Survey Research Methods (2016)
Vol. 10, No. 2, pp. 119-142
doi:10.18148/srm/2016.v10i2.6305

ISSN 1864-3361
http://www.surveymethods.org

Comparing Wealth – Data quality of the HFCS

Anita Tiefensee
Hertie School of Governance
Berlin, Germany

Markus M. Grabka
DIW Berlin & TU Berlin
Berlin, Germany

The Household Finance and Consumption Survey (HFCS) provides information about household wealth (real and financial assets as well as liabilities) from 15 Euro-countries around the year 2010 (first wave). The survey will be the central dataset in this topic in the future. However, several aspects point to potential methodological constraints regarding cross-country comparability. Therefore the aim of this paper is to get a better insight in the data quality of this important data source. The framework for our analysis is the "Guidelines for Micro Statistics on Household Wealth" from the OECD (2013). We have two main focuses: First, we present a synopsis of cross-country differences, which is the core of the paper. We compare the sampling processes, the interview modes, the oversampling techniques, the unit and item non-response rates and how it is dealt with them via weighting and imputation as well as further points which might restrict cross-country comparability of net wealth. We classify the individual country behavior and evaluate the impact on net wealth. Second, we give a first insight in the selectivity of item non-response in a cross-national setting. We make use of logit models to identify differences in item non-response patterns across countries as well as between households within countries.

Keywords: Household Finance and Consumption Survey (HFCS); data quality; cross-country comparability; wealth; item non-response
JEL-code: D31, N30, C83

1 Introduction

In spring 2013 the European Central Bank (ECB) released the Household Finance and Consumption Survey (HFCS, 2010). The first wave of the HFCS provides information about household wealth, income and indicators of consumption and credit constraints from (nearly) all Euro-countries[1] around the year 2010. The survey is of general interest because for the first time it is possible to compare real and financial assets as well as liabilities on the household level between Euro-countries.[2] For several countries this was not even possible on a national level before. The survey will therefore be the central dataset in this topic in the future.

The release of the data caused a lot of attention and was followed by several discussions because the bigger picture drawn by the numbers was somehow surprising. The figures (all ECB, 2013a) showed that in comparison with the other investigated countries the households in Luxemburg and Cyprus have the highest median wealth (397,800 Euro and 266,900 Euro). On the other side German households hold only 51,400 Euro, this is the lowest value, followed by Slovak households (61,200 Euro). The median over all surveyed Euro-countries is 109,200 Euro. The explanations of the ECB for these differences ranged from structural differences like household sizes or age patterns, over different macroeconomic dynamics to varying historical, cultural and institutional factors like intergenerational transfers, land ownership or allocation of household wealth between real and financial assets (ECB, 2013b). The public debate quickly added additional explanations like wars, the German reunification, transition processes in eastern countries or tax systems (Fessler, 2013). Furthermore the survey only collects private pension wealth while wealth accrued from public pension schemes is not provided by the HFCS. The latter may affect wealth accumulation dependent on the generosity of public pension systems (Fessler & Schürz, 2015; OECD, 2013).

Another important but not widely discussed source for differences between countries might be due to methodological

Contact information: Anita Tiefensee, Hertie School of Governance, Friedrichstr. 180, 10117 Berlin, Germany, (tiefensee@hertie-school.org)

[1] Current countries: Austria, Belgium, Cyprus, Finland, France, Greece, Germany, Italy, Luxembourg, Malta, the Netherlands, Portugal, Slovakia, Slovenia and Spain. Additional countries in the future: Estonia, Ireland and Latvia.

[2] Other projects like the Luxembourg Wealth Survey (LWS) try to make independent wealth surveys comparable with each other via an ex-post harmonization. The HFCS is intended for comparison from the start.

reasons. A look into the data documentation (ECB, 2013c) reveals further restrictions for comparison. Reference periods are not the same in all countries, only some countries oversampled the wealthy households, which for them can increase precision, while those without oversampling may suffer from coverage errors. Some countries did not survey all mandatory variables and Finland estimated a lot of information from registers. Very low initial response rates in some countries are another challenge for cross-country comparability. Furthermore the item non-response rate is a serious problem in lots of surveys especially if they deal with a sensitive and difficult subject like wealth (Frick, Grabka, & Marcus, 2010a; Kennickell, 2011). The potential underlying selectivity of non-response needs to be considered in a proper imputation otherwise it influences survey estimates.

All of these aspects point to potential constraints when making cross-country analyses regarding wealth based on the HFCS. Therefore the aims of this paper are to get a better insight in the data quality of the first wave of this important data source to help users to understand and interpret their results better as well as to make a contribution to improve data quality further. We first define the term quality by applying the "Guidelines for Micro Statistics on Household Wealth" from the OECD (2013). Then we go through the seven proposed criteria institutional environment, relevance, coherence, timeliness, accessibility, comparability and accuracy (section 2). The main focus will be on the last two points. Therefore we present a synopsis of cross-country differences which is the core of the paper (section 3). We compare the sampling processes, the interview modes, the sample sizes, and the unit and item non-response rates and how it is dealt with them via weighting and imputation. In addition we show which countries oversampled wealthy households based on which data, compare the survey periods as well as further points which might restrict country comparability. This part is mainly based on the documentations of the ECB and the national central banks; what we add is further literature on the individual topics to classify the individual country behavior and to evaluate the impact on net wealth and its components. Under the characteristic "accuracy" we focus on non-response and in particular on item non-response in a cross-national setting (section 4). We make use of logit models to identify differences in item non-response patterns for different wealth components across countries as well as between households within countries and thus give a first insight in the selectivity of item non-response in a cross-national setting. This approach is to our knowledge completely unique for this set of countries. In section 5 we summarize our results and make suggestions for improvements for the dataset.

2 Definition of Data Quality

The framework for our analysis is the internationally agreed "Guidelines for Micro Statistics on Household Wealth" from the OECD (2013) which provides "guidelines on best practice methods of assessing quality" (OECD, 2013, p. 191). There, in accordance with the International Organization for Standardization (ISO 9100), quality is defined as the "degree to which a set of inherent characteristics fulfills requirements" (OECD, 2013, p. 191). Before this backdrop the OECD defines the following seven criteria to describe data quality: institutional environment, relevance, coherence, timeliness, accessibility, comparability, accuracy. Hereinafter we apply all these criteria to the HFCS.

2.1 Institutional Environment

Institutions producing the data should be "impartial, objective, independent from political and other institutional pressures and free of potential conflicts of interest" (OECD, 2013, 192f). In addition they need to be "adequately resourced to produce the statistics of interest" and have a "mandate to collect the relevant data" (OECD, 2013, p. 202). In the case of the HFCS the survey is coordinated by the ECB and carried out by the national central banks and in three cases by the national statistical institutes (France, Finland and Portugal).[3] The main operational regulations of these institutions can be found in the Treaties of the European Union, the Statute of the European System of Central Banks as well as in the national bank acts/laws on national statistic institutes which ensure the required points.

2.2 Relevance

Relevance defines the "degree to which statistics meet the needs of actual and potential users ... thus it depends upon coverage of the required topics and the use of appropriate definitions or concepts" (OECD, 2013, p. 193). The HFCS surveys an extensive balance sheet (see figure 1) and some variables about income, consumption and credit constraints. An extension to this balance sheet can be claims on public pension funds although there are discussions whether this should be a component of total private household wealth or not. The literature tends to recommend to analyze it not as standard component but alongside private net wealth (Davies & Shorrocks, 2000; OECD, 2013).[4]

[3] Together all these institutions build the Household Finance and Consumption Network (HFCN).

[4] Public pension funds are not tradable or acceptable as collateral (OECD, 2013). Further, there is no standard market interest rate (such as interest and dividends from capital), limits to bequeathing (which goes beyond survivors pensions) as well as the issue of liquidation/immediate availability and finally pension wealth is not associated with economic power (as compared to high net wealth). The OECD (2013) recommends to exclude entitlements of all social

Main residence
\+ Other real estate property
\+ Investments in self employed businesses
\+ Vehicles
\+ Valuables

= *Real assets*

Sight accounts
\+ Saving accounts
\+ Saving plans with building and loan associations
\+ Life insureance policies
\+ Mutual funds
\+ Debt securites
\+ Publicly trated stocks
\+ Money owned to household
\+ Other

= *Financial assets*

Debt by main residence
\+ Debt by other real estate property

= *Collateralized debt*

Bank overdrafts
\+ Credit card debt
\+ Other uncollateralized loans

= *Uncollateralized debt*

Real assets + Financial assets = Gross Wealth

Coll. debt + Uncoll. debt = Debt

Gross Wealth - Debt = Net wealth

Figure 1. Households' balance sheet in the HFCS (Source: Fessler, Mooslechner, & Schürz, 2012)

2.3 Coherence

Coherence concerns the "adequacy [that the data can] be reliably combined in different ways and for various uses" (OECD, 2013, p. 199). Internal coherence refers to "coherence between different economic variables collected in the same cross-section or inferable from the longitudinal component of the survey" (OECD, 2013, p. 199). So far for most countries in the HFCS only the first part is relevant. It is among other things achieved via the editing and imputation process as well as the survey mode CAPI (computer assisted personal interviews), which automatically recognizes inconsistencies (Banca d'Italia, 2012). External coherence is related to the "coherence with external sources of information, such as the national accounts or population census" (OECD, 2013, p. 199). Net wealth levels in the HFCS are lower than in the national accounts and range between 50 and over 90 percent. However, there exist significant differences between the two concepts related to methodology, coverage etc. (ECB, 2013c).[5] In some countries another possibility is to check for external coherence trough comparison with existing wealth surveys.[6]

2.4 Timeliness

Timeliness is the "interval of time between publication and the period to which the data refer" (OECD, 2013, p. 201). The ECB released the HFCS in spring 2013, before that, extensive data preparations were made. The reference periods of ten countries are between spring 2010 and summer 2011. For the other five they are however between the winters 2008 and 2009, thus a time-lag of five years can be on hand.

2.5 Accessibility

Accessibility refers to the "degree to which users are able to use the data. The concept of accessibility spans physical requirements for access, structure of the data files, tools available for access, restrictions placed on accessing the data, adequacy of supporting documentation" (OECD, 2013, p. 201) With an academic affiliation the micro data is accessible with a manageable effort.[7] The data is already multi-

security schemes, however, "primarily for practical reasons and to maintain consistency with the SNA [System of National Accounts] definition of financial assets" (OECD, 2013, p. 71). In many countries reliable estimates may not be available (yet). However, they also acknowledge that in several countries claims on public pension funds are a highly relevant wealth element and "without some measurement of this asset, any estimate of total wealth is an underestimate of the true wealth of the household" (OECD, 2013, p. 119). As Frick and Grabka (2013) have shown the net present value of all public pension entitlements for example in Germany nearly doubles standard aggregate net wealth and thereby significantly reduces wealth inequality.

[5]For example both reflect a different target population. Compared with the HFCS SNA also include non-profit institutions serving households, like churches, trade unions or political parties. In addition they cover persons living in institutions (further details on that regarding the HFCS see table 2). This will certainly lead to higher wealth aggregates in the SNA. Another caveat for comparison is that in surveys the valuation of assets is based on the self-assessment of the households and in the SNA it is based on estimated market values (ECB, 2013c).

[6]This is in principle possible in case of e.g. Germany. However, here too differences in concepts need to be considered. While in the Socio-Economic Panel (SOEP) information about the value of vehicles is not collected (Grabka & Westermeier, 2014), the German Income and Expenditure Survey (EVS) provided by the Federal Statistical Office do not ask for business assets (Frick, Grabka, & Hauser, 2010b).

[7]Researchers have to fill in a form in which they explain the ECB who they are, what they want to do with the data, how they will store it and that they ensure confidentiality. After examination trough the ECB the researcher will get a link for download of the data set (available formats: SPSS, Stata, ASCII).

ple imputed and contains survey weighting factors as well as information for calculating the variance (bootstrap replicate weights, which contains for example sample design information). However, some variables like the geographical location of the households are not (centrally) available for all countries and have to be requested individually at each central bank. The ECB also provides files and explanations how to work with the individual files (household, personal etc.) and implicates (1-5).[8] Still, the data is quite complex to deal with. The ECB meanwhile tries to account for that for example via a google group.[9] Paradata is in general not accessible.

3 Comparability Issues of the HFCS

Comparability refers to the "degree to which data can be compared over domains, across countries, and over time" (OECD, 2013, p. 198). To get a better understanding which countries are comparable with each other in which dimensions or under which conditions regarding net wealth table 2 summarizes main comparability issues (ECB, 2013c, 2013d). In addition to the extensively methodological report of the ECB some countries reveal further more or less detailed information about their procedure (see Banca d'Italia, 2012; Bover, 2011; Caruana & Pace, 2013; Mathä, Porpiglia, & Ziegelmeyer, 2012; OeNB, 2012; Statistics Finland, 2015; Tzamourani, 2012; von Kalckreuth, Eisele, Le Blanc, Schmidt, & Zhu, 2012).

3.1 Sampling, sampling frames and target population

In the first wave in total 62,521 households where surveyed (see figure 2). Slovenia has the smallest net sample size consisting of 343 households, which is therefore "not (be) deemed fully representative for the country" (ECB, 2013c, p. 9), followed by Malta (843 households) and Luxembourg (950 households). In the last two countries analyses for small subgroups tend to be hindered due to the small sample size. On the other side France surveyed the most households (15,006) followed by Finland (10,989) and Italy (7,951). However, even for those countries analyses at a detailed regional level seem to be not reasonable. In general larger samples reduce sampling errors (see also section 4) (Fowler, 2014) and allow more precision (lower variance) when estimating unknown population parameters.

All surveys except for Slovakia have a probabilistic design. This means each household in the sample frame has a positive probability of being drawn into the sample. However, Slovakia used a quota sampling for the first wave (based on the income distribution of EU-SILC). Therefore correct sampling and standard errors as well as confidence intervals are impossible to calculate (Fessler & Schürz, 2013).[10] Types of sampling frames differ across countries. In most countries units were drawn from some sort of population or dwelling register; in Belgium from telephone register and in

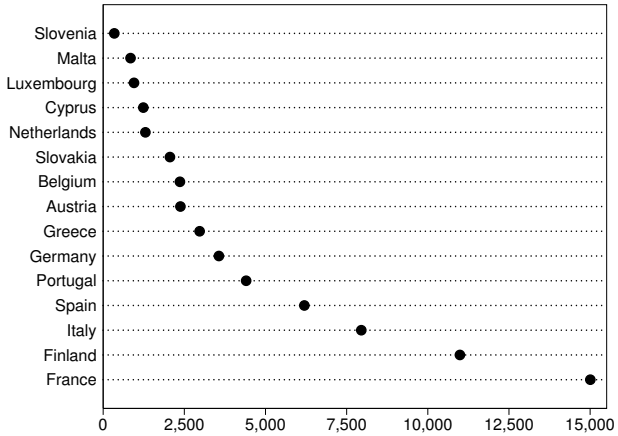


Figure 2. Number of surveyed households in HFCS by country (in total 62.521 households) (Source: ECB, 2013c)

Cyprus from the customer register of the electricity authority. The stratification criteria as well as the number of stages also differ between the countries. The target population of the HFCS consists of all members of private households residing in the national territory at the time of data collection. Persons living in collective households and institutions as well as homeless are excluded in most of the countries. How well the sampling frames represent this target population is not clear for each country. In particular a telephone register may not cover the total population given that some households do not have a telephone or there are telephone numbers which are protected and thus not available (Häder, Häder, & Kühne, 2012).[11] In the Netherlands people who do not speak Dutch and also blind people were excluded from the target population from the beginning, which most likely bias mean net wealth upwards, given that migrants' wealth is typically lower than average wealth (Cobb-Clark &

[8] Further accompanying documents are the survey questionnaire, variables descriptions and a central methodological report. Here it would be nice to also have methodological reports (in English) for the individual countries, explaining their procedure in more detail (see for for example OeNB, 2012).

[9] https://groups.google.com/forum/#!forum/hfcs-users (02.02.2016)

[10] Slovakia will have a probabilistic design from the second wave on.

[11] Unfortunately information about the existence of a telephone in the surveyed households is not available in the HFCS. Thus we make use of the SOEP to show effects on net wealth (Wagner, Frick, & Schupp, 2007). Those households in Germany who do not report to have a telephone have only one third of net wealth compared to those who have both a landline telephone and a mobile phone (59,000 Euro vs. 182,000 Euro in 2012). Thus it can be assumed, that ceteris paribus in Belgium net wealth most likely is significantly overestimated.

Hildebrand, 2006).[12] In Greece smaller villages where excluded (comprising about seven percent of the total number of households).[13]

3.2 Survey modes and interviewer training

The survey mode is consistent in most of the countries. They mainly used Computer Assisted Personal Interviews (CAPI) – only Cyprus, Finland and the Netherlands mainly/only used Paper- and Pencil Interview (PAPI), Computer Assisted Telephone Interviews (CATI) or Computer Assisted Web Interviews (CAWI). The literature shows that CAPI is the most reliable method for data collection (Honkkila & Kavonius, 2013). Face to face surveys have notably higher response rates and lower item non-response rates than those without (Tourangeau, Rips, & Rasinski, 2000), but also construct more socially desirable answers (De Leeuw, 1992, 2008). Therefore especially the (item) non-response behavior of the Netherlands which mainly used CAWI has to be investigated in more detail (see section 4).[14] Finland in addition drew a lot of information from registers or by a modeling process based on previous survey data (Statistics Finland, 2015) which might be problematic with regard to cross-country comparability as has been stressed by Lohmann (2011). In general administrative register data have two major advantages, they usually cover the whole population and accuracy is typically ensured. Disadvantages are systematic error, which can occur for example from the modeling process based on previous survey data, and conceptual differences for example due to a different definition of certain variables like in the case of business.[15] Furthermore in some cases (as for fiscal purposes) households have an incentive to minimize their values (OECD, 2013). Unfortunately, the impact of different survey modes on net wealth cannot be analyzed with the HFCS. However, the so-called CHINTEX-project compared mean equivalence income for Finland using survey data and register data from the ECHP for the very same persons. Here the result was a significant underreporting in the survey data of nearly 8 percent for the total population in 1995 (Rendtel, Nordberg, Jäntti, Hanisch, & Basic, 2004). For the top income decile this underreporting further increases to almost 20 percent. If this finding is also true for wealth then the Finnish data of the HFCS tend to be systematically higher and may show less wealth inequality for the whole population than if the same survey mode had been used.[16]

Further differences are found regarding the length of interviewer training. In the majority of the countries it is at most eight hours. In France and Spain interviewers were trained almost 30 hours. Taking into account that these countries continued preexisting wealth surveys it can be assumed that the interviewers in these countries are much more experienced than in others and might therefore have a positive impact on e.g. item non-response behavior and response quality in general.

3.3 Unit non-response and weighting

The response rates (see figure 3) in the 15 countries range from only 18.7 percent in Germany to almost 70 percent in France, where participation is compulsory like in Portugal (64.1 percent) "though participation is never enforced" (41 ECB, 2013c).[17] In Finland, France, Italy, the Netherlands and Spain a preexisting wealth survey was adjusted and continued.[18] The response rates for the countries with a preexisting panel component where on average higher because people are already used to the survey content and the interviewer. It is also well known that panel surveys are affected by learning effects (Haunberger, 2011) and by selective panel attrition (Kroh, 2014). Thus households from cross-sections may most likely differ from those of existing panel studies.[19]

Survey weights are used to adjust for the unit non-response behavior; which is assumed not to be random due to the sensitive topic of the survey (see for example Kennickell and Woodburn, 1999 and section 4). Therefore non-response weights are calculated to reduce bias (Fessler &

[12] Once again SOEP data can be used to give an idea about the relevance of such an assumption. Those migrants who state that they speak the national language only fairly bad or not at all show an individual net wealth which only achieves 22 percent of the net wealth of the total population. However, the affected population in Germany is rather small with a share of roughly 1 percent.

[13] Again one would expect an upward bias, given that the value of property wealth is usually lower in the countryside than in city regions. For Germany with SOEP data this presumption can be confirmed. Households living in small villages with less than 2,000 inhabitants, show a net wealth of nearly 90 percent of that of the total population.

[14] It is also known from the experiences of EU-SILC that the use of different survey modes may influence data quality. E.g. Germany was the only country with self-administered interviews in EU-SILC while other EU-SILC countries performed predominantly CAPI. However, self-administered interviews with cover letters only in the local language tend to discourage in particular migrants to take part in such a survey which in fact happened in Germany (Hauser, 2007).

[15] In Finland the variable "business" does not contain the value of non-self-employment not publicly traded businesses (ECB, 2013c).

[16] A comparison of income inequality supports this assessment, given that the Finnish register data show a smaller Gini-coefficient and less income poverty than survey data (tab 4 Rendtel et al., 2004)

[17] Finland even has a higher response rate (82.2 percent), but due to the fact that Finland drew a lot of information from registers, it is not completely comparable.

[18] In Cyprus and Portugal an existing wealth survey was discontinued and replaced by the HFCS.

[19] Frick and Grabka (2005) have shown, that first time respondents have a significantly higher share of item-non response than panel participants from subsequent waves. The effect on net wealth then heavily depends on a proper imputation technique to control for this type of measurement error.

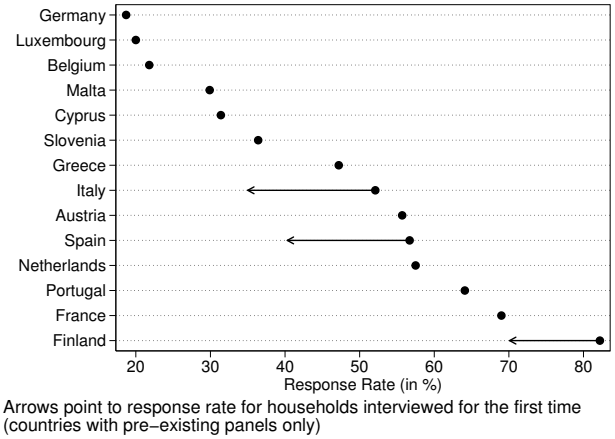


Figure 3. Initial response rates in the HFCS (Source: ECB, 2013c)

Schürz, 2013). This is done in a similar way in each country. In a first step design weights are calculated as the inverse probability of being selected into the sample. In a second step the design weights are adjusted to coverage issues and non-response behavior and are calibrated to external sources. From the documentation of the ECB it is not completely clear in which ways the calculation processes differ between the countries. Information is available on calibration variables like age, gender, household size, region and some other variables – all from external sources – as well as on the existence of weight trimming or limitations for weight adjustment factors. However, just a few country documentations identify which information is available on non-respondents and/or if additional information collected from the interviewer is used. Therefore the quality of the weights cannot be judged with the available HFCS data here.

3.4 Oversampling

The goal of oversampling is to increase precision of estimates for wealth in general and for those (financial) assets which are only owned by a small fraction of households (Fowler, 2014; Kennickell, 2007). Theoretically oversampling has no effect on the expected values of measured wealth. The empirical fact that countries without oversampling of the wealthy might be more likely to underestimate wealth (especially at the top) and wealth inequality (Bover, 2011; Kennickell, 2007) comes from the connection of how the oversampling is often practically implemented. As a priori information about the wealth distribution (or some correlated variable) is needed, countries with an accurate oversampling scheme mostly achieve higher coverage than those without. The method for oversampling as well as the range in which people were identified as wealthy was different in almost each country (see Table 1). Spain and France oversampled wealthy households based on individual information about net wealth from a wealth tax register. Finland and Luxembourg used individual income information – Cyprus household information about the electricity bill. Greece, Belgium and Germany applied geographical information, in the first case regional real estate prices and in the two others regional-level income information.[20] However, this can lead to increased estimation variance if variance within areas is high (OECD, 2013). Austria, Portugal, Slovenia and also Germany oversampled big cities and Italy, Malta, the Netherlands and Slovakia did not oversample at all. It is obvious that pure regional information, income or electricity bills are only weak proxies to identify top wealth households. The effective oversampling rate demonstrates the degree to which the share of wealthy households in the sample exceeds their share in the population. The effective oversampling rate of the top ten percent (explanation see table 1) clearly shows that Spain and France have the highest values with 192 and 129 respectively (based on individual wealth information). Vermeulen (2014) tries to compensate the undercoverage of top wealth holders in the HFCS by a simulation of the respective population using information from the Forbes list. He showed that the wealth share hold by the top 1 percent significantly increase for those HFCS-countries who do not implement a proper oversampling. The strongest effect can be observed for Austria where this share increased from 23 percent to almost 36 percent. But even also in other countries this increase amounts to 9 percentage points in Germany, 8 for the Netherlands or 6 percentage points for Italy, while for Spain no effect of this simulation of top wealth holders has been shown. Therefore it can be expected that those surveys with a proper oversampling get a good coverage of the top wealth holders. For countries with no or weak oversampling it can be assumed that they underestimate the true degree of wealth inequality, wealth levels and aggregates.[21]

[20] In Germany this is only done for municipalities < 100,000 inhabitants. Big cities (> 100,000 inhabitants) were divided into wealthy sections and others based on information about the quality and type of the building and purchasing power indicators.

[21] The authors tried to approximate the degree of bias on wealth levels and inequality by excluding the top wealth holders of the oversample. However, the HFCS data do not provide any indicator variable to differentiate between "normal" sample members and those from the oversample. It would be helpful to find such a variable in a next release of the HFCS. In order to reflect the relevance of such an oversampling the SOEP can be used exemplary. In 2002 a top income sample was drawn to improve capturing wealthy households. When excluding this oversample mean net wealth would drop by more than 6 percent (based on own calculations).

Table 1
Oversampling strategy

Country	Oversampling wealthy households	Basis for oversampling	effective oversampling rate of the top 10%[a]
Spain	Yes	individual information from taxable wealth	192
France	Yes	individual information from net wealth	129
Finland	Yes	individual information from income + socio-economic status	68
Luxembourg	Yes	individual information from income	55
Cyprus	Yes	household information from electricity bills	81
Germany	Yes	geographic income and other information	117
Belgium	Yes	geographic income information	47
Greece	Yes	geographic real estate price information	−2
Slovenia	Partly	geographic information (Ljublijana, Maribor)	22
Portugal	Partly	geographic information (Lisbon, Porto)	16
Austria	Partly	geographic information (Vienna)	1
Netherlands	No	-	87
Italy	No	-	4
Malta	No	-	−5
Slovakia	No	-	−11

Source: Based on ECB (2013c)
[a] Explanation taken from ECB (2013c, p. 37): "if the share of rich households in the net sample is exactly 10%, the effective oversampling rate of the top 10% is 0. If the share of households in the wealthiest decile is 20%, the effective oversampling rate is 100, meaning that there are 100% more wealthy households in the sample than would be if all households had equal weights".

3.5 Item non-response and imputation

Another important issue of data quality is the share of item non-response (INR) and how it is dealt with (Bover, 2010; Zagorsky, 1999). The share of INR rates differ significantly between different assets and liabilities and also within one component between the countries (see section 4). Except for Finland and Italy all countries used multiple imputation by chained equations (MICE) to estimate the missing values (five implicates).[22] The number of covariates used for the multiple imputations greatly differs between countries as well as for assets and liabilities. The goal is not to use the same set of variables for all countries. From the literature we know that a more detailed set of covariates may better capture the selectivity of the non-response behavior than only a very limited set of covariates (Barceló, 2008; Mathä et al., 2012). In Spain 239 covariates were used to impute missing values of the household main residence, Malta only used four, the Netherlands six. For the most important mortgage for the household main residence Greece used 154 variables, Slovenia only four. Which variables were used or how the INR patterns look like in the individual countries and for the wealth components or other indications for the imputation quality (such as trace plots or comparisons of the distributions of observed and imputed data values) are not documented for most of the countries.[23] Therefore the quality of the estimations and if "variance was traded against bias" or the other way round[24] cannot be judged completely (Fessler & Schürz, 2013, p. 47).[25]

Single imputation provides only one value and does not reflect the uncertainty of the imputation. Therefore the standard error of the variable might be underestimated (Rubin, 1987). Due to very low rates of INR two countries make use of single instead of multiple imputations. Finland estimated a lot of information from registers and Italy had a special agreement with the survey company, which only considers

[22] The same procedure is used in the Survey of Consumer Finances in the USA (Kennickell, 1998) and in the Spanish Survey of Household Finances (Barceló, 2006), which served as prototypes. The later one is now part of the HFCS.

[23] It would be highly welcomed if this kind of documentation would be made available by the ECB.

[24] If variance is traded for bias, estimations will more often be significant "even though they may have larger true bias" (Fessler & Schürz, 2013, p. 47). Fessler and Schürz (2013) provide more details about the bias variance trade off with regard to the HFCS.

[25] To estimate the effect of different imputation techniques on net wealth a simulation strategy would be necessary as has been used by e.g. Frick and Grabka (2005). They applied the very same imputation method to two different surveys and compared the effect on income levels and inequality to the original imputation strategy, finding significant differences.

interviews below a certain level of INR as completed. On the one hand this approach may maximize superficial data quality. On the other hand it may influence the incentives of the interviewers or the respondents; in consequence they might report unreliable values. In addition it can be assumed that such a precondition may yield to a selective sample of respondents and can lead to selectivity bias in the survey estimates (OECD, 2013).[26]

3.6 Reference periods

The reference periods for the assets and liabilities also differ between the countries and thus impair cross-country comparability (see figure 4). They range from 2008 to 2011 (almost three years between the first and the last one), but for most countries they are between 2010 and 2011.[27] Especially for Spain comparability issues might occur due to the financial crisis and its effects. Here the reference period already starts in November 2008. Estimates from the OECD show that house prices declined by more than 10 percentage points (real) from 2008 to the beginning of 2011.[28] Considering the deteriorating house prices the mean of real estate in Spain would ceteris paribus only account to 221,000 Euro instead of 251,000 Euro and net wealth would then only amount to on average 252,000 Euro instead of 291,000 Euro (significant to the five percent level).[29] In addition one should also account for inflation and the interest rate. Despite from Spain this should also be done in Greece and France where the reference periods also start before 2010. Another challenge is the long lasting collection periods in some countries. Sierminska and Medgyesi (2013) show that during the collection period in many countries the stock-market index fluctuated between 10 percent and 20 percent. The house-price index already changed by 11-18 percent within the collection period in Austria, Germany, Spain and Portugal. While financial assets are usually only hold by a minority of the households, housing makes up to two-third of the wealth portfolio. Summing up, a fixed reference point (e.g. the 31.12. of a given year as has been done in three countries) may improve cross-country comparability instead of rather long reference periods.

3.7 Questionnaire and variable catalogue

The individual national banks agreed on a common blueprint questionnaire for the HFCS, which served as a basis for the national questionnaires. This means it is not directly translated (not input harmonized) to account for the diversity of (financial) institutions in Europe as well as to accommodate pre-existing wealth surveys (von Kalckreuth et al., 2012). The questionnaire is divided into three parts: (1) harmonized data, which is output harmonized and collected in every country (so called core variables), (2) harmonized data, which is also output harmonized however not collected in every country (so called non-core variables) and

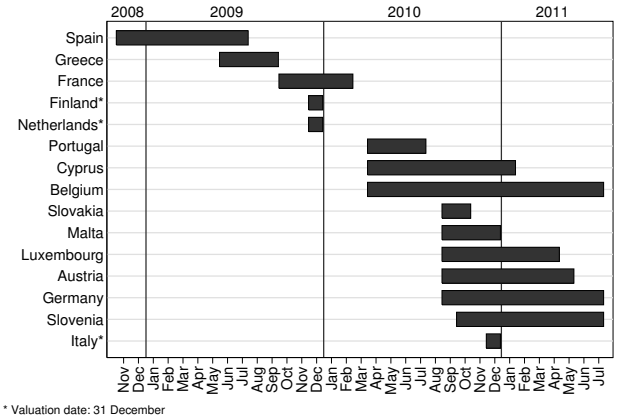


Figure 4. Reference periods for assets and liabilities: 2008-2011 (Source: ECB, 2013c)

(3) country specific data, which is not harmonized. Real and financial assets as well as liabilities fall into category (1). In each household a reference person[30] answered the very detailed and extensive questions about household's assets and liabilities as well as some information about intergenerational transfers, gifts and consumption patterns. Information about income, pensions and insurances policies as well as employment are available for each person in the household older than 16 years. Demographic characteristics (age, gender, country of birth, since when living in current country, relationship to the reverence person) are available for all household members.

[26] For illustration: Austria and Italy have comparable average values for household main residences. In Austria the average value is 258,000 Euro and drops to 245,000 Euro if imputed values are not considered. This is significant to the five percent level.

[27] The reference periods will be further harmonized and will coincide from wave three onward (Tzamourani, 2012).

[28] Information about house prices are used from OECD (doi: ~10.1787/hsprice-table-2014-1-en). House price indices are often based on current transaction prices and not on self-estimation like it is done in surveys. Therefore differences may accrue and the estimations can only be seen as an approximation.

[29] Spain already surveyed the second wave of the HFCS in 2010. Maybe this will serve as a better basis to compare net wealth between the countries.

[30] For selection criteria see ECB (2013c, pp. 16-17).

Table 2
Methodological differences across countries in the HFCS

	RR	Sampling Design				Int. T.[a]	Weighting		Oversampling		Imputation[c]		
n	(%)	Frame	Strata	Excl. groups	Mode	h	Trim.	Lim.[b]	Basis	Details	Hmr	Mort	Save
Austria													
2380	56	List of enumeration districts; register of postbox addresses	NUTS III region, population of municipality	Homeless, all institutionalized population	CAPI	7	no	no	Geographic areas	some oversampling in Vienna because of higher expected non-response rate	104	51	133
Belgium													
2364	22	Telephone register and street register	NUTS I region and average income by neighborhood of residence	Homeless, prisoners	CAPI	6	no	no	Geographic information about average income	Neyman allocation, based on the standard deviation of income in stratum and stratum size	46	31	49
Cyprus[d,e]													
1237	31	Customer register of the Electricity authority of Cyprus	Census districts divided into urban and rural	Homeless, prisoners, population of the areas of Cyprus not under the effective control of the Government of Cyprus	CAPI (12%) PAPI (88%)	5	no	no	Household information about electricity bill	61% of the gross sample was drawn from households within the top 10% according to electricity consumption	50	38	48

Continues on next page

Continued from last page

RR		Sampling Design				Int. T.[a]	Weighting		Oversampling		Imputation[c]		
n	(%)	Frame	Strata	Excl. groups	Mode	h	Trim.	Lim.[b]	Basis	Details	Hmr	Mort	Save
Finland[f,g]													
10989	82[h]	Central population register using master sample of 50,000 persons 16+ and members living in the same household-dwelling unit	Socio-economic criteria of the highest income-earner	All institutionalized population	CAPI (3%) CATI (97%)	40[i]	yes	yes	Individual information about income and socio-economic status	from population register (High-income employees, self-employed and farmers)	Single imputation		
France[g]													
15006	69[j]	List of geographical units (based on Census); list of dwellings	Region, regional population; socio-economic criteria	All institutionalized population	CAPI	27	no	n.a.	Individual information about net wealth	Four strata have been made. For each primary unit and each stratum, an allocation proportional to main residences is computed. Then, a systematic selection is made within each couple stratum-primary unit	17	12	21

Continues on next page

Continued from last page

RR		Sampling Design				Int. T.[a]	Weighting		Oversampling		Imputation[c]		
n	(%)	Frame	Strata	Excl. groups	Mode	h	Trim.	Lim.[b]	Basis	Details	Hmr	Mort	Save
Germany													
3565	19	Clusters of addresses from municipalities (NSI); list of street sections, population registers of municipalities	Demographic size, average taxable income of municipalities; additionally wealth-related parameters of street sections for large municipalities	Homeless, all institutionalized population	CAPI	11	no	yes	Geographic information about taxable income and other information	HH in smaller municipalities (<100,000 inhabitants) were oversampled using income tax statistics. A municipality is declared as "wealthy" if a share of more than a fixed percentage of taxpayers with a total taxable income was above a fixed threshold. In larger municipalities (>100,000 inhabitants) the oversampling of wealthy street sections was based on information about the quality and type of the building and purchasing power indicators (Knerr, Chudziak, Gilberg, & Kleudgen, 2012)	84	10	17

Continues on next page

Continued from last page

RR		Sampling Design				Int. T.[a]	Weighting		Oversampling		Imputation[c]		
n	(%)	Frame	Strata	Excl. groups	Mode	h	Trim.	Lim.[b]	Basis	Details	Hmr	Mort	Save
Greece													
2971	47	List of municipalities (Census); random routing for secondary sampling units	NUTS II region, degree of urbanization	Homeless, all institutionalized population, smaller villages (comprising about 7% of the total number of households)	CAPI	8	yes	no	Geographic information about real estate prices	The sampling rate for Athens and Thessaloniki is proportional to the real estate prices of each cluster	233	154	49
Italy[g,k]													
7951	52[h]	List of municipalities; resident lists from municipalities	NUTS II region and population of the municipality	Homeless, all institutionalized population	CAPI (85%) PAPI (15%)	8	no	no	-	-	Single imp. (excep. save: 10)		
Luxembourg													
950	20	Addresses of fiscal households from social security register	Individual income, nationality, employment status	Diplomats, non-resident citizens, homeless, international civil servants and in general households where no individual is entitled to be registered in the social security register, all institutionalized population	CAPI	6	no	yes	Individual information about personal income subject to social contributions	20% of the gross sample was drawn from the top income decile according to the social security register and the self-employed-headed fiscal household subpopulation	86	118	31

Continues on next page

Continued from last page

RR		Sampling Design				Int. T.[a]	Weighting		Oversampling		Imputation[c]		
n	(%)	Frame	Strata	Excl. groups	Mode	h	Trim.	Lim.[b]	Basis	Details	Hmr	Mort	Save
Malta													
843	30	Dwelling register of the NSI	Statistical region	Diplomats, non-resident citizens, armed forces, homeless, civilians living in military institutions, prisoners	CAPI (81%) PAPI (19%)	9	yes	yes	-	-	4	10	14
Netherlands[g]													
1301	58	Postal addresses	-	Blind people, people who do not speak Dutch, all institutionalized population	CAWI	- .	no	no	-	-	6	7	7
Portugal[e]													
4404	64[j]	List of geographical areas; list of private dwellings, from Census	NUTS II region	All institutionalized population, homeless, people living in military area	CAPI	16	no	no	Geographic areas	Metropolitan areas of Lisbon and Porto oversampled, 50% of gross sample drawn from these areas	16	23	17
Slovakia[l]													
2057	-	List of municipalities, households chosen by random walk.	NUTS III region, population of municipality. In each stratum, ten income quotas were prescribed, which interviewers had to fulfill	Homeless, all institutionalized population	CAPI	4	no	yes	-	-	102	31	69

Continues on next page

Continued from last page

RR		Sampling Design				Int. T.[a]	Weighting		Oversampling		Imputation[c]		
n	(%)	Frame	Strata	Excl. groups	Mode	h	Trim.	Lim.[b]	Basis	Details	Hmr	Mort	Save
Slovenia[m]													
343	36	List of districts from Census; list of persons 16+ from population register	Population of the municipality, with adjustments for expected non-response	All institutionalized population, diplomats, homeless, non-citizens, armed forces, civilians living in military area	CAPI	7	no	yes	Geographic areas	Municipalities of Ljubljana and Maribor were oversampled, as higher non-response rates were expected	47	4	14
Spain[g]													
6197	57[h]	Municipal census (list of addresses) supplemented by tax office information; list of addresses	Population of the municipality, taxable wealth	All institutionalized population	CAPI	28	no	yes	Individual information about taxable wealth	Eight wealth strata were defined and were oversampled progressively at higher rates	239	104159	

Source: if not otherwise noted ECB (2013c, d).

Note: Estonia, Ireland and Latvia are not part of the first.

[a] Interviewer Training [b] Limits for weight adjustment factors [c] Imputation technique: multivariate imputation by chained equations, MICE, unless otherwise noted. Number of covariates used for main variables: value of household main residence (hmr), outstanding amount of most important hmr mortgage (mort), value of savings accounts (save). [d] Statement of the ECB: "The data for Cyprus appears not to be comparable with those for other Euro-area countries in a number of dimensions and should therefore be interpreted with caution." (ECB, 2013c, p. 4) [e] Existing wealth survey was discontinued and replaced by the HFCS.

[f] Register and estimated data. [g] Preexisting wealth survey continued. [h] Repsonse rate for HH interviewed for the first time are as follows: Finland 70; Italy 35; Spain 40 [i] Includes general interviewer training modules of the NSI. [j] Participation compulsory.

[k] Only interviews with a level of item non-response below a certain threshold were considered [l] Quota sampling for the first wave; all other countries probabilistic design. In the second wave the country will adopt a probabilistic design.

[m] Reduced sample size; "not [be] deemed fully representative for the country" (ECB, 2013c, p. 9).

A closer look into the data documentation and the variable catalogue reveals nevertheless some comparability issues regarding the core variables (ECB, 2013c, 2013d, 2013e). The biggest deviations are in Finland: Several core variables are not provided at all: valuables, non-self-employment not publicly traded business, additional assets in managed accounts, money owed to the household, other assets, outstanding credit line/overdraft balance and outstanding credit cards balance.[31] It is obvious that net wealth in Finland is biased downwards given these restrictions. The average share of these missing wealth components from total assets is in the other countries almost nine percent. For liabilities it is a bit over one percent. In consequence this would mean a significant increase of net wealth ceteris paribus from 161,500 Euro to 179,000 Euro in Finland. Other variables are only available in an "aggregated form". This means for example for mortgages on the household main residence Finland only provides one variable with all mortgages on the household main residence whereas all other countries asked for the first, the second, the third and all additional mortgages on the household main residence (all together maximal four variables per household). This practice is also applied in several other countries for some assets and liabilities categories (see Figure A1). Therefore the variables could be underestimated because people might tend to forget about a small for example mortgage if not asked separately for it. In addition analysis with all countries cannot be done separately for all single components for example mortgages. Furthermore the variable "mutual funds" is not collected in a uniform way over the countries. Taken together researchers should check carefully depending on the individual research question if the chosen variables are really comparable between the countries.

4 Accuracy

Accuracy is linked to the "degree to which the data correctly allow estimation of the population characteristics they are designed to describe" (OECD, 2013, p. 193). Usually it is described in terms of sources of errors. The total survey error (TSE) refers to the "accumulation of all errors that may arise in the design, collection, processing, and analysis of survey data" (OECD, 2013, p. 193). The sources of error can be divided into sampling and non-sampling error. The first one refers to an "inaccuracy that arises because data is collected only from a sample that may not be fully representative of the total population of interest" (OECD, 2013, p. 202). The second one "mainly relate[s] to measurement, data collection and processing" (OECD, 2013, p. 193). It can be classified in specification error (relevant variables are missing or are only approximated), coverage error (sampling does not completely cover reference population) non-response error (households do not participate in the survey or do not answer all applicable questions), measurement error (errors made by the interviewer or the respondents) and processing error (inaccuracy during data entry, editing etc.). Typically accuracy can be improved more or less in every survey. A big challenge to data quality in cross-sectional surveys comes from the response process; it can induce bias to the estimates. Therefore we will focus in the following analysis on non-response and especially on item non-response in the HFCS.[32]

4.1 Item Non-Response in the HFCS

A common problem in population surveys and in particular in surveys dealing with wealth is the failure to collect complete information. While a refusal to the total questionnaire is named unit non-response (UNR), item non-response (INR) refers to single questions that are not answered. The UNR behavior can be adjusted for through weighting of the data and INR is typically corrected through imputation (also see section 3). A survey with a higher response rate is assumed to produce a better and less biased sample than one with lower rates (Fowler, 2014). INR may be caused by a respondent's reservation to answer a question that appears to be too sensitive, i.e. it affects confidentiality and privacy or simply from the fact that the correct answer is not known (given the underlying complexity of the surveyed construct). In general, simple demographic information such as gender, age or marital status is not very sensitive to ask for, leading to low incidence of INR. Wealth or income questions, however, are typically associated with higher rates of INR (see for example Grabka & Westermeier, 2014; Riphahn & Serfling, 2005). Furthermore the survey mode (self-administered vs. conducted by interviewers), the question structure (for example open-ended questions) and the interviewer's characteristics (like experience or character) can have an effect on INR (Groves, Dillman, Eltinge, & Little, 2001). There is increasing literature which explicitly acknowledges the INR phenomenon in micro-econometric research as a specific form of measurement error (see for example Cameron & Trivedi, 2005, p. 193). Most importantly, INR on wealth questions has been found to be selective with respect to inequality (see for example Frick et al., 2010b) and thus can lead to biased results.

As long as the missing process of INR is completely missing at random (MCAR) the potential bias could be disregarded (see Rubin, 1987). However, it is typically assumed that INR follows a missing at random process (MAR), which means that the missing data depend on observed information in a data set. Another type of missing data is called missing

[31] The last liability is also not surveyed in France due to institutional differences (ECB, 2013c, p. 83).

[32] Another reason for our choice is that due to confidential reasons users of the HFCS are not able to use paradata or information about non responding units etc. Therefore the possibilities of investigations are rather limited.

Figure 5. Relevance of imputation – Weighted sum of all components of the aggregate that were imputed divided by the weighted sum of the aggregate variable (Source: ECB, 2013c, p. 57)

not at random (MNAR). Here the missing data cannot be explained by observed characteristics and may be for example dependent on missing values itself. The latter both missing mechanisms are non-ignorable and need to be carefully considered. In general older people and those with less education have a higher probability for INR (Groves et al., 2001). It has been shown for example in the SOEP that the probability for missing wealth information is lower for males, persons with higher education levels and civil servants. It is higher for self-employed (Frick et al., 2010a). A proper imputation has to consider the missing process and thus the underlying selectivity.

Based on the imputation method applied in the HFCS the relevance of the imputed values is almost 30 percent for gross real assets in Austria and more than 40 percent for gross financial assets in the Netherlands. For the other countries which make use of multiple imputation the respective shares vary between five and 30 percent (see figure 5). Hence INR and the respective imputation have a significant impact on wealth levels and inequality.

We will therefore analyze the INR patterns in the HFCS for selected assets and liabilities in all countries. We want to know, whether the selectivity of INR is uniform across countries or if there are structural differences – which one could interpret as cultural discrepancy of INR. As Couper and De Leeuw (2003) argue, non-response in cross-national studies has so far not been extensively researched, and this is still the case. However, differential response rates and patterns between countries can threaten the validity of cross-national comparisons (Couper & De Leeuw, 2003). In case of sensitive information such as wealth with rather high INR the problem of cross-country comparability may be of important relevance.

In order to reduce complexity we will focus on assets and liabilities with a high incidence and those with a high quantitative relevance.[33] As assets we choose the variables *"household main residence"* (real asset) and *"saving accounts"* (financial asset); both have an incidence greater than 50 percent. Regarding relevance (measured by the mean) *"business 1"* (real asset) will be added; the variable has quite a high relevance and even incidence in some countries – especially in the southern part of Europe (Malta, Portugal, Cyprus, Italy). For the liabilities the further investigation is based on *"mortgage of the household main residence 1"* and *"non-collaterised loan 1"* – they both have an incidence around 20 percent. Regarding relevance (measured by the mean) no additional variable is added.

The HFCS provides flag variables which give information about potential reasons for the non-response. In total 16 different values are available. There is a category for edited values and one for estimated ones. The imputed category has five different characteristics. One can differentiate between the responses "Don't know" and "No answer". Furthermore the categories "Originally not collected due to missing answer to a previous question", "Originally collected from a range or from brackets" and "Collected value deleted or value not collected due to a CAPI or interviewer error" can be identified. Furthermore there are different categories for missing values, which were not imputed.[34] In addition one can see of course which variables were collected as complete observations and which were not applicable (recorded as missing). For the following analysis we will concentrate on the edited and imputed ones.

Figure 6 illustrates the share of INR as well as the reason for it for the selected assets and liabilities by country.[35] It is obvious that the shares differ not only significantly between the components but also between and within countries. Especially France and Italy have conspicuous response patterns compared with the other countries. In the case of France the value of household main residence and business 1 has been completely imputed. In the latter case respondents were only asked for a range, which is a slightly different approach than in the other countries where respondents were first asked for the "exact" value and in a second step, if they had difficulties

[33] Regarding assets mutual funds and private pensions/life insurances are excluded from the analysis. The first one is further divided in subgroups in several countries however not in all and the second one is collected on an individual and not on household level. Finland and Slovenia will not be part of the analysis. The first one does not really have item non-response because of the use of register information and the second one has too few households to investigate.

[34] They were not used extensively but only in some countries, which may be a hint of problems in cross-country comparability. However, the documentation gives no information for the rationale of this procedure.

[35] The variables refer to the question of the value of the respective wealth component not the holding of the asset or liability.

answering the question, for a range. On the other side Italy has in almost all observed variables no imputed or edited data, which is due to the already mentioned agreement with the survey company. Relying on this information the percentage of imputed cases of the value of saving accounts above 50 percent is quite surprising. Malta and Austria have – compared with the other countries – quite high shares of INR. High numbers of imputed values introduce, given a proper imputation, broader confidence bands for these values and in consequence a poorer estimator. Countries with a general low rate of INR are Germany[36] and Portugal.

Looking at the different reasons for INR one finds that edited values are scarce. Imputations due to "previous question missing" (filter information whether a specific wealth component is held by the household) just play a minor role, although it arises more often in the Netherlands (especially for business 1). The category "Don't know" is frequently filled for business 1, given that it is rather problematic for a respondent to give a precise valuation of their own enterprise. One can interpret a "No answer" as a strong refusal by the interviewees. However, the respective share is usually not much above 10 percent (exception: "non-collaterised loan 1" in the Netherlands).[37] Deleted and then imputed values (considered incorrect or unreliable) are also especially present for business 1.[38] Most of the imputed values are collected from a range or from brackets, which means that respondents do not know the exact amount of their asset or liability but subsequently unfolding brackets are offered to narrow the value down into ranges, which eases the imputation process. In addition this procedure reduces complete item non-response (Heeringa, Hill, & Howell, 1995; Vazquez Alvarez, Melenberg, & van Soest, 2001). However, from a data user point of view, it yields to missing information which need to be handled carefully.

4.2 Estimation Strategy and Results for Item Non-Response

After the general descriptions of the different INR patterns for the chosen assets and liabilities as well as the differences between the countries we will now analyze the similarities and differences due to characteristics. The multivariate part consists of a logit model with the following equation:

$$p_j(w) = F(\alpha + \beta X_j + \varepsilon_j) \quad ,$$

where p_j denotes INR probability of households in country j for a particular wealth component w, α is an intercept, ε_j a random error term. X_j is the matrix of all explanatory variables which include predominately socio-demographic information: age, gender, work status and education of the reference persons, income of the household, the value of its assets and liabilities,[39] its size and if children under 14 years are present.[40] We will do this for the Euro-countries as a whole (pooled analysis) and for each country separately.[41] France, Italy and Portugal are not part of all regressions.[42]

The average marginal effects of the pooled logit estimations, which calculate INR probabilities for the selected assets and liabilities for the whole Euro-area with dummy variables for the individual countries[43] (see table A1) show that for all chosen wealth components men have a lower INR probability than women. These points into the direction that men tend to know on average their wealth portfolio better or are just more certain of it. The same holds for the liabilities of the household. The higher they are the less likely the household has INR and is therefore on average better informed about its wealth portfolio. The opposite is the case for the assets of the households (only for household main residences the assets follow the same pattern as for the liabilities). Four out of the five components show that people in the first and second income quintile (compared with the third) tend to have higher INR probabilities. In respectively two of the investigated cases the following holds: people over

[36] The low item non-response rate might be a counterpart to the high unit non-response rate. The interviewed household were few but cooperative (von Kalckreuth et al., 2012).

[37] Studies confirm these findings: "questions that require more cognitive effort to be answered receive more don't knows" (here: business) and "more sensitive questions get more refusals" (here: non-collaterised loan 1) (Shoemaker, Eichholz, & Skewes, 2000, p. 1).

[38] It is not fully clear on which basis this decision has been made. The ECB should provide more information why this has been done.

[39] It is assumed that with a higher value of assets and/or liabilities the wealth portfolio gets more complex therefore more questions have to be answered. The variables hence also serve as controls for complexity.

[40] More methodological variables which give information about the interview process (such as information about characteristics of the interviewer, given that there exist interviewer-interviewee responding effects, see Pickery, Loosveldt, & Carton, 2001) are unfortunately not available. This could be a valuable extension of the HFCS data in the future.

[41] Code from the ECB (2013f) and OeNB (2012) has been used to merge the five data files together.

[42] In France the variables household main residence and business 1 are completely imputed. In Italy the variables household main residence, its most important mortgage and the most important loan do not have imputed values. In Portugal the variable business 1 has too few imputed values to be analyzed here.

[43] We choose Germany as a reference category. Slovenia has a very small sample size, Malta and Luxembourg have rather small ones as well. Slovakia used a quota sampling. Finland used register data. The Netherlands applied CAWI, Cyprus PAPI. France and Italy have for some variables 100 or 0 percent imputation. Portugal has for some variables too few imputed observations. We are left with: Austria, Belgium, Germany, Spain and Greece. From this countries Germany has on average the lowest INR-rates and a middle size sample size.

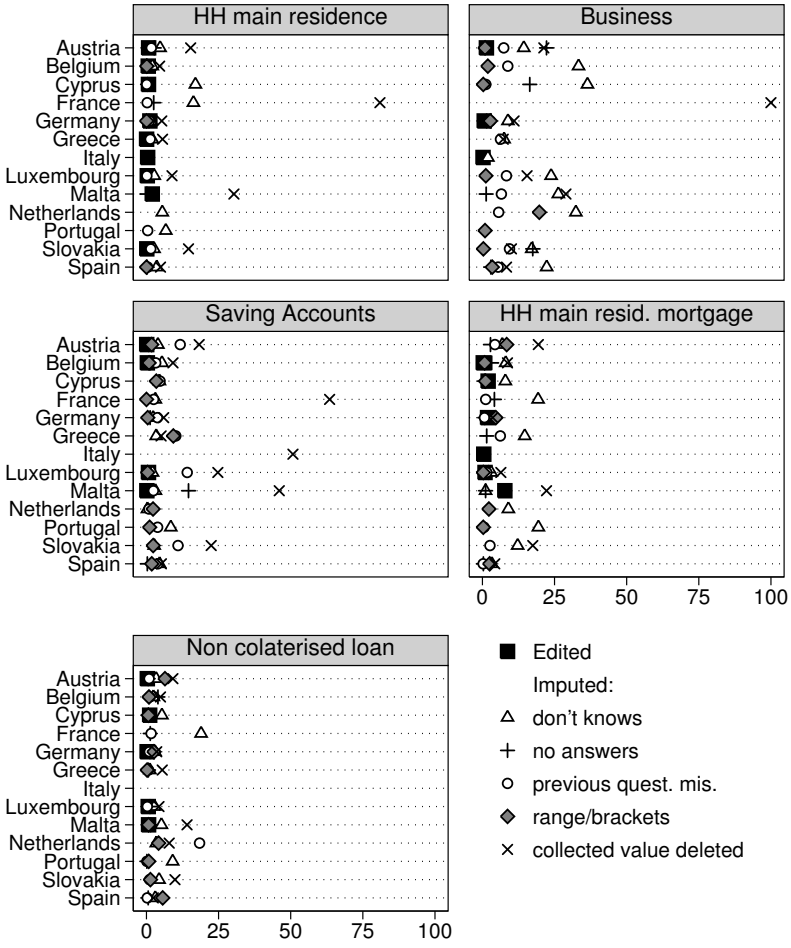


Figure 6. Information from flag-variables for selected assets and liabilities – only those holding the respective wealth/liability component (Source: Own calculations based on HCFS. The data for business in Belgium is based on a corrected version from the national Bank of Belgium)

the age of 65 tend to have higher INR than middle aged ones; self-employed have on average higher non-response shares than employed. This seems intuitively right because in most countries they have to make provisions for their pensions on their own which makes their wealth portfolio more complicated. Persons with primary education tend to have higher probabilities for INR than those with secondary education. The more people live in a household the more complex a wealth portfolio usually is. Thus the probability for INR for people in a two person household is lower than for larger households.[44] In general the results found here follow the patterns described in the previous literature (see for example Frick et al., 2010a; Groves et al., 2001). Controlling for the various countries Austria, Malta and Slovakia show for all five considered wealth components positive effects compared with the reference category Germany.[45]

The effects for the separate country regressions do not show clear patterns (see table A2). Significant effects on the Euro-area level are not generally confirmed on the country level. This can be partly explained by small sample sizes at the country level. But we do also find opposing effects. This does apply for asset and liability levels when looking at the value of housing main residence. At the Euro-area level these covariates point to a significantly lower probability for INR – which is confirmed for at least three HFCS countries – while

[44]The described results are significant at least at the 10 percent level. The estimations are done without weights. Including them only leads to minor changes in the values but not in sign.

[45]Further tests show that the differences between the single countries are in the majority of the cases significant. Further research especially with paradata is needed to analyze differences between countries further and to identify clear structures.

Slovakia stands out with significant positive effects. When looking at saving accounts again Slovakia and Malta show opposite effects for asset levels.[46] A common pattern across countries however can be found for gender. Although this covariate is not significant in all HFCS countries, we find the general effect, that if women are the reference person they have a significantly higher probability for INR. We also confirm for several countries the effect for the household size. The more people are living in a household the more complex is the wealth portfolio and thus the higher is the probability for INR. Altogether there is not a harmonized item non-response pattern across the HFCS countries.[47] This implies the necessity for well aligned imputation models. But as indicated above the documentation is not very precise how and with which imputation models item non-response has been handled. In addition further analyzes which take unit non-response patterns into account would be desirable, however are not possible with the user data set.

5 Conclusion

The HFCS micro dataset is a milestone for cross-country comparisons of private wealth in the Euro-area. The core questionnaire and also the survey methodology was largely pre-harmonized, however there are significant differences across country surveys which impair cross-country comparability of net wealth and inequality, and thus should be carefully taken into account by researchers. The aims of this paper are to get a better insight in the data quality of the first wave of this important data source to help users to understand and interpret their results better as well as to make a contribution to improve data quality further. Based on the "Guidelines for Micro Statistics on Household Wealth" from the OECD (2013) we defined the term quality. We then went through the seven criteria institutional environment, relevance, coherence, timeliness, accessibility, comparability and accuracy (with the main focus on the last two points) and checked how the HFCS implemented these points. We present a synopsis of methodological differences in the HFCS dataset to shed some light on cross-country comparability and thus on potential restrictions for wealth comparisons. We find that net wealth is most likely biased in Finland due to a deviating survey mode and more importantly the absence of various wealth components which may lead to an underestimation of assets (9 percent) and liabilities (1 percent). The Netherlands also deviates with respect to the survey mode, however a quantification of this effect is not possible here. For a wealth survey oversampling of rich households is crucial to reduce potential coverage error. As has been shown by Vermeulen (2014) the lack of oversampling, like in the Netherlands and Italy, lead to a systematic undervaluation of mean net wealth and the wealth share of the top 1 percent. When researchers are interested in subgroup analyses, they should prescind from looking at Slovenia, Malta and Luxemburg due to rather small sample sizes. Finally Slovakia should not be used so far, given that only a quota sample has been used to survey the population, which does not fulfill accepted quality requirements.

In addition under the point accuracy incidence and selectivity of item non-response in a cross-national setting are investigated, which gives a first insight in different non-response patterns for the chosen assets and liabilities as well as for the individual countries. Strong refusals when respondents are not willing to give an answer are acceptably low in the HFCS with a share of less than roughly 10 percent, while UNR play a larger role in the HFCS. Nevertheless, imputation took place up to 100 percent for all those holding a wealth component in France for housing main residence and business assets. After controlling for demographic characteristics via a pooled and country specific logit regressions we in principle confirm the results from individual country cases in the literature but cannot find harmonized item non-response patterns across countries. However, due to the lack of more methodological information (such as interviewer characteristics) or about the unit non-response process we are not able to investigate this aspect further.

Taken together the HFCS is at the moment the best dataset for cross country comparisons of wealth levels and inequality in the Euro-area and it is definitely a first (big) step into the right direction. Nevertheless some improvements are very helpful. First, we would suggest publishing detailed methodological reports for all countries in English, in addition to the methodological report from the ECB. Second, methodological differences which are not based on country specific differences should be reduced or better even vanish. Desirable points to work on are for example the full (output) harmonization of the collected and provided wealth and liability components (which is essential for cross-country comparability of wealth levels and inequality), the application of more harmonized sampling frames, the reduced sample size in Slovenia, the survey modes in Cyprus, the Netherlands and in Finland, a harmonization and shortening (in some cases) of the reference periods, and even a more harmonized procedure with respect to the oversampling of top wealth households, given that these households have a pronounced effect when looking at the skewed wealth distribution. An oversample identifier could also ease analyses about the relevance of such a methodologial add-on. Third, necessary country specific differences like in the case of weighting or imputation should be documented in more detail for example has para-

[46] Again contrarious effects can be found for Slovakia when considering liability levels for mortgages of household main residence.

[47] We also applied a decomposition method suggested by Fairlie (1999, 2005) to identify structural (cultural) differences in the item non-response missing process. However, we could not find unified structures. Results can be found in the Working Paper Tiefensee and Grabka (2014).

data been used for the construction of weights and what covariates are used for the imputation, which would allow other researchers to go further than the study at hand to analyze the data quality of the HFCS. If it is not possible to make some information publicly available due to data protection, one could examine the possibility of a protected platform for data users. Additionally, countries with a very low initial response rate like Germany should make endeavors to raise the willingness of the respondents to take part in such a survey, not only to reduce potential bias in a cross sectional, but more importantly in a longitudinal sense. Further, exemptions such as Italy, that achieved very low INR by a special agreement with the survey company, to only consider interviews below a certain level of non-response as completed, should be avoided to ease comparability. The ECB should also reconsider to survey public pension entitlements in the HFCS and to provide this information in a separate variable to enable data users to decide whether this information might be considered in wealth analyses. Finally, it should be checked whether paradata could be made available for external researchers to better separate substantial cross-country differences from methodological distinctiveness for example for investigating INR patterns further.

Acknowledgments

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.

The authors like to thank Arthur Kennickell, Pirmin Fessler, two anonymous referees and participants at the NOeG conference and the Conference on Crises and the Distribution for their comments and suggestions. The authors gratefully acknowledge funding from the Hans Böckler Foundation (project number: S-2012-610-4).

References

Banca d'Italia. (2012). *Household income and wealth in 2010*. Supplements to the Statistical Bulletin, Banca d'Italia, Rome.

Barceló, C. (2006). *Imputation of the 2002 wave of the Spanish survey on household finances (eff)*. Documentos Ocasionales No 0603, Banco de España, Madrid.

Barceló, C. (2008). *The impact of alternative imputation methods on the measurement of income and wealth: evidence from the Spanish Survey of Household Finances*. Documentos de Trabajo No 0829, Banco de España, Madrid.

Bover, O. (2010). Wealth inequality and household structure. *Review of Income and Wealth*, *52*, 259–290.

Bover, O. (2011). *The Spanish Survey of Household Finances (EFF). Description and methods of the 2008 wave*. Documentos Ocasionales N. 1103, Banco de España, Madrid.

Cameron, A. & Trivedi, P. (2005). *Micro-econometrics, methods and applications*. New York: Cambridge University Press.

Caruana, K. & Pace, C. (2013). *HFCS in Malta: main results of 2010 exercise*. Valletta: Bank Centrali ta'Malta.

Cobb-Clark, D. & Hildebrand, V. (2006). The wealth and asset holdings of US-born and foreign-born households: evidence from SIPP data. *Review of Income and Wealth*, *52*, 17–42.

Couper, M. & De Leeuw, E. (2003). Nonresponse in cross-cultural and cross-national surveys. In J. Harkness, F. van de Vijver, & P. Mohler (Eds.), *Cross-cultural survey methods.* (pp. 157–177). Hoboken, New York: John Wiley & Sons.

Davies, J. & Shorrocks, A. (2000). The distribution of wealth. In A. Atkinson & F. Bourguignon (Eds.), *Handbook of income distribution* (pp. 605–675). Amsterdam: Elsevier.

De Leeuw, E. (1992). *Data quality in mail, telephone and face to face surveys*. Amsterdam: TT-Publikaties.

De Leeuw, E. (2008). *The effect of computer-assisted interviewing on data quality: a review of the evidence*. Department of Methodology and Statistics, Utrecht University.

ECB. (2013a). *Statistical tables*. Frankfurt: European Central Bank.

ECB. (2013b). *Results from the first wave*. Statistics Paper Series No 2/April, Frankfurt.

ECB. (2013c). *Methodological report for the first wave*. Statistics Paper Series No 1/April. Frankfurt: European Central Bank.

ECB. (2013d). *HFCS country surveys metadata information*. Frankfurt: European Central Bank.

ECB. (2013e). *HFCS core variables catalogue*. Frankfurt: European Central Bank.

ECB. (2013f). *Using the HFCS with stata*. Frankfurt: European Central Bank.

Fairlie, R. (1999). The absence of the African-American owned business – an analysis of the dynamics of self-employment. *Journal of Labor Economics*, *17*, 80–108.

Fairlie, R. (2005). An extension of the Blinder-Oaxaca decomposition technique to logit and probit models. *Journal of Economic and Social Measurement*, *30*, 305–316.

Fessler, P. (2013). *Eurosystem Household Finance and Consumption Survey – challenges with regard to cross country comparability*. Presentation in the Research Seminar in Economics, FU Berlin.

Fessler, P., Mooslechner, P., & Schürz, M. (2012). *Eurosystem Household Finance and Consumption Survey - first results for Austria*. Monetary Policy and the Economy Q3/12, Österreichische Nationalbank, Wien.

Fessler, P. & Schürz, M. (2013). *Cross-country comparability of the Eurosystem Household Finance and Consumption Survey*. Monetary Policy & the Economy Q2/13, Österreichische Nationalbank, Wien.

Fessler, P. & Schürz, M. (2015). *Private wealth across European countries: the role of income, inheritance and the welfare state*. Working Paper Series No 1847, European Central Bank, Frankfurt.

Fowler, F. (2014). *Survey research methods*. Los Angeles: Sage.

Frick, J. & Grabka, M. (2005). Item-non-response on income questions in panel surveys: incidence, imputation and the impact on the income distribution. *Allgemeines Statistisches Archiv*, *89*(1), 49–61.

Frick, J. & Grabka, M. (2013). Public pension entitlements and the distribution of wealth. In J. C. Gornick & M. Jäntti (Eds.), *Income inequality: economic disparities and the middle class in affluent countries* (pp. 362–388). Stanford University Press.

Frick, J., Grabka, M., & Hauser, R. (2010b). *Die Verteilung der Vermögen in Deutschland. Empirische Analysen für Personen und Haushalte*. Forschung aus der Hans-Böckler-Stiftung Nr. 118. Berlin: Edition Sigma.

Frick, J., Grabka, M., & Marcus, J. (2010a). *Editing und multiple Imputation der Vermögensinformation 2002 und 2007 im SOEP. Data Documentation.* Berlin: DIW.

Grabka, M. & Westermeier, C. (2014). *Anhaltend hohe Vermögensungleichheit in Deutschland*. DIW Wochenbericht 9/2014, Berlin.

Groves, R., Dillman, D., Eltinge, J., & Little, R. (Eds.). (2001). *Survey nonresponse*. New York: John Wiley & Sons.

Häder, S., Häder, M., & Kühne, M. (Eds.). (2012). *Telephone surveys in europe: research and practice*. Berlin, Heidelberg: Springer.

Haunberger, S. (2011). *Teilnahmeverweigerung in Panelstudien*. Wiesbaden: VS Verlag für Sozialwissenschaften.

Hauser, R. (2007). *Probleme des deutschen Beitrags zu EU-SILC aus der Sicht der Wissenschaft – Ein Vergleich von EU-SILC, Mikrozensus und SOEP*. RATSWD working paper No. 3, Berlin.

Heeringa, S., Hill, D., & Howell, D. (1995). *Unfolding brackets for reducing item nonresponse in economic surveys*. PSID Technical Series Papers #95-01. University of Michigan.

HFCS. (2010). *Household Finance and Consumption Survey*. Data file edition from 2014. European Central Bank, Frankfurt.

Honkkila, J. & Kavonius, I. (2013). *Micro and macro analysis on household income, wealth and saving in the Euro area*. Working Paper Series No 1619, European Central Bank, Frankfurt.

Kennickell, A. (1998). *Multiple imputation in the Survey of Consumer Finances*. Proceedings of the Section on Survey Research Methods, Joint Statistical Meetings, Dallas.

Kennickell, A. (2007). *The role of over-sampling of the wealthy in the Survey of Consumer Finances*. Survey of Consumer Finances Working Papers, Federal Reserve Board, Washington, D.C.

Kennickell, A. (2011). *Look again: editing and imputation of SCF panel data*. Prepared for the 2011 Joint Statistical Meeting, Miami.

Kennickell, A. & Woodburn, R. (1999). Consistent weight design for the 1989, 1992 and 1995 SCFs, and the distribution of wealth. *Review of Income and Wealth*, *45*, 193–215.

Knerr, P., Chudziak, N., Gilberg, R., & Kleudgen, M. (2012). *Methodenbericht – Vermögenssurvey. Erste Erhebungswelle 2010/2011*. Infas.

Kroh, M. (2014). *Documentation of sample sizes and panel attrition in the 7th german Socio Economic Panel (SOEP) (1984 until 2012)*. SOEP Survey Papers 177, DIW Berlin.

Lohmann, H. (2011). Comparability of EU-SILC survey and register data: the relationship among employment, earnings and poverty. *Journal of European Social Policy*, *21*(1), 1–18.

Mathä, T., Porpiglia, A., & Ziegelmeyer, M. (2012). *The Luxembourg HFCS (LU-HFCS): introduction and results*. Working Paper No 73, Banque Centrale du Luxembourg, Luxembourg.

OECD. (2013). *OECD guidelines for micro statistics on household wealth*. Paris: OECD Publishing.

OeNB. (2012). *HFCS 2010 – methodological notes for Austria*. Monetary Policy & The Economy Q3/12, Österreichische Nationalbank, Wien.

Pickery, J., Loosveldt, G., & Carton, A. (2001). The effects of interviewer respondent characteristics on response behavior in panel surveys. *Sociological Methods and Research*, *29*, 509–523.

Rendtel, U., Nordberg, L., Jäntti, M., Hanisch, J., & Basic, E. (2004). *Report on quality of income data*. CHINTEX working paper #2.

Riphahn, R. & Serfling, O. (2005). Item non-response on income and wealth questions. *Empirical Economics*, *30*(2), 521–538.

Rubin, D. (1987). *Multiple imputation for nonresponse in surveys*. New York: John Wiley & Sons.

Shoemaker, P., Eichholz, M., & Skewes, E. (2000). Item nonresponse: distinguishing between don't know and

refuse. *International Journal of Public Opinion Research*, *14*, 193–201.

Sierminska, E. & Medgyesi, M. (2013). *The distribution of wealth between households*. Research Note 11/201, European Commission, Brussels.

Statistics Finland. (2015). Retrieved from http://www.stat.fi/meta/til/vtutk%5C_en.html

Tiefensee, A. & Grabka, M. (2014). *Comparing wealth – data quality of the HFCS*. DIW Discussion Papers 1427.

Tourangeau, R., Rips, L., & Rasinski, K. (2000). *The psychology of survey response*. New York: Cambridge University Press.

Tzamourani, P. (2012). *Coherence assessment of the Greek 2009 HFCS with comparable data sources*. Athen: Bank of Greece.

Vazquez Alvarez, R., Melenberg, B., & van Soest, A. (2001). *Nonparametric bounds in the presence of item nonresponse, unfolding brackets, and anchoring*. ISSC Discussion Paper Series. University College Dublin.

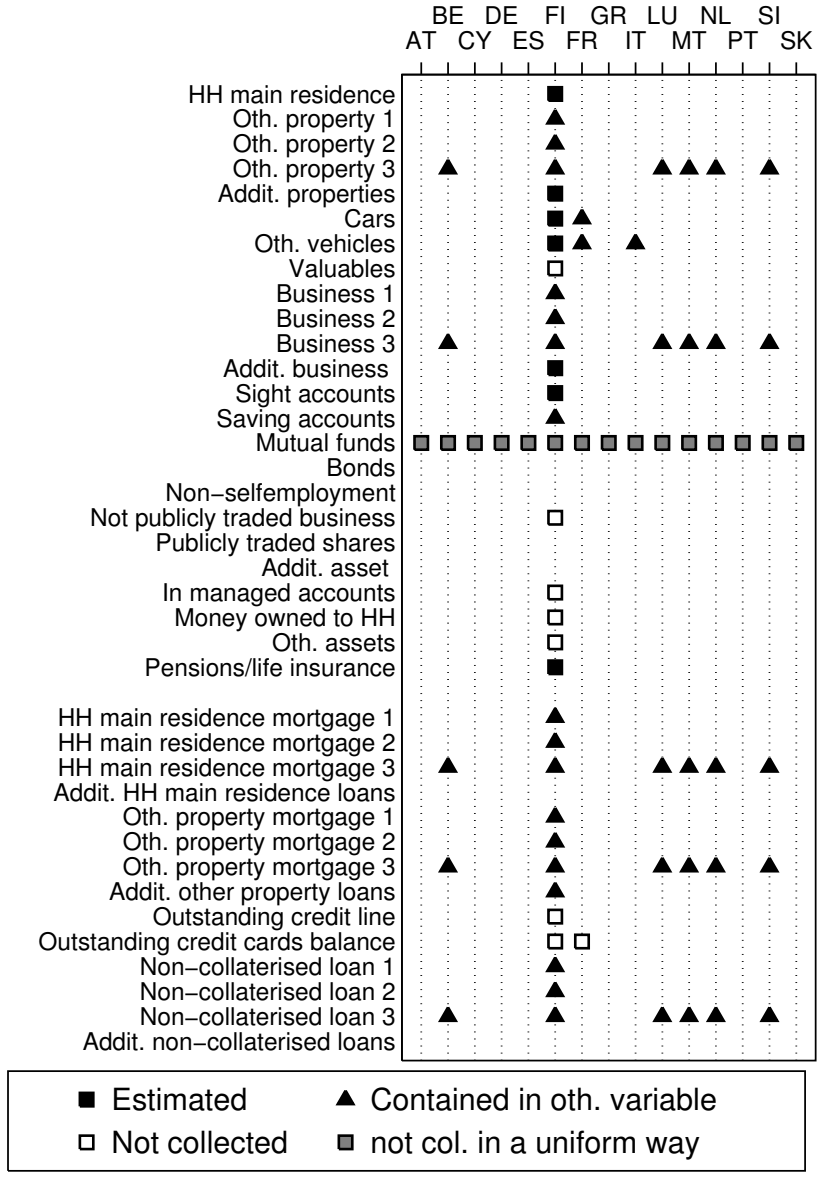
Vermeulen, P. (2014). *How fat is the top tail of the wealth distribution?* Working Paper Series No 1692, European Central Bank, Frankfurt.

von Kalckreuth, U., Eisele, M., Le Blanc, J., Schmidt, T., & Zhu, J. (2012). *The PHF: a comprehensive panel survey on household finances and wealth in Germany*. Discussion Paper No 13/2012, Deutsche Bundesbank, Frankfurt.

Wagner, G., Frick, J., & Schupp, J. (2007). The German Socio-Economic Panel Study (SOEP) – scope, evolution and enhancements. *Schmollers Jahrbuch*, *127*, 139–169.

Zagorsky, J. (1999). Young baby boomers' wealth. *Review of Income and Wealth*, *45*, 135–156.

Appendix
Figures and Tables



Source: HFCS (2010)

Figure A1. Differences in variables collection

Table A1
Average marginal effects of the pooled logit estimations

	HMR		Business 1		Saving account		HMR mortgage 1		Loan 1	
	b	std. err.	b	std. err.	b	std. err.	b	std. err.	b	std. err.
Men	−0.031***	0.005	−0.074***	0.014	−0.022***	0.006	−0.036***	0.009	−0.022**	0.009
Age group (reference: age 45-54)										
16-34	0.020**	0.009	0.029	0.023	0.009	0.010	0.020	0.014	−0.017	0.015
35-44	0.015*	0.008	0.030*	0.018	0.003	0.009	0.014	0.012	0.000	0.013
55-64	0.012	0.008	−0.003	0.018	−0.011	0.009	−0.019	0.015	0.009	0.014
65plus	0.032***	0.010	0.000	0.028	−0.010	0.012	−0.038	0.027	0.045**	0.022
Employment status (reference: employed)										
self-employed	0.031***	0.008	−0.011	0.015	0.025***	0.009	0.002	0.012	−0.004	0.013
unemployed/other	0.008	0.008	0.048**	0.024	0.000	0.010	0.005	0.015	0.004	0.014
retired	0.000	0.009	0.035	0.029	−0.014	0.010	0.012	0.021	−0.044**	0.018
Education (reference: secondary)										
primary	0.007	0.007	−0.013	0.019	0.013*	0.007	0.018	0.012	0.025**	0.011
tertiary	0.010*	0.006	−0.003	0.015	−0.015**	0.007	−0.010	0.010	−0.014	0.011
Income quintiles (reference: 3rd)										
1st	0.026***	0.008	0.003	0.035	0.040***	0.010	0.050**	0.020	0.010	0.018
2nd	0.009	0.007	0.056**	0.027	0.003	0.009	0.030*	0.016	−0.005	0.015
4th	−0.007	0.007	−0.021	0.021	−0.004	0.008	−0.004	0.013	−0.019	0.013
5th	0.002	0.007	−0.023	0.020	−0.002	0.008	0.007	0.013	−0.013	0.013
Log assets	−0.010***	0.003	0.012**	0.005	0.006***	0.002	0.019***	0.006	0.011***	0.003
Log liabilities	−0.003***	0.001	−0.003**	0.001	−0.004***	0.001	−0.024***	0.003	−0.021***	0.002
HHousehold size (reference: 2 persons)										
1	0.001	0.007	−0.051**	0.024	−0.048***	0.007	0.007	0.015	−0.006	0.014
3	0.001	0.007	0.005	0.018	0.044***	0.008	−0.009	0.014	−0.006	0.013
4	−0.001	0.008	0.009	0.019	0.059***	0.010	0.015	0.014	0.013	0.015
5plus	0.025**	0.010	0.019	0.023	0.075***	0.013	0.010	0.018	0.024	0.018
HH with child. < 14yrs.	−0.008	0.008	−0.026	0.018	−0.046**	0.009	−0.016	0.012	0.004***	0.013
Mortgage	0.008	0.009								
Country (reference: Germany)										
Austria	0.097***	0.010	0.272***	0.029	0.230***	0.013	0.214***	0.022	0.137***	0.029
Belgium	−0.035***	0.012	0.132***	0.031	0.080***	0.015	0.084***	0.021	0.019	0.029
Cyprus	0.079***	0.012	0.186***	0.027	−0.048*	0.029	−0.002	0.026	−0.060*	0.034
Spain	−0.027***	0.009	0.101***	0.021	0.004	0.016	−0.052**	0.022	−0.024	0.023
France					0.485***	0.009	0.115***	0.018	0.148***	0.019
Greece	−0.041***	0.011	−0.084***	0.030	0.125***	0.039	0.076***	0.024	−0.074**	0.033
Italy			−0.478***	0.035	0.315***	0.013				
Luxembourg	0.034**	0.014	0.146***	0.040	0.273***	0.017	0.014	0.029	−0.034	0.036
Malta	0.125***	0.012	0.218***	0.044	0.430***	0.018	0.143***	0.038	0.126***	0.040
Netherlands	−0.052***	0.017	0.219***	0.046	−0.284***	0.035	−0.038	0.027	0.333***	0.028
Portugal	−0.055***	0.011			−0.025	0.016	0.074***	0.021	−0.040	0.028
Slovakia	0.041***	0.011	0.187***	0.027	0.222***	0.019	0.158***	0.028	0.059**	0.029
N	19,959		4,618		29,353		8,479		8,990	
R_P^2	0.06		0.23		0.20		0.05		0.06	

Source: Own calculations based on HFCS (2010)
* $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$