Is the Top Tail of the Wealth Distribution the Missing Link Between the Household Finance and Consumption Survey and National Accounts

Robin Chakraborty, Ilja Kristian Kavonius, Sébastien Pérez-Duarte, and Philip Vermeulen
(European Central Bank)

Robin Chakraborty[*]
European Central Bank

Ilja Kristian Kavonius
European Central Bank

Sébastien Pérez-Duarte
European Central Bank

Philip Vermeulen
European Central Bank

Is the Top Tail of the Wealth Distribution the Missing Link between the Household Finance and Consumption Survey and National Accounts?

DRAFT VERSION 31.07.2016. PLEASE DO NOT CITE

Note:

ABSTRACT

Using data from the first wave of the Household Finance and Consumption Survey (HFCS) and financial accounts (FA) data from the euro area, this paper makes a comparison between both sources for Austria, Germany, France, Spain and Finland. Specifically, the focus of the comparison is on financial assets. Although wealthy households are oversampled in some countries in the HFCS, there is evidence that the top tail of the wealth distribution is not entirely covered by the surveys. Since wealth is concentrated, a limited coverage of the top tail of the distribution will lead to an underestimate of total wealth. As the analysis in Vermeulen (2016) has shown, adding the Forbes data to the survey observations can increase the top wealth shares for some countries significantly. What is missing so far in the literature is an estimate of how much of the gap between the micro and macro data is caused by the underestimation of the top tail of the wealth distribution. Knowledge of the size of this effect would be important to later combine distributional figures from the survey with FA. Therefore, this paper attempts to fill this research gap by following the approach of Vermeulen (2016) and fitting a distribution to the upper tail of the wealth distribution. This is done in three main steps: First, a naïve comparison which indicates gaps between the micro and the macro aggregates is performed. Second, the data of HFCS and FA are reconciled by identifying conceptual differences and coverage issues, which have been partly identified already in the literature. Finally, we focus on the wealthiest households and estimate how much of the gap can be attributed to this group. Therefore, the purpose of this work is not only to identify any of the differences, but to have a first estimate on the size of these sources. In the longer run it is important to understand the main reasons of these gaps to be able to develop further uses for linking micro and macro data. This paper brings us one step further in this direction.

Key words: financial accounts, HFCS, wealth inequality, Pareto distribution, households

1. INTRODUCTION

Inequality and welfare are high on the political and economic research agenda. Stiglitz, Fitoussi and Sen (2009) and Piketty (2014) illustrate the importance of distributional information of wealth in analysing economic progress. Central banks are also increasingly interested in the distributional issues as these have been recognised to interact with monetary policy.[1]

Household wealth surveys are heretofore the most important source of distributional wealth information but are only conducted in longer time intervals (every two to three years). Financial accounts (FA) provide timely information on household financial wealth. Several European and international groups have been established with the underlying motivation to include distributional measures in national accounts (NA) as well as having timely distributional data. Survey information is likely to be used as one input. Our paper contributes to the infant literature which attempts to reconcile these different sources of information. The purpose of this paper is therefore to investigate the differences between the first wave of the Household Finance and Consumption Survey (HFCS) and FA. Research has shown that the upper parts of the wealth distribution are often missing, leading to difficulties in reconciling information from surveys and NA. One focus of this paper will therefore be on estimating how much of the gap between the micro and macro data is caused by the underestimation of the top tail of the wealth distribution.

Our work can be seen in light of the following initiatives. In the beginning of 2016, the European Central Bank (ECB) established an Expert Group on Linking Micro and Macro Household Data (EG-LMM). The focus is on linking FA balance sheet data with the HFCS. The results in section 2 of this paper benefit from the discussions of that group. At the first stage the idea is to understand the differences between these two data sources. At the second stage, once the differences are understood and quantified, the intention is to further investigate the possible practical applications for this linkage. Similar kind of work has been done in the U.S.[2]

In order to establish the micro-macro linkage and disparities in NA, the OECD and European Commission (EC) established an Expert Group on Disparities in NA to investigate this linkage and other possible ways to disaggregate NA. The first group has now finished its work and the second group established by the OECD has started to continue the work on income, consumption and savings. The focus is to improve the methodology as well as to have an estimate on the time series for these accounts.[3]

---

[1] The IMF/FSB report to the G-20 Finance Ministers and Central Bank Governors' data gap initiative emphasised in particular need for including distributional information in macro data.
See more on the G20 Data Gap Initiative and micro data and its interest for macroprudential and monetary policies: Tissot (2015), pp 5-6.
[2] Dettling, Devlin-Foltz, Krimmel, Pack and Thompson (2015). Additionally, Henriques and Hsu (2012) and Antoniewich (2000) have written on the linkage between Flows of Funds Accounts and the Survey of Consumer Finances.
[3] The results of the first group are reported in the following working papers: Fesseau and Mattonetti (2013) and Fesseau, Wolff and Mattonetti (2013).

Independent of these international initiatives, a certain development in the area of linking these two data sets can be identified. The development can be divided into two waves. In the first wave which already started in the 1980s there were several attempts in investigating the linkage between income as well as consumption surveys and NA data.[4] At that time the focus was on income and consumption as wealth data and in particular wealth surveys were uncommon. However, in the U.S. the first Survey of Consumer Finance (SCF) was already conducted in 1946. For the SCF, comparisons of wealth data have already been investigated in the 1980s.[5] In the second wave, there has been a growing interest in wealth and distributional information. The consequence of this is the increasing availability of wealth data like the HFCS which was for the first time published in 2013. The focus has not only been on investigating the linkage between FA and the HFCS[6] but also on developing practical applications where the linkage is also used for analytical purposes.[7]

However, linking the data has also raised the need to understand the differences between the two statistics. In order to link the two data sets properly, it is essential to know the source of any differences and what kind of distributional effects they might have. The key question is whether the distribution of the survey represents the distribution of the total. Therefore, it would be valuable to have an estimate of how much a possible under-representation of wealthy households in the survey explains in terms of the difference between the survey and FA.[8]

The paper aims at quantifying the differences, i.e. to analyse whether these are caused by statistical and definitional reasons and by the missing tail of the wealthiest households. Thereby the paper contributes in analysing the plausibility of both sources as well as in understanding the results of the comparison. This is essential when practical applications based on the linkage will be developed as different types of discrepancies most likely have different distributional impacts on the results.

This paper is organised as follows: the following chapter briefly summarises the framework applied and analyses the generic statistical and conceptual differences between the two sources. Based on this analysis, we develop two different adjusted concepts for financial assets with the intention to base the comparison only on those instruments that are included in both sources and that are conceptually comparable. For the two adjusted concepts we indicate the differences between the HFCS and the FA. The third chapter focuses on the methodology to estimate the tail of the wealth distribution based on Vermeulen (2016). In the fourth chapter we finally analyse how the estimated tail based on the Pareto distribution changes the remaining gaps for one of the adjusted concepts developed in chapter 2. The final chapter summarises the main conclusions.

---

[4] See for instance: Atkinson and Micklewright (1983).
[5] See for instance: Avery, Elliehausen, and Kennickell (1988).
[6] See: Andreasch, Fessler and Lindner (2013); ECB (2013); Kavonius and Honkkila (2013).
[7] See: Ampudia, Pavlivkova, Slacelek and Vogel (2014) and Kavonius and Honkkila (2016).
[8] See: Bach, Thiemann and Zucco (2015); Dalitz (2016); Eckerstorfer, Halak, Kapeller, Schütz, Springholz and Wildauer (2015); Ogwang (2011); Vermeulen (2016); Vermeulen (2016b).

2. Conceptual linkages, Generic statistical differences and adjusted concepts

This chapter is divided into three parts. The first part discusses the conceptual linkages, i.e. the linkage between different assets and liabilities in the HFCS and FA. The second part focuses on generic differences that have a potential effect on the coverage ratio of the survey. Finally, in the third part we derive different adjusted concepts of financial assets which aim to provide a more comparable picture of financial assets than a purely naïve comparison can provide.

The purpose of this section is to quantify generic conceptual and statistical differences. Many of these differences in this section have already been discussed e.g. in Kavonius and Törmälehto (2010). Therefore, we keep the discussion short and focus rather on new aspects.

2.1 Conceptual linkages

The HFCS is a triennial survey which provides individual household data on wealth and some income items. It is collected in a harmonised way in 15 euro area countries for a sample of more than 62,000 households. Contrary to the HFCS, the FA includes only aggregated macro level balance data for institutional sectors including the households sector. The FA in the euro area becomes available on a quarterly basis three to four months after the end of the reference quarter.

Although the HFCS commonly uses definitions which are aligned to FA where possible, the concepts and definitions differ sometimes to fit the purpose of the questionnaire as data have to be collected in a way that households can understand the questions and provide the appropriate information.[9] This might involve asking about assets or debts that do not fit the FA breakdowns, or skipping some items entirely, for confidentiality concerns for example. To analyse the conceptual linkages of the HFCS it is important to derive adjusted concepts of financial assets as well as in the future for non-financial assets, liabilities and finally net wealth that are comparable across both sources. Especially those assets and liabilities that are not covered in either of the two sources have to be first eliminated to make both sources as comparable as possible. But also assets and liabilities that are hardly comparable would have to be excluded to not distort the coverage ratios on an aggregated level.

Table A 1 in the Annex shows the linkages on an instrument level between the HFCS and the FA. As Kavonius and Törmälehto (2010) have already examined the linkages between the HFCS and FA we limit our discussion to different approaches, refinements of their linkages, and point out changes that came into place through the change from ESA95 to ESA2010. The table also gives an indication on the instruments that we excluded from the adjusted concept of net wealth that will follow in section 2.3. In the following we list the differences and changes we have introduced to the linkage table presented by Kavonius and Törmälehto (2010):

Deposits
In regards to further detail of information in FA, there is now a split of 'F.2 Deposits' in 'F.22 Transferable deposits' and 'F.29 Other deposits' available, which corresponds to the survey

---

[9] Kavonius and Törmälehto (2010).

definitions of sight accounts and saving deposits. Conceptually, there is only a minor difference for savings accounts: deposit-like instruments held with other financial intermediaries (e.g. insurance companies) are recorded as short-term loans in FA whereas in the survey they are included in HD1200 savings accounts. However this has no effect to the comparability of the aggregated financial assets in either of the concepts.

Debt securities and financial derivatives
In the ESA95/SNA95 debt securities were broken down to "F33 securities other than shares expect financial derivatives" and "F.34 financial derivatives". In ESA2010/SNA2008, financial derivatives are classified under F.71. In ESA2010/SNA2008 debt securities are classified as F.3 and match the HFCS item "Bonds".

Equity
In FA, 'F.51 Equity' is split into 'F.511 Listed shares', 'F.512 Unlisted shares' and 'F.519 Other equity'. Listed shares are equity securities listed on an exchange whereas unlisted shares are accordingly equity securities not listed on an exchange. Other equity comprises all forms of equity other than those classified in the sub-categories listed shares and unlisted shares, e.g. equity in limited liability companies whose owners are partners and not shareholders.[10] The HFCS also collects the value of publicly traded shares (HD1510) which can be linked to F.511 Listed shares'. But contrary to the classification in FA ('Unlisted shares' and 'Other equity'), in the HFCS the classification is based on the household's activity in the enterprise. If the household is self-employed or has an active role in running the business, it is classified as a 'Self-employed business'. If the household is just invested in the business e.g. as a silent partner without having an active role in running the business and there are no publicly traded shares, then it is classified as a 'Non-self-employment not publicly traded business'. In the HFCS, the value of self-employed businesses is regarded as real wealth, whereas any investments in non-self-employed businesses are regarded as financial assets in the survey classification. To match the categorisation of financial assets in FA we reclassify the value of self-employed businesses to financial assets (other equity).

Insurance and pensions
The SNA2008 introduced new breakdowns for F.6 insurance, pensions and standardised guaranteed schemes which allow to better link the concepts between the HFCS and NA. F.61 Non-life insurance technical reserves are not covered by the HFCS wealth concept. F.62 Life insurance and annuity entitlements correspond with the HFCS item voluntary pensions/whole life insurance schemes. The data in FA is typically based on actuary information on technical reserves reported by insurance corporations. HFCS figures are based on the premium paid by households (and may thus include for instance service fees). F.63 Pension entitlement corresponds with the HFCS item "current value of all occupational pension plans that have an account". However, as the concept in FA not only covers pensions that have an account balance and as the stock of occupational pensions of households that are already retired is not included in the survey (and in the FA they are), we exclude the pension entitlements in the adjusted concept

---

[10] For further explanations see ESA2010, pp 142-144.

of financial assets. F.64 Claims of pension funds on pension managers, F.65 Entitlement to non-pension benefits and F.66 Provision for calls under standardised guarantees are not included as it is not considered to be relevant for the comparison.

2.2 GENERIC DIFFERENCES

This section focuses on generic differences i.e. it goes briefly through the following differences: (1.) population differences; (2.) timing; (3.) potential measurement errors in the FA; (4.) underreporting and item non-response; and (5.) Differences caused by the treatment of sole-proprietors/partnerships and quasi-corporations. Many of these items are already discussed in other articles and therefore, we keep the discussion short.

<u>Population</u>
In the comparisons of NA and the HFCS there are potentially two generic differences in regards to the population: (1.) the difference caused by the fact that non-profit institutions serving households (NPISH) are reported in FA in the same aggregate with households. However, in the euro area countries this is less of an issue as most of the countries transmit the households separately from the NPISH. This is also the case in the countries which are discussed in this paper. (2.) Differences in the definition of the household sector and the HFCS population. FA/NA have a resident approach, covering all households that plan to stay for at least one year, and irrespective of periods spent abroad of less than one year. In the HFCS non-resident citizens are not excluded in all countries. In the HFCS persons living in institutions e.g. in prisons or retirement homes are excluded in most countries; persons with the intention of staying less than 6 months in the country are also excluded from the target population. Therefore, the household weights, which are designed to represent the population, do not include these specific excluded groups in most countries. Any comparison has to take this into account and the country totals of the survey or FA have to be adjusted. As a first estimate using per capita amounts seems reasonable.[11]

Table 1: Comparison of population between FA (ESA 95 population data) and HFCS[12]

Country Code	Population NA (historical vintage)	Population HFCS	Difference total	Difference in %
Austria	8,388,130	8,021,945	366,185	4%
Germany	81,629,370	81,085,984	543,386	1%
Spain	45,456,960	45,632,180	-175,220	0%
Finland	5,336,910	5,271,534	65,376	1%
France	64,444,520[13]	62,464,244	1,980,276	3%

Table 1 compares the population numbers between FA and the HFCS. The number for FA is based on the last available vintage of the ESA95 population data that corresponds to the

---

[11] This assumes that the excluded groups have the same average wealth as the rest of the population, which may not be the case. For instance, people living in retirement homes may have a higher than the average per capita wealth.
[12] Please see Table 2 for the definition of the time periods.
[13] French overseas territories included in the Financial Accounts, whereas the HFCS only included the "metropolitan France".

reference year of the fieldwork period. Because of the above mentioned excluded groups the population in the HFCS should be lower than the one for the whole population. This is the case for all countries except for Spain.

Timing and frequency
The primary drawbacks of the HFCS are the biennial to triennial frequency and the lag between data collection and data release. Furthermore, the different fieldwork periods may raise concerns about comparability on an aggregated level. The first wave of the HFCS was carried out from 2008/2009 to 2011. For the comparison of the FA with the HFCS, FA data which are closest to the fieldwork period for each country are used. At a first stage this is based on annual (year-end) figures as some EU countries do not yet provide quarterly FA backdata for ESA2010, which would better match the fieldwork period of the first wave. Table 2 gives an overview of the different fieldwork periods and the annual end date which is currently taken for the comparison. The timing can contribute to any observed difference as the value of assets and liabilities may change between the time the survey was conducted and the period taken for FA.

Table 2: Fieldwork period and time periods for comparison[14]

Country	Fieldwork	Assets & Liabilities	FA (annual end)
Austria	09/10 – 05/11	Time of interview	Q42010
Germany	09/10 – 07/11	Time of interview	Q4/2010
Spain	11/08 – 07/09	Time of interview	Q4/2008
Finland	01/10 – 05/10	31/12/2009	Q4/2009
France	10/09 – 02/10	Time of interview	Q4/2009

Potential measurement errors in the FA data
As the FA is based on other statistical sources and the validation of primary statistics, it is possible that errors are inherited from source statistics. Additionally, as the FA is a closed and balanced system, it is possible that some of the household aggregates are adjusted by adding balancing adjustments. In some cases balance sheet items can even be based on residual estimations. However, in the euro area countries and in particular in countries we analyse in this paper the FA balance sheets are mostly based on counterpart information and thus the data can be considered relatively accurate.[15]

Underreporting and item non-response in the HFCS
Item non-response refers to the problem that for some assets and liabilities the household may not report any value. There are several approaches to alleviate this issue. In the HFCS the problem of item non-response is tackled by multiple imputation. Multiple imputation is the leading method to tackle the problem of item non-response.[16] This means that the HFCS, instead of providing one imputed value for each missing one, is giving a set of values drawn from the

---

[14] Source of Fieldwork period, A&L and Income: ECB (2013).
[15] More concerning potential measurement errors in financial accounts: Kavonius and Törmälehto (2010) and Kavonius and Honkkila (2013).
[16] Rubin (2004).

distribution of values, conditional on the characteristics of the household and the other variables. A full dataset for the main instruments without missing values is provided .[17] This reduces the overall coverage problem between the survey and NA for these items as the imputed values increase the total amounts of the survey accordingly. One measurement problem that remains apart from item non-response is that the household still may not accurately estimate the value of some assets or liabilities, or denies that it possesses the instrument. This might also be one reason for discrepancies between the HFCS and FA.

<u>Differences caused by the treatment of sole-proprietors/partnerships and quasi-corporations</u>
FA distinguishes between producer households (to be classified within the household sector/S.14) and quasi-corporations (to be classified within the non-financial corporations sector S.11). This distinction is relevant because it affects the gross wealth of the household sector and the composition of the household balance sheet. In the FA framework it depends whether the business is a separate institutional unit or not: "Quasi-corporations are unincorporated enterprises that function as if they were corporations. Quasi-corporations are treated as corporations: that is, as separate institutional units from the units to which they belong in recognition of their distinct economic and financial behaviour."[18] Unincorporated enterprises are part of the household sector and are classified as producer households if they are not considered as a separate institutional unit as described above. Financial and non-financial assets as well as financial liabilities of these unincorporated enterprises (classified in S.14) are spread over the various items of the household balance sheets and it is not possible to distinguish between wealth of the unincorporated enterprise and wealth of the household. In this case there is no value of net equity recorded in 'F.519 Other Equity'. If, however, the economic activity is considered to be a separate unit[19], any property rights are classified in NA as equity participation held by the household (other equity).

The survey definition of self-employed businesses (including sole-proprietorships and partnerships) ideally[20] enables to identify values for the net value of the business separately from other non-business related positions of the household. This conceptual difference implies that for producer households there is a net value collected in the survey whereas in FA the assets and liabilities of these producer households are spread over the different instruments.[21] The question is which instruments are affected by this difference and to what extent. Real assets and liabilities may as well be affected as financial assets. To have a measure on the size of this difference for each of the instrument would need separate accounts for sole-proprietorships and partnerships. This might account for part of the difference in the coverage ratios of many instruments as well

---

[17] ECB 2013.
[18] ESA2010, p. 422.
[19] This separate institutional unit is then classified in S.11 or S.12.
[20] The question in the survey suggests that respondents may actually also include some parts of business assets or liabilities in other household balance sheet items ("Aside from any assets and debts connected with this business that I may have already recorded … what is the net value of the business?"). Therefore, it cannot be ruled out that the recording of values for businesses differs across survey respondents.
[21] For legal forms other than sole-proprietorships and partnerships (e.g. limited liability companies) the household holds a net equity position in the business both in the FA and in the HFCS.

as on an aggregated level for each component of net wealth (financial assets, real assets and liabilities).

Table 3: Recording of businesses and inclusion in the different concepts

Case	Type of business	Net/Gross Value		Included/Excluded in the HFCS[22]			Comment
		HFCS	FA	Naïve comparison	Adjusted concept 1	Adjusted concept 2	
1	Sole proprietorships and partnerships which are <u>not</u> classified as quasi-corporations in FA	Net value	Gross Values: Recorded in the household sector indistinguishable from the "private part of the household". The assets and liabilities of the business part are distributed across the household balance sheet (including financial assets, real assets and liabilities)	Excluded	Included	Excluded	The net value might include real assets
2	Sole proprietorships and partnerships which are classified as quasi-corporations in FA.	Net value	Net value (other equity)	Excluded	Included	Excluded	In principle comparable but quasi-corporations would be difficult to identify in the HFCS based on the information provided in the survey.
3	Limited liability companies and other incorporated businesses	Net value	Net value (other equity)	Excluded	Included	Included	

Table 3 provides an overview of the different types of businesses and how they are recorded in the HFCS and FA. As can be seen, the main comparability issue arises only for those sole-proprietors and partnerships which are not classified as quasi-corporations and hence are recorded in the household sector in FA indistinguishable from the "private part of the household" (case 1). For these, there is a net value for the business provided in the HFCS whereas in the FA the assets and liabilities of the business are spread across the balance sheet of the household

---

[22] In the FA we keep F.5 Equity the same in all concepts.

sector including real assets and liabilities.[23] Hence, for this part of the sole proprietors and partnerships it is not known if the net value of the business provided in the HFCS should be allocated to financial assets, real assets or liabilities in FA. For quasi-corporations (case 2), there is a net value provided in the HFCS and also a net value recorded in the FA. The same applies to the other incorporated businesses: there is in principle no difference in the recording, as both in the HFCS and in the FA there is a net value provided.

Table 3 furthermore provides an overview which of the described cases is included or excluded in the HFCS in each of the concepts described in the next section. For the other instruments Table A 1 in the annex provides an overview which instruments are excluded from both sources in the adjusted concepts.

2.3 ADJUSTED CONCEPTS OF FINANCIAL ASSETS

The aim of this section is to derive an adjusted concept of financial assets. The intention to go from a naïve comparison to an adjusted concept is to base the comparison on those instruments which are included in both sources and are conceptually comparable. The adjusted concept allows providing a more reliable coverage ratio than a purely naïve comparison between both sources would be able to. The coverage ratio indicates the per capita amount of financial assets covered by the survey, e.g. a value of 98% would imply that the per capita amount of the HFCS is only 2% below the per capita amount in FA. We provide the coverage ratios differently for the household sector (S14) and for the sector where households are aggregated with the NPISH (S1M) to give an impression on the magnitude. As expected, the coverage ratio of the household sector (S14) is slightly higher compared to the ones when households are combined with NPISH.

Table 4: Coverage ratios of financial assets – Naïve comparison vs. adjusted concepts

Country	Naïve comparison		Adjusted concept 1		Adjusted concept 2		Financial assets of FA covered in the adjusted concepts (same in both concepts)
	S1M	S14	S1M	S14	S1M	S14	
Austria	35%	35%	96%	98%	45%	46%	87%
Germany	41%	43%	83%	86%	64%	67%	77%
Spain	34%	34%	74%	75%	58%	59%	82%
Finland	35%	37%	51%	55%	41%	45%	83%
France	37%	38%	58%	59%	50%	51%	90%

Naïve Comparison

The naïve comparison just takes the concepts of financial assets as they are in the HFCS and in the FA. This serves a benchmark but this concept includes also non-comparable instruments (e.g. F.21 Currency, which is not covered by the HFCS) and uses different classifications (e.g. of the value of self-employed) which distort the picture of the actual coverage ratios. The HFCS concept of financial assets does not include the value of self-employed businesses as well as the value of occupational pension plans which are accordingly also not included in the naïve

---

[23] For quasi-corporations both in the FA and in the HFCS the household holds a net equity position, which would not cause any comparability issues between the HFCS and FA.

comparison. Therefore it is not surprising that the naïve comparison shows relatively low coverage ratios of 34% to 43% for financial assets (S14).

Adjusted concept 1

For the adjusted concept 1 we include on the survey side the value of self-employed businesses in the comparison (reclassification from real assets to financial assets) and we exclude the amount owed to the household (HD1710) as well as the other financial assets (HD1920). In the FA we exclude F.21 Currency, F.4 Loans (Assets), F.7 Financial derivatives and F.8 Other accounts receivable. For pensions we only include F.62 Life insurance and annuity entitlements and exclude the other sub-categories (F.61, F.63-F.66) as they are not comparable to the survey (see discussion above). As can be seen in Table 4, going from a naïve comparison to the adjusted concept 1 significantly increases the coverage ratio for financial assets (to 55% - 98%). The adjusted concept 1, as it includes all self-employed businesses most likely overstates the coverage ratio as the value for sole proprietors and partnerships may also include real assets (see discussion above about the delineation between sole-proprietors and quasi-corporations).

Adjusted concept 2

In the HFCS the value for self-employed businesses can be broken down by legal status (see Table A 1). Therefore the distinction between sole-proprietorships, partnerships and other incorporated businesses is possible. While the adjusted concept 1 includes the net values of all legal forms of self-employed businesses in the HFCS (including sole proprietors and partnerships), the adjusted concept 2 excludes sole proprietors and partnerships from the value of self-employed businesses in the survey. For FA we keep F.5 Equity the same in both concepts. The intention of this adjusted concept is that it serves as a lower benchmark as it only comprises the net value of those legal forms in the survey that are in the FA recorded in the non-financial corporations' sector and consequently the household only holds a net equity position in the business (other equity). Thus, for the legal forms included in this concept both in the FA and in the HFCS the household holds a net equity position.

As can be seen the coverage ratios for the adjusted concept 2 are higher compared to the naive comparison but significantly lower compared to the adjusted concept 1 where all legal forms of self-employed businesses were included. Certainly, this concept underestimates the coverage ratios as it excludes all financial assets of sole-proprietorships and partnerships from the survey.

For improving the concepts further the following information would be needed: first, an estimate of sole proprietorships and partnerships included in the HFCS that are classified as quasi-corporations in the FA (case 2 in Table 3). Second, for the sole-proprietors and partnerships which are recorded in the household sector, one would need the breakdown to financial assets, real assets and liabilities (case 1 in Table 3).

3.   THE WEALTH DISTRIBUTION AND METHODOLOGY TO ESTIMATE THE TAIL

In this section, we first discuss the general problem of wealth surveys, i.e. the wealth distribution and differential unit non-response and which oversampling strategies are used by countries to mitigate this issue in the HFCS. In the second part we explain the methodology to estimate the top tail of the wealth distribution by a Pareto distribution. Our approach and discussion builds on

Vermeulen (2016). The third part discusses the Forbes list and its consistency with the statistical data. These data are used in part for the estimations of the Pareto tail in Vermeulen (2016).

3.1 OVERSAMPLING, WEALTH DISTRIBUTION AND DIFFERENTIAL UNIT NON-RESPONSE IN THE HFCS

In general the bias in the HFCS caused by unit non-response is reduced by weight adjustments (Pérez-Duarte et al. 2010). But as the wealth distribution is often skewed, unit non-response of the wealthiest households, or the fact that extremely wealthy households are rarely included in the survey sample, might still be a problem. To cover this bias, some countries have oversampled wealthy households in the HFCS. Table 5 gives an overview which strategy the countries included in our analysis have used to oversample wealthy households in order to reduce unit non-response bias and to obtain better estimates of wealth at the top of the distribution. Germany used an oversampling strategy based on geographical areas which resulted in a less effective oversampling than in France and Spain which used net wealth or taxable wealth. To compare FA with survey data, oversampling is important as not doing so will reduce the estimates of the aggregated survey values. Without oversampling, income and wealth concentrations are likely to be underestimated using survey data, as there is a high concentration of wealth in the top quintile and the response rates of especially this quintile is usually lower. For the top tail of the wealth distribution there is some evidence on how response rates correlate with the amount of wealth owed by a household. Based on the Survey of Consumer Finance from the US, Kennickell and Woodburn (1999) have documented the following response rates based on different strata (differential unit non- response): 34% for USD 1 million to 2.5 million, 14 % for USD 100 million to 250 million. For the stratum which likely includes the wealthiest households Kennickell (2008) observe an overall response rate of 10%.

Table 5: Oversampling in the first wave of the HFCS by country[24]

Country	Oversampling wealthy households	Basis for oversampling	Effective oversampling rate of the top 5%
Austria	No	n.a	4
Germany	Yes	Geographical areas	148
Spain	Yes	Taxable wealth	314
Finland	Yes	High-income employees, self-employed and farmers	85
France	Yes	Net wealth	208

However, for the HFCS the amount of wealth owned by the top tail varies from country to country and available evidence suggests that the response rates declined to a different extent in different countries. For the 2011 wave, the Spanish survey of household finances documented the following response rates by wealth strata: Stratum 5 (0.9 to 2 million) 31%, Stratum 6 (2 to 6

---

[24] Effective oversampling rate of the top 5%": (S95 – 0.05)/0.05, where S95 is the share of sample households in the wealthiest 5%. Wealthiest households are defined as having higher net wealth than 95% of all households, calculated from weighted data (ECB 2013).

million) 26%, Stratum 7 (6 to 25 million) 21%, Stratum 8 (wealth above 25 million EUR) 21%.[25] Contrary, in Finland – although response rates varied across different strata, age groups, regions and education level, non-response rates did not increase along the level of taxable wealth for the Finnish Household Wealth Survey 2004.[26]

Even with oversampling, one question that will remain is how much of the wealth of the wealthiest households is actually covered by the survey. This in turn might be one reason for part of the gap between the amounts of FA and aggregated amounts from the survey. In terms of the coverage ratio, capturing the value of assets from these wealthiest households might be even more relevant for specific instruments, as there are particular financial assets which are largely possessed by a small fraction of the wealthier households. The idea here is to replace the observations above a certain threshold of wealth per household by an estimated Pareto distribution and see which impact this has on the coverage ratio of the HFCS in comparison to FA. The methodology to estimate the Pareto tail is explained in the following section and based on the same approach as in Vermeulen (2016). Therefore, we keep the explanation here short.

3.2   METHODOLOGY

Wealth is heavily skewed at the top and the literature has reached a consensus that the top of the wealth distribution is well approximated by a Pareto distribution.[27] The Pareto distribution has two parameters, the tail exponent α and the Threshold parameter T. The distribution is given by the following complementary cumulative distribution function (ccdf):

$$(1) \quad P(W > w) = (\frac{T}{w})^\alpha$$

The Pareto distribution is defined on the interval [T,∞) and $\alpha > 0$. The threshold T is the lower bound of the distribution. Estimating a Pareto distribution on a simple random sample is fairly straightforward. The maximum likelihood estimator of α from a random sample of n observations drawn from a Pareto distribution with a given threshold T is given by:

$$(2) \quad \alpha_{ml} = \left[\sum_i \frac{1}{n} \ln \frac{w_i}{T}\right]^{(-1)}$$

Alternatively, the tail exponent has been estimated in the literature using linear regression on ranked data. Let i be the rank of the observation (with rank 1 being the highest observation). The Pareto tail exponent α can be estimated by:

$$(3) \quad \ln(i - 0.5) = C - \alpha * \ln(w_i)$$

Where the "subtract 0.5 from the rank" is suggested in Gabaix and Ibragimov (2011).

---

[25] Bover et al. (2014). The survey has a panel-component for which the response rate drops from 74% to 62% for these wealth strata.
[26] Pérez-Duarte et al. (2010).
[27] Davies and Shorrocks (1999).

However, wealth survey data does generally not consist of a simple random sample. In particular, sample observations have weights. Vermeulen (2016) shows that taking into account the weights can be done in the regression method above, using the ranked n highest observations:

$$(4) \quad \ln(i - 0.5)\frac{N_{fi}}{\bar{N}} = C - \alpha * \ln(w_i)$$

where N_{fi} is the average weight of the highest i sample points and $\bar{N}$ is the average weight of all n highest sample points. This regression method can be used in two ways. First, to estimate α using only the survey data (i.e. the highest n observations). Alternatively these observations can be pooled with data of rich lists which contain datapoints that are higher than the highest observation in the survey (this joint dataset is then ordered first). Using this regression method works particularly well when combining the survey data with such extraneous data points.

A particular problem is the choice of the threshold T. There is no clear-cut way in finding a "correct" threshold. However the Pareto distribution has the interesting property that a distribution with tail exponent α and threshold T, when restricted above $T^*>T$ remains a Pareto with the same tail exponent. It therefore seems prudent with survey data to take a high threshold. This avoids taking lower observations that are not Pareto distributed. However, there is a trade-off: A higher threshold T* implies using less data to estimate α. It is probably best to estimate α using different thresholds of the data and check for sensitivity.

After estimating the α for a given threshold T, the n observations can be replaced by the estimated Pareto distribution. The mean of a Pareto distribution is given by $\frac{\alpha}{\alpha-1}T$, so that we can say that the total wealth in the Pareto tail is given by $n*\bar{N}\left(\frac{\alpha}{\alpha-1}\right)T$, where $n*\bar{N}$ is the total sum of weights of the highest n observations in the survey sample.

We use the thresholds 500T, 1M and 2M to estimate the alpha and we use the same thresholds to replace the survey observations by the estimated Pareto tail. The Pareto distribution is estimated using the above described methods: (1.) using the maximum likelihood[28]; (2.) using the regression method excluding data from the Forbes and (3.) using the regression method including data from the Forbes.

3.3   FORBES DATA

The wealth concept of the Forbes list does not strictly follow any defined concept and therefore, it should be interpreted as a proxy. Four conceptual issues related to using these estimates in statistical estimations can be identified. First, the estimations are based either on interviews of billionaires themselves or their handlers, employees, rivals etc. This implies that it is impossible

---

[28] Specifically we use the pseudo-maximum likelihood estimator which has the same form as the maximum likelihood estimator but takes into account the weights of the sample observations in the survey (see Vermeulen 2016).

to cover all the asset types or to have similar type of market valuation to FA or household surveys. The methodological report of the Forbes states that "not that we pretend to know what is listed on everyone's private balance sheet, though some folks do provide that information. We do attempt to vet these numbers with all billionaires. Some cooperate, others don't."[29] Almost all the Forbes list families from the countries analysed in this paper have earned their money in businesses and therefore, it can be assumed that the majority of their wealth is in equity. For the Forbes list, the privately owned businesses have been valued by coupling estimates of revenues or profits with prevailing price-to-revenues or price-to-earnings ratios for similar companies.[30] The method can be considered similar to the methods used in the valuation of the unlisted equity in the FA.

Second, the wealth concept does not cover all asset types as these are partly based on external estimations. Additionally, the wealth concept also covers items which are defined as durable goods in the NA (such as yachts). Third, sometimes the fortune is distributed to the different family members and sometimes it is not and a large number of family members is aggregated (Dolan 2016). The starting point in statistics and in particular in the HFCS is that the applied unit is the household. In the case of the Forbes lists it is very well possible that the applied family concept covers several households or contrary one person e.g. the head of the household.

Fourth, the Forbes list covers families by nationality and it does not correspond with the residence concept applied in the HFCN and NA. The families living outside of the country of their citizenship should not be included in the HFCN population but they are included in the Forbes list. A brief analysis proved that the majority of these families are actually resident in the countries of their citizenship. For instance in the case of Finland all six persons who are on the list are also residents in Finland. In larger countries where the number of billionaires is also higher, there are some families which live outside the country of their citizenship. In the future work, it can be considered to allocate these types of families to their resident countries. Even though there are these drawbacks in using the Forbes list, the data are the best proxy for the very top tail of the wealthiest households.

---

[29] Dolan (2016).
[30] Dolan (2016).

4.   RESULTS

4.1   ESTIMATES OF THE PARETO COMPARED WITH THE HFCS

The results for the Pareto tail index are provided in Table 6.[31] In general lower alphas lead to higher tail wealth and higher total wealth. As described in the methodology we replace the tail wealth of the survey observations above each of the thresholds by the estimated wealth from the Pareto distribution using the same method as Vermeulen (2016) and thus, the same alphas. Hereby we assume that the weights allocated in the HFCS to those households having wealth above these thresholds are correct. Table 7 gives an overview of the weights in % of the population. Obviously, increasing the threshold decreases the weights allocated to these households. Nevertheless, the weights allocated to households above each of the threshold, varies from country to country, e.g. the weight of households in Austria having wealth above 2M is 1.9% compared to Finland with only 0.3%. Table 9 to Table 11 provide an estimate using the different methods to estimate the tail and for each estimate a comparison in terms of the HFCS tail (Pareto tail divided by the HFCS tail). Furthermore, the tables provide an estimate in terms of actual net wealth of the HFCS when the tail is replaced by the Pareto estimate (estimated net wealth when tail is replaced by the Pareto divided by the actual net wealth of the HFCS).

Using the pseudo maximum likelihood method without the Forbes list seems not to increase the tail significantly for those countries that used an effective oversampling strategy (Spain and France). But especially in Austria and Germany having less effective oversampling strategies the estimated Pareto tail seems to increase the tail as well as total net wealth. The total effect on net wealth is lower compared to the effect on the tail as the weight of those households having wealth above each of the thresholds is taken into account.

The regression method excluding the Forbes is similarly not adding much to the HFCS tail using either of the thresholds for those countries with an effective oversampling strategy. Generally the estimates get more imprecise the higher the threshold is as fewer sample observations from the survey can be taken for the analysis. This seems to be especially prevalent for those countries using a less effective oversampling strategy. The undercoverage of the survey tail by the estimated Pareto tail above the 2M threshold for Austria and Germany is most likely based on this fact and the results need to be interpreted with caution. However, using a lower threshold generally bears the risk to include observations in the estimate that may not be Pareto distributed.[32] Finally, including the Forbes data into the regression method yields the highest estimates for the tail as well as for net wealth in line with the results from Vermeulen (2016). It even adds wealth for those countries that used an effective oversampling strategy although to a minor extent compared to those countries with a less effective oversampling strategy.

---

[31] These are the same estimates of the Pareto tail index as used in Vermeulen (2016).
[32] Vermeulen (2016).

Table 6: Pareto tail index (alpha)[33]

	Pseudo max.likelihood			Regression method excl. Forbes			Regression method incl. Forbes		
	≥2M	≥1M	≥500T	≥2M	≥1M	≥500T	≥2M	≥1M	≥500T
Austria	1.67 (0.42)	1.42 (0.30)	1.34 (0.16)	1.87 (0.72)	1.65 (0.45)	1.44 (0.26)	1.47 (0.06)	1.47 (0.05)	1.43 (0.08)
Germany	1.41 (0.26)	1.43 (0.17)	1.61 (0.10)	1.87 (0.35)	1.64 (0.23)	1.54 (0.13)	1.38 (0.04)	1.39 (0.02)	1.40 (0.01)
Spain	1.71 (0.27)	2.05 (0.18)	1.85 (0.08)	1.67 (0.13)	1.76 (0.11)	1.87 (0.08)	1.59 (0.05)	1.69 (0.05)	1.80 (0.05)
Finland	2.01 (0.23)	2.47 (0.18)	2.26 (0.06)	1.94 (0.57)	2.13 (0.23)	2.27 (0.10)	1.60 (0.14)	1.88 (0.13)	2.16 (0.08)
France	1.65 (0.09)	1.84 (0.08)	1.75 (0.04)	1.67 (0.13)	1.78 (0.08)	1.83 (0.06)	1.50 (0.02)	1.63 (0.03)	1.73 (0.03)

Table 7: Weights below and above threshold

	≥2M		≥1M		≥500T	
	Below	Above	Below	Above	Below	Above
Austria	0.981	0.019	0.954	0.046	0.887	0.113
Germany	0.991	0.009	0.974	0.026	0.918	0.082
Spain	0.992	0.008	0.964	0.036	0.865	0.135
Finland	0.997	0.003	0.986	0.014	0.937	0.063
France	0.992	0.008	0.970	0.030	0.896	0.104

Table 8: Wealth below and above threshold HFCS (bn EUR)

	≥2M		≥1M		≥500T		Total
	Below	Above	Below	Above	Below	Above	
Austria	673	327	528	472	357	643	1,000
Germany	5,907	1,836	4,945	2,798	3,489	4,254	7,743
Spain	4,273	685	3,637	1,321	2,475	2,483	4,958
Finland	384	25	349	60	267	142	409
France	5,466	1,036	4,620	1,883	3,200	3,303	6,503

REPLACEMENT OF HOUSEHOLDS IN THE SURVEY ABOVE THRESHOLD BY THE PARETO TAIL

Table 9: Replacing wealth above threshold using Pseudo max.likelihood (bn EUR)

	≥2M			≥1M			≥500T		
	Above threshold	Tail	Net wealth	Above threshold	Tail	Net wealth	Above threshold	Tail	Net wealth
Austria	354	108%	103%	590	125%	112%	842	131%	120%
Germany	2,536	138%	109%	3,496	125%	109%	4,304	101%	101%
Spain	672	98%	100%	1,213	92%	98%	2,503	101%	100%
Finland	26	106%	100%	58	96%	99%	142	100%	100%
France	1,064	103%	100%	1,820	97%	99%	3,374	102%	101%

---

[33] Mean over implicates and standard errors in parentheses.

Table 10: Replacing wealth above threshold using Regression method excluding Forbes (bn EUR)

	≥2M			≥1M			≥500T		
	Above threshold	Tail	Net wealth	Above threshold	Tail	Net wealth	Above threshold	Tail	Net wealth
Austria	305	93%	98%	443	94%	97%	699	109%	106%
Germany	1,585	86%	97%	2,694	96%	99%	4,651	109%	105%
Spain	696	102%	100%	1,438	109%	102%	2,472	100%	100%
Finland	27	110%	101%	65	108%	101%	142	100%	100%
France	1,045	101%	100%	1,896	101%	100%	3,188	97%	98%

Table 11: Replacing wealth above threshold using Regression method including Forbes (bn EUR)

	≥2M			≥1M			≥500T		
	Above threshold	Tail	Net wealth	Above threshold	Tail	Net wealth	Above threshold	Tail	Net wealth
Austria	444	136%	112%	546	116%	107%	710	110%	107%
Germany	2,678	146%	111%	3,747	134%	112%	5,708	134%	119%
Spain	752	110%	101%	1,521	115%	104%	2,587	104%	102%
Finland	35	141%	103%	74	123%	103%	148	104%	101%
France	1,258	121%	103%	2,149	114%	104%	3,427	104%	102%

In the next section we analyse how replacing the tail by the Pareto distribution changes the coverage ratios of the adjusted concepts discussed above. We limit the analysis to the last estimation method including the Forbes list and the adjusted concept 1. For the calculations we take the mean over the five implicates provided in the HFCS.

4.2   COMPARISON WITH THE ADJUSTED CONCEPT OF FINANCIAL ASSETS

The estimation described in the previous section estimates the tail by taking net wealth as the underlying concept. To make these estimates comparable to the adjusted concept of financial assets discussed in section 2.3 we need to allocate the estimated tail wealth (net wealth) to financial assets, real assets and liabilities as well to an instrument level.

As a first estimate we take the aggregated share for those households that have wealth above each of the thresholds and based on these shares allocate the Pareto tail to an instrument level. To give an indication how this allocation changes with net wealth, Table 12 shows the shares above each of the thresholds that are included in the HFCS. The share of financial assets is increasing while the share of real assets is decreasing with a higher threshold of net wealth for all countries included in the study. For these households, liabilities play a minor role (1% - 6%). For the breakdowns provided in Table 12 we have already reclassified self-employed businesses to financial assets. In Table 13 we show the breakdown of financial assets to the instrument level for the households in the survey having wealth above the threshold of 2M. Here it can be seen that the large part of net wealth for these households consist of the value of self-employed businesses (28% - 51% of net wealth). Most likely the share of financial assets and specifically

equity is further increasing for the wealthier households that are not included in the survey. After allocating net wealth to an instrument level, we apply the same procedure to derive the adjusted concept 1 – reallocate self-employed businesses to financial assets and exclude again those instruments from the adjusted concept that are not or hardly comparable. The effect of this procedure on the coverage ratios is shown in Table 14.

Table 12: Total share of assets and liabilities for households above different thresholds (in % of net wealth)[34]

	≥2M			≥1M			≥500T		
	Financial Assets	Real Assets	Liab.	Financial Assets	Real Assets	Liab.	Financial Assets	Real Assets	Liab.
Austria	59%	42%	1%	57%	44%	1%	50%	52%	2%
Germany	57%	47%	3%	49%	56%	5%	43%	63%	6%
Spain	49%	53%	2%	39%	65%	3%	31%	74%	4%
Finland	67%	39%	6%	46%	59%	6%	31%	76%	6%
France	57%	45%	3%	48%	55%	3%	39%	65%	5%

Table 13: Total share of financial assets for households with wealth above 2M (in % of net wealth)

	Austria	Germany	Spain	Finland	France
DA2101 Deposits	3%	2%	5%	3%	2%
DA2102 Mutual Funds	2%	2%	2%	7%	2%
DA2103 Bonds	2%	2%	0%	1%	1%
DA1140 Value of Self-Employment Businesses	51%	46%	33%	33%	28%
DA2104 Value of Non Self-Employment Private Business	0%	0%	5%	0%	4%
DA2105 Shares, Publicly Traded	0%	2%	2%	22%	6%
DA2106 Managed Accounts	0%	0%	0%	0%	0%
DA2107 Money Owed To Households	1%	0%	1%	0%	0%
DA2108 Other Assets	0%	1%	0%	0%	1%
DA2109 Voluntary Pension/Whole Life Insurance	0%	2%	1%	1%	13%
Total Financial Assets	59%	57%	49%	67%	57%

Table 14: Adjusted concept 1 (financial assets) if tail is added using regression method including Forbes [35]

Country	Adjusted concept 1 (≥2M) S14	Increase (vis-à-vis Table 4)	Adjusted concept 1 (≥1M) S14	Increase (vis-à-vis Table 4)	Adjusted concept 1 (≥500T) S14	Increase (vis-à-vis Table 4)
Austria	110%	(+12%)	105%	(+7%)	103%	(+5%)
Germany	100%	(+14%)	100%	(+14%)	104%	(+18%)
Spain	78%	(+3%)	81%	(+6%)	78%	(+3%)
Finland	59%	(+4%)	59%	(+4%)	56%	(+1%)
France	63%	(+4%)	63%	(+4%)	60%	(+1%)

---

[34] In this breakdown, the value of self-employed businesses has already been classified within the financial assets. The percentages show the total share of assets and liabilities for those households having wealth above each threshold.

[35] The brackets show the change in the coverage ratio to the adjusted concept 1 for the household sector (S14) without the estimation of the tail (see Table 4).

We only apply this procedure using the regression method including the Forbes and the adjusted concept 1. We take the aggregated portfolio structure above each threshold of those households included in the survey. Thus, a lower threshold also implies a lower percentage of net wealth allocated to financial assets as can be seen from Table 12. The intention is to point out one further measurement problem that arises when breaking down the estimated tail net wealth to financial assets and real assets. As the wealth of the Forbes is almost always originated from listed or unlisted incorporations, most likely the large bulk of their wealth is invested in equity. So the estimates gained here can only be understood as an indication of the portfolio allocation of the top tail, but most likely the net wealth estimated by the Pareto could be allocated even more to financial assets/equity reducing the gap for financial assets even further.

As can be seen, adding the estimated tail increases the coverage ratio for all countries but to a larger extent for countries with a less effective oversampling strategy. For Spain, Finland and France the increase from adding the Pareto tail including the Forbes is not sufficient to reduce the gap to FA. For Austria and Germany when taking the threshold of 2M the coverage ratio is increased significantly. To see why this is the case: first, the two countries seem have a lower estimated alpha, hence a bigger estimated tail. Second, the weights allocated to households above the 2M threshold is highest in Austria and Germany. Third, the share of financial assets in Austria is relatively high. Taking a lower threshold implies here apart from taking different estimated alphas and weights also taking a lower portfolio share for financial assets. The share of financial assets for those households having wealth above ≥500T most likely underestimates the share of financial assets of the Pareto tail and thus also underestimated the coverage ratios. The adjusted concept 1 seems to work particular well for Austria and Germany, but one has to keep in mind that two opposing influences still have an impact which have not been estimated here. On the one hand the value of real assets of sole-proprietors may be included in the adjusted concept 1 in financial assets. Excluding these real assets would lead to a lower coverage ratio for financial assets. On the other hand, taking a higher portfolio share of financial assets would lead to an even higher coverage ratio for financial assets. This higher portfolio share can be assumed from the discussion on the Forbes and when taking into account the development of the share in financial assets when increasing the threshold (see Table 12).

5.   SUMMARY AND WAY FORWARD

Using data from the HFCS and FA, this paper makes a thorough comparison between both sources for financial assets for Austria, Germany, France, Spain and Finland. We have briefly reviewed the linkages between both sources on an instrument level. Furthermore, we have pointed out and partly estimated basic statistical differences between both sources that have a potential effect on the coverage ratio.

By developing an adjusted concept of financial assets we have shown that a large part of the gap in comparison to a naïve comparison can already be explained by conceptual differences and by a reclassification of self-employed businesses from real assets to financial assets. Correcting for non-comparable items is essential before being able to actually calculate more reliable coverage ratios.

One challenge in deriving adjusted concepts for financial assets is the treatment of self-employed businesses. Here the issue is which part of sole-proprietors and partnerships included in the survey are assigned in FA to the household sector and which ones are classified as quasi-corporations and hence are recorded in the non-financial corporations sector. In the latter case the household only holds a net equity position in the business (other equity). Contrary, if the sole-proprietors and partnerships are recorded in the household sector the assets and liabilities may be spread over the balance sheet of the household sector and the net value recorded in the survey may very well include real assets and liabilities. Although, this does not have an effect on the coverage ratios in terms of net wealth, it has an effect on the coverage ratio of financial assets on an aggregated level as well as on an instrument level.

Focusing on the wealthiest households we have adapted the estimation procedure from Vermeulen (2016) and replaced those observations in the survey (households) above three different thresholds (500T, 1M and 2M) by an estimated Pareto tail. Hereby we allocate the same weights to the estimated tail as allocated to households above these thresholds in the HFCS. Using the estimates from the Pareto we have shown the effect on the tail itself and the effect on net wealth. For the countries using already an effective oversampling strategy the estimates without the Forbes list seem not to add much to net wealth. For countries with a less effective/no oversampling strategy the Pareto estimates seems to increase the tail and net wealth. One of the contributions of this paper is that we analysed how the coverage ratios for an adjusted concept of financial assets changes when the top tail (above 500 T, 1M and 2M) is replaced by a Pareto distribution including the Forbes and which factors are of importance.

It seems that for counties with an effective oversampling strategy the increase in the coverage ratio is lower than for countries with a less effective oversampling strategy. Apart from oversampling three factors are relevant for the final results: first, the estimated alpha is crucial as *a lower estimated alpha leads to a bigger estimated tail*. Second, *the weight allocated to households having wealth above the thresholds is different* from country to country and hence leads to a different effect on net wealth. Third, *the portfolio allocation to financial assets is relevant when net wealth is broken down to its components (financial assets, real assets and liabilities).* Households having higher net wealth seem to be more invested in financial assets. The analysis shows that it is reasonable to assume that the largest part of financial assets of the wealthiest households is equity. This matters for the estimated coverage ratios for financial assets as a higher portfolio share in financial assets implies that a larger part of the estimated tail wealth is allocated to financial assets.

In the future we need to continue to work on adjusting the concept of net wealth including real assets and liabilities. Hence, we will also develop an adjusted concept of real assets and liabilities to finally derive an adjusted concept of net wealth. For the estimation of the coverage ratio of the different components it would be valuable to have an estimate on the share of financial assets, real assets and liabilities held by sole-proprietors and partnerships as this would give an estimation by how much the adjusted concept for financial assets (adjusted concept 1) overestimates the coverage ratio.

The analysis shows that the threshold for estimating the alpha might be of importance as the outcome of the Pareto index might be quite different when taking different thresholds. Generally there is a trade-off as increasing the threshold decreases the number of households on which the estimates are based on. But taking a lower threshold bears the risk of including observations (households) that are not Pareto distributed. The threshold is of equal importance for taking the portfolio shares of net wealth allocated to financial assets, real assets and liabilities as this has an impact to the coverage ratio for each component of net wealth. In the analysis of this paper we have kept the thresholds for estimating the alpha and the portfolio shares the same. A sensitivity analysis when varying these, e.g. estimating the alpha based on the 500T threshold but varying the share of financial assets held by these households, would be one way to further analyse the effect on the coverage ratios. Although the regression method with the Forbes shows on average lower alphas and hence a bigger tail, the coverage ratios are still dependent on the weight allocated to the tail in the survey. Thus varying the weight and see the effect on the coverage ratio would be worth examining as the weight differs quite a bit between the countries.

Finally returning back to our initial question stated in the title 'Is the Top Tail of the Wealth Distribution the Missing Link between the Household Finance and Consumption Survey and National Accounts?' we have shown that the estimated Pareto tail might explain part of the coverage ratio for financial assets but to a less extent than we initially expected. For those countries that have a less effective oversampling strategy a larger part of the gap to FA seems to be explained by the estimated top tail. But apart from the applied oversampling strategy the change in the coverage ratio seems to depend on the distribution of wealth in each country (leading to different alphas), the weight allocated to households in the survey and the portfolio allocation of the wealthy households. Finding the 'correct' estimates for each measurement problem is a difficult task. The question remains for some countries in our analysis why the coverage ratios using the adjusted concept are still relatively low and further explanations have to be found.

REFERENCES:

Ampudia, M., A. Pavlivkova, J. Slacelek and E. Vogel (2014): "Household Heterogenity in the Euro Area since the Onset of the Great Recession", Working Papers 1705, European Central Bank, Frankfurt am Main 2014.

Andreasch, M., P. Fessler and P. Lindner (2013): "Linking Microdata and Macrodata on Austrian Household Financial Wealth Using HFCS and Financial Accounts Data". Statistiken, Special Issue, pp. 14-23, Oesterreichische Nationalbank.

Antoniewicz, R. (2000):" A Comparison of the Household Sector from the Flow of Funds Accounts and the Survey of Consumer Finances", Federal Reserve Board of Governors Survey of Consumer Finances Working Paper.

Atkinson, A. B and Micklewright, J. (1983): "On the reliability of income data in the family expenditure survey 1970-1977", Journal of the Royal Statistical Society, Series A (General), 146(1): 33-61.

Avery, R. B., G. E. Elliehausen, and A. B. Kennickell (1988): "Measuring Wealth with the Survey Data: AN Evaluation of the 1983 Survey of Consumer Finances", Review of Income and Wealth 37(4): 409-432.

Bach, S., Thiemann, A., and Zucco, A. (2015): The Top Tail of the Wealth Distribution in Germany, France, Spain, and Greece. Discussion Paper 1502, DIW Berlin.

Bover, O., Coronado, E., & Velilla, P. (2014): The Spanish survey of household finances (EFF): description and methods of the 2011 wave. Banco de España occasional Paper, (1407).

Dalitz, C. (2016): Estimating Wealth Distribution: Top Tail and Inequality. Technischer Bericht 2016-01, Hochschule Niederrhein, Fachbereich Elektrotechnik und Informatik.

Davies, J.B. and A. Shorrocks (1999): "The Distribution of Wealth," in A.B. Atkinson and F. Bourguignon, eds., Handbook of Income Distribution: Volume 1, 605-675, North-Holland, Amsterdam.

Dettling, Lisa J., Sebastian J. Devlin-Foltz, Jacob Krimmel, Sarah J. Pack, and Jeffrey P. Thompson (2015): "Comparing Micro and Macro Sources for Household Accounts in the United States: Evidence from the Survey of Consumer Finances," Finance and Economics Discussion Series 2015-086. Washington: Board of Governors of the Federal Reserve System, http://dx.doi.org/10.17016/FEDS.2015.086.

Dolan, Kerry A. (2016): "Methodology: How We Crunch the Numbers", Forbes www-sides, referred on 21. July 2016, http://www.forbes.com/sites/kerryadolan

ECB (2013): The Eurosystem Household Finance and Consumption Survey: Methodological Report for the First Wave. European Central Bank Statistics Paper Series, No. 1.

Eckerstorfer, P., Halak, J., Kapeller, J., Schütz, B., Springholz, F., and Wildauer, R. (2015): Correcting for the Missing Rich: An Application to Wealth Survey Data. *Review of Income and Wealth*.

European System of Accounts 1995 (ESA95): Council (EC) Regulation No 2223/96 of 25 June 1996 on the European system of national and regional accounts in the Community, Official Journal of the European Union L 310/1-469.

European system of accounts (ESA2010), Eurostat/European Commission, Publications Office of the European Union, Luxembourg 2013.

Fesseau, M. and M. L. Mattonetti (2013): "Distributional Measures Across Household Groups in a National Accounts Framework: Results from an Experimental Cross-country Exercise on Household Income, Consumption and Saving", OECD Statistics Working Papers, No. 2013/04, OECD Publishing. http://dx.doi.org/10.1787/5k3wdjqr775f-en

Fesseau, M., F. Wolff and M. L. Mattonetti (2013): "A Cross-country Comparison of Household Income, Consumption and Wealth between Micro Sources and National Accounts Aggregates", OECD Statistics Working Papers, No. 2013/03, OECD Publishing. http://dx.doi.org/10.1787/5k3wdjrnh7mv-en

Gabaix, X. and R. Ibragimov (2011): "Rank -1/2: A Simple Way to Improve the OLS Estimation of Tail Exponents," Journal of Business and Economic Statistics, 29, 24-39.

Henriques, A. M., J. W. Hsu (2012): "Analysis of Wealth Using Micro and Macro Data: A Comparison of the Survey of Consumer Finances and Flows of Funds Accounts", Federal Reserve Board of Governors Survey of Consumer Finances Working Paper.

IMF/FSB report to the G-20 Finance Ministers and Central Bank Governors, http://www.financialstabilityboard.org/publications/r_091107e.pdf (referred on the 23. June 2016)

Kavonius, I. K. and J. Honkkila (2016): "Deriving Household Indebtedness Indicators by Linking Micro and Macro Balance Sheet Data", Statistical Journal of the IAOS (International Association for Official Statistics), Issue X/32, pp XX-XX, IOS Press, Amsterdam 2016 (to be published, accepted).

Kavonius, I. K. and J. Honkkila (2013): "Reconciling Micro and Macro Data on Household Wealth: A Test Based on Three Euro Area Countries", Journal of Economic and Social Policy: Vol. 15: Iss. 2, Article 3.

Kavonius, I. K. and Törmälehto, V-M (2010): "Integrating Micro and Macro Accounts – The Linkages between Euro Area Household Wealth Survey and Aggregate Balance Sheets for Households", IARIW 31st General Conference, St - Gallen, Switzerland, August 22-28, 2010, http://www.iariw.org/papers/2010/7akavonius.pdf

Kennickell, A. B. (2008): The Role of Over-sampling of the Wealthy in the Survey of Consumer Finances. Irving Fisher Committee Bulletin, 28, 403-408.

Kennickell, A. B., & Woodburn, R. L. (1999). Consistent Weight Design for the 1989, 1992 and 1995 SCFs, and the Distribution of Wealth. Review of Income and Wealth, 45(2), 193-215.

Ogwang, T. (2011): Power laws in top wealth distributions: evidence from Canada. *Empirical Economics*, *41*(2), 473-486.

Pérez-Duarte, S., Sánchez-Muñoz, C., & Törmälehto, V. M. (2010): Re-weighting to reduce unit non-response bias in household wealth surveys: a cross-country comparative perspective illustrated by a case study. In Conference on European Quality in Statistics, Helsinki.

Piketty, T. (2014): "Capital in the Twenty-First Century", The Belknap Press of Harward University Press, Cambridge and London 2014.

Rubin, D. B. (2004): Multiple imputation for nonresponse in surveys (Vol. 81). John Wiley & Sons.

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

System of National Accounts 1993 (SNA93), Commission of the European Communities, International Monetary Fund, United Nations, World Bank, Brussels/Luxembourg, New York, Paris, Washington D.C., 1993.

System of National Accounts 2010 (SNA2010), Commission of the European Communities, International Monetary Fund, United Nations, World Bank, New York 2009.

Tissot, B. (2015): "Closing Information Gaps at the Global Level – What Micro Data Can Bring", Irving Fisher Committee Workshop, Warsaw 14-15 December 2015, http://www.bis.org/ifc/events/ws_micro_macro/tissot_paper.pdf

Vermeulen, P. (2016): How fat is the top tail of the wealth distribution?.*Review of Income and Wealth* (forthcoming).

Vermeulen, P. (2016b): Estimating the top tail of the wealth distribution. American Economic Review: Papers & Proceedings 106(5)

Annex

Table A 1: Correspondence table: Financial wealth in HFCS and FA

ESA2010 Code	FA/ Instrument Name	HFCS Variable Code(s)	HFCS Variable	Adjusted concept	Remarks
	Assets				
F.2	**Currency and deposits**				
F.21	Currency	N/A	N/A	Excluded	FA: holdings by households included but estimated due to the lack of direct sources. HFCS: Not collected.
F.22	Transferable deposits	HD1110	Sight accounts	Included	Specific conceptual differences exist but are unlikely to be significant (e.g. deposit like instruments with non-deposit taking corporations are included in the HFCS which are classified as short term loans in FA.
F.29	Other deposits	HD1210	Savings accounts	Included	
F.3	**Debt securities**	HD1420	Bonds	Included	Conceptual differences are not known.
F.4	**Loans**	HD1710	Amount owed to household	Excluded	Not fully comparable, loans between households missing from FA in practice for most countries (conceptually wise they are included).
F.5	**Equity and investment fund shares**			Included	
F.511	Listed shares	HD1510	Publicly traded shares	Included	
F.512	Unlisted shares	HD1010	Investment in non-self-employment not publicly traded shares (ownership only as an investor or silent partner)	(Partly) Included dependent on adjusted concept	- In the HFCS, classification is based on the household's activity in the enterprise. - FA value includes assets that are classified as real wealth in the HFCS (value of self-employment businesses) and has to be reallocated to financial wealth. - The split between 'Unlisted shares' and 'Other equity' cannot be made in the survey. Investments in self-employed businesses could be included in 'Unlisted shares' or 'Other Equity'. - The value of sole proprietorships or partnerships are spread over the different instruments in FA if it is not considered as a separate legal entity (quasi-corporation). - In the HFCS, the value can be provided for the different legal status, although the legal status is not imputed in all countries. ("Unknown" category).
F.519	Other equity	HD0200	Investments in Self-Employment Businesses 1 - Sole proprietorship / independent professional 2 - Partnership 3 - Limited liability companies 4 - Co-operative societies 5 - Non-profit making bodies		

ESA2010 Code	FA/ Instrument Name	HFCS Variable Code(s)	HFCS Variable	Adjusted concept	Remarks
			6 - All other Forms (Spain) 7 - Unknown (not imputed)		
F.521	Money Market Fund shares/units	1320c	Investments in mutual funds c - Funds predominantly investing in money market instruments	Included	Value dependent on fund type not imputed in every country. The breakdown by type of fund may not be available and only the total HD1330 is imputed in all countries. Hence the distinction between MMF and non-MMF funds may not be made in these countries.
F.529	Non-MMF Fund shares/units	HD1320x	a - Funds predominantly investing in equity b - Funds predominantly investing in bonds d - Funds predominantly investing in real estate e - Hedge funds f - Other fund types (specify)	Included	
F.6	**Insurance, pension and standardised guaranteed schemes**				
F.61	Non-life insurance technical reserves	N/A	N/A	Excluded	Non-life included in NA. Assets in non-life (e.g. health insurance, term insurance) can be significant.
F.62	Life insurance and annuity entitlements	DA2109 (Sum of PF0920 over household members)	Voluntary pension/whole life insurance schemes	Included	
F.63	Pension entitlements	Sum of PF0700 over household members	Current Value Of All Occupational Pension Plans That Have An Account	Excluded	It is not clear if defined benefit plans are included in this variable in the HFCS. Furthermore, pensions are prone to measurement problems in surveys.
F.64	Claims of pension funds on pension managers	N/A	N/A	Excluded	F.64-F.66 likely to be irrelevant for households.
F.65	Entitlements to non-pension benefits	N/A	N/A	Excluded	

ESA2010 Code	FA/ Instrument Name	HFCS Variable Code(s)	HFCS Variable	Adjusted concept	Remarks
F.66	Provision for calls under standardised guarantees	N/A	N/A	Excluded	
F.7	**Financial derivatives**	HD1920	Other financial assets	Excluded	Financial derivatives are not a separate item in the HFCS and are included in 'Other financial assets'. Definition of 'Other accounts receivable/payable' not comparable to 'Other financial assets', different definitions.
F.8	**Other accounts receivable/payable**				
		HD1620	Managed accounts	Included	May be spread over the FA balance sheet of the household depending on set up of the management and dependent on the assets invested in. Does not, however, affect comparability of total financial assets.