of the proportional allocation therefore depends on the heterogeneity of the discrepancies across the wealth instruments and the households' portfolio compositions.

Table 9 shows the coverage of wealth instruments for the raw HFCS and after the inclusion of missing wealthy households.[33] This difference in coverage between instruments can help explain the different changes in inequality in the final proportional allocation. For Germany, after capturing the missing wealthy, we face a substantial over-coverage in business wealth (an instrument typically held by wealthy households) and an under-coverage of deposits, while all other instruments show a good coverage. Business wealth is therefore decreased proportionally across all households that have holdings in business wealth, i.e. the wealthy households. Deposits, on the contrary, is an instrument held by almost all households, and therefore proportionally increased among all of them. As these two instruments represent some of the largest wealth instruments, the impact of the proportional allocation is mostly explained by the changes in these two instruments. Thus, in Germany the proportional allocation leads overall to a decrease in inequality.

[33]The inclusion of missing wealthy households may lead to a coverage above 100% for wealth instruments that are typically held by wealthy households, due to the portfolio composition assumed in Section 2.3 and Table 3. Conversely, even if no gap between HFCS and NtlAs is observed, it appears that a plausible portfolio for the added rich still needs to be added.

Table 9: Coverage ratio by instrument for the four largest economies in the euro area.
Note: The columns "Wealthy added" show the coverage ratios after capturing the missing wealthy (as described in section 2.3).

	Germany		Spain		France		Italy	
	HFCS	Wealthy added	HFCS	Wealthy added	HFCS	Wealthy added	HFCS	Wealthy added
Net wealth	86%	95%	73%	78%	57%	65%	57%	73%
Deposits	51%	55%	46%	49%	45%	51%	34%	44%
Debt securities	42%	101%	15%	101%	22%	102%	37%	100%
Listed shares	63%	101%	98%	101%	46%	102%	37%	100%
Investment fund shares	41%	85%	33%	68%	24%	75%	19%	84%
Life insur. and vol. pensions	60%	79%	86%	109%	32%	38%	8%	35%
Business wealth	124%	158%	48%	58%	45%	64%	19%	50%
Housing wealth	97%	101%	89%	91%	75%	79%	84%	90%
Mortgages	85%	97%	101%	102%	77%	95%	53%	89%
Other liabilities	30%	83%	50%	88%	58%	92%	18%	80%

For Spain and France, the impact of the proportional allocation is less pronounced on the inequality measures due to a more homogeneous distribution of the remaining discrepancies. France, however, also shows a substantial under-coverage of net wealth, which leads to the greatest increase in the number of households with net wealth above EUR 1 million. Compared to Spain and France, Italy faces a more substantial under-coverage of liabilities jointly with an under-coverage in business wealth and life insurance and voluntary pensions (both instruments typically held by wealthy households). This, overall, leads to relative shift in net wealth towards the wealthy, thus showing a greater increase in inequality measures.

3 Development of inequality and wealth structures in the euro area

The methodology presented in Section 2 enables a detailed analysis of the development of inequality and wealth structures over time in a consistent NtlA framework. By linking the granular data set to the macro statistics, the analysis can move beyond measurements of the mean, which is currently all that is possible based on NtlA alone, to provide information on variations in the ownership of instruments within the household sector. It should be highlighted that the country data presented in this section reflect the implementation of the methodology outlined in Section 2, but do not take into account country specific features, for example, weaknesses in a given source for a specific instrument or additional information available at the national level.

Using the quarterly times series distributional data, we compute a number of inequality measures on individual country level and for the euro area. These measures shed light on the evolution of wealth inequality over the period 2009 - 2020 Q3, firstly by providing figures for cross country and intertemporal comparisons, and secondly, by enabling a deeper level of analysis into the causes of these changes and how this impacts the distribution of wealth. Section 3.1 starts by providing a contemporary overview of inequality indicators, including shares of net wealth, percentage of households with net wealth over EUR 1 million, and the Gini coefficient. Section 3.2 proceeds to demonstrate how the derived time series of DNW allows an analysis of the evolution of the Gini coefficient over the period 2009 - 2020 Q3. Section 3.3 identifies which wealth instruments drive the observed increase in inequality, by analysing the portfolio composition of different sectors of the net wealth distribution, and how these have changed over time. Section 3.4 explores the distribution of households by nationality, calculating the representation of countries in different sections of the euro area distribution of net wealth. Lastly, in section 3.5 the euro area distribution is used to produce net wealth per capita for each decile group over the analysed period, showing how growth rates vary across the distribution and which households have been most affected by changes over time.

3.1 Inequality indicators

A first overview of the most common inequality indicators for 2020 Q3 for all euro area countries and the euro area aggregate is provided in Table 10.[34] Overall, about 8.25 million households in the euro area (which is approximately 5.40% of the population) have a net wealth over EUR 1 million. The subset of those wealthy households that belong to the top 1% owns approximately 28% of the total net wealth of the euro area household sector. The sizable disproportion of the distribution of net wealth is further highlighted by the top 10%, which owns 58% of the total net wealth, while the bottom 50% jointly only owns less the 4%. This disparity is also reflected in the Gini coefficient of 0.731 for the euro area.

[34] As explained above, these results have been obtained by applying the standard methodology outlined in Section 2, and do not take into account country specific features, for example weaknesses of a given source for a specific instrument, or additional information available at the national level.
Note that the sum of households with a net wealth above EUR 1 million across all euro area countries equals 8,157,601, which is slightly below the number reported for the euro area. This occurs due to matching the euro area accounts (see section 2.7).

Table 10: Inequality indicators on household level for euro area countries for 2020 Q3
Notes: For example, the top 1% contains the richest 1% of the country's households in terms of net wealth. To ensure comparability of the Gini coefficients, we follow the normalisation proposed by Raffinetti et al. (2015), which ensures that the Gini coefficient is bounded between 0 and 1.

	Number of households with a net wealth above EUR 1 million	Top 1% share of net wealth	Top 5% share of net wealth	Top 10% share of net wealth	Bottom 50% share of net wealth	Net wealth Gini coefficient
AT	205,310	40%	56%	66%	3%	0.785
BE	381,931	33%	50%	59%	6%	0.722
CY	15,223	33%	55%	67%	3%	0.774
DE	2,193,443	22%	42%	56%	2%	0.738
EE	6,372	36%	55%	66%	3%	0.773
ES	1,151,972	21%	23%	41%	8%	0.676
FI	65,658	41%	41%	57%	1%	0.790
FR	2,315,695	27%	44%	56%	5%	0.717
GR	46,484	28%	43%	55%	8%	0.681
IE	148,252	25%	44%	58%	5%	0.720
IT	1,079,974	35%	50%	60%	7%	0.721
LT	8,621	36%	51%	61%	8%	0.696
LU	66,945	24%	41%	54%	6%	0.688
LV	9,829	32%	50%	61%	6%	0.735
MT	7,136	24%	45%	56%	7%	0.681
NL	331,200	29%	47%	59%	3%	0.736
PT	101,899	28%	45%	57%	8%	0.697
SI	11,377	29%	46%	57%	8%	0.684
SK	10,280	14%	32%	45%	10%	0.591
EA	8,268,409	28%	46%	58%	4%	0.731

A measure that everyone can easily relate to and that offers a first insight into the distribution and development of net wealth is the median net wealth per capita, as shown in Figure 10.[35] The median net wealth per capita of the euro area increases from EUR 45,000 in 2009 Q1 to EUR 52,360 in 2020 Q3, with the growth occurring predominantly since 2017. Both the median net wealth per capita and the growth rates can be seen to vary across the euro area countries

[35] In order to compute per capita measures, the household weights are multiplied by the number of household members, while all considered instrument holdings are divided by the number of household members. Since the number of household members of the added wealthy households (see Section 2.3) is not available, we approximate this value by the average of the number of household members in the respective country and year. The average number of household members ranges from 2.02 in Germany (2014 Q3) to 2.85 in Malta (2009 Q1).

over the last decade.

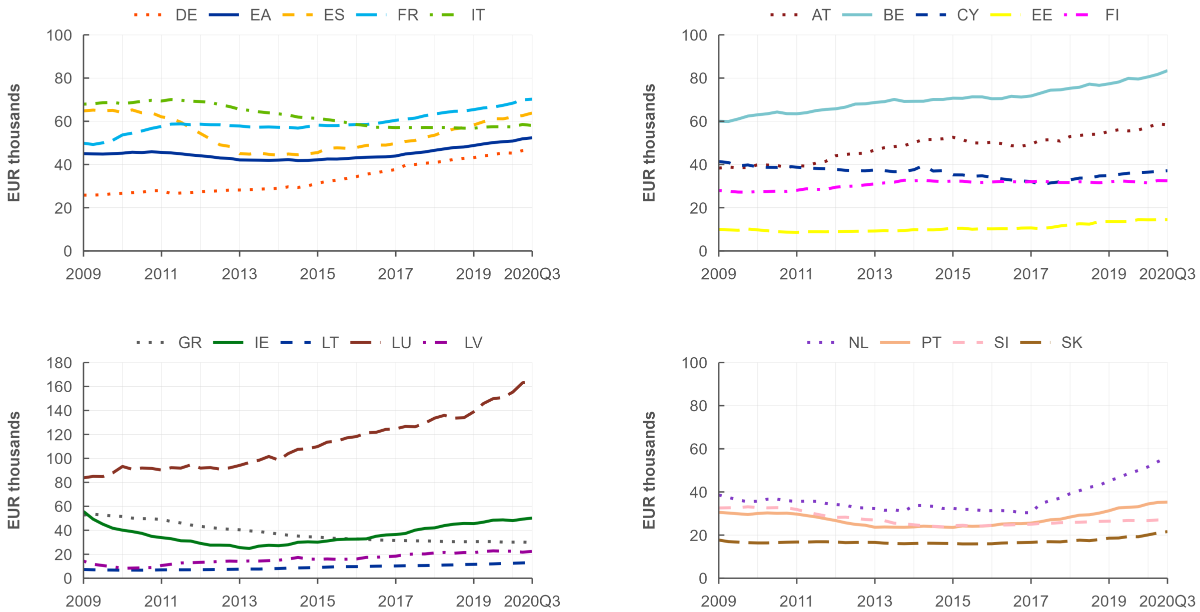

Figure 10: Median net wealth per capita for all euro area countries and the euro area (EA) aggregate from 2009 Q1 to 2020Q3.
Note: The vertical axes display different ranges.

Among France, Germany, Italy, and Spain, only Germany has a median net wealth per capita below the euro area median. Luxembourg is considerably higher than other euro area countries, increasing from EUR 83,760 in 2009 Q1 to EUR 164,100 in 2020 Q3, while Lithuania has a lower net wealth per capita compared with other countries, increasing from EUR 7,250 to EUR 13,000 over the same period.

3.2 Gini coefficients

One of the most well-known inequality measures is the Gini coefficient, which is based on the Lorenz curve and bound by 0 and 1, where a value of 0 is achieved for perfect equality and a value of 1 in the case of perfect inequality. As the net wealth of some households and individuals reaches negative numbers (because they hold more debt than assets), we follow the reformulation of the Gini coefficient by Raffinetti et al. (2015), who derived a normalisation to reestablish the 0 and 1 bounds for the Gini coefficient in the presence of negative numbers, ensuring cross-country comparability. The development of the net wealth per capita Gini coefficients from 2009 Q1 to 2020 Q3 is shown in Figure 11.[36] Overall, we can see that inequality increased in the euro area,

[36] Please see considerations highlighted in footnote 34.

even though several countries saw declines in net wealth inequality.

Table 11 shows the change in the net wealth Gini coefficient for all euro area countries over the observed period and the income Gini coefficients derived by the OECD.[37] While net wealth inequality estimates in 2020 Q3 range between 0.611 and 0.796, income inequality values range between 0.24 and 0.36. The comparatively low income Gini coefficients indicate that higher net wealth inequality is driven more by households' consumption and saving behavior and investment decisions than income earned. Section 3.3 goes on to analyse differences in the portfolio allocation of households in different net wealth quantiles.

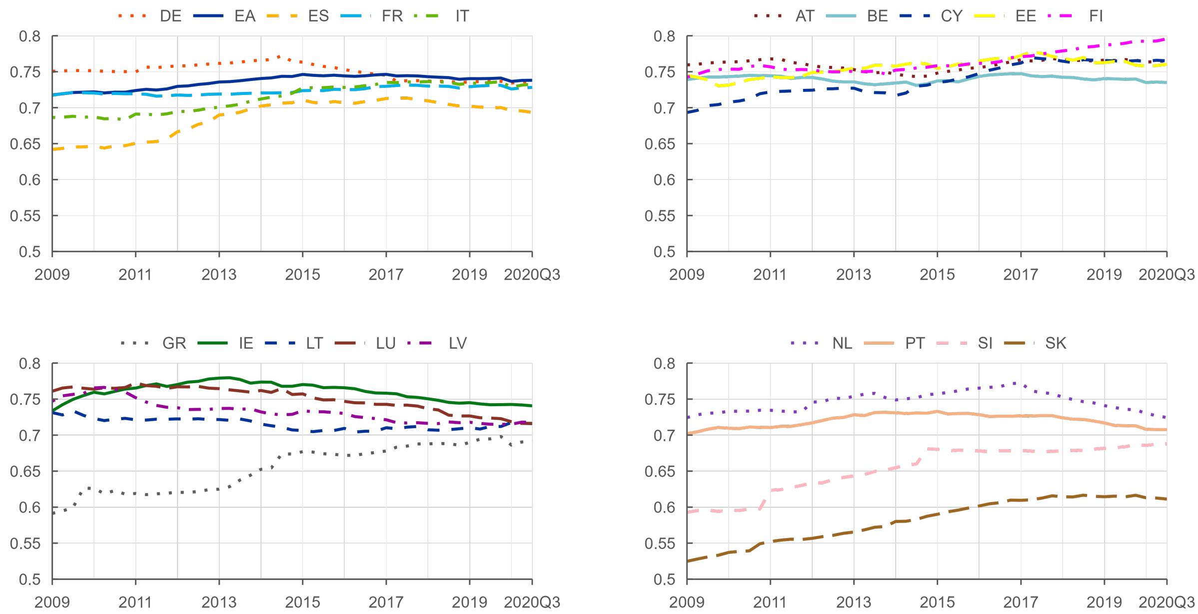

Figure 11: Development of the net wealth per capita Gini coefficients for all euro area countries and the euro area (EA) aggregate from 2009 Q1 to 2020Q3.

The Gini coefficients of the reconciled DNW corroborate findings of related literature. We see that the euro area Gini coefficient increases from the beginning to the end of the observed period. Inequality increases mainly during that period from 2009 to 2015. Thereafter, there is a slow decline in the euro area Gini coefficient, which is not enough to offset the previously observed increase. It should be highlighted that the compiled data set shows significant developments between the period 2017 - 2020, even though the distributional information has been assumed to be stable over this period. As explained in section 2.6, this reflects developments in the NtlA over this period. In particular, the increase in housing prices as from 2017 has benefited homeowners.

[37] See the OECD data warehouse `https://data.oecd.org/inequality/income-inequality.html`.

Taking into account that home ownership is more equally distributed than ownership of financial instruments such as equity, this has led to a slight decrease in inequality in recent years.

We observe an increase of 0.02 between 2009 Q1 and 2020 Q3, compared with a 0.013 increase reported in the HFCS between wave 1 and wave 3. Differences in the growth and the level, which the HFCS shows to increase from 0.682 to 0.695, are affected by differences in the concept of net wealth, in the timing of country-level data, and in the countries included in the HFCS in each wave, as well as in the Pareto adjustment applied in this paper and the final closure of the micro-macro gaps.

Table 11: Changes in the net wealth per capita Gini coefficients from 2009 Q1 to 2020 Q3 in descending order with respect to 2020 Q3.

	Gini coefficient			Income		Gini coefficient			Income
	2009 Q1	2020 Q3	change	inequality		2009 Q1	2020 Q3	change	inequality
FI	0.742	0.796	0.053	0.269 (2018)	NL	0.724	0.724	0.000	0.285 (2016)
AT	0.759	0.765	0.006	0.280 (2018)	LV	0.747	0.718	−0.029	0.351 (2018)
CY	0.693	0.765	0.072	NA	LU	0.761	0.716	−0.045	0.318 (2018)
EE	0.745	0.760	0.016	0.305 (2018)	LT	0.731	0.716	−0.015	0.361 (2018)
IE	0.734	0.741	0.007	0.295 (2017)	MT	0.581	0.709	0.128	NA
EA	0.718	0.738	0.020	NA	PT	0.702	0.708	0.005	0.317 (2018)
DE	0.751	0.736	−0.014	0.289 (2017)	ES	0.642	0.694	0.052	0.330 (2018)
BE	0.739	0.735	−0.004	0.258 (2018)	GR	0.591	0.691	0.100	0.306 (2018)
IT	0.686	0.733	0.046	0.334 (2017)	SI	0.593	0.688	0.095	0.249 (2018)
FR	0.717	0.728	0.011	0.301 (2018)	SK	0.525	0.611	0.086	0.236 (2018)

The Gini coefficients of the reconciled DNW corroborate several findings of related literature. Chakraborty and Waltl (2018) adjust the HFCS (wave 2) for missing wealthy households, finding Gini coefficients of 79.6% for Austria and 79.7% for Germany, where we find 77.2% and 78.2% respectively. Bach et al. (2019) calculate Gini coefficients for Germany, France, and Spain using adjusted HFCS data for wave 1 and wave 2. They find Gini coefficients of 78%, 71%, and 63% in wave 1 and 79%, 74%, and 63% in wave 2 for Germany[38], France and Spain[39] respectively. The DNW estimates fall within 3 percentage points of these results with the exception of Spain in wave 2, where the HFCS data have since been revised.

[38]Schröder et al. (2020) demonstrate that the German SOEP (Socio-Economic Panel) misses the upper tail of the distribution, leading to a clear underestimation of inequality. The authors estimate a Gini coefficient for Germany of 83% when extending the SOEP with additional survey data from surveys conducted on households at the upper tail of the distribution and taking into account the Manager Magazine rich list.

[39]Countries with the best survey data in qualitative terms, i.e. Estonia and Finland, show some of the highest inequality, measured by the Gini coefficient. Both countries use additional national data sources (e.g. credit and business registers) to refine the construction of their HFCS.

3.3 Composition of wealth

As presented in Section 3.2, inequality measures in the euro area have increased over the last decade. However, the approaches used in this paper enable researchers to go beyond these measures of inequality to examine the household data underlying the distribution and explore how variations could explain the causes of these changes. Therefore, in order to better understand the drivers of the increases in inequality, we examine the composition of different sections of the distribution.

Figure 12 shows an overview of the evolution of the different net wealth percentiles for the period from 2009 Q1 to 2020 Q3. While the total net wealth of the euro area economy, represented by the blue line, has increased by around 34%, the net wealth of the bottom 99% has increased by only around 28%; the group with the worst performance within the bottom 99% was the bottom half of the distribution, with an increase of around only 20%. In the meantime, the net wealth of the top 1% of the distribution has increased by almost 50%.

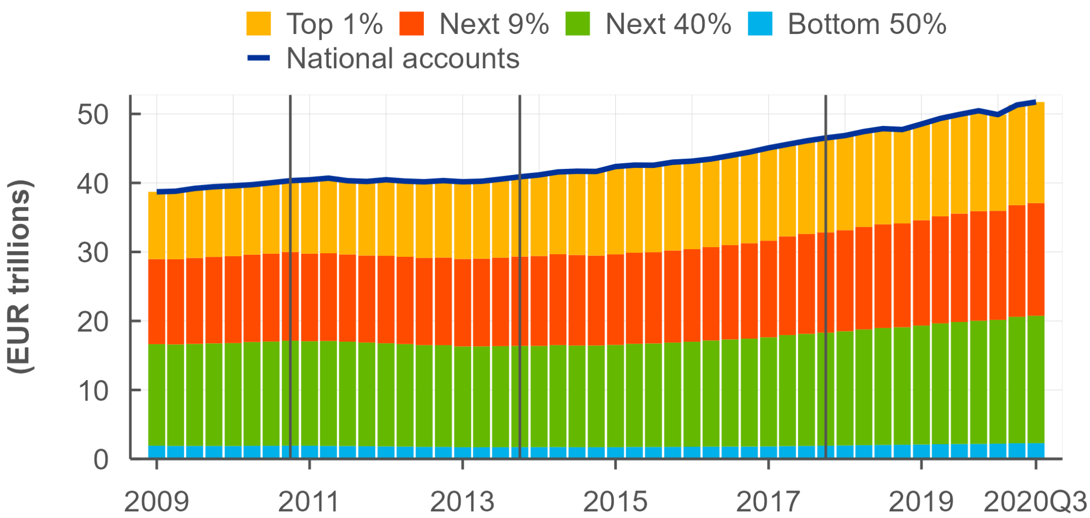

Figure 12: Net wealth of euro area households

A long-run analysis of the United States in the post-war period highlights the systematic difference in the portfolios in the wealth distribution and how this shaped wealth dynamics during that period (see Kuhn et al., 2020). This fact is also observable, to some extent, in the euro area; if we look at the portfolio composition along the net wealth distribution in the euro area, as shown in Figure 13, we can see these differences across the wealth distribution.

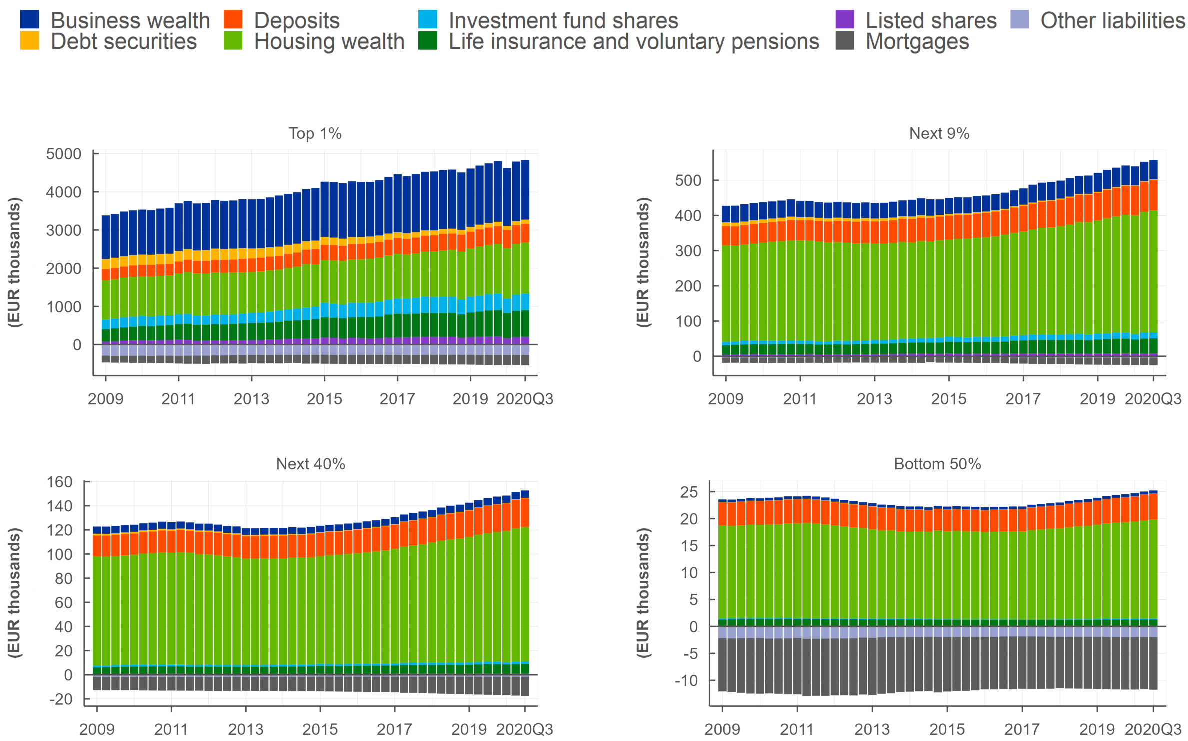

Figure 13: Portfolio composition of an average person, belonging to different net wealth percentiles, in the euro area.

While the portfolios of the top 1% are dominated by corporate and non-corporate equity (e.g. business wealth, investment fund shares, listed shares, etc.)[40], the portfolios of the bottom 99% mainly contain housing wealth and deposits, and this observation is even more pronounced for the bottom 90%.[41] Furthermore, the bottom of the distribution is relatively highly leveraged, especially in the case of the bottom 50%. Looking in closer detail at the portfolio composition of the bottom 20% of the net wealth distribution (see Figure 14), we observe how these individuals have a very large mortgage debt. Liabilities at this end of the distribution can even exceed the value of the assets held. By contrast, the top 10% (see Figure 13) debt holdings are negligible compared with their assets.

In the World Inequality Database, the evolution of portfolios in the household sector over the wealth distribution was analysed by Garbinti et al. (2021) for France in 2012, and by Martínez-Toledano (2017) for Spain in 2015. In both cases, they show how the distribution features an increasing share of net housing (i.e. housing wealth minus total liabilities), with the share

[40]The top 1% of the wealth distribution contains a large portion households from rich lists or added synthetic households and therefore reflects the portfolio allocation applied in Section 2.3.

[41]The bottom 90% refers to the entire population, but the richest 10%.

declining within the top 20% of the net wealth per capita distribution as non-deposit financial assets ownership increases. A key difference between our work and the WID is the treatment of debt, which the WID nets with housing. This causes a lack of housing wealth within the bottom 20% of the distribution. We show that housing wealth within the bottom 10% of the distribution makes up a greater proportion of the portfolio than in the bottom 20%, as seen in Figure 14, but that the liabilities can exceed the total value of assets, resulting in negative wealth.

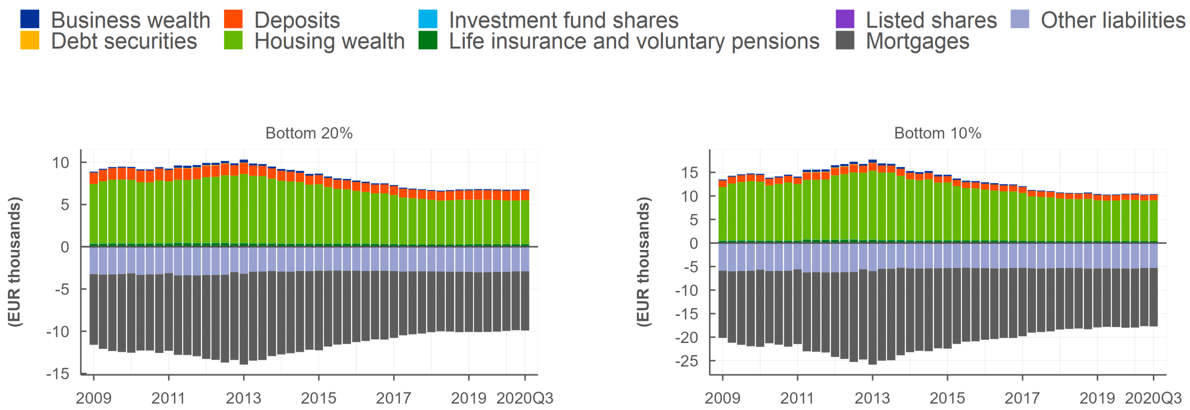

Figure 14: Portfolio composition of an average person, belonging to different net wealth percentiles, in the euro area.

The spread between stock returns and the risk-free rates of the return (i.e. equity premium) has been documented in previous literature (see, for example, Constantinides, 1990; Mehra, 2007). This spread could help explain why the difference in portfolios between the top 1% and the rest of the distribution is driving up inequality in the long run. While for certain assets (equity vs. deposits) this could be explained by the equity premium, for others, such as housing wealth, it is less clear whether housing wealth or equity produces higher returns. However, if housing is financed mainly with mortgages, which are also highly concentrated in the bottom 90%, and they represent about one-third of the housing wealth in the bottom 50%, they could partially offset an increased profitability of housing wealth if it has to be paid with an increased leverage.[42] In Figure 15 we can observe how mortgages are owned mainly by the bottom 90%, which is in line with the previously mentioned higher leverage of these households.

[42] Causa et al. (2019) provide a wide-ranging analysis of the impact of housing wealth on the wealth distribution, evaluating the evidence and implications of ownership and value over the distribution. In addition, Adam and Tzamourani (2016) compare how varying portfolio compositions between the bottom, the middle and the top of the distribution impact the distribution of wealth. Future research could consider comparing the distribution of homeowners to non-homeowners, as shown in Kaas et al. (2019), or examine the equalizing effect for net wealth inequality that homeownership could have, as in Bezrukovs (2013).

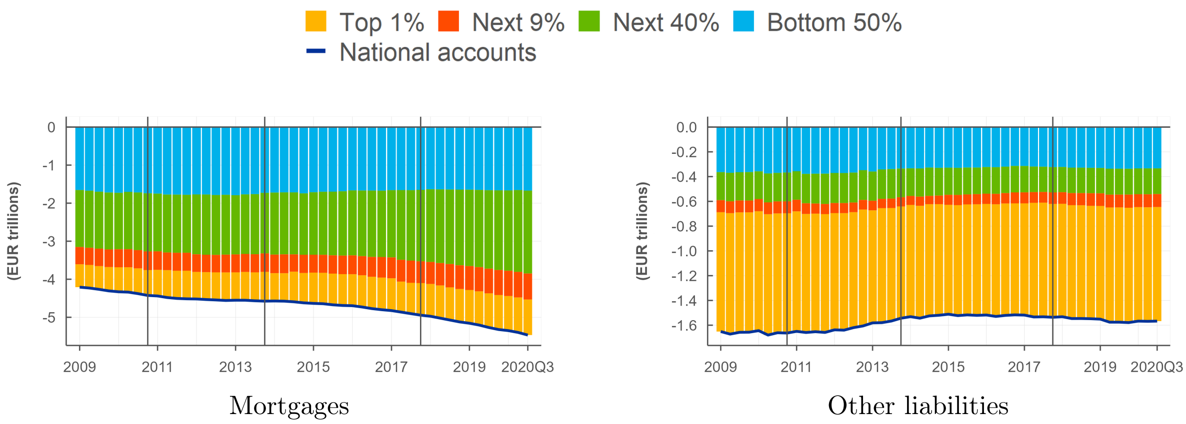

Figure 15: Liabilities by net wealth percentile of euro area households.

There have been many studies on the lower participation in the stock market (see, for example, Vissing-Jørgensen, 2002). In that paper, Vissing-Jørgensen finds that a participation cost of only USD 50 is sufficient to explain the choice to stay out of the stock market of half of those not participating. This might explain the lower participation rates at the bottom of the net wealth distribution. It is also true that the widespread use of the internet and other new technologies has contributed to the so-called "democratisation of finance", and that an increase in trading activity has been documented, for example, in Hvide et al. (2021). However, looking at the portfolio composition of the bottom of the distribution, as shown in Figure 13, increased ease of access to financial markets has not yet resulted in increased participation.[43]

3.4 Distribution of households in the euro area

This section takes a closer look at the representation of euro area member states at different levels of net wealth deciles. A histogram of the net wealth per capita is provided in Figure 16 for 2020 Q3.[44] An accumulation of net wealth can be identified at a value of around EUR 60,000. A second and third accumulation of net wealth, though to a lesser extent, can be identified at a value of zero and at a debt level of around (negative) EUR 3,000.

[43]Future research could be considered along these lines, as the reasons behind the lack of participation in the equity markets are still not clear.

[44]The shape of the distribution stays relatively stable over the analysed period from 2009 Q1 to 2020 Q3.

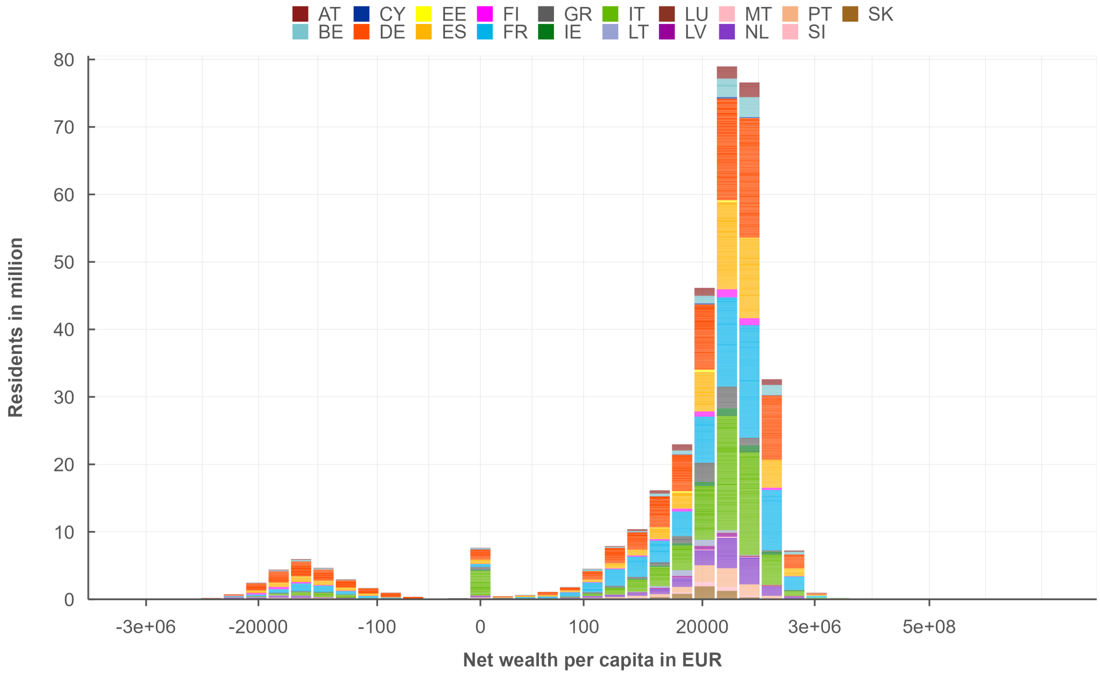

Figure 16: Histogram of net wealth per capita across all euro area countries in 2020 Q3. The horizontal axis is in log scale.

Figure 17 zooms in at different deciles of the distribution and compares the representation of the biggest four euro area economies with their population share. Regarding the richest decile (D10) France shows a clear over-representation (w.r.t. its population share) over the entire observed period. While Germany went from an under-representation in the richest decile in 2009 Q1, meeting its population share in 2014, to an over-representation in 2020 Q3. Italy shows the reverse development, starting from an over-representation in 2009 Q1 regarding the richest decile D10, meeting its population share in 2015, to under-representation in 2020 Q3. Spain went from over-represented in the richest decile in 2009 Q1 to an under-representation in 2013 and is now close to its population share. Within the next four wealthy deciles (D6 to D9), Germany is clearly under-represented, France is close to its population share, and Spain and Italy are clearly over-represented. Moving to the bottom half of the population (D1 to D5), Spain, France and Italy show an under-representation. In contrast, Germany was over-represented in the deciles D2 to D5 until 2016, but is now close to its population share. However, Germany is clearly over-represented in the poorest decile (D1) over the entire observed period.

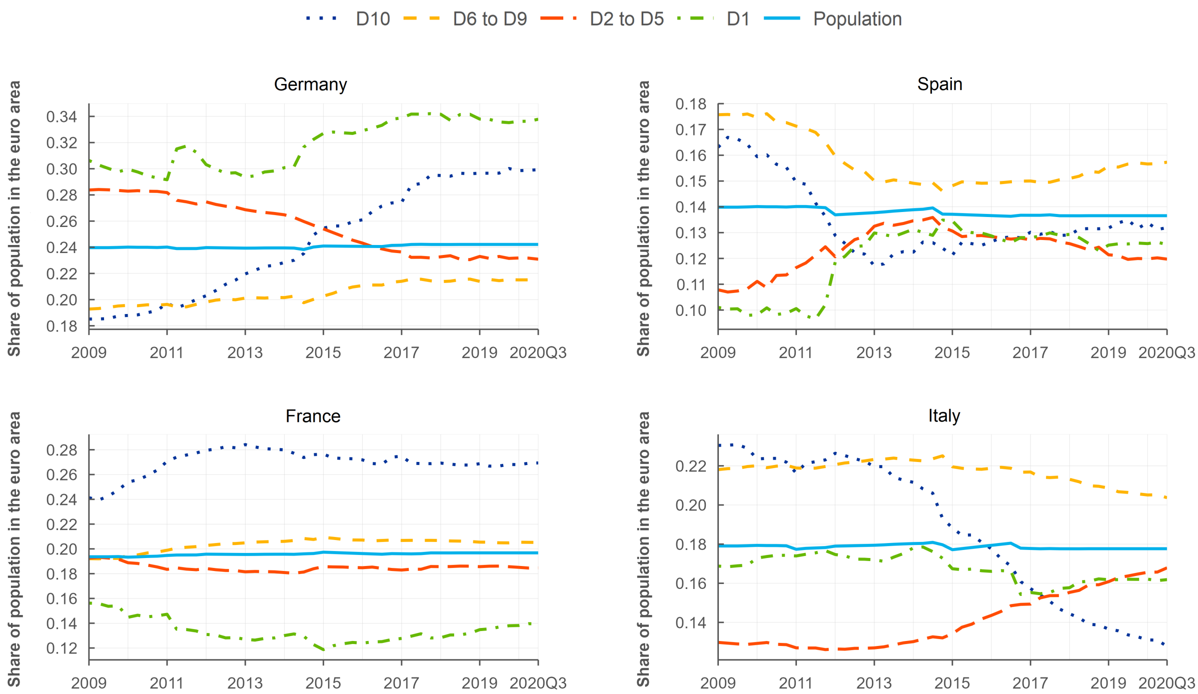

Figure 17: Comparison of the countries' share of population in the euro area with their share in the richest decile (D10), the following four deciles (D6 to D9), the following four deciles (D2 to D5), and the poorest decile (D1).

3.5 Analysis of net wealth decile - entry and exit wealth levels within the distribution

The derived granular data set allows a detailed analysis of the minimum and maximum amount of net wealth that an individual needs to have to belong to each net wealth decile respectively. The minimum, the maximum and the mean net wealth of each decile for the euro area from 2009 Q1 to 2020 Q3 are shown in Figure 18. We see that the median (D5 upper threshold) series starts at around EUR 45,000, while in 2020 Q3 at the end of the series, it has exceeded EUR 52,000. Over the same period, the top 10% (i.e. D10) lower threshold moves from below EUR 240,000 to over EUR 300,000, with similar movements found for the top 20% (i.e. D9 and above) thresholds. Common trends are seen in the majority of threshold and mean series.

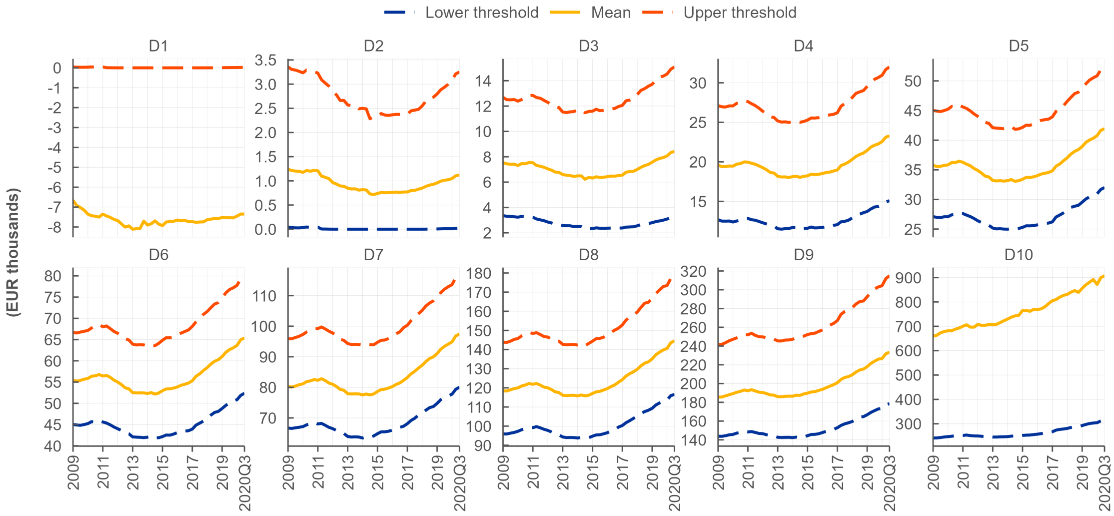

Figure 18: Upper, lower and mean net wealth per capita threshold of the euro area deciles.
Notes: The upper threshold and the lower threshold of the following decile coincide, as they represent the breaking point. The lower threshold for D1 and the upper threshold for D10 are omitted to improve readability of the y-axis.

The majority of the threshold growth series follow a cyclical pattern, mirroring business cycle movements. This is due to households holding large amounts of housing wealth in the portfolios across the distribution. Martínez-Toledano (2017) discusses how the inclusion of housing wealth in measurements of net wealth reduces the concentration of wealth at the top of the distribution and acts to smooth wealth inequality. This was shown in Section 3.3, where we see that, although the value of housing wealth in the top 10% is much greater than in the bottom 10%, the portfolio of the bottom 10% shows a higher concentration of housing wealth. However, this has not translated into equal levels of growth in net wealth.

As discussed in Section 3.3, owing to the different composition of the portfolio along the distribution, households at the top of the net wealth distribution can retain a greater proportion of the increase in value of housing wealth. By contrast, the returns at the bottom of the distribution are offset by the interest to be paid on their liabilities, resulting in very little growth. Low rates of return on deposits compound this effect, as the return on their portfolio of instruments becomes negligible.

This effect can be seen to have a dramatic impact on growth in net wealth at the bottom of the distribution. Different rates of growth suggest a separation within the lower half of the distribution between those at the bottom, who experience no growth in net wealth, and those closer to the upper threshold of D5. We find that the two lowest deciles of individuals have experienced no growth in net wealth in the last decade, with the thresholds of the bottom 20% (i.e. D2 and below) and bottom 10% (i.e. D1) being at a similar level in 2020 Q3 to where they were in 2009 Q1, and with the mean for each decile decreasing over the same period. In Figure 18, we see that the thresholds and mean net wealth for D1 and D2 decreased from 2009 Q1 to 2020 Q3. Moreover, with a mean annual inflation rate of 1.25% in the euro area[45] between 2009 and 2020, inflationary pressures are likely to have driven down net wealth for the majority of households in the lower half of the distribution.

4 Conclusion

Based on the HFCS microeconomic data, reconciled quarterly DNW is derived via the steps presented in this paper: (i) appropriately matching the country-specific fieldwork periods of the HFCS with the quarterly NtlA reference dates; (ii) correcting the differences in the concepts of business and housing wealth between the HFCS and NtlA; (iii) estimating and adding the missing wealthy based on a Pareto distributed upper tail of the net wealth distribution; (iv) aligning the underlying population definitions, as persons living in institutions (e.g. prisons and retirement homes) are included in the NtlA population but excluded from the HFCS; (v) eliminating the remaining discrepancies in the instrument totals between the HFCS and NtlA via a proportional allocation; (vi) constructing a time series of DNW via an inter- and extrapolation model; and (vii) computing euro area aggregates from the national results. A useful feature of the presented approach is that it relies almost completely on publicly available data, besides the HFCS, which is however available to researchers upon request via the Household Finance and Consumption Network (HFCN).[46]

The proposed method derives a granular quarterly micro data set that is in line with NtlA aggregates and enables a detailed analysis of per household and per capita net wealth, including the portfolio composition, the development over time, and the evolution of various inequality indicators. The method can also be expanded to incorporate further data sources and refined using country-specific expertise and information on households, macro aggregates or the missing wealthy households. In addition to presenting the developed methodology, this paper provides a first analysis, leading to the following insights.

[45] Monthly updates of the data can be found at: `https://sdw.ecb.europa.eu/reports.do?node=10000051`.

[46] See `https://www.ecb.europa.eu/pub/economic-research/research-networks/html/researcher_hfcn.en.html`

First, our results corroborate the general perception and findings of related literature with regard to a high level of inequality. The (per capita) Gini coefficient of the euro area amounts to 0.738 in 2020 Q3. The richest 1% of households own 28% of the total net wealth of the euro area household sector, the richest 10% own 58% of the total net wealth, while the bottom 50% jointly own less than 4% of the total net wealth.

Second, measuring net wealth inequality in terms of the Gini coefficient shows that inequality increased in the majority of euro area countries (14 out of 19) as well as in the euro area aggregate, over the observed window from 2009 Q1 to 2020 Q3.

Third, to gain further insight into these findings, the composition of wealth at different percentiles is subsequently analysed. We find that over the last decade the top 1% of households increased their net wealth by almost 50%, while the remaining 99% of households saw only a 28% increase in net wealth. The net wealth of the bottom 50% increased by only 20%. Overall, housing wealth is the largest asset group held by the household sector for most countries, representing around 50% of the total assets. A further analysis of the portfolio composition of individuals in different net wealth percentiles reveals that the observed disparity of net wealth can be explained by differences in investment behaviour. The portfolio of the richest 1% is dominated by equity investments (e.g. business wealth, investment fund shares, listed shares). Moving towards the lower percentiles, the portfolio instead becomes dominated by deposits and housing wealth, while at the same time being increasingly leveraged with mortgages. Within the bottom 50% of individuals, financial assets other than deposits are almost entirely absent. Therefore, the spread between stock returns and the risk-free rates of return leads to the wealthy further increasing their wealth, while the poorer households see a smaller increase in net wealth, which in turn results in an increase in inequality.

Fourth, by calculating the minimum, the maximum and the mean net wealth of each decile of the euro area net wealth distribution, we find that there are differences in the growth rates of net wealth over the distribution. We find that the net wealth for the bottom 20% of the net wealth distribution decreased between 2009 Q1 and 2020 Q3. This contrasts with the positive growth seen elsewhere in the distribution, with the median net wealth growing by 16% over the same period, while the top 20% and top 10% thresholds grew by 25% and 31% respectively. This shows that growth in household sector net wealth has been concentrated in the upper areas of the distribution.

The construction of regular and timely reconciled DNW goes some way to fulfilling the request of policymakers, regulators and politicians for distributional statistics and greatly expands the use of NtlA data. The European System of Central Bank's development of distributional statistics provides further results on the evolution and the drivers of inequality in the household sector and across euro area countries that are consistent with the estimates of aggregate wealth in the national accounts. By providing these additional dimensions to regular reporting, economic models can be expanded upon to incorporate the distribution of wealth, improving the possibilities for analysis and providing valuable information on the impact of policies on inequality.

References

Adam, K. and Tzamourani, P. (2016). Distributional consequences of asset price inflation in the Euro Area. *European Economic Review*, 89(C):172–192.

Adam, K. and Zhu, J. (2016). Price-level changes and the redistribution of nominal wealth across the euro area. *Journal of the European Economic Association*, 14(4):871 – 906.

Ampudia, M., Pavlickova, A., Slačálek, J., and Vogel, E. (2014). Household heterogeneity in the euro area since the onset of the great recession. *Working Paper Series* 1705, European Central Bank.

Arrondel, L., Lamarche, P., and Savignac, F. (2015). Wealth effects on consumption across the wealth distribution: empirical evidence. *Working Paper Series* 1817, European Central Bank.

Atkinson, A. B. (2017). Pareto and the upper tail of the income distribution in the UK: 1799 to the present. *Economica*, 84(334):129–156.

Bach, S., Thiemann, A., and Zucco, A. (2019). Looking for the missing rich: tracing the top tail of the wealth distribution. *International Tax and Public Finance*, 26(6):1234–1258.

Batty, M., Bricker, J., Briggs, J., Friedman, S., Nemschoff, D., Nielsen, E., and Volz, A. H. (2020). The Distributional Financial Accounts of the United States. Measuring and understanding the distribution and intra/inter-generational mobility of income and wealth. *University of Chicago Press*.

Batty, M., Bricker, J., Briggs, J., Holmquist, E., McIntosh, S., Moore, K., Nielsen, E., Reber, S., Shatto, M., Sommer, K., Sweeney, T., and Henriques Volz, A. (2019). Introducing the Distributional Financial Accounts of the United States. *Finance and Economics Discussion Series* 017, Board of Governors of the Federal Reserve System, United States.

Bezrukovs, D. (2013). The role of housing in wealth inequality in Eurozone countries.

Blanchet, T., Chancel, L., and Gethin, A. (2019). How unequal is Europe? Evidence from Distributional National Accounts. *Working Paper* 6, World Inequality Database.

Blanchet, T., Fournier, J., and Piketty, T. (2017). Generalized Pareto curves: theory and applications. *Working Paper* 3, World Inequality Database.

Blanchet, T., Garbinti, B., Goupille-Lebret, J., and Martínez-Toledano, C. (2018). Applying generalized Pareto curves to inequality analysis. *AEA Papers and Proceedings*, 108:114–118.

Bruil, A. and Koymans, M. (2014). Measuring inequalities in the Dutch household sector. *Discussion Paper* 13, Statistics Netherlands.

Cantarella, M., Neri, A., and Ranalli, M. G. (2021). Mind the wealth gap: a new allocation method to match micro and macro statistics for household wealth. *Occasional Papers (Questioni di economia e finanza)*, (646).

Carroll, C. D., Slačálek, J., and Tokuoka, K. (2014). The distribution of wealth and the MPC: implications of new European data. *American Economic Review*, 104(5):107 – 11.

Causa, O., Woloszko, N., and Leite, D. (2019). Housing, wealth accumulation and wealth distribution: Evidence and stylized facts. *Working Papers* 1588, Organisation for Economic Co-operation and Development.

Chakraborty, R., Kavonius, I. K., Perez-Duarte, S., and Vermeulen, P. (2019). Is the top tail of the wealth distribution the missing link between the Household Finance and Consumption Survey and national accounts? *Journal of Official Statistics*, 35(1):31–65.

Chakraborty, R. and Waltl, S. R. (2018). Missing the wealthy in the HFCS: micro problems with macro implications. *Working Paper Series* 2163, European Central Bank.

Chow, G. C. and Lin, A. (1971). Best linear unbiased interpolation, distribution, and extrapolation of time series by related series. *The Review of Economics and Statistics*, 53(4):372 – 375.

Constantinides, G. M. (1990). Habit formation: a resolution of the equity premium puzzle. *Journal of Political Economy*, 98(3):519–543.

Credit Suisse Research Institute (2019). Global Wealth Report 2019. Credit Suisse Group AG.

Cussen, M., Lydon, R., and O'Sullivan, C. (2018). Macro and micro estimates of household wealth. *Research Technical Papers* 11, Central Bank of Ireland.

Disslbacher, F., Ertl, M., List, E., Mokre, P., and Schnetzer, M. (2020). On top of the top-adjusting wealth distributions using national rich lists. *INEQ Working Paper Series* 20, Vienna University of Economics and Business.

Dossche, M., Slačálek, J., and Wolswijk, G. (2021). Monetary policy and inequality. *Economic Bulletin*, European Central Bank, Frankfurt am Main.

Expert Group on Linking macro and micro data for the household sector (2020). Understanding household wealth: linking macro and micro data to produce distributional financial accounts. *Statistical Paper Series* 37, European Central Bank.

FSB and IMF (2009). The Financial Crisis and Information Gaps. Basel/Washington DC, October.

FSB and IMF (2019). G20 Data Gaps Initiative (DGI-2) The Fourth Progress Report—Countdown to 2021. Basel/Washington DC, October.

Garbinti, B., Goupille-Lebret, J., and Piketty, T. (2021). Accounting for wealth-inequality dynamics: methods, estimates, and simulations for France. *Journal of the European Economic Association*, 19(1):620–663.

Honkkila, J., Kavonius, I. K., and Reynaert Lefebvre, L. (2018). Linking macro and micro household balance sheet data - time series estimation . *The 35th General Conference of the International Association for Research in Income and Wealth, Copenhagen.*

Household Finance and Consumption Network (2020a). The Household Finance and Consumption Survey: methodological report for the 2017 wave. *Statistics Paper Series* 35, European Central Bank.

Household Finance and Consumption Network (2020b). The Household Finance and Consumption Survey: results from the 2017 wave. *Statistics Paper Series* 36, European Central Bank.

Hvide, H. K., Meling, T. G., Mogstad, M., and Vestad, O. L. (2021). Broadband internet and the stock market investments of individual investors. *Discussion Papers* 946, Statistics Norway, Research Department.

Kaas, L., Kocharkov, G., and Preugschat, E. (2019). Wealth inequality and homeownership in europe. *Annals of Economics and Statistics*, (136):27–54.

Kavonius, I. K. and Honkkila, J. (2013). Reconciling micro and macro data on household wealth: a test based on three euro area countries. *Journal of Economic and Social Policy*, 15(2).

Kavonius, I. K. and Honkkila, J. (2016). Deriving household indebtedness indicators by linking micro and macro balance sheet data. *Statistical Journal of the IAOS*, 32(4):693 – 708.

Kuhn, M., Schularick, M., and Steins, U. I. (2020). Income and wealth inequality in America, 1949–2016. *Journal of Political Economy*, 128(9):3469–3519.

Lakner, C. and Milanovic, B. (2016). Global income distribution: from the fall of the Berlin wall to the Great Recession. *World Bank Economic Review*, 30(2):203–232.

Lenza, M. and Slačálek, J. (2018). How does monetary policy affect income and wealth inequality? Evidence from quantitative easing in the euro area. *Working Paper Series* 2190, European Central Bank.

Martínez-Toledano, C. (2017). Housing bubbles, offshore assets and wealth inequality in Spain (1984-2013). *Working Paper*, World Inequality Database.

Mathä, T. Y., Porpiglia, A., and Ziegelmeyer, M. (2018). Wealth differences across borders and the effect of real estate price dynamics: evidence from two household surveys. *Journal of Income Distribution*, 26(1):1–35.

Mehra, R. (2007). The equity premium puzzle: a review. *Foundations and Trends(R) in Finance*, 2(1):1–81.

Meriküll, J. and Rõõm, T. (2021). Are survey data underestimating the inequality of wealth? *Empirical Economics*.

Neri, A. and Ranalli, M. G. (2012). To misreport or not to report? The measurement of household financial wealth. *Temi di Discussione (Working Papers)* 870, Banca d'Italia.

OECD (2013). *OECD Guidelines for Micro Statistics on Household Wealth*.

OECD (2017). *Understanding Financial Accounts*.

Ohlwerter, D. (2021). Contributing to estimating the distribution of household wealth - solving under-reporting via optimization problems with invariant Gini coefficient, [unpublished manuscript].

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

Piketty, T., Postel-Vinay, G., and Rosenthal, J. (2006). Wealth concentration in a developing economy: Paris and France, 1807-1994. *American Economic Review*, 96(1):236 – 256.

Piketty, T. and Saez, E. (2003). Income inequality in the United States, 1913–1998. *The Quarterly Journal of Economics*, 118(1):1–41.

Piketty, T. and Zucman, G. (2014). Capital is back: wealth-income ratios in rich countries 1700–2010. *The Quarterly Journal of Economics*, 129(3):1255 – 1310.

Raffinetti, E., Siletti, E., and Vernizzi, A. (2015). On the Gini coefficient normalization when attributes with negative values are considered. *Statistical Methods & Applications*, 24(3):507–521.

Ruiz, N. and Woloszko, N. (2016). What do household surveys suggest about the top 1% incomes and inequality in OECD countries? *Working Papers* 1265, Organisation for Economic Co-operation and Development.

Schröder, C., Bartels, C., Göbler, K., Grabka, M. M., and König, J. (2020). MillionärInnen unter dem Mikroskop: Datenlücke bei sehr hohen Vermögen geschlossen - Konzentration höher als bisher ausgewiesen. *DIW Wochenbericht*, 87(29):511–521.

The Economist (2018). How the super-rich invest - the rise of the family office. *The Economist 15 December*, pages 22 – 24.

Törmälehto, Veli-Matti (2019). Reconciliation of EU statistics on income and living conditions (EU-SILC) data with national accounts. *Statistical Working Papers*, Eurostat.

UBS/Campden Research (2018). The Global Family Office Report.

Vermeulen, P. (2016). Estimating the top tail of the wealth distribution. *American Economic Review*, 106(5):646 – 50.

Vermeulen, P. (2018). How fat is the top tail of the wealth distribution? *Review of Income and Wealth*, 64(2):357 – 387.

Vissing-Jørgensen, A. (2002). Towards an explanation of household portfolio choice heterogeneity: non-financial income and participation cost structures. *Working Papers* 8884, National Bureau of Economic Research.

Vogel, E. (2014). Optimal level of government debt-matching wealth inequality and the fiscal sector. *Working Paper Series* 1665, European Central Bank.

Waltl, S. R. (2020). Multidimensional wealth inequality: a hybrid approach toward distributional national accounts in Europe. *NEQ Working Paper Series*, Vienna University of Economics and Business.

Zwijnenburg, J., Bournot, S., and Giovannelli, F. (2017). Expert Group on Disparities in a National Accounts Framework: Results from the 2015 Exercise. *Statistics Paper Series* 2016/10, Organisation for Economic Co-operation and Development.

Annexes

A Reference quarter identification methods

There are two main valuation methods used when compiling the HFCS results. Table 12 gives a detailed picture of the methods and the reference quarter.

Table 12: Identification method for determining the reference quarter that is closest to the HFCS fieldwork period by wave and country.

	Wave 1		Wave 2		Wave 3	
	Reference quarter	Identification method	Reference quarter	Identification method	Reference quarter	Identification method
AT	2010 Q4	Mid-point quarter	2014 Q4	Mid-point quarter	2017 Q1	Mid-point quarter
BE	2010 Q3	Mid-point quarter	2014 Q3	Mid-point quarter	2017 Q2	Mid-point quarter
CY	2010 Q3	Mid-point quarter	2014 Q2	Mid-point quarter	2017 Q2	Mid-point quarter
DE	2011 Q1	Mid-point quarter	2014 Q3	Mid-point quarter	2017 Q2	Mid-point quarter
EE	-	-	2013 Q2	Mid-point quarter	2017 Q2	Mid-point quarter
ES	2011 Q4	Mid-point quarter	2014 Q4	Mid-point quarter	2017 Q4	Mid-point quarter
FI	2009 Q4	Valuation date	2013 Q4	Valuation date	2016 Q4	Valuation date
FR	2009 Q4	Mid-point quarter	2014 Q4	Mid-point quarter	2017 Q4	Mid-point quarter
GR	2009 Q3	Mid-point quarter	2014 Q3	Mid-point quarter	2018 Q2	Mid-point quarter
IE	-	-	2013 Q2	Mid-point quarter	2018 Q3	Mid-point quarter
IT	2010 Q4	Valuation date	2014 Q4	Valuation date	2016 Q4	Valuation date
LT	-	-	-	-	2016 Q4	Valuation date
LU	2010 Q4	Mid-point quarter	2014 Q3	Mid-point quarter	2018 Q3	Mid-point quarter
LV	-	-	2014 Q2	Mid-point quarter	2017 Q3	Mid-point quarter
MT	2010 Q4	Mid-point quarter	2013 Q4	Valuation date	2016 Q4	Valuation date
NL	2009 Q4	Valuation date	2013 Q4	Valuation date	2017 Q2	Mid-point quarter
PT	2010 Q2	Mid-point quarter	2013 Q2	Mid-point quarter	2017 Q3	Mid-point quarter
SI	2010 Q4	Mid-point quarter	2014 Q4	Mid-point quarter	2017 Q3	Mid-point quarter
SK	2010 Q3	Mid-point quarter	2014 Q1	Mid-point quarter	2017 Q1	Mid-point quarter

Of the countries listed in Table 12, only Italy, Finland, and Lithuania consistently use valuation dates that coincide with the end of a quarter. Malta and the Netherlands also use valuation dates but have at least one wave in which, a mid-point reference date is used to link the HFCS to the NtlA.

Using the variation in valuation methods for the survey data allows us to examine whether coverage ratios are greater when valuation dates are used rather than mid-point reference dates. A comparison of all countries using a valuation date with all countries using the mid-point quarter, as in Figure 19, shows that having a fixed date on which assets are valued does not systematically improve asset coverage. This would suggest that the coverage ratio is not affected by synchronising the valuation dates of the survey data and the NtlA data, and so the mid-point quarter method is likely to produce similar results to those countries for which a valuation date is specified for instrument holdings.

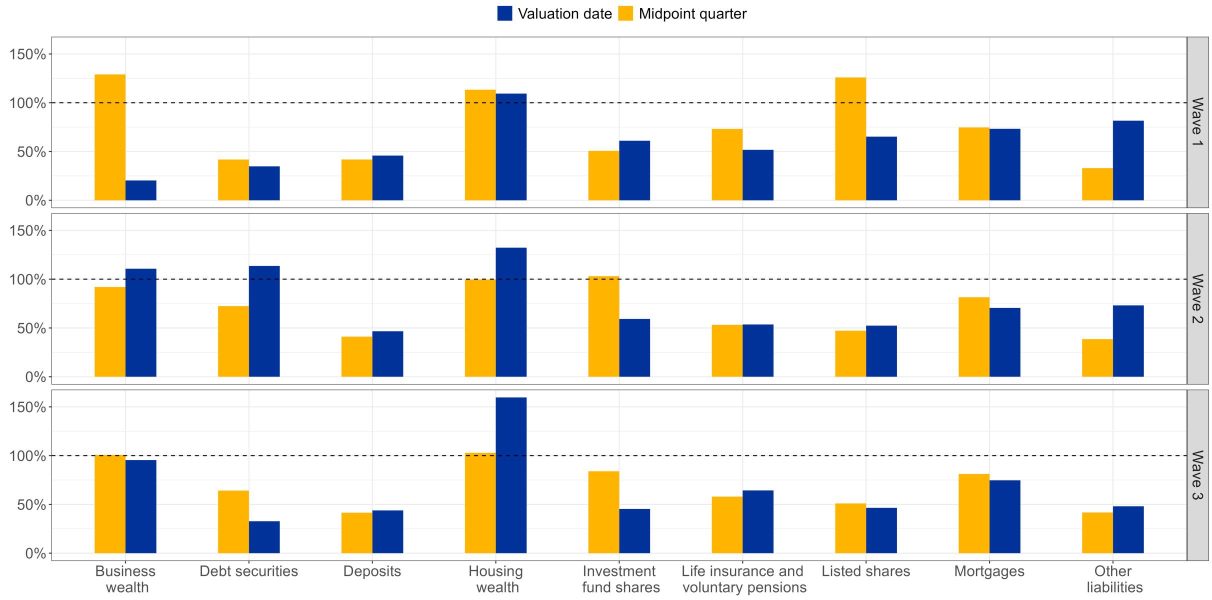

Figure 19: Coverage ratios by date selection method.

We also consider Malta and the Netherlands as these two countries use different valuation methods for different waves. This controls for discrepancies at the national level in the survey and NtlA administration, but not for changes across time. Looking at Figure 20, we can see again that there is no pattern that would suggest that one method outperforms the other. We therefore conclude that the timing in the HFCS does not systematically affect the coverage ratio.

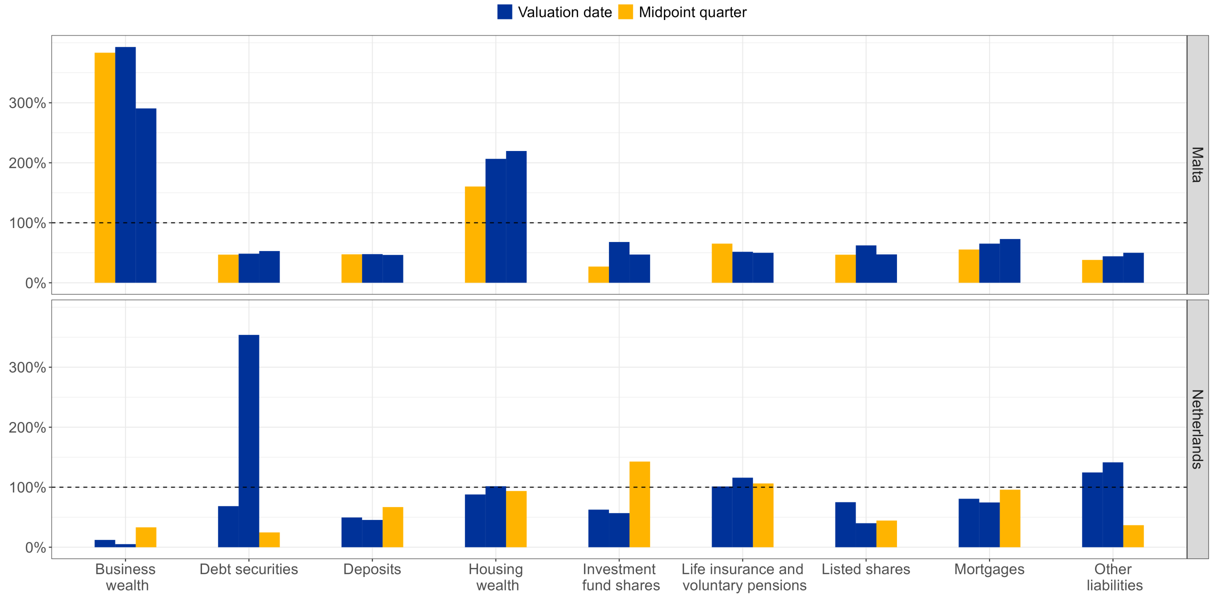

Figure 20: Coverage ratios by date selection method for countries that change methods.

Having concluded that the selection of the NtlA quarter corresponding to the HFCS has no significant impact on the coverage ratio, a final test is conducted that looks at the maximum possible range of values for the coverage ratio owing to variation in the NtlA quarter. For this purpose, the coverage ratio for every quarter covered by the HFCS fieldwork period for each country in each wave is computed, which provides a sensitivity analysis for the coverage ratio. The results in Figure 21 suggest there is little evidence that changing the selected date of the NtlA quarters to other quarters in the fieldwork period would have a substantial effect on the coverage ratio. While NtlA variation over time means that dates close or within the fieldwork period should be used, the evidence presented suggest that choices between the values reported within these periods have very little impact on the coverage ratio.

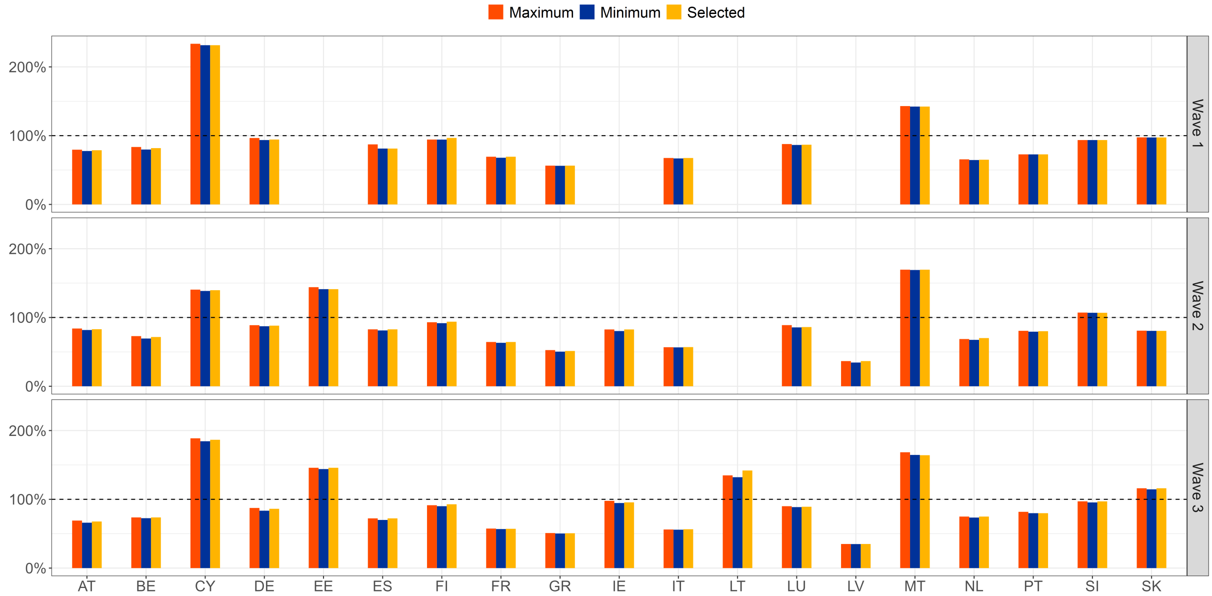

Figure 21: Selected, maximum and minimum possible net wealth coverage ratios for fieldwork period quarters by wave and country.
Note: For those selected reference quarters that refer to a valuation date, the coverage ratio may be below the minimum or above the maximum coverage ration computed for the field work period, as the valuation date may be outside the fieldwork period (see Figure 1).

Overall, the analysis presented shows that the use of valuation or mid-point reference dates does not affect the coverage ratio. Furthermore, selecting other end quarter values during the fieldwork period has no significant effect on the net wealth coverage ratio.

B Aligning the population

In order to have a consistent definition of the covered population, a proportional adjustment is applied by rescaling all household weights. More precisely, let $\mathbf{d} = (d_1, \ldots, d_n) \in \mathbb{R}^n_{>0}$ denote the household weights of the n households participating in the HFCS in a certain country. Moreover, let $\mathbf{m} = (m_1, \ldots, m_n) \in \mathbb{N}^n_{>0}$ denote the corresponding number of household members (i.e. the number of people within each household) provided by the HFCS and P_{NtlA} the population total given by the NtlA. The new household weights $\mathbf{d}^{(\text{new})} = \left(d_1^{(\text{new})}, \ldots, d_n^{(\text{new})}\right) \in \mathbb{R}^n_{>0}$ are computed by

$$d_i^{(\text{new})} = d_i \frac{P_{\text{NtlA}}}{\sum_{j=1}^n d_i m_i}, \qquad \text{for } i = 1, \ldots, n. \tag{28}$$

Table 13 provides an overview of the population totals in the HFCS and NtlA at the time periods of the three HFCS waves.[47]

Table 13: Population totals of the HFCS and the NtlA, in thousands, as well as the coverage ratio.

	Wave 1			Wave 2			Wave 3		
	HFCS	NtlA	HFCS/NtlA	HFCS	NtlA	HFCS/NtlA	HFCS	NtlA	HFCS/NtlA
AT	8,022	8,371	95.8%	8,263	8,576	96.4%	8,403	8,777	95.7%
BE	10,819	10,908	99.2%	11,151	11,214	99.4%	11,321	11,368	99.6%
CY	837	831	100.7%	837	845	99.1%	837	858	97.6%
DE	81,086	80,229	101.1%	80,042	81,024	98.8%	81,693	82,622	98.9%
EE	-	-	-	1,286	1,320	97.4%	1,302	1,316	98.9%
ES	45,931	46,813	98.1%	45,993	46,420	99.1%	46,163	46,593	99.1%
FI	5,272	5,349	98.6%	5,371	5,450	98.6%	5,419	5,502	98.5%
FR	62,464	64,812	96.4%	64,398	66,420	97.0%	65,527	67,143	97.6%
GR	10,861	11,110	97.8%	10,735	10,884	98.6%	10,476	10,735	97.6%
IE	-	-	-	4,594	4,615	99.5%	4,843	4,869	99.5%
IT	60,310	59,928	100.6%	60,783	60,801	100.0%	60,243	60,610	99.4%
LT	-	-	-	-	-	-	2,677	2,868	93.3%
LU	463	512	90.4%	508	560	90.7%	544	611	89.0%
LV	-	-	-	1,975	1,997	98.9%	1,925	1,939	99.3%
MT	410	415	98.8%	416	429	97.0%	431	459	93.9%
NL	16,366	16,568	98.8%	16,580	16,829	98.5%	16,833	17,112	98.4%
PT	10,638	10,570	100.6%	10,487	10,464	100.2%	10,310	10,296	100.1%
SI	2,000	2,050	97.6%	2,063	2,063	100.0%	2,067	2,066	100.0%
SK	5,412	5,432	99.6%	5,217	5,416	96.3%	5,257	5,435	96.7%

C Numerical example – housing and business wealth refinements

This annex presents a numerical example to illustrate the housing and business wealth refinements.[48]

C.1 Housing wealth

Our starting point is equation 1. In order to provide a more illustrative example, let us assign a set of numbers to the different variables that are available either from NtlA, the HFCS, or the added wealthy households:

[47]In rare cases the NtlA population can be slightly lower than the HFCS population, this could be due to inconsistencies in the population data sources used at national level.

[48]The numbers are not linked to any country. They are only used as an illustration of the method.

	Housing wealth *HFCS*	Housing wealth *Wealthy households*[49]	Dwellings *NtlA – S.1M*	Land *NtlA – S.1M*	Real estate abroad *NtlA*
Wave 1	800	100	700	400	10
Wave 2	850	120	730	420	11
Wave 3	900	125	760	450	11

According to the assumptions and constraints derived in Section 2.2.1, we have that $b = 0.99a$, $c = 0.86$, $d = 0.85$ and $a \in [0.96, 1]$. With that information, we can replace equation 1 with the example values to obtain the following:

$$\begin{bmatrix} Gap_1 \\ Gap_2 \\ Gap_3 \end{bmatrix} = \underbrace{\begin{bmatrix} 1,040.56 \\ 1,087.588 \\ 1,143.13 \end{bmatrix}}_{X} a - \underbrace{\begin{bmatrix} 891.50 \\ 960.65 \\ 1,015.65 \end{bmatrix}}_{Y}$$

With the aim of minimizing the gap for all waves and a constant value of a, the problem can be thought of as a constraint least squares problem with the following functional form:

$$\begin{aligned} \min_{a} \quad & L(a) = \sum_{i=1}^{3} (Gap_i)^2 = \sum_{i=1}^{3} (x_i a - y_i)^2 \\ \text{s.t.} \quad & a \in [0.96, 1] \end{aligned} \tag{29}$$

The unconstrained solution of equation 29 would be $\hat{a} \approx 0.88$. This result would imply that NPISH own approximately 12% of the dwellings in the economy. This value, given the available data from several countries, is not plausible. Thus, we introduce the constraint and our new estimation takes the corner solution of $\hat{a} = 0.96$.

C.2 Business wealth

Analogous to the previous example in which we derive an estimation of the dwellings owned by NPISH, we can also provide an example in which we estimate the amount of non-financial assets owned by NPISH, as stated in equation 2. In order to do that, let us assign a set of numbers (in addition to those previously defined for housing wealth) to the different variables that are available either from NtlA, the HFCS or the added wealthy households:

[49]For more information on the method used to add wealthy households see Section 2.3

	Business wealth *HFCS*	Business wealth *Wealthy households*[50]	Non-financial assets *NtlA – S.1M*	Unlisted shares and other equity *NtlA – S.14*
Wave 1	300	50	1,000	40
Wave 2	320	60	1,030	50
Wave 3	330	65	1,040	55

According to the assumptions and constraints derived in Section 2.2.2, we have that $b = 0.99a$, $c = 0.86$, $d = 0.85$ and $e \in [0.95, 1]$. In addition, we can use the value of $a = 0.96$ derived in the example for the housing wealth refinement. With that information, we can replace equation 2 with the example values to obtain the following:

$$\begin{bmatrix} Gap_1 \\ Gap_2 \\ Gap_3 \end{bmatrix} = \underbrace{\begin{bmatrix} 1,000 \\ 1,030 \\ 1,040 \end{bmatrix}}_{X} e - \underbrace{\begin{bmatrix} 937.2776 \\ 984.26648 \\ 1,019.0748 \end{bmatrix}}_{Y}$$

With the aim of minimizing the gap for all waves and a constant value of e, the problem can be, once again, thought of as a constraint least squares problem with the following functional form:

$$\begin{aligned} \min_{e} \quad & L(e) = \sum_{i=1}^{3} (Gap_i)^2 = \sum_{i=1}^{3} (x_i e - y_i)^2 \\ \text{s.t.} \quad & e \in [0.95, 1] \end{aligned} \tag{30}$$

The unconstrained solution of equation 30 would be $\hat{e} \approx 0.96$. This result would imply that NPISH own approximately 4% of the non-financial assets in the economy. This value, given the available data from several countries, is plausible and, thus, the constraint is not binding. The solution to the problem in this case is $\hat{e} \approx 0.96$.

[50]For more information on the method used to add wealthy households see Section 2.3

D Euro area instrument by wealth decile

The evolution of net wealth is presented in Figure 22 for the euro area aggregate. The adjusted micro data set was converted from a household unit to a per capita unit.[51] Moreover, the population is split into deciles according to their net wealth (i.e. each decile comprises one-tenth of the population). The poorest decile is denoted by D1 and represents the average net wealth of an individual belonging to the poorest decile. Likewise, the richest decile is denoted by D10 and represents the average net wealth of an individual belonging to the richest decile. The share of net wealth of the richest decile increased more over time than in the case of the other deciles, which indicates an increase in inequality.

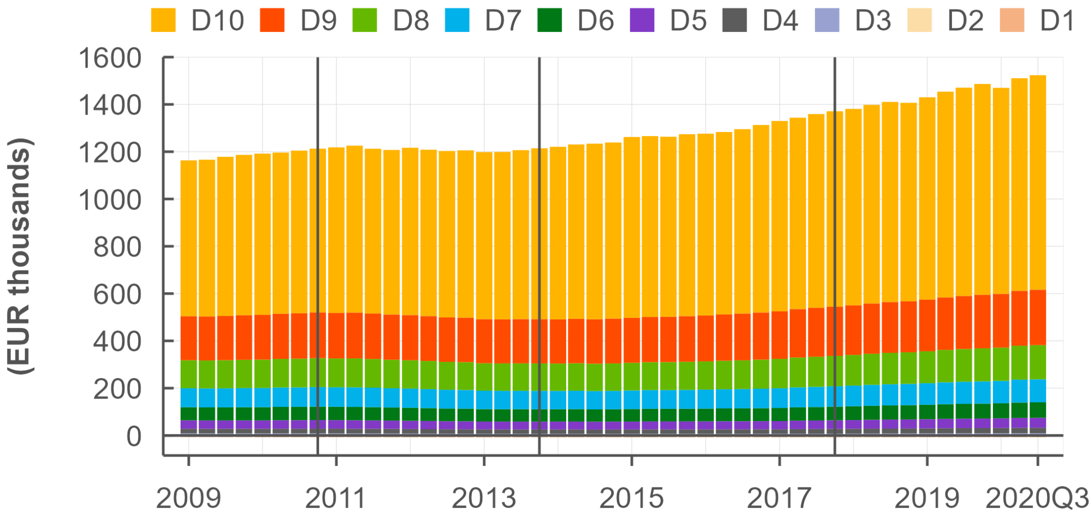

Figure 22: Net wealth per capita of euro area households

The dynamics over the last years of different wealth instruments (i.e. deposits, housing wealth, investment fund shares, listed shares, life insurance and voluntary pensions, debt securities, mortgages, other liabilities and business wealth) and net wealth are also presented in Figure 23 for the euro area, where the population is split into deciles according to their net wealth. An analysis with a different split (i.e. top 1%, next 9%, next 40% and bottom 50%) is available in Section 3.3, while a different split is provided here for additional information.

[51] More precisely, the household weights are multiplied by the number of household members, while all considered instrument holdings are divided by the number of household members. Since the number of household members of the added wealthy households (see Section 2.3) are not available, we approximate this value with the average of the number of household members in the respective country and quarter. The average number of household members ranges from 2.02 in Germany (2014 Q3) to 2.85 in Malta (2009 Q1).

Figure 23: Instrument holdings and net wealth of euro area households by net wealth-deciles.

Acknowledgements

We are grateful for the discussions we had at the meetings of the Household Finance and Consumption Network and the Expert Group for Distributional Financial Accounts while the drafting of the paper was ongoing. This works builds on previous work of the Expert Group on Linking Macro and Micro Data for the Household Sector (EG-LMM) — The EG-LMM has now become the Expert Group on Distributional Financial Accounts (EG-DFA). This Group included members from 12 EU countries over the period when this paper was developed. We also acknowledge the insightful comments of Henning Ahnert, Michele Cantarella, Alexander Erler, Maria Mercedes Frech, Juha Honkkila, Ilja Kristian Kavonius, Luís Teles Morais, Andrea Neri, Dennis Ohlwerter, Malgorzata Osiewicz, Matthias Scherer, Veli-Matti Törmälehto, Panagiota Tzamourani, Francesco Vercelli, and Philip Vermeulen, as well as those of the referees. However, any errors are entirely our own. The authors carried out large parts of this work during their employment at the European Central Bank (ECB), but the views expressed in this paper are those of the authors and are not representative of the views of the ECB or European System of Central Banks.

Disclaimer: This paper uses data from the Eurosystem Household Finance and Consumption Survey. The results published and the related observations and analysis may not correspond to results or analysis of the data producers.

Janina Engel

European Central Bank, Frankfurt am Main, Germany; email: janina.engel@ecb.europa.eu

Pau Gayà Riera

European Central Bank, Frankfurt am Main, Germany; email: pau.gaya.riera@gmail.com

Joseph Grilli

email: joe@grilli.co.uk

Pierre Sola

European Central Bank, Frankfurt am Main, Germany; email: pierre.sola@ecb.europa.eu

Postal address 60640 Frankfurt am Main, Germany
Telephone +49 69 1344 0
Website www.ecb.europa.eu

PDF ISBN 978-92-899-5271-2 ISSN 1725-2806 doi:10.2866/412495 QB-AR-22-052-EN-N