universität wien

DISSERTATION

Titel der Dissertation

"Essays in Survey-Based Economic Household Analysis"

Verfasser

Mag. Peter Lindner MSc

angestrebter akademischer Grad

Doctor of Philosophy (PhD)

Wien, 2013

Studienkennzahl lt. Studienblatt:	A 094 140
Dissertationsgebiet lt. Studienblatt:	Volkswirtschaftslehre (Economics)
Betreuer:	Univ.-Prof. Dipl.-Ing. Dr. Robert M. Kunst

Contents

0.1 Introduction

This dissertation consists of four chapters each containing a complete research paper using survey based household level micro data in order to tackle research questions that are laid out in each part separately. Chapters are ordered in a chronological way so that the first chapter was mostly written in 2010 and the last chapter contains the most recent work.

Under the title "Decomposition of Wealth and Income using Micro Data from Austria" chapter 1 analyses how the different sources of wealth and income are connected with the overall distribution of wealth and income respectively. A decomposition procedure according to factor sources is applied to data from several household level surveys in Austria for an understanding of the connection between these two distinct distributions and its foundations in the contributing factors. One finds that the various types of wealth holdings and income sources have distinct characteristics, connections, and correlation Austrian economy. This chapter describes the results of the decomposition of wealth and income in combination for the first time in Austria. One of the main drawbacks is given by the lack of a combined dataset that includes the hole balance sheet of households in Austria which is further developed on in chapter 4 making use the in the meantime collected data from the Households Finance and Consumption Survey (HFCS).

In chapter 2, which is based on a joint paper with Pirmin Fessler and Maximilian Kasy, we look deeply into the field of survey methodology in the economically relevant terrain. How does the way one collects information on income, i.e. applying a computer-assisted personal interview in comparison to a computer-assisted telephone interview, affect results on the full income distribution that are estimated with survey data. In doing so, the emphasis is put on the estimation and interpretation of a causal effect of the interview mode on the item non-response of the survey and the level as well as the distribution of income. Lacking an a priory by the researcher

designed experiment, the case for causality of the results is established through the application - apart from standard regression techniques with truly exogenous control variables - of matching estimation techniques from the treatment effect literature and of the recentered influence function in order to estimate the effect on the whole distribution. The results show that applying a personal interview mode instead of the telephone reduces item non-response by about a third and the distribution of income covers a wider range of values. In other words using a telephone interview mode distorts the estimated distribution of income measured by the Gini coefficient significantly downwards in statistical as well as economic terms.

Chapter 3 consists of a paper that is joint with Pirmin Fessler and Esther Segalla. It followed the wide discussion about cross country comparability of the results from the Households Finance and Consumption Survey published by the European Central Bank and tries to address the problem of differences in the household structure if one want to compare household level results from this survey. Using and specifying a particularly fine granulated non-parametric re-weighting approach for a thorough control of differences in the household structure in terms of its size and within household age patterns of up to four household members, we are able to show that a sizeable proportion of the results on net wealth can be attributed to these structural discrepancies. However, there still remain cross country differences in the level and distribution of (net) wealth and the components of it once the unit of collection is controlled for.

The single-authored work that is contained in the 4^{th} and last chapter again makes use of the HFCS and extends the work from chapter 1. Applying now several decomposition techniques from the literature to wealth instead of income, to which variable the techniques has been commonly applied, this part discusses the differences in wealth distributions of euro area countries. Additionally, there is a new decomposition method laid out that is based on the recentered influence function. The results of this chapter do not contradicting the previous findings for the Austrian

case (see chapter 1). That is, that there are some forms of the household wealth portfolio, especially the households main residence, which are major contributing factors to the overall distribution of wealth, while others can explain relatively little (e.g. risky financial investments). From these findings one can also add to the discussion of cross country comparability in the sense that it points towards the important factors that have to be further investigated in order to understand why households display different wealth patters in the countries of the euro area. There are, however, diverging findings over the countries of the elasticity of inequality with respect to a certain component, pointing towards varying impacts of policy measures across the euro area.

Paper specific appendices including additional results and analytical derivations are put at the end of each chapter. For the convenience of the reader and in order to save space the bibliography of all chapters is provided at the end of the thesis, where all references are listed in a alphabetical order. Additionally, abstracts - both in English and German language - are provided in an appendix at the end of the dissertation, which also contains an up to date version of the curriculum vitae. Obviously all the work would not have been possible without the support of family, friends, as well as colleagues and hence acknowledgements round up the thesis. Clearly, I am solely responsible for all potentially remaining errors and the content of the dissertation provides solely the assessment of the author and does not necessarily reflect the opinion of my employer which is at the time of writing the Oesterreichische Nationalbank.

Chapter 1

Decomposition of Wealth and Income using Micro Data from Austria

1.1 Introduction

The decomposition of distributional measures has a long-standing theoretical as well as empirical background in economic analyses. More or less closely connected approaches for the decomposition of income inequality measures in various forms have been discussed in the literature. Researchers from different backgrounds interested in the set-up of the economy looked at this problem. The distribution of wealth and income separately as well as the change of these indicators over time, has also been studied by many authors. The goal of this paper is to put these strains together and investigate the wealth and income distribution in Austria by decomposing them into various contributing components/factors.

Lambert and Aronson (1993), for example, analyse how overall inequality (of household expenditures in India) can be attributed to different subgroups in the population[1] and how the resulting coefficients can be appropriately interpreted using the geometrical representation of the Lorenz curve. A special emphasis is put on the within group, between group and overlapping components of the Gini coefficient. So, this approach decomposes inequality according to socio-economic characteristics of the population under consideration. Decomposition procedures were also applied by labour economists to approach several key questions in the distribution as well as differences in income.[2] These informative studies concentrate on income distribution, but have little to say about wealth distribution. It is also not possible to analyse questions about wealth with this type of regression-based decomposition procedures using the income generating function since there is no theoretical foundation for a similar regression estimation to this income generating process for wealth.

Decomposing wealth and comparing the results to the ones of income, however, is an interesting investigation. Already the documentation of wealth distribution, its

[1] See also Silber (1989) for a theoretical approach along similar lines.

[2] See e.g. Shorrocks (1982), Juhn, Murphy, and Pierce (1993), Fields (2003), and Yun (2006) for studies using U.S. data; Gunatilaka and Chotikapanich (2009), and Morduch and Sicular (2002) for Sri Lanka and rural China respectively; and Cowell and Fiorio (2009) as well as Frosini (2012) for a theoretical discussion with an international empirical application.

development, and connections to poverty has attracted a huge variety of investigations.[3] For the decomposition exercise conducted here we turn to a line of research that decomposes the Gini Coefficient (henceforth abbreviated by GC), a well known and widely used measure of the distribution, according to various sources. This line of thought has its beginning with the publications of Gastwirth (1970), Gastwirth (1972), and Gastwirth and Glauberman (1976), where the authors looked very closely at the Lorenz-Curve and the calculation of the GC. The ideas were further developed by Dorfmann (1979) and Lerman and Yitzhaki (1985). We take the latter approach and apply it to the distribution of both wealth and income in Austria. The sort of question one tries to answer with this method is, how does a change of the distribution of a certain type of wealth holding or income source affect overall wealth or income distribution, respectively. Another interesting aspect, that is analysed here as well, is how certain ways of holding financial/housing wealth or generating income are correlated with total wealth holdings and income. Furthermore, we can see how important certain investment or income aspects are for total financial/housing wealth and income. And, finally, we will also have a close look at the distribution of the various investment possibilities and income sources. By doing so, we document for the first time in Austria various specific aspects of income and wealth distribution.

The results show that there are huge differences in how financial assets are held. More sophisticated forms of the portfolio choice menu (e.g. bonds or business participations) are largely restricted to certain parts in the the population. Along these lines the changes in different ways financial wealth is held have a significantly different impact on the overall distribution. A similar finding holds for earnings. From several sources of income, that are under consideration, only some forms are received by a large fraction of the sample population (like income from employment), others, however, generate a stream of income for only a small part of the population

[3]Se e.g. Davies, Sandström, Shorrocks, and Wolff (2009), Jenkins and Jäntti (2005), Peichl, Schäfer, and Scheicher (2010), Wolff and Zacharias (2009), and Dell, Piketty, and Saez (2005).

(e.g. income from rent and lease) and hence there are different consequences for the overall income distribution.

The article is organised as follows. The next section lays out the conceptual and methodological background of the wealth/income concept and the decomposition procedure applied. Then the data, that are used in the analysis, are introduced, which is followed by the main part of the article that discusses the results. The final section gives concluding remarks and some ideas for further research.

1.2 Background

1.2.1 Wealth and Income Concepts

The clear definition of wealth and income is itself a complicated endeavour. There are various issues that have to be accounted for.

On the one hand wealth consists of a wide variety possible forms ranging from holdings in the different forms of financial assets (including e.g. money holdings, savings, investments, stocks, and the like); over real assets like housing, cars or other real valuables; to social and human capital. Also income usually comes in several forms like income from (self-)employments, rents, or return on investments (see e.g. Smeeding and Weinberg (2001) who investigate the possibility of a uniform definition of income). The distinction and hence collection of data on the various forms of both income and wealth is a difficult task. This might explain why there are so few data sources available that include both income and wealth at the level of detail that is needed for the decomposition presented in this paper.[4] In Austria, there is so far (at the time of writing in 2011) no source of data available that includes information on both wealth as well as income.[5]

[4] Some notable exceptions include the Survey of Consumer Finance (SCF) (Federal Reserve) for the United States, the Survey of Household Finances (EFF) (Banco de España) for Spain, and the Survey on Household Income and Wealth (SIHW) (Banca D'Italia) for Italy.

[5] This lack of information will hopefully be filled once the data from the Household Finance and Consumption Survey (HFCS) (see e.g. `www.hfcs.at`) become available.

On the other hand the level of aggregation is not clear as income is (at least in the most common forms) received on a personal level and wealth typically is held or rather shared (as is income in a family with a main earner) on a household level. On top of that it is a priori not clear what constitutes a household.[6] Several attempts are in the literature[7] to account for the household structure in order to get some form of equivalence income. The most well know is using the OECD equivalence scale[8] giving the household head a weight of 1, each additional adult member a weight of 0.5 and the weight of 0.3 to each children, and calculate the equivalence income of the household. The equivalence scales, however, also have their drawbacks. Sierminska and Smeeding (2006) give a recent summary on the common forms used in the literature, noting that

> "[p]recisely because the person assessing the wealth distribution has no idea about how the owners plan to utilize the wealth that is being measured, equivalence scales are difficult to determine with any precision." (Sierminska and Smeeding (2006), page 2)

Although equivalence scales are commonly used for the aggregate of the total income at the household level, it is not applied in wealth studies (see e.g. Azpitarte (2008)). Applying no equivalence scale is implicitly assuming perfect returns to scale, which is what we do in this exercise.

In order to keep the analysis as simple, clear, and comparable between wealth and income as possible as well as due to the lack of better data, we concentrate on the decomposition of financial wealth (housing wealth is provided in the appendix) into its various forms of holding and net / disposable income into the different forms it is received (gross income is provided in the appendix) in separate data-sets. Doing so we abstract away from equivalence scales and cannot incorporate (due to

[6] See e.g. Smeeding and Weinberg (2001)

[7] See for example Buhmann, Rainwater, Schmaus, and Smeeding (1988), Ebert and Moyes (2003), or Koulovatianos, Schröder, and Schmidt (2005).

[8] See e.g. `http://www.oecd.org/LongAbstract/0,3425,en_2649_33933_35411112_119669_1_1_1,00.html` [accessed on 12th September 2011].

data constraints) an overall decomposition of total household wealth including and combining information on financial wealth, real assets, and income using one dataset including all necessary information. This analysis is certainly a fruitful endeavour once the data from the HFCS become accessible.

1.2.2 Methodology

One of the most widely used measure of inequality is the GC. In order to tackle the questions posed in the introduction we look closely at this statistic. As is well known the GC is defined as the area between the Lorenz curve and 45-degree line. There are many different ways to calculate this value describing the unevenness of any empirical distribution function. In a recent survey by Xu (2004) they are described in detail.

The most intuitive way to think about it can be written as the following integral of the above described area

$$GC = 1 - 2\int_0^1 L(u)du,$$

where $L(u)$ captures the Lorenz Curve. If we assume that the cumulative distribution function $F(Y)$ of total wealth (or income) Y is piecewise differentiable, zero for all negative values[9] of Y and has mean $\bar{Y}$, the GC can also be written in form of the covariance between Y and its distribution, $F(Y)$, normalized by the arithmetic mean. All the derivations are provided in the appendix (see appendix 1.A).[10] So the alternative calculation can be stated as

$$GC = \frac{2}{\bar{Y}} cov[Y, F(Y)].$$

Now, we can think of total wealth or income consisting of several parts, e.g., as in Lerman and Yitzhaki (1985), total (household) income is the sum of the income for

[9] That means $F(Y) = 0$ for $Y < 0$

[10] References of the original studies are given in the appendix, and can also be found in Xu (2004).

the head of the household, the income of other family member, transfers, etc. or one can also think of the total income consisting of income from wages, from letting a flat/house, from interest and dividends payments and so on. In the following empirical exercise we apply this line of thought also to wealth. To be more precise we think of total financial wealth being held in various forms of assets such as bank accounts, savings accounts, stock shares, bonds, life insurance and so on. Also real wealth can be held in the primary housing unit, secondary housing, building land and so on (see appendix 1.C). So, total wealth or income can be held in or received from different sources such that $Y = \sum_{k=1}^{K} y_k$ where k denotes the source of wealth and income, respectively, and K is the total number of components / sources. Noting the additive separability of the co-variance,[11] one gets

$$GC = \frac{2}{\bar{Y}} \sum_{k=1}^{K} cov[y_k, F(Y)].$$

In appendix 1.A the derivations of the decomposition method are laid out, so that it can be written as

$$GC = \sum_{k=1}^{K} R_k G_k S_k \tag{1.2.1}$$

where R_k is the relative Gini correlation between wealth (income) source k and total wealth (income), G_k is the relative Gini component of contributing factor k, and S_k is $k's$ share of the total wealth (income). Thus we can see how the source of wealth (income) is related to the total in R_k; i.e. it explains the correlation between the component k and the overall cumulative distribution. The estimate on S_k tells us how important the specific source is for total wealth (income); i.e. it is the relative share of component k in the total. And G_k gives us the GC of the specific source. The most interesting feature of this decomposition when it comes to an empirical analysis is that one can calculate the impact, i.e. the percentage change of total income distribution given a change in the source k. In a sense one can think of

[11] I.e. for any two random variables $X = \sum_{i=1}^{N} x_i$ and Z, one can put the summation out of the calculation of the co-variance to get $cov[X, Z] = cov[\sum_{i=1}^{N} x_i, Z] = \sum_{i=1}^{N} cov[x_i, Z]$.

the elasticity of the overall distribution of Y with respect to a change in one of its components y_k. For an infinitesimal change in the wealth (income) source k such that it is $\varepsilon_k y_k$ for a small ε_k after the change. This is done by taking the partial derivative (applying the chain rule) and dividing the result by the overall Gini coefficient, which yields

$$\frac{\frac{\partial GC}{\partial \varepsilon_k}}{GC} = \frac{S_k R_k G_k}{GC} - S_k. \tag{1.2.2}$$

Using this formula gives us an estimate of how the overall GC changes if a specific way of holding an asset or certain sources of income change. It depends, resulting from the application of the decomposition equation (equation 1.2.1), on the GC of source k, the correlation between the component and the overall distribution, and how important factor k is relative to the other parts. Intuitively all three aspects seem clear; e.g. the effect of component k on the overall inequality will depend on the fraction of households having this item, the distribution of it and the correlation between this factor and the overall distribution. We will find that both a change of the wealth and the income distribution is correlated differently to some investments opportunities and sources of income then to others (i.e. the elasticity of the GC of financial wealth with respect to current and savings account as well as life insurance in Austria is negative whilst other ways of holding wealth like shares, bonds and NPTBS have the opposite sign). In what follows, we apply this way of decomposing to Austrian data on financial/housing wealth and income. A similar approach was for example also taken by Azpitarte (2008). We will see that it provides interesting insights.[12] For example, we will also see that income inequality is mostly explained by the distribution of income from work whilst social benefits in Austria have the desired reducing effect on inequality.

[12]López-Feldmann used this decomposition procedure in several studies, see, for example, López-Feldman (2006) and Taylor, Mora, Adams, and López-Feldman (2005), and granted access to the .ado-program in STATA, which is used later in some of the estimations. For the documentation see López-Feldman, Mora, and Taylor (2006). Their contribution is gratefully acknowledged.

1.3 Data Description

The focus in this empirical exercise lies in the structural features of the wealth and income distribution in Austria. We use two separate sources of data, the Survey on Financial Wealth conducted by the Oesterreichische Nationalbank (OeNB) and the EU-SILC survey administered by the National Statistical Offices, for the Austrian data by Statistics Austria, in coordination with EUROSTAT.[13] In the calculation of the descriptive statistics and the measures of inequality appropriate weights are used to account for the survey structure.

1.3.1 OeNB: Survey of Financial Wealth in Austria

The survey on financial wealth[14] from the OeNB (Household Survey on Financial Wealth, abbreviated HSFW) was carried out in summer and autumn 2004.

In total 2,556 households answered in either a personal or written interview extensive questions on the households balance sheet, detailing financial wealth holdings and some forms of credit; socio-economic characteristics of the household, including size, location, education, age, etc.; and working status. The questions were answered by a household member that was considered being knowledgeable about households' financial matters. The sample is based on multi-stage random sampling procedure in order to reach a representative sample of the whole population in Austria. However, there was no special oversampling of certain parts of the population (e.g. the wealthy), which would be desirable to have in order to draw clear results about wealth for the part of the population that plays a distinctive role in aggregate developments of financial wealth issues. The data are weighted taking into account geographic as well as socio-economic indicators.[15] Also this survey unfortunately

[13] For the results provided in the appendix survey data on households' housing wealth from OeNB collected in 2008 and again the EU-SILC data are used.

[14] A good overview on the data that were collected can also be found in Beer, Mooslechner, Schürz, and Wagner (2006).

[15] Since there are no missing values in the data on financial assets, due to the way of the data collection process, where households were required to provide all necessary information for a suc-

Table 1.1: Definition of the HSFW Variables

Variable	**Accronym**	**Definition**
Total Financial Wealth	TFW	Sum off all possible forms in which financial wealth is held. It is gross financial wealth in the sense that no liabilities are subtracted.
Current Account	CA	Value in the account that is used for transaction purposes. The questionnaire asked for ranges and the data were imputed with the mean of these ranges.
Life Insurance	LI	Sum of premium already invested in life insurances.
Savings Account	SA	Sum of wealth in various forms of savings accounts, building savings, and bonus scheme savings.
Bonds	BO	Sum of wealth invested in various forms of bonds.
Shares	SH	Sum of wealth invested in publicly traded shares.
Investment Certificates	IN	Sum of wealth invested in certificates and funds.
Business Shares	NPTBS	Sum of wealth held in **not** publicly traded shares of businesses.

Notes:
(i) *Source:* OeNB HSFW (2004).

does not cover real assets (apart from not publicly traded business participations which are often defined as real assets) and hence, although being fully aware of the importance of real assets in households total wealth, we have to restrict the analysis due to data constraints to the financial part of households wealth.[16] Furthermore, lacking real assets, especially housing wealth, it is not possible to calculate net worth, defined as total gross assets minus total liabilities. Due to the fact that leaving out real assets and still subtracting liabilities would distort the estimates because a big proportion of debt is for the financing of housing, we restrict the empirical exercise

cessful participation in the survey, no special imputation methods are used to account for item non response.

[16]The appendix (see appendix 1.C) gives a comparable exercise using data on real assets. In this survey, however, there is no information on financial assets. A complete picture can only be derived once the data from the Household Finance and Consumption Survey (HFCS) will become available where all the necessary information is included.

to the decomposition of gross financial wealth.

For the analysis conducted here we are using the variables on a common set of possible forms in which financial wealth is held at the level of households. For the definition see table 1.1. We take a close look at total financial wealth (TFW) consisting of money held in bank (current) accounts, in savings accounts (including building savings), in bonds, in shares, investments certificates (including various types of funds), not publicly traded business shares (NPTBS), and life insurances.

1.3.2 Eurostat: EU-SILC

The second source of data is the European Statistics on Income and Living Conditions[17] (EU-SILC), which is a household survey for comparative statistics on income and social exclusion in the European Union.

It has both a cross sectional and longitudinal structure and includes in depth variables on income at the individual and the household level. In this analysis we consider the data for Austria surveyed in 2005,[18] which is close to the time of, and hence comparable to, the above introduced survey on financial wealth from the OeNB (and the survey on housing wealth from 2008 considered in appendix 1.C). The survey is conducted using the computer assisted personal interviewing technique (CAPI) and yields data for a representative sample of households in Austria including more than 5,000 households. The imputations provided in the data set, where Statistics Austria used hot deck procedures to account for item non response, are used in the analysis, so that there are no remaining missing observations. The assumption of correct hot deck imputations seems not too strong for most variables since the number of imputed items is (far) below 1%. However, total income and income from capital investment show a high rate of imputed items (about a third of the obser-

[17] See `http://epp.eurostat.ec.europa.eu/portal/page/portal/living_conditions_and_social_protection/introduction/income_social_inclusion_living_conditions` for further informations [accessed on 7th April 2010].

[18] An analysis with the data from 2004 reproduced very similar results.

Table 1.2: Definition of the EU-SILC Variables

Variable	**Accronym**	**Definition**
Total Disposable Income	TDI	Sum of all personal gross income of all household members **minus** income tax, social insurance contributions, interest payments for mortgage, regular taxes for wealth, and regular inter-household cash transfer.
DI excluding Social Benefits	IEXSB	TDI **minus** unemployment benefits, sickness benefits, disability benefits, education related allowances, family/children related allowances, housing allowances, and social exclusion related but not elsewhere classified.
NI form Rent and Lease	NIRL	Income received from renting a property minus its cost (mortgage interest payments, repairs, maintenance, insurance, etc.), the tax and social insurance contribution.
Accommodation Allowance	AW	Means tested transfers from the public authority to tenants and owners to meet housing expenses.
NI from Private Transfers	NIPT	Regular cash received from inter-household transfers, such as alimony or monetary help for students minus tax and social insurance contribution.
NI from Capital Investment	NICI	Interest, dividends, and other profits from capital investment minus tax and social insurance contribution.
NI from Family/Children related	NIFC	Sum of net family and children related income (HY050n), and other net income against social exclusion (HY060n).
NI from Pension	NIP	Proxy estimated from the household income including pensions but excluding social benefits (HY022) minus household income excluding both, pensions and social benefits (HY023).
NI from Work	NIW	Total net income from employment and self-employment summed over all household members(py010n and py050n).
Residual	RES	Total disposable household income minus the sum of all specific forms of income mentioned above.

Notes:
(i) *Source:* EU-SILC 065/04: Description of target variables; EUROSTAT.

vations are imputed). We use, again in accordance to the OeNB survey, household level data on income. Total disposable income (DI) (variable code HY020) and total gross income (GI) (variable code HY010) consist of various types of individual and households level income sources. From these a subset is taken and the following further variables generated in order to conduct the empirical analysis. They are an aggregate for social benefits, denoted as family/children related income, which is the sum of both family and children related as well as ohter income against social exclusion; a proxy for income from pension calculated through total household income excluding social benefits minus total household income excluding benefits and pensions; the income generated through employment; and the residual, i.e. total household income minus the sum of all specific forms. Net income is abbreviated by NI. We include the variables defined and described in table 1.2 in the analysis.

In the appendix in table 1.8 we report the socio-economic structure of the households that participated in the surveys and that reflect the structure of the population in Austria at the time. Both the absolute number and the appropriately weighted percentage share of the survey sample are reported. Essentially the overall structure of households is very similar in all data sources. It seems worth noting, however, that the households in the wealth surveys are more often employed, on average somewhat more educated and live more often in households with more household members. These facts at least tends to the direction that these surveys covered a comparable sample and if anything the wealth surveys included more households with a higher financial capability (due to e.g. higher education and employment rate).

1.4 Results

1.4.1 Descriptive Statistics

First, we look at some general overview of the data that are used in the decomposition. Table 1.3 shows that, given the 2,556 observations on each of the indicators, how much financial wealth is held by a household in each form of asset.

We can see that the mean is always higher than the median, indicating a distribu-

Table 1.3: Descriptive Statistics: Financial Wealth

Variable	**Mean**	**Median**	**Variance**	**Min**	**Max**	**% Pop**
Financial Wealth						
Total	54,666	23,575	152,507	0	4,487,290	99.5%
Current Account	1,589	1,275	1,277	0	4,200	91.3%
Saving Account	28,226	13,200	50,682	0	1,000,000	96.8%
Bonds	3,132	0	17,908	0	590,000	10.6%
Life Insurance	9,013	1,200	17,938	0	312,000	53.1%
Investment Certificates	2,900	0	17,460	0	500,000	11.4%
Shares	4,069	0	25,039	0	1,000,000	15.7%
Not publicly traded business shares	5,737	0	113,488	0	4,300,082	2.6%

Notes:
(i) *Number of Observations:* 2,556.
(ii) *Source:* OeNB HSFW (2004).

tion that is skewed to the right of the wealth spectrum. Another interesting point is that the various forms of financial investment (stocks, bonds, investment certificates, NPTBS) are only used by more affluent households, so that no member of the population with the poorest 50% (of each asset component) is participating in this market. Thus the median is given with zero. One can also see that the variation, measured by the variance, in the variables is relatively high. As an example consider total wealth, where the variance is almost three times the mean. The minimum value of each variable is zero. Although it is worth noting again that one would rather look at total net wealth, including real assets and the liability side of households, the

data do not include this information on real wealth, so that we concentrate here on gross values. This still implies, however, that there are households in Austria that do not have any wealth in each of the described asset categories. These households, however, do not have to be the same households, meaning that, for example, there are some that do not hold shares and others that have no life insurance, but may own shares. The participation rate of households for Current and Savings Account are close to 100% and hence there are some households which have an account but no wealth in it.

The last column presents the weighted share of the population that owns a certain type of asset. Although almost all people have some form of financial wealth (99.5%),[19] only some possible ways to invest financial capital are used by the majority of the sample population. These are current and savings accounts and also life insurances are quite common, which are owned by about 53% of the population. All the other forms of investments are relatively concentrated in a small group. The most extreme case are business shares that are not publicly traded, where only 2.6% of the population own this portfolio item. We have to keep this fact in mind for the interpretation of the decomposition.

Turning now to the data in income, we can identify a similar pattern, reported in table 1.4, in the EU-SILC data as well. Using the data for 5,148 households in Austria, the mean of the various income variables is higher than the median and the variation (measured by the variance) is rather high; i.e. about twice the mean in the indicators for income from pensions and social benefits (Family related), and even relatively higher for the others except total income and income from employment which are less volatile. Note, that in the EU-SILC data it is possible to have a negative disposable income value since there are several expenses subtracted from

[19] Note that this participation rate is calculated on having a positive amount in the specific asset. This fact also explains the participation rate for the current account, which is expected to be closer to full coverage.

Table 1.4: Descriptive Statistics: Income Sources

Variable	**Mean**	**Median**	**Variance**	**Min**	**Max**	**% Pop**
Disposable / Net income						
Total	32,313	27,915	22,120	-103,012	330,345	100%
Excl. Social Benefits	29,250	24,732	21,983	-103,012	323,490	99%
Rent and Lease	351	0	3,461	-12,000	121,600	4%
Accom. Allowance	57	0	334	0	3,600	4%
Private Transfer	274	0	1,470	0	28,800	6%
Capital Income	294	66	1,721	0	58,650	78%
Family related	1,571	0	2,933	0	28,771	33%
Pension	7,699	0	13,211	0	180,805	39%
Employment	20,846	18,000	22,072	-103,142	270,690	70%
Residual	1,221	0	4,706	-28,800	75,240	

Notes:
(i) *Number of Observations:* 5,148.
(ii) *Source:* EU-SILC (2005).

income.[20] Also note that, for the same reason, the minimum value of income from employment and rental of property and land may be negative.[21] There are exactly two negative observation regarding total disposable income. Both of them had a substantial negative income from self-employment which were solely due to extraordinary high tax (and social insurance contribution) payments. Although these observations with high negative income are legally possible, they are considered to be unreliable outliers and are not used in the calculations of several inequality measures reported further down.[22]

Also for income we see in the last column in table 1.4 that only a minor fraction of the population receives income from rental of land and property as well as from private transfers, and accommodation allowance (about 4%). The very large fraction (almost 80%) who gets some income from capital investment is due to the fact that every person as soon as it has positive bank deposit receives some interest payment.

[20] These are interest paid on mortgage; regular taxes on wealth; regular intra-household cash transfers.

[21] For the latter income source, there are several income reducing factors subtracted such as mortgages interest repayments, maintenance costs minor repairs, and other charges.

[22] In the literature there is also no common approach on how to adjust some of the used inequality measures for negative income. See e.g. Stich (1996) for one possible way.

Table 1.4 shows, however, that the variation in this income variable is very high, i.e about six times its mean. Note also that the fraction of population that benefits from social plans, denoted in the table as family related, is relatively high (about a third of the population). This reflects the generally generous social system in Austria, as, for example, children support ('Familienbeihilfe' in German) which every family with a child gets independent of other factors. Further note that the proportion of households receiving income from employment is relatively high, reflecting the fact that at least one household member generate income from employment.
We can see how each of these wealth and income variables is distributed in the society by looking at the Lorenz-curve. This is done in the next subsection.

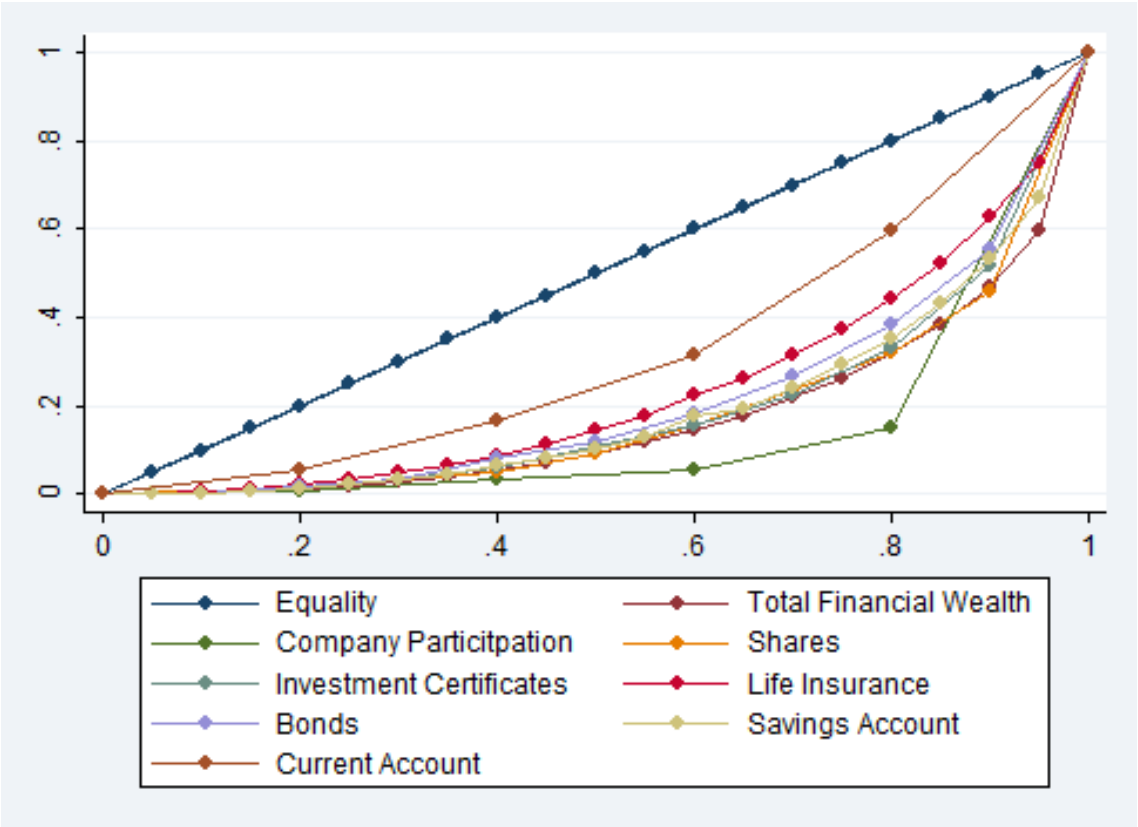
1.4.2 Inequality Measures

Now, we take a look at how the distribution of wealth holdings of the Austrian population shows up in our data.

The Lorenz-curves of the various possible ways to hold financial assets are given in figure 1.1. It can be seen that, as expected, the two possibilities to invest money in a current or savings account are most equally distributed. This is due to the fact the most households have a bank account and on average higher wealth levels are more often invested in other forms of investments. It can also be seen from the data, see e.g. table 1.3, that the participation rate for several portfolio choices is relatively small. Only households with considerable wealth tend to invest in a broad range of investment opportunities including shares, bonds and other more complex financial instruments. Thus the more wealth in total is invested the more complex and diversified portfolios get.[23] Notice that this representation takes only positive values into account, i.e. it is the within group distribution of the part of the population that holds any of the according financial assets. These curves would look even (much) more extreme, if we had not excluded the part of the population, who does not hold

[23] This is in line with various findings in the literature; e.g. Andreasch, Fessler, and Schürz (2009) or Fessler and Schürz (2008).

Figure 1.1: Lorenz-Curves for the Financial Wealth Holdings



Notes:
(i) *Source:* OeNB HSFW (2004).

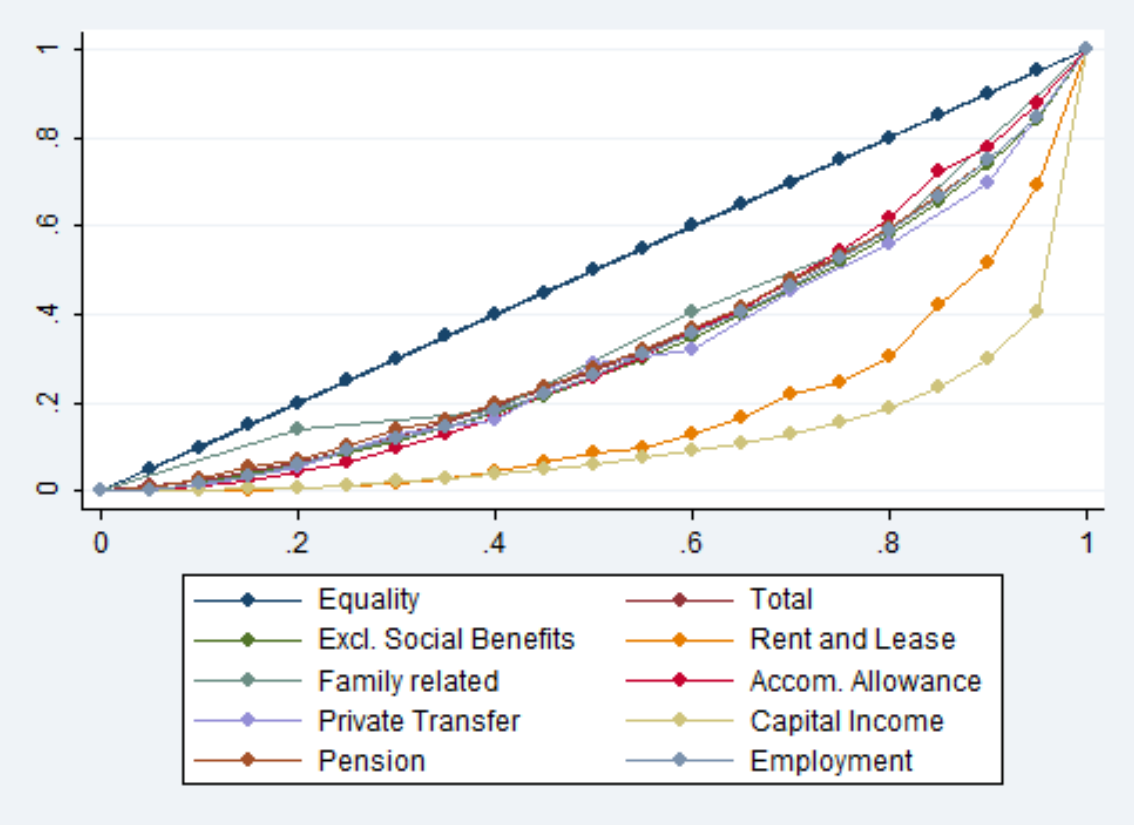
any of the more complex assets (stocks, bonds, investment certificates, shares, and life insurance), in the calculation of the respective Lorenz curve. We will come back to this issue later comparing measures of inequality.

For the EU-SILC data the Lorenz curves are as one would expect. The distribution is fairly similar across income sources with the exception of income from rental of property and land and from interest and dividends. This is not so surprising since these assets are relatively unequally distributed and thus only a small portion of the population gets any income from these sources.[24] Note that the two groups receiving the two different income sources might be quite distinct from each other.

[24]For further information on the distribution of holdings of property and land in Austria see e.g. Fessler, Mooslechner, Schürz, and Wagner (2009).

This result is robust to whether net or gross income measures are used, as can be

Figure 1.2: Lorenz-Curves of Disposable / Net Income Sources



Notes:
(i) *Source:* EU-SILC (2005)).

seen in in the appendix in figure 1.4. Once again, we did not use observations with the value of zero in the calculation of the according shares to get the Lorenz curves for the following income sources: total disposable income with and without social benefits; net income from rental of land and property; income from social benefits (family/children related allowances and other benefits against social exclusion); housing allowances; net income from private transfers; net income from pensions; net income from employment; net income from interest and dividend payments.[25] An even deeper understanding of the underlying distributional characteristics can be found from various measures of inequality. In order to understand the exact fea-

[25] In appendix 1.C the same holds for gross income from rental of land and property, gross income from social benefits, housing allowances, gross income from private transfers, gross income form employment, and gross income from interest and dividend payments.

tures, one has to recall the fraction of the population who holds or receives a certain form of asset or income respectively. It can be seen from tables 1.3 and 1.4.
The results are given for both estimations, first, called within inequality, the one excluding the households that do not hold or receive the asset or source of income, respectively, and, second, called overall inequality, the one including these households.

Table 1.5: Inequality Measures

Variable	Overall (including observations with zero)				Within (excluding non-positive observations)			
	GC	GE(2)	AkI	$\frac{P90}{P50}$	GC	GE(2)	AkI	$\frac{P90}{P50}$
Financial Wealth								
Total	0.66	3.89	.[i]	4.5	0.66	3.87	0.37	4.5
Current Account	0.44	0.32	.	3.3	0.39	0.25	0.13	2.9
Saving Account	0.63	1.61	.	4.9	0.61	1.54	0.33	4.8
Bonds	0.96	16.34	.	.	0.62	1.28	0.32	4.5
Life Insurance	0.75	1.97	.	22	0.53	0.82	0.24	3.4
Investment Certificates	0.96	18.12	.	.	0.63	1.63	0.34	5.0
Shares	0.95	18.92	.	.	0.67	2.54	0.37	5.0
Not publicly traded business shares	0.995	195.58	.	.	0.82	4.58	0.59	22.2
Disposable / Net Income								
Total	0.33	0.23	.	2.0	0.33	0.23	0.09	2.0
Excl. Social Benefits	0.37	0.28	.	2.2	0.36	0.27	0.11	2.2
Rent and Lease	0.99	48.7	.	.	0.66	1.41	0.38	6.3
Accom. Allowance	0.97	17.0	.	.	0.35	0.20	0.11	2.1
Private Transfer	0.96	14.4	.	.	0.41	0.43	0.14	2.2
Capital Income	0.83	17.2	.	7.1	0.78	13.2	0.53	6.0
Family related	0.78	1.74	.	.	0.35	0.24	0.10	2.3
Pension	0.74	1.47	.	.	0.33	0.26	0.09	2.0
Employment	0.54	0.56	.	2.6	0.35	0.24	0.11	2.1

Notes:
(i) All the values denoted with '.' cannot be calculated (e.g. due to the fact, that the fiftieth percentile still holds nothing and thus one would have to divide by zero).
(ii) *Statistics reported:* Gini coefficient (GC), measure of the General Entropy Class with $\varepsilon = 2$ (GE(2)), Atkinson Index with $\alpha = 0.5$ (AkI), and the ratio of the 90^{th} to the 50^{th} percentile.
(iii) *Source:* EU-SILC (2005).

In table 1.5 one can inspect various measures of inequality for all the variables under consideration. We present the well known Gini coefficient (GC), ranging from 0 to 1 where 0 implies complete equality and inequality is increasing with the value of the GC; a member of the General Entropy Class (GE(2)) where a higher value implies a higher inequality, which is more sensitive to the top and usually defined as $GE(\varepsilon = 2) = \frac{1}{N\varepsilon(1-\varepsilon)} \sum_{i=1}^{N} \left[\left(\frac{w_i}{\bar{w}} \right)^{\varepsilon} - 1 \right]$; the Atkinson Index[26] (with $\alpha = 0.5$ and hence equal weight on both the upper and the lower tail of the distribution) which is defined as $AkI(\alpha = 0.5) = 1 - \frac{1}{\bar{w}} \left(\frac{1}{N} \sum_{i=1}^{N} w_i^{1-\alpha} \right)^{\frac{1}{1-\alpha}}$; and the 90th to 50th percentile ratio, which simply gives the factor by which wealth and income is higher for the 90^{th} percentile in terms of the median wealth and income, respectively. First, note that it is not always possible to calculate the value, since for the calculation of the ratio the median has to be non zero otherwise it is not defined, and the Atkinson index is only defined for non zero values as well. Second, one can see that in general all the measures point in the same direction.

If we compare the left (using the whole sample of the population, i.e. also people that have income or wealth of zero, respectively) and the right (including only strictly positive values) half of the table, it is clear that inequality measured in all various ways decreases. Given the description above, this finding is not further surprising, but, concentrating on the within group inequality now, even the distribution within the group of holders is remarkably high. In line with other investigations[27] we see, for example, that wealth held in NPTBS is overall highly unequal distributed with a Gini coefficient of above 0.8 and the 90^{th} percentile owns more than 22 times of the wealth held by the 50^{th} percentile. This is obviously the most extreme example in the table, but also all other ways to invest financial wealth apart from the current account and life insurances have a GC of above 0.6 and the 90^{th} percentile possesses about five times the median amount. The estimates are broadly comparable to other

[26] The Atkinson Index ranges from 0 to 1, giving total equality in the sense of every household in the society holding the same amount of wealth or income an index value of 0 whereas total inequality takes the value of unity.

[27] See e.g. Andreasch, Fessler, and Schürz (2009) and Fessler and Schürz (2008)

findings in the literature. Sierminska, Brandolini, and Smeeding (2006) for example calculate the GC of household net worth for Canada, Finland, Italy, Sweden and the United States raging from 0.6 to 0.89. If one takes into account the distribution of housing wealth, which is included in the definition of net worth, our estimate of 0.66 for total financial wealth seems to lie in this range.

It can also be seen from the table that income in general is more equally distributed than wealth. With two slight exceptions in income from rental of land and property, and income from capital investment the GC is relatively low at around 0.35. There is no apparent difference between the gross and net income variables.[28] Indicating that the tax burden does not severely change the distribution of income among households in Austria.

Close inspection of the table also yields another interesting finding. As expected, however in public debates sometimes doubted, we find that in the EU-SILC data social benefits in Austria reduce inequality. This can be seen from the two variables on total disposable income including (TDI) and excluding social benefits (IEXSB). Let us turn now in the following section to the main part of this empirical exercise, that is the decomposition of total wealth and income, respectively, according to its sources using the above introduced method.

1.4.3 Decomposition of Wealth by Ways of Holdings

In this section of the article we look at the decomposition of wealth based on survey data that were gathered by the OeNB in 2004. Essentially, the distribution of total financial wealth can be decomposed in its various possible types of investment laid out in the data section. This means that that GC reported in the left column in table 1.5 for total financial wealth (and total income in the next section) are decomposed with regard to their respective contribution factors. We use a bootstrapping procedure with 500 replications to get standard errors and thus confidence inter-

[28] See appendix 1.C for the estimates of the gross income variables.

vals for the most interesting estimate; i.e. what a change of the distribution in the particular variable means for the distribution of wealth or income, respectively (last column in the following tables).

Recall from section 1.2.2 above the methodological definitions / meaning of these

Table 1.6: Decomposition of the Financial Wealth Distribution

Variable	S_k	R_k	Share	% Change (Std. Err.)
Current Account	0.027	0.421	0.008	-0.019** (0.001)
Saving Account	0.529	0.938	0.484	-0.045** (0.009)
Bonds	0.066	0.890	0.086	0.020** (0.003)
Life Insurance	0.165	0.703	0.135	-0.030** (0.003)
Invest Certificates	0.060	0.882	0.077	0.018** (0.003)
Shares	0.086	0.896	0.112	0.026** (0.004)
NPTBS	0.067	0.956	0.099	0.031* (0.013)

Notes:
(i) *Significance levels (bootstrapped standard errors):* †: 10%, ∗: 5%, ∗∗: 1%.
(ii) *Reported results:* Share of wealth in source k (S_k); correlation of source k–inequality and cumulative inequality (R_k); share of total inequality that is explained by source k (Share); and how a change in source $k's$ wealth component affects total inequality (% Change).
(iii) *Source:* OeNB HSFW (2004).

empirical results. S_k is k's share of the total wealth, and hence it gives us an indicator for the importance of the particular way to invest financial wealth. R_k is the relative Gini correlation between wealth source k and cumulative distribution function of total financial wealth. That means, it tells us what is the correlation between a chosen investment k and the overall GC of financial wealth. In the penultimate column, denoted by 'Share',[29] we see the proportion of total wealth inequality that

[29]This measure is exactly the same as for example in López-Feldman (2006) or Taylor, Mora, Adams, and López-Feldman (2005) called "share in total income inequality" see e.g. table 3 on

is explained by component k. This implies that the values in column 'Share' sum up to 100% (apart from rounding). It is distinguished from S_k in so far that the former gives us how important a particular way of investment is in terms of the amount of money relative to total wealth and the latter (i.e. 'Share') how much the money invested in a certain way contributes to total wealth inequality.

In table 1.6 we see the decomposition of total financial wealth in Austria according to several ways of capital investment. In the second column it can be seen that the biggest part of wealth is held in savings accounts (about 53% of total wealth), followed by life insurances (about 17%) and than a group consisting of stocks, shares, investment certificates, and NPTBS (each around 7%). In the current account only a small fraction of total wealth (about 3%) is held. This seems to reflect the preference for security, since more secure investments are held more extensively, and a reasonable choice because the savings account and life insurances give better returns than the current account. Apart from the wealth held in the current account the correlation of the distribution of the various ways to invest and total distributional pattern is very high. This comes from the fact that the GC of the current account holdings is rather low compared to all the other investment opportunities and total distribution.

Comparing the second and the fourth column it can be seen that the contribution of more sophisticated ways of investment (stocks, shares, investment certificates and NPTBS) is relatively higher than is their respective share of wealth. This result is not very surprising since these types of assets are mainly in the hands of households that invest a relatively high level of wealth and hence are the households that drive total distributional issues.

Finally, in the last column we see how a change of the holdings of one way of investment will affect the distribution of total wealth holding everything else being held constant (i.e. ceteris paribus).[30] First, note that all the estimates are signifi-

page 29 in López-Feldman (2006).

[30]For the calculation see equation (1.2.2).

cantly different from zero at the 1%-level using the bootstrapping method explained above, except the estimate on the NPTBS which is statistically significant only at the 5%-level. We also see that an increase in the amount of money held in certain investment opportunities tends to increase inequality whilst for others it tends to decrease it. This is very much along the lines of how many households own a particular type of investment. So would an increase in the money held in the current account, the savings account,[31] and in life insurance to be correlated with a decreased inequality, and on the other hand the elasticity with respect to an increase in the remaining types (stocks, shares, NPTBS and certificates) is positive. The size is relatively small, i.e. a 10% increase in the wealth in NPTBS, holding everything else constant, would lead to an increase of overall GC of 0.31%, for example.

Azpitarte (2008) reports in his study about wealth distribution for Spain a similar contribution to wealth inequality of the various sources in which financial wealth is held. Although the Azpitarte's separation in the specific factors is slightly different then what is analysed here, we see that also in Spain wealth held in the more sophisticated financial instruments tend to increase overall inequality whilst on the other hand wealth held in bank account (current and saving) has a reverse effect. Furthermore, to get a more complete picture of households total wealth one can look at real property (housing, collectables, cars, etc.) a household owns. For Austria there is so far only one micro-data source on housing wealth, i.e. the OeNB Survey on Housing Wealth (HSHW 2008). For the results of the decomposition exercise the reader is referred to appendix 1.A. Again, comparable to the findings for Spain, one can see in table 1.12 that the result is driven by the primary housing property which has an equalizing effect. All the other ways seem to have a minor, often statistically insignificant, effect. An overall decomposition with all the various forms of wealth be taken into account can only be done once data from the HFCS become available.

[31] The estimation was also run for current and savings account combined to fit the aggregate data from ESA95 (definitions AF.22 and AF.29). The qualitative result does not change and the point estimate increases for this combined factor to -0.064.

1.4.4 Decomposition of Income according to Sources of Income

Having analysed the decomposition of wealth, let us now turn to a similar exercise for income. Here we decompose, firstly, disposable income, and later in Appendix B, gross income according to sources of it. According to table 1.7, one can see that the incomes from employment and pensions are considerable shares of total household income. They amount to about 64% and 23% of total income, respectively. Each other sources of income constitute a significantly lower share to total income.

One can also see that an increase in private transfers, social benefits, and pensions is correlated with a decrease in income inequality, whilst an increase in income from rent and lease of property and land, as well as capital investment and income from work displays a tendency to increase inequality. These results are quite intuitive, since private transfers are often made to help poorer members of the family or poorer friends, and hence this money has an equalizing effect as well as social benefits and housing allowance. Since property and financial capital is fairly unequally distributed the income from these assets tend to increase overall income inequality as well. For pensions the case is also clear, as it is the case in Austria that (public) pensions are reducing income inequality to a significant, in economic as well as statistical term, effect. This is due to the fact that there is a lower (minimum pension) and an upper limit to public pension schemes and thus income receiver in this group do neither fall into the extremely low nor the extremely high earner part of the population. As expected as well, the main driving force of income inequality comes from the part that is explained through income from work. What seems to be more surprising is that the effect of an increase in income from work changes overall income inequality by approximately the same magnitude as a change in income from pension but in the reverse order; i.e. a 10% increase received from work increases

Table 1.7: Decomposition of Net / Disposable Income Distribution

Variable	S_k	R_k	Share	% Change (Std. Err.)
Rent and Lease	0.012	0.686	0.024	0.013** (0.003)
Accom. Allowance	0.002	-0.402	-0.002	-0.004** (0.000)
Private Transfer	0.009	0.021	0.001	-0.008** (0.001)
Capital Income	0.009	0.595	0.014	0.005** (0.001)
Family related	0.056	0.338	0.043	-0.012** (0.002)
Pension	0.229	0.129	0.068	-0.161** (0.009)
Work	0.644	0.800	0.843	0.199** (0.009)
Residual	0.041	0.057	0.009	-0.031** (0.004)

Notes:
(i) *Significance levels (bootstrapped standard errors):* †: 10%, ∗: 5%, ∗∗: 1%.
(ii) *Reported results:* Share of wealth in source k (S_k); correlation of source k–inequality and cumulative inequality (R_k); share of total inequality that is explained by source k (Share); and how a change in source $k's$ wealth component affects total inequality (% Change).
(iii) *Source:* EU-SILC (2005).

inequality by roughly 1.99% whereas the same change in pensions decreases inequality by about 1.61%.

As a robustness check the exercise is repeated using data on gross income from EU-SILC. All the estimates for gross income are comparable to the ones for net/disposable income and thus the result seem to be fairly robust.[32] We want to conclude this section with two final remarks. First, as far as there are comparable results in the literature, our findings are similar to these. Lerman and Yitzhaki (1985) find for the U.S. that private transfers tend to decrease inequality, though their estimated effect is larger, and income from rent and lease of land and property has the tendency to

[32]See Appendix 1.C table 1.16.

increase overall inequality. Also Taylor, Mora, Adams, and López-Feldman (2005) and López-Feldman, Mora, and Taylor (2006) report an inequality decreasing effect of remittances, a form of private transfers, in Mexico. Second, in the light of the analysis we have to bear in mind that there is a huge difference[33] in the distribution of wealth and income. If one wants to discuss issues of equality of a society, certainly both have to be taken into account.

1.5 Conclusions

In this paper we analysed wealth holdings and income of the households in Austrian. Based on the surveys of financial wealth conducted by the OeNB and on living standards (EU-SILC) from Statistics Austria and Eurostat, we analysed the distribution of financial wealth and income.

Starting from the methodological foundation of the decomposition of income first proposed by Lerman and Yitzhaki (1985) we provide the possibility to apply it to data on financial wealth as well as on income in Austria. We find, in line with the existing literature,[34] that certain types of investment possibilities for financial wealth are only held by a very small fraction of the population. The contribution of these types to the total distribution is seen to be significant. A ceteris paribus analysis gives the result that an increase in these types of investment typically tends to increase overall inequality keeping everything else constant. It turns out that the more conservative modes, in which financial wealth is held by almost all members of the population, are in this analysis inequality decreasing.

The article also provides an according analysis for the income pattern of households in Austria. We find, both for disposable and gross income, that income from rent and lease of land and property as well as income from capital investment, and income from employment increases inequality whilst income from social benefits,

[33] Compare the upper and the lower half of table 1.5.

[34] See, for example, in the case of shares Fessler and Schürz (2008), and in the case of not publicly traded business shares Andreasch, Fessler, and Schürz (2009).

private transfers, and pensions tend to decrease it. In terms of policy relevance, one can get a clear indication on how certain policy decisions would affect overall financial wealth and income distribution.

There still remain several issues to be resolved. For example, one could have a closer look at total net worth including both financial, housing wealth together and possibly add the debt burden of households to get the complete picture. Here it would also be interesting to see which aggregate component exerts what kind of effect on total net worth of a household. Other possible way towards greater understanding would be to track the compositional pattern over time or the decomposition of income (possibly with panel data) according to the income generating function.

1.A Technical Derivations of the Applied Decomposition Method

Different Methods to calculate the Gini Coefficient

As is widely known, the GC is defined as the area between the Lorenz curve and 45-degree line, it is normalized between zero and one, and lower values indicating lower inequality. The Lorenz curve, $L(u)$, is defined on the interval between zero and one, and $u \in [0,1]$ represents the cumulative distribution of an empirical probability measure (proportion of households). In its graphical representation with the proportion of households on the x-axis and the proportion of total wealth or income on the y-axis, it shows us how much of the variable under consideration is owned by what percentage of the population. Assuming piecewise differentiability of the cumulative distribution function $F(Y)$ of an indicator Y (usually income or wealth), that is zero for all negative values has mean $\bar{Y}$, the GC is given by.[35]

$$GC = 1 - 2\int_0^1 L(u)du$$

Dorfmann (1979) is the first who derives this formulation formally and shows that it can be rewritten to

$$GC = 1 - \frac{1}{\bar{Y}}\int_0^{Y^*}(1 - F(Y))^2 dY,$$

where Y^* is the upper bound of the wealth or income and $F(Y)$ is the cumulative distribution function.

In a slightly different way, we can also derive the GC in the form of the absolute mean difference, Δ, of the variable under consideration. Gastwirth (1972)[36] shows that the GC is the ratio of this measure of dispersion, Δ, and twice the mean of the

[35] See for example Gastwirth (1972).

[36] The author in turn relates to the classic work of Kendall and Stuart (1963).

variable, such that $GC = \frac{\Delta}{2\bar{Y}}$ where the absolute mean difference is given by

$$\Delta = 2\int_0^{Y^*} F(Y)(1 - F(Y))dY.$$

Following the outline from Lerman and Yitzhaki (1985), we can use this formula to decompose the GC according to its sources. First, one uses integration by parts with $u = F(Y)(1 - F(Y))$ and $v = Y$ to get

$$\Delta = 2\int_0^{Y^*} Y[F(Y) - \frac{1}{2}]f(Y)dY.$$

Note that $f(Y)$ is the probability distribution function of Y. Then defining the inverse function $Y(F) = F^{-1}(Y)$ it is possible to rewrite[37] it by transforming the variables to

$$\Delta = 4\int_0^1 Y(F)[F - \frac{1}{2}]dF.$$

Now, one can use the definition of the covariance and that F is uniformly distributed between zero and one with mean of $F = \frac{1}{2}$ in order to express this integral as the covariance of wealth or income Y and its cumulative distribution $F(Y)$. Thus we get

$$\Delta = 4cov[Y, F(Y)].$$

Putting it back into $GC = \frac{\Delta}{2\bar{Y}}$ yields exactly the expression stated in the text.

[37] This result also goes back to Gastwirth (1972).

Derivation of the Decomposition Method

Recollecting from above that $GC = \frac{\Delta}{2\bar{Y}}$ with $\Delta = 4cov[Y, F(Y)]$ we have

$$\begin{aligned} GC = \frac{\Delta}{2\bar{Y}} &= \frac{2}{\bar{Y}} \sum_{k=1}^{K} cov[y_k, F(Y)] \\ &= \sum_{k=1}^{K} (cov[y_k, F(Y)]) \left(\frac{cov[y_k, F(y_k)]}{cov[y_k, F(y_k)]} \right) \left(\frac{\bar{y}_k}{\bar{y}_k} \right) \left(\frac{2}{\bar{Y}} \right) \\ &= \sum_{k=1}^{K} \left(\frac{cov[y_k, F(Y)]}{cov[y_k, F(y_k)]} \right) \left(\frac{2cov[y_k, F(y_k)]}{\bar{y}_k} \right) \left(\frac{\bar{y}_k}{\bar{Y}} \right) \\ &= \sum_{k=1}^{K} R_k G_k S_k. \end{aligned}$$

This is the decomposition method discussed in the text.

1.B Socio-demographic Characteristics

Table 1.8: Socio-Demographic Characteristics

	HSFW		EU-SILC		HSHW	
	# HH	% of Pop.	# HH	% of Pop.	# HH	% of Pop.
Sex						
Female	916	36.5%	2,027	40.5%	1,100	46.8%
Male	1,640	63.5%	3,121	59.5%	981	53.2%
Age						
up to 29	157	7.8%	324	7.8%	259	14.0%
30 – 39	418	19.0%	905	18.5%	381	20.3%
40 - 49	602	25.1%	1,108	20.6%	491	25.5%
50 - 59	483	16.6%	948	16.2%	307	16.2%
60 - 69	578	17.3%	946	16.1%	239	12.1%
70 and above	318	14.3%	917	20.8%	230	12.0%
Family status						
Single	410	17.4%	819	19.6%	489	25.4%
Married	1,582	59.5%	3,015	52.9%	1,141	50.4%
Divorced	309	12.1%	582	11.8%	276	14.7%
Widowed	255	11.0%	724	15.6%	175	8.4%
Education						
Primary	308	13.6%	1,215	24.3%	356	17.2%
Secondary	1,864	74.0%	3,255	61.6%	1,510	72.2%
Tertiary	384	14.4%	678	14.1%	215	10.6%
Employment status						
Employed	1,428	60.5%	2,670	51.5%	1,306	63.2%
Unemployed	56	1.5%	173	3.8%	55	2.8%
Pension	922	31.2%	1,951	38.1%	553	26.4%
Other	150	6.4%	354	6.6%	167	7.7%
Household size						
1 Member	772	32.3%	1,529	36.7%	599	35.1%
2 Members	844	28.6%	2,501	43.4%	728	28.6%
3 Members	406	15.3%	718	12.5%	310	14.9%
4 or more	534	23.8%	400	7.3%	444	21.3%

Notes:
(i) *Statistics reported:* Absolute number of households (# HH) in the according survey, and weighted frequency representing the proportion of population (% of Pop.) (Very few "no Answer" in EU-SILC are not reported).
(ii) *Source:* OeNB HFWS (2004), EU-SILC (2005), and OeNB HSHW (2008).

1.C Decomposition of Housing Wealth and Gross Income

As in the paper first the used variables are introduced. The descriptive statistics including various measures about the distribution are followed by the results from the decomposition analysis.

Housing Wealth

The second major way in which households accumulate wealth is in housing wealth (real asset side). Unfortunately, due to lack of data on both, financial and housing wealth, in one comprehensive data source, it is not possible at this point in time to decompose inequality of both financial and housing wealth at the same time. This will be possible once the data from the European Household Finance and Consumption Survey (HFCS) are out, which is expected in 2013. One can, however, have a closer look only at wealth held in real assets using the Household Survey on Housing Wealth (HSHW) from the OeNB conducted with computer assisted personal interview method in 2008. For the calculations we used the imputed data.[38] For the calculations of the descriptive statistics appropriate estimation techniques for multiply imputed data were used applying also household weights. For the structural part, however, we only use collected information.

Housing wealth can be separated into various forms (see table 1.9). In this exercise we use housing wealth held in the primary residence (where the household lives most of the time), secondary housing wealth (including holiday homes), further housing wealth (i.e. further private property owned), building ground, property used for agricultural purposes (short: agricultural), and other housing wealth (including hotels, offices, etc.).

The socio-demographic characteristics of the HSHW are reported together with the

[38]The data were imputed by the method called Multiple Imputations by Chained Equations. Further details on the method can be found in Albacete (2011).

Table 1.9: Definition of Variables on Housing Wealth

Variable	**Accronym**	**Definition**
Total Housing Wealth	THW	Sum of value of the following various real (housing) assets.
Primary Housing	PR	Current value of housing unit that is occupied as main residence by the owner.
Secondary Housing	SE	Current value of the house or apartment that is described as the second unit; i.e. for an owner of primary housing it is the second housing unit that is stated and for a renter of primary housing it is the first reported housing unit that is owned.
Further Housing	FH	Sum of current values of all further houses and apartments (third or more for an owner of PR and second or more for a renter of PR).
Building Ground	BG	Sum of current value of all building grounds.
Agricultural	AG	Sum of current value of further housing wealth (over and above the primary housing unit) that is used for agricultural purposes.
Other	OT	Sum of current value of further housing wealth (over and above the primary housing unit) that is used for various other means (e.g. hotel, office, etc.).

Notes:
(i) *Source:* OeNB HSHW (2008).

respective statistics of the other surveys (see table 1.8). They look very similar to the HFSW survey. Descriptive statistics and distributional information are reported in tables 1.10 and 1.11 as well as figure 1.3. There are various interesting aspects of the descriptive statistics of housing wealth. For example, each and every form of housing wealth is kept by less then 50% of the population in Austria. This, however, is not the focus of the paper and the interested reader is referred to these already well documented facts.[39]

Table 1.10: Descriptive Statistics: Housing Wealth

Variable	**Mean**	**Median**	**Variance**	**Min**	**Max**	**% Pop**
Total	247,949	100,000	16,279	0	10,440,118	59%
Primary	152,397	0	7,881	0	4,903,659	50%
Secondary	38,964	0	6,945	0	8,752,059	11%
Further	10,363	0	5,314	0	9,115,563	2%
Building Ground	30,436	0	9,567	0	10,440,118	7%
Agricultural	12,312	0	3,914	0	5,123,126	6%
Other	4,414	0	1,633	0	3,586,417	2%

Notes:
(i) *Number of Observations:* 2,081.
(ii) *Source:* OeNB HSHW (2008).

As we can see in table 1.12, only two components seem to influence overall inequality in a statistically significant way (at the 1%-significance level). From these a increase in housing wealth held in the primary housing unit tends to decrease inequality whereas the effect is reversed for 'Further Housing'. This is due to the fact that only a small fraction of wealth households own additional to the primary and secondary housing wealth further property. All the other components are only marginally or not at all significant and thus are not interpreted here.

[39]See e.g. Fessler, Mooslechner, Schürz, and Wagner (2009).

Figure 1.3: Lorenz-Curves for various Forms of Real Assets

Notes:
(i) *Source:* OeNB HSHW (2008).

Table 1.11: Inequality Measures: Housing Wealth

Variable	Overall (including observations with zero)				Within (excluding non-positive observations)			
	GC	GE(2)	AkI	$\frac{P90}{P50}$	GC	GE(2)	AkI	$\frac{P90}{P50}$
Housing Wealth								
Total	0.73	2.4	.	.	0.46	0.82	0.19	2.4
Primary	0.75	2.6	.	.	0.43	0.85	0.16	2.0
Secondary	0.97	19.6	.	.	0.57	1.06	0.28	3.2
Further Housing	0.99	84.0	.	.	0.69	1.03	0.50	44.8
Building Ground	0.98	69.9	.	.	0.69	3.51	0.41	7.8
Agricultural	0.99	81.4	.	.	0.80	4.00	0.55	11.8
Other	0.99	105.4	.	.	0.74	1.81	0.50	11.4

Notes:
(i) *Statistics reported:* Gini coefficient (GC), measure of the General Entropy Class with ε = 2 (GE(2)), Atkinson Index with $\alpha = 0.5$ (AkI), and the ratio of the 90^{th} to the 50^{th} percentile.
(ii) All the values denoted with '.' cannot be calculated (e.g. due to the fact, that the fiftieth percentile still holds nothing and thus one would have to divide by zero).
(iii) Only collected information used in the estimation.
(iv) *Source:* OeNB HSHW (2008).

Table 1.12: Decomposition of Housing Wealth Distribution

Variable	S_k	R_k	Share	% Change (Std. Err.)
Primary	0.783	0.946	0.756	-0.027** (0.007)
Secondary	0.107	0.813	0.116	0.010* (0.005)
Further Housing	0.012	0.942	0.016	0.004** (0.001)
Building Ground	0.036	0.767	0.038	0.002 (0.004)
Agricultural	0.054	0.887	0.066	0.012* (0.006)
Other	0.008	0.695	0.007	-0.000 (0.001)

Notes:
(i) *Significance levels (bootstrapped standard errors):* †: 10%, ∗: 5%, ∗∗: 1%.
(ii) *Reported results:* Share of wealth in source k (S_k); correlation of source k–inequality and cumulative inequality (R_k); share of total inequality that is explained by source k (Share); and how a change in source $k's$ wealth component affects total inequality (% Change).
(iii) Only collected information used in the estimation.
(iv) *Source:* OeNB HSHW (2008).

Gross Income

As a robustness check to the results for net/disposable income all the calculations for the income decomposition where done with gross income data as well. Tables 1.13 to 1.16 and figure 1.4 report the results. Note that for gross income a separation of pensions is slightly differently since there is no possibility to disentangle it from total gross income in the household file and thus had to be taken from the personal data file. Socio-demographic statistics are identical to the one reported in the main body (see table 1.8). These estimates can also be seen as a check whether any results discussed in the paper where derived from negative minimum values in some of the income variables. This seems not to be the case.

Table 1.13: Definition of the EU-SILC Variables

Variable	**Accronym**	**Definition**
Total gross income	TGI	Sum of all personal gross income of all household members.
GI from rent and lease	GIRL	Income received from renting a property minus its cost (mortgage interest payments, repairs, maintenance, insurance, etc.).
Accommodation allowance	AW	Means tested transfers from the public authority to tenants and owners to meet housing expenses.
GI from private transfers	GIPT	Regular cash received from inter-household transfers, such as alimony or monetary help for students.
GI from capital investment	GICI	Interest, dividends, and other profits from capital investment.
GI from family/children related	GIFC	Sum of gross family and children related income (HY050g), and other gross income against social exclusion (HY060g).
GI from pension	NIP	Household gross income from private and public pension plans (py080g and py100g summed over all household members).
GI from work	NIW	Total gross income from employment and self-employment summed over all household members(py010g+ py050g).
Residual	RES	Total gross household income minus the sum of all specific forms of income mentioned above.

Notes:
(i) *Source:* EU-SILC 065/04: Description of target variables; EUROSTAT.

Table 1.14: Descriptive Statistics: Gross Income Sources

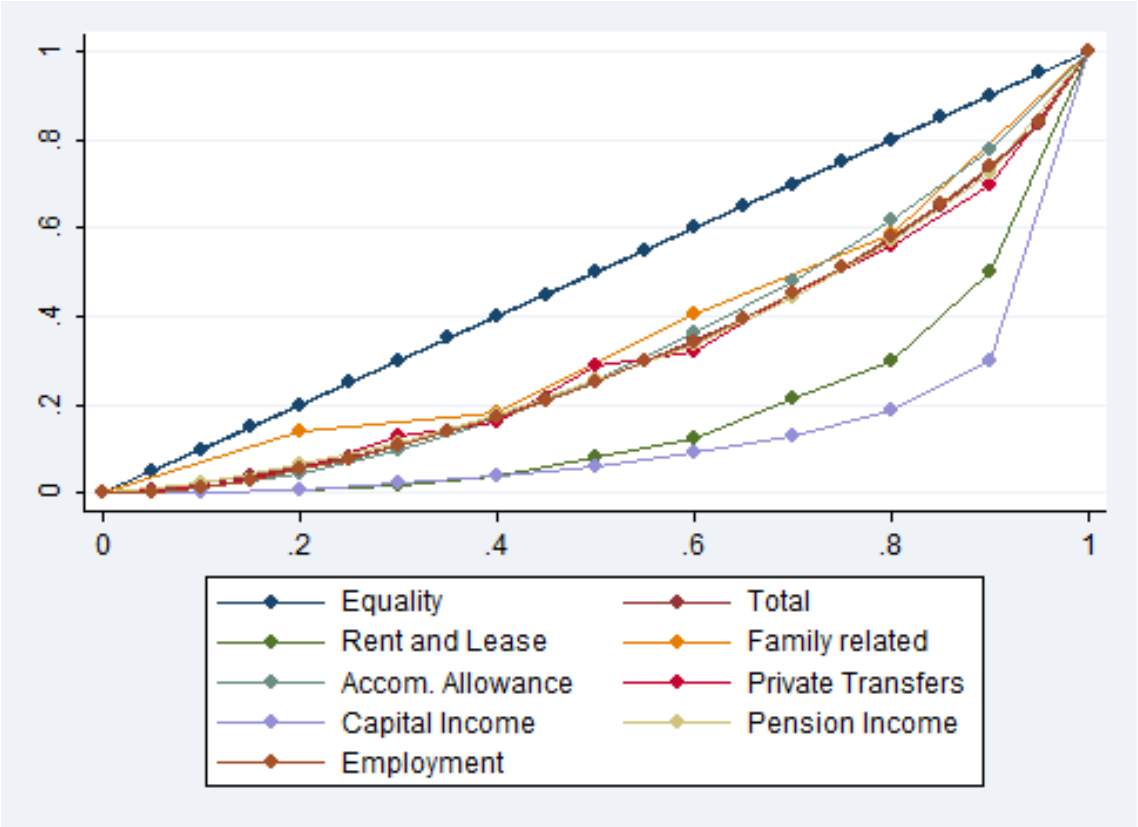
Variable	**Mean**	**Median**	**Variance**	**Min**	**Max**	**% Pop**
Gross Income						
Total	43,600	36,425	32,873	1	501,935	100%
Rent and Lease	493	0	4,975	0	175,456	4.0%
Accom. Allowance	57	0	334	0	3,600	4.0%
Private Transfer	274	0	1,470	0	28,800	6.3%
Capital Income	367	82	2,151	0	73,313	77.5%
Family related	1,571	0	2,933	0	28,771	33.3%
Pension	9,524	0	18,215	0	300,191	37.3%
Employment	29,567	24,920	32,331	0	420,869	70.3%
Residual	1,745	0	5,527	-31,200	115,146	

Notes:
(i) *Number of Observations:* 5,148.
(ii) *Source:* EU-SILC (2005).

The estimates[40] of the change in all sources of income are significant, but relatively small. So, for example, would a 10% increase in the income from rent and lease of property and land increase total income, measured with the GC, of 0.11%. All the estimates for gross income are comparable to the ones for net/disposable income.

[40] See table 1.16.

Figure 1.4: Lorenz-Curves of Gross Income Sources



Notes:
(i) *Source:* EU-SILC (2205).

Table 1.15: Inequality Measures: Gross Income

Variable	Overall (including observations with zero)				Within (excluding non-positive observations)			
	GC	GE(2)	AkI	$\frac{P90}{P50}$	GC	GE(2)	AkI	$\frac{P90}{P50}$
Gross Income								
Total	*	*	*	*	0.36	0.28	0.11	2.2
Rent and Lease	0.99	50.9	.	.	0.67	1.56	0.39	5.8
Accom. Allowance	0.97	17.0	.	.	0.35	0.20	0.11	2.1
Private Transfer	0.96	14.4	.	.	0.41	0.43	0.14	2.2
Capital Income	0.83	17.2	.	7.1	0.78	13.2	0.53	6.0
Family related	0.79	1.74	.	.	0.35	0.24	0.10	2.3
Pension	0.76	1.83	.	.	0.37	0.37	0.11	2.1
Employment	0.55	0.60	.	2.7	0.37	0.27	0.12	2.2

Notes:
(i) *Statistics reported:* Gini coefficient (GC), measure of the General Entropy Class with ε = 2 (GE(2)), Atkinson Index with $\alpha = 0.5$ (AkI), and the ratio of the 90^{th} to the 50^{th} percentile.
(ii) All the values denoted with '*' have no observation with zero value in this variable, thus only the estimates for the right part of the table are given.
(iii) All the values denoted with '.' cannot be calculated (e.g. due to the fact, that the fiftieth percentile still holds nothing and thus one would have to divide by zero).
(iv) *Source:* EU-SILC (2005).

Table 1.16: Decomposition of Gross Income Distribution

Variable	S_k	R_k	Share	% Change (Std. Err.)
Rent and Lease	0.012	0.690	0.023	0.011** (0.003)
Accom. Allowance	0.001	-0.490	-0.002	-0.003** (0.000)
Private Transfer	0.007	-0.072	-0.001	-0.008** (0.001)
Capital Income	0.008	0.592	0.012	0.003** (0.001)
Family related	0.042	0.278	0.025	-0.017** (0.001)
Pension	0.210	0.143	0.065	-0.144** (0.008)
Employment	0.679	0.840	0.877	0.199** (0.009)
Residual	0.042	0.013	0.001	-0.041** (0.004)

Notes:
(i) *Significance levels (bootstrapped standard errors):* †: 10%, ∗: 5%, ∗∗: 1%.
(ii) *Reported results:* Share of income in source k (S_k); correlation of source k–inequality and cumulative inequality (R_k); share of total inequality that is explained by source k (Share); and how a change in source $k's$ wealth component affects total inequality (% Change).
(iii) *Source:* EU-SILC (2005).

Chapter 2

Survey Mode Effects on Item Non-response and Income Inequality Measurement[1]

[1] This chapter comprises a paper that is joint work with Pirmin Fessler and Maximilian Kasy.

2.1 Introduction

Intertemporal and international comparisons of economic inequality are typically based on inequality measures calculated from survey data; see for instance a recent report by OECD (2011). The surveys most widely used for such calculations include the Survey of Consumer Finances (SCF), the Panel Study of Income Dynamics (PSID), and the EU - Statistics on Income and Living Conditions (EU-SILC). Across different surveys and across different waves of the same survey different interview modes are used. In particular, some of these surveys are conducted in person (computer assisted personal interview, CAPI), while others are conducted via telephone (computer assisted telephone interview, CATI). For a comparison of the different EU-SILC surveys see table 2.1.

Existing evidence (e.g. De-Leeuw, 1992; Lohmann, 2011) suggests that in such surveys nonresponse and misreporting of income data is a concern. This is of particular importance in the upper and lower tails of the distribution, and both of these problems might be even more severe in telephone interviews. Nonresponse and misreporting are important problems in the context of inequality measurement because many common inequality measures, such as the Gini coefficient, are not robust[2] (c.f. Cowell and Victoria-Feser, 1996), which implies that minor data contaminations can have a large impact on measured inequality. The possibly large influence of minor mismeasurement stands in contrast to comparably small differences in inequality measures across countries and time, see table 2.1, which suggests that comparisons across different surveys might be quite problematic.

In this paper, we use quasi-experimental variation of interview methodology in the Austrian EU-SILC 2008 survey to provide causal estimates of the effect of the interview mode (CAPI vs. CATI) on the estimated income distribution measured in several ways as e.g. the Gini coefficient and the 90/10 percentile ratio. Our estimates exploit the panel structure of the EU-SILC survey conducted in Austria

[2]A statistic is called robust if it has a bounded influence function, see Huber (2003).

in 2007 and 2008 and control for a rich set of covariates from the baseline survey. We find that a switch from CAPI to CATI leads to major changes in response behaviour, which imply large differences of estimated inequality measures. Selective item-nonresponse in the tails is significantly higher for CATI (on average by about 20%-30%), and incomes for CATI tend to be closer to the mean income. A switch from CAPI to CATI in particular decreases the Gini coefficient of household income by roughly 10%, and has a statistically insignificant effect on the 90/10 percentile ratio (see table 2.9). The smaller effect on the latter statistic is likely due to the fact that it is robust, whereas the Gini coefficient is not. A "Back of the Envelope" calculation in the last column of table 2.1 illustrates these possible effects and leads to severe re-ranking of several countries. These calculations suggest that countries with considerable proportions of CATI interviews would typically rank higher in rankings of income inequality (e.g. 7 positions higher in the case of Denmark and 5 positions higher in the case of Finland).

We draw the following general conclusions from these findings. First, it seems that CAPI yields more reliable measures of inequality, and should be used where possible. Second, when making international or inter-temporal comparisons of inequality, we should make sure to "compare apples with apples". Comparisons of inequality measures based on surveys using different methodologies might be quite misleading. Third, given the issues with survey data in general and surveys using CATI in particular, it might be advisable to focus on robust inequality measures such as quantile contrasts when conducting inequality comparisons. We view our estimates as the lower bound of the effect as coverage effects of the interview mode cannot be covered.

In our analysis of the effects of modes on the distribution of measured incomes we are using reweighting and regression methods developed in the literature on distributional decompositions, in particular by DiNardo, Fortin, and Lemieux (1996) and Fortin, Lemieux, and Firpo (2009). Our paper deviates from this literature in

Table 2.1: Overview of EU-SILC Surveys and their Mode of Data Collection

	PAPI	CAPI	CATI	Self-administrated	Gini 2007	adjusted Gini
Belgium	0.0	100.0	0.0	0.0	26.3	-
Czech Republic	99.7	0.0	0.0	0.3	25.3	-
Denmark	0.0	0.0	94.2	5.8	25.2	27.6
Germany	0.0	0.0	0.0	100.0	30.4	-
Estonia	2.2	97.6	0.2	0.0	33.4	33.4
Ireland	0.0	100.0	0.0	0.0	31.3	-
Greece	80.8	14.9	2.1	2.3	34.3	34.4
Spain	0.0	92.9	7.1	0.0	31.2	31.4
France	0.0	100.0	0.0	0.0	26.6	-
Italy	100.0	0.0	0.0	0.0	32.3	-
Cyprus	0.0	100.0	0.0	0.0	29.8	-
Latvia	11.3	81.2	7.5	0.1	35.7	36.0
Lithuania	95.3	0.0	3.8	0.9	33.8	33.9
Luxembourg	100.0	0.0	0.0	0.0	27.4	-
Hungary	100.0	0.0	0.0	0.0	25.6	-
Malta	0.0	100.0	0.0	0.0	26.3	-
The Netherlands	0.0	0.0	100.0	0.0	27.6	30.4
Austria	0.0	94.0	6.0	0.0	26.2	26.4
Poland	100.0	0.0	0.0	0.0	32.2	-
Portugal	8.0	92.0	0.0	0.0	36.8	-
Slovenia	0.0	44.5	55.5	0.0	23.2	24.5
Slovakia	99.4	0.0	0.0	0.7	24.5	-
Finland	0.0	3.4	96.6	0.0	26.2	28.7
Sweden	0.0	0.0	100.0	0.0	23.4	25.7
United Kingdom	0.0	100.0	0.0	0.0	32.6	-
Iceland	0.0	0.0	100.0	0.0	28.0	30.8
Norway	0.0	0.6	99.4	0.0	23.7	26.1

Notes:
(i) This table shows percent shares of EU-SILC surveys in 2007 conducted by paper assisted personal interview (PAPI), computer assisted personal interview (CAPI), computer assisted telephone interview (CATI), and self administered.
(ii) Gini coefficients are based on household disposable equivalence income.
(iii) The adjusted Gini coefficients is the "Back of the Envelope" calculation accounting for the effect (decrease of 10%) of the CATI interviewing technique from the RIF-regression.
(iv) *Source:* Eurostat: Comparative Intermediate EU Quality Report 2007. Version 5, and Eurostat website for Gini coefficients.

studying effects on measured, rather than actual, incomes. Survey methodology is not part of the standard economics curriculum, but economists become increasingly aware of the importance of careful data collection. This is partly due to the rise of experimental methods, and the consequent collection of primary data by development economists (see Duflo, Glennerster, and Kremer, 2008) and labor economists (see List and Rasul, 2011), but also true for economists using "subjective" data from large pre-existing surveys, see for instance Conti and Pudney (2011). This paper is intended to contribute to the rising awareness of data issues among economists.

The rest of this paper is structured as follows. In the next section we introduce the hypotheses in the background of the analysis and some more literature for it. It helps to clarify the mechanism that are at play in this specific part of the data production process and embeds the analysis in a wider context. The underlying data as well as the quasi-experimental set-up are laid out in the following part of the paper. It gives a detailed account on the switch of the interviewing mode and the measures that are used. Section 2.4 is dedicated to giving a short overview of the identification strategy. Additionally, we introduce the specific implementation of the estimation strategy. The main part of the paper, section 4.5, discusses the results reached from the investigation. We split it into the discussion of the effect of the interviewing mode on item-nonresponse, reported income level, and the distribution of income. Finally, section 4.6 sums up the results, draws some conclusions and concludes the paper.

2.2 Interview Modes

Evidence on the impact of the interview mode is mixed. In a meta-study De-Leeuw (1992) compared face-to-face interviews, telephone surveys, and self-administered mail questionnaires. By employing 67 mode-comparison-papers she showed that the main difference lies between self-administrated and interviewer-based survey

modes, the differences between interviewer-based modes, and therefore also between CATI and CAPI seem less pronounced. While there was no significant difference in response validity and social desirability bias, Face-to-Face interviews lead to slightly less item non-response.
CATI is the oldest of the computer assisted interviewing methods and is heavily used especially in market research. CAPI is widely used especially for complex surveys of governmental statistic agencies and universities (De-Leeuw (2008)).

> "Because of the increased availability of other survey modes, face-to-face interviews are typically reserved for the most difficult and longest surveys that place the greatest burden on respondents. These are all kinds of surveys for which the other modes are not so likely to perform well. Face-to-face surveys also tend to be reserved for surveys that are most important to society, for which sponsors are willing to pay the cost." (De-Leeuw, Hox, and Dillman (2008), page 164)

The main advantage of face-to-face interviews is that they are more flexible than telephone interviews. The interviewer can use response cards, visual scales etc. but also explain things better by being physically present which allows for a broader range of communication and interaction between the interviewer and the respondent. Via telephone the respondent can only rely on his/her memory when answering questions with multiple answer possibilities.
Also interview length is an important determinant of the quality of the interviews given a certain mode. While CAPI and face to face in general is usually used for longer interviews (>30min) telephone interviewing is less suitable for longer interviews. Major data collection organizations in the US are refusing to conduct telephone interviews which are expected to last longer than 18 minutes (De-Leeuw, Hox, and Dillman (2008)).

We hypothesize that there are three different mechanisms at work how the mode influences measured income inequality.

First, on the level of unit non-response. We find that even though higher income households tend to be selected more often towards a CATI interview the highest values are found for households interviewed via CAPI. The same is also true for the lower end of the distribution. Therefore it seems that via CATI the households at the tails of the income distribution can not be reached as good as via CAPI (see table 2.3).

Second, on the level of item nonresponse. In addition to the effect of leading to generally higher item nonresponse the effect of CATI is especially pronounced at the tails of the distribution. Especially at the top this is very worrisome as the share of total income held by the top income earners is a multiple of their population share.

Third, on the level of the income values. We find that CATI leads to a positive effect on income values reported which is relatively higher for lower income groups than for very high income groups and leads to significantly lower measures of income inequality. This might be due to two mechanisms. On the one hand, the questionnaires in these surveys are rather complicated and so especially financially less literate (low end of the income distribution) households as well as households with very complicated income structures (high end of the income distribution) might be more more likely to exhibit measurement error over the telephone than in a personal interview. On the other hand, it might be simply easier to lie over the phone than in a personal conversation.

We can only quantify the third effect in terms of measures of income inequality and the second in terms of a causal estimate of the interviewing mode on the item non-response. However, we also find strong evidence that the first effect exist and is non-negligible. Therefore it is very likely that we underestimate the total effect of the interview mode on observed inequality measures as both the first and second effect also seem to compress the observed income distribution. In general people

with very low income as well as people with very high income tend to report values biased towards the mean. Over all three effects this measurement error seems to be larger when CATI instead of CAPI is used as an interview mode. The decision to not report at all (unit non-response), selectively not report (item non-response), or report values closer to the mean might be easier via the phone than in a face to face situation.

2.3 Quasi-Experiment and Data

For this empirical analysis we use Austrian EU-SILC data[3] of the waves 2007 and 2008. The documentation on the data is provided by Statistik Austria.[4] Additionally to the data available in the standard user dataset we obtained directly from Statistik Austria an indicator for both years about whether a specific household was interviewed using the CAPI or the CATI interviewing mode. We use both waves since the CATI option was first (after a test period for 2007 with households that are excluded as is described below) introduced in 2008. Only households that already were interviewed in 2007 could choose between a personal or telephone interview. The first choice of the data producers was to use the cheaper CATI option and first contact was - wherever possible - established via phone. We model the selection (see section 2.4.1) in order to control for it and identify the remaining effect due to the interview mode.[5] Table 2.2 reports how many households can be used for the identification of the effect of the interview mode. There are 5,711 households interviewed in 2008 of which about 30% (i.e. 1,710 households) are surveyed via a CATI interview. Only two thirds are panel households in the sense that they were

[3]See `http://epp.eurostat.ec.europa.eu/portal/page/portal/microdata/eu_silc` for further informations [accessed on 13th of November 2012].

[4]See Statistik (2010) for a version in German.

[5]According to the handbook (see Statistik (2010)) there were about 160 and 40 interviewer respectively for the CAPI and CATI part of the survey. Information on which household was interviewed by which interviewer, to control for possible interviewer effects, however, is unfortunately not part of the EU-SILC data. There seem to be enough interviewers, however, for the possible interviewer bias being rather small.

also interviewed in the previous wave. Thus the effective sample size is reduced to 3,377 households which could be (self-) selected to one of both interview modes. 42.6% of these actually were interviewed via telephone while the rest opted for a personal survey.[6]

Table 2.2: Number of Households in Different Modes and Waves

	Number	%-Share
Total Number of Households 2008	5,711	100%
CATI-Interview-Mode	1,710	29.94%
CAPI-Interview-Mode	4,001	70.06%
Panel Households 2007/2008	3,772	66.05%
CATI-Test Households 2007	395	6.92%
Effective Sample Size	3,377	59,13%
From the Effective Sample Mode was:		
CATI	1,438	42.58%
CAPI	1,939	57.42%

Notes:
(i) This table reports the household sample size for the 2007/2008 waves of EU-SILC.
(ii) Effective sample size is "Panel Households" minus "Test Household 2007".
(iii) *Source:* EU-SILC 07/08.

In EU-SILC 2008 selection towards CAPI is characterized by the following rules.[7] Households that re-participate, i.e. already took part in a previous wave, in the 2008 EU-SILC-Survey were generally intended to answer the questionnaire via the telephone. This tendency was implemented through a first contact (after the sending of an information letter) by calling the respondent and asking the household to participate via telephone. The change to the CAPI-mode was available for households

[6] There is a minor technical issue with 15 households that were not part of the 2007 wave, were, however, interviewed over the telephone in 2008. From the information provided by Statistik Austria in a personal correspondence this is due to the fact the households that were already part of the panel component in 2006, but could not be reached in 2007 (wave non-response), were possibly interviewed in the CAPI mode. Since this is only the case for the 2008-wave and was changed in subsequent waves to completely drop those households, we leave out these 15 households in the analysis. Furthermore as there exists no 2007 information on these 15 households we can not use them in our empirical exercise.

[7] This information is based both on the official documentation (see Statistik (2010)) and personally transmitted material from Statistik Austria.

that indicated that they refuse to take part in the survey over the phone, could not respond on the phone due to illness or age related reasons, or could either not be reached over the phone (i.e. wrong number, not available on the phone, etc.) or set dates were postponed by the respondent for several times.[8] All CATI-interviews were conducted during the weekdays (Monday to Friday) from 4pm to 8pm and additionally on Tuesdays between 9am and 1pm.[9] In general there seems to be a tendency to push households to conduct interviews on the phone while no unit-non-response is allowed due to the wish of a respondent to conduct a CAPI-interview.

To sum up the quasi-experiment at hand: All households were interviewed in 2007 via CAPI. All of them were targeted to be reinterviewed via CATI in 2008. Due to accessibility problems via phone and the possibility to opt for the CAPI mode roughly 40% of them were interviewed via CATI and 60% via CAPI. We are able to control for the non-random part in this treatment assignment with the information which was gathered in 2007 for all households with the same (CAPI) mode, including income and item nonresponse information, to estimate the effects of the treatment - the interview mode - on the observed income distribution in 2008.

Data from the 2007 wave (instead of the 2008 wave) is used because it is predetermined and hence by construction not exposed to a mode effect. Thus the controls are no outcome of the selected mode and can be considered strictly exogenous to the mode. Descriptive statistics of the controls by sub-samples are given in table 2.3.

2.3.1 Measures

The main focus is laid on disposable income. Household disposable income is constructed by summing up all of the households income sources. EU-SILC provides flag variables to distinguish three categories with regard to necessary imputation of this

[8] Information from Statistik Austria: "CAPI interview mode was offered only if there is a refusal of the phone interview, there are health related difficulties, or age related reasons that make a telephone interview impractical."

[9] No general rules were set for interviewing time of CAPI-interviews.

variable, (i) not imputed, (ii) partly imputed, (iii) completely imputed. Whereas "not imputed" implies that all items the aggregate household level variable consists of, were provided by the respondents, "partly imputed" means that one or more items where missing and had to be imputed and "completely imputed" means that all items where missing and had to be imputed.
Using this information we construct a household income item non-response dummy (HINR) being 1 if the household disposable income flag is indicating that household disposable income is "partly imputed" or "completely imputed" and 0 otherwise. The dummy therefore indicates if there was missing information with respect to the main household income variable.[10]

Disposable income[11] is taken as it is provided in the data. Table 2.3[12] reports in the upper part the mean of the variables in the panel households and the CAPI and CATI sub-samples. Already here it becomes clear that the average income is lower for households surveyed in a CAPI interview and the item non-response is 20%-30% lower for these households. Furthermore, we can see that most of the income-poor households, i.e. 75% of the households below an income of 10,000€, and all of the income rich households (above 200,000€) are interviewed via CAPI. This fact already indicates the problems originating from the telephone interview mode in terms of a possible compression of the income distribution via unit- and item nonresponse. CATI seems to increase coverage problems espcially at the tails of the distribution. Obviously none of the very high as well as very few low income households could be reached via CATI. Furthermore, it can be seen that both disposable and gross income increased from 2007 to 2008 on average 6% and 9% respectively. The differ-

[10]Furthermore, as a robustness check, we repeated the analysis (results can be provided upon request) using a differently constructed measure of item non-response. The construction is based on the individual income components and is calculated as the percentage of components imputed (i.e. missing in the first place) for the total household income. The results are qualitatively very similar and left out due to space constraints.

[11]As a robustness check we also worked with gross income.

[12]These estimates are not weighted since we are not interested in population but only sample averages at this stage of the analysis.

ence in the change between the two groups, however, seems small compared to the difference of 2008 income, i.e. 2 percentage points relative to about 20% (non log).

Table 2.3 also provides the mean for all household level[13] and personal level control variables from the EU-SILC wave of 2007. There are some albeit generally small differences in the averages between the two groups. Most noticeable, more affluent households in terms of income (see the difference of log income in 2007) as well as wealth (see as two indicators the percentage of owning the primary residence and the size of the primary residence) seem to be (self-) selected to CATI. Additionally, phone coverage (see land line and mobile phone coverage) and education seem to be lower for households with a CAPI-interview and women answer the questionnaire more often over the telephone. These differences indicate the possibility of households with certain characteristics to be (self-) selected into one or the other interviewing mode, and hence we model this process and control for it in the estimations (see sections 2.4, and 2.4.1). Another insight into how strong the selection into the different interviewing modes is can be gained from the statistics in table 2.4. According to income deciles in 2008 we see in the second column that the participation rate in a CAPI interview is decreasing with income. Stated otherwise households with higher income tend to opt at a higher rate for CATI-interviews. Also it is reported that item nonresponse is higher for higher income households. However, there is a enormous difference of more than 10 percentage points in the two top deciles. At the lowest end, in the first decile the difference is almost as high, and in between income from households with CATI-interviews always have higher item nonresponse, although the difference is not stable over the deciles. As can be seen in the last two columns, the percentage of households with missing income data does not follow the same trend in 2007. Here the two groups are defined according

[13]We also included regional indicators for of households as controls, but do not report them in table 2.3 due to space constraints. The partitioning is approximately as follows: Burgenland 4%; Kaernten 8%; Niederoesterreich 20%; Oberoesterreich 20%; Salzburg 7%; Steiermark 14%; Tirol 8%; Vorarlberg 5%; and Wien 17%.

Table 2.3: Mean of Dependent and Explanatory Variables

	Panel	CAPI	CATI
Dependent variables: 2008			
Mean household disposable income	10.24	10.16	10.37
Mean household gross income	10.51	10.42	10.66
Household Item nonresponse	21.84	18.08	27.53
Share of interviews below 10k	16.52		
Interviewed by		74.80	25.20
Share of interviews above 200k	0.07		
Interviewed by		100.00	0.00
Mean difference disposable income to 2007	0.06	0.07	0.05
Mean difference gross income to 2007	0.09	0.10	0.08
Control variables at the household level: 2007			
Household size	2.35	2.28	2.45
Households with kids	31.06	29.27	33.65
Single family home	39.34	37.16	42.48
Home-owners	53.51	49.40	59.43
Size of Flat in sqm	97.96	92.60	105.70
Land line coverage	66.16	58.14	77.74
Mobile phone coverage	87.30	85.18	90.36
Mean household disposable income	10.18	10.09	10.32
Mean household gross income	10.42	10.32	10.58
City	38.44	38.18	38.81
Urban	26.17	26.95	25.04
Rural	36.03	35.73	36.46
Control variables at the level of the household head: 2007			
Female	45.87	44.03	48.52
Blue collar	43.44	42.46	44.85
Self-employment	6.21	5.77	6.83
Jobless	50.36	51.77	48.32
Weekly working hours	19.91	19.84	20.02
Married	49.69	45.52	55.71
Education: secondary school	21.35	25.91	14.76
Education: apprenticeship	52.54	52.68	52.34
Education: higher secondary school	16.04	13.13	20.25
Education: university	10.07	8.28	12.66
Age	52.39	51.86	53.14

Notes:
(i) This table shows the mean of the variables on the consideration as well as the full set of control variables.
(ii) Each statistic is provided for the panel component of the 2008 wave, and the CAPI and CATI sub-samples of the panel component.
(iii) All income variables are reported after taking the natural logarithm.
(iv) The percentage of households in each region is left out due to space constraints, but the partitioning is approximately as follows: Burgenland 4%; Kaernten 8%; Niederoesterreich 20%; Oberoesterreich 20%; Salzburg 7%; Steiermark 14%; Tirol 8%; Vorarlberg 5%; and Wien 17%.
(vi) *Source:* EU-SILC 07/08.

to the 2008 interview mode choice and we report item nonresponse in 2007 when both groups (CAPI as well as CATI) were still interviewed via CAPI. Again the percent of households with missing income data increases with income level. The difference between the groups however is not stable already suggesting that there might be indeed a interview mode effect at play. In the top decile the income of households that were interviewed via CAPI in 2008 was missing more often than for the other group.

Table 2.4: Structure of Missingness of Income over Deciles Conditional on Interview Mode

		2008		2007	
Interview mode		CAPI or CATI		CAPI	
Interview mode in 2008	% CAPI Int	CATI	CAPI	CATI	CAPI
Total	.	0.271	0.181	0.271	0.234
First Decile	0.691	0.233	0.135	0.223	0.157
Second Decile	0.715	0.124	0.108	0.124	0.188
Third Decile	0.629	0.174	0.141	0.215	0.161
Fourth Decile	0.596	0.241	0.143	0.158	0.173
Fifth Decile	0.557	0.203	0.118	0.250	0.237
Sixth Decile	0.592	0.223	0.214	0.304	0.260
Seventh Decile	0.556	0.278	0.190	0.311	0.243
Eighth Decile	0.520	0.295	0.267	0.301	0.233
Ninth Decile	0.467	0.341	0.236	0.358	0.344
Tenth Decile	0.441	0.435	0.316	0.330	0.424

Notes:
(i) This table shows the average use of the CAPI interviewing mode over deciles (column 2).
(ii) Additionally, one can see the average rate of missing data on the income variable for 2008 and 2007 over the deciles (column 3 and 4, as well as 5 and 6 respectively).
(iii) *Source:* EU-SILC 07/08.

2.4 Identification and Estimation Strategy

We approach the problem of estimating the effect of the interview mode (CAPI versus CATI) on the income distribution by referring to recent microeconometric literature on causal inference and program evaluation as it best fits the quasi experimental setting of our data. The background of this approach can be found in the books of Morgan and Winship (2007), Angrist and Pischke (2009), Abbring and

Heckman (2007), or the papers of Angrist and Krueger (1998), Blundell and Dias (2002), and Imbens and Wooldridge (2009).

As we cannot assume random assignment of the interview mode in the case of EU-SILC 2008 we need to control for selection in order to justify a causal interpretation of the estimated effects. In order to check the validity of the Conditional Independence Assumption and robustness of the results in general we use various ways (OLS, OLS with controls, fully interacted model, propensity score matching, and coarsened exact matching techniques) to control for the selection into the different interview modes when estimating the average treatment effects (ATE) on item non-response and income levels.

Note that even small effects on different parts of the income distribution might accumulate to relevant overall effects on income inequality measures. That is why we turn in a second step towards an analysis of the mode effect on income inequality measures. After evaluating in a naive way simply the differences of various statistics about the income distribution, we follow the approach suggested by Fortin, Lemieux, and Firpo (2009) and use recentered influence functions (RIF) to estimate the ATE of a a bivariate variable (i.e. CAPI vs. CATI) on any distributional statistic $\nu(F(Y))$, where $F(Y)$ is the unconditional distribution of household income. The results concerning the effect of the interview mode on the distribution reported below are based on these calculation of the RIF[14] and use appropriate weights in order to estimate the effect on the population estimate of the distribution. For the standard errors the delta method is used. A robustness check with bootstrapped standard errors, however, yields very similar results. For the readers convenience the analytical tools are summed up in appendix 2.A.

[14]For the calculation of the RIF the reader is also referred to Essama-Nssah and Lambert (2011).

2.4.1 (Self-) Selection with Relation to the Mode

In order to model the selection into interview mode we apply a logit model where the dependent variable is the interview mode (CATI=0, CAPI=1) and the explanatory variables are the controls.[15] Table 2.5 shows the average marginal effects calculated from this model. Additionally, figure 3.16 displays the common support of the propensity scores given by the logit model. One can see that almost the whole support of both, CATI- and CAPI-households' propensity scores, is included in the common support. Selection bias can hence be expected to be relatively small. However, the chance of (self-) selection to CAPI decreases (statistically) significant on a 5% level with household size, household disposable income, the availability of a telephone line and a mobile phone in the household as well as the main income earner being female, being married and living together as well as having higher educational attainment.

2.4.2 Methods used to Control for Selection Bias

As a first step we estimate logit regressions using HINR as dependent variable and the CAPI (mode) dummy as independent variable as well as two sets of controls, i.e. only household-level and household- and personal-level controls. Our next step in terms of flexibility is to estimate a fully interacted linear model (FILM),[16] which allows the effect of the mode to vary over all controls. A further step in terms of flexibility is to estimate treatment effects based on propensity score matching (PSM).[17] While FILM allows the mode effects to vary over all controls, propensity score matching additionally allows (i) to impose common support based on the overlapping regions of the propensity scores (see figure 3.16) and (ii) does - due to its

[15] We also control for regional differences (not reported). The estimates of which are not significant.

[16] We use *film*, a STATA program provided by Edwin Leuven and Barbara Sianesi, see `http://www.ifs.org.uk/publications/2712` [accessed on 19th of November 2012].

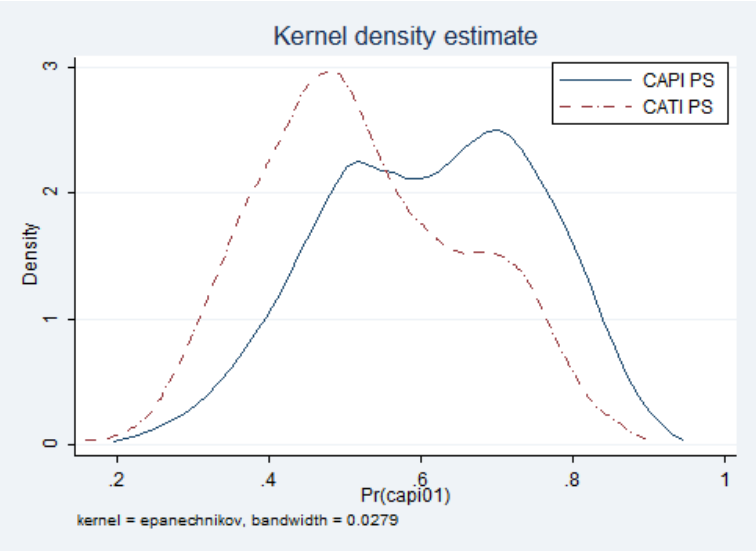
[17] We use *psmatch2*, a STATA program provided by Edwin Leuven and Barbara Sianesi, see `http://www.ifs.org.uk/publications/2684` [accessed on 19th of November 2012].

Table 2.5: Logit-regression of the (Self-)Selection of the Mode on Control Variables

Selection towards CAPI	
	Mode Selection
household characteristics	
Household size	0.031
	(0.012)
Household with kids	-0.041
	(0.028)
Single family home	0.023
	(0.022)
Owner occupier	-0.028
	(0.021)
Living space in sqm	-0.001
	(0.000)
Land line	-0.178
	(0.020)
Mobile phone	-0.097
	(0.029)
Log-disposable income	-0.052
	(0.018)
Personal characteristics of household head	
Female	-0.072
	(0.018)
Self-employed	0.009
	(0.037)
Jobless	0.079
	(0.037)
Weekly working hours	0.002
	(0.001)
Married and living together	-0.068
	(0.022)
Education: apprenticeship	-0.090
	(0.023)
Education: higher sec. school	-0.191
	(0.030)
Education: university	-0.131
	(0.036)
Age	-0.001
	(0.001)
N	3376

Notes:
(i)This table shows average marginal effects (AME) of household characteristics of the 2007 EU-SILC wave on being interviewed by CAPI in the 2008 EU-SILC wave. All average marginal effects are calculated from a logistic regression using an CATI(0)/CAPI(1) as dependent variable. Furthermore we controlled for federal states and regional population density. The coefficents of both are not significant.
(ii) Standard errors calculated by the delta method are given in parentheses.
(iii) *Source:* EU-SILC 07/08.

Figure 2.1: Propensity Score Density of CAPI and CATI



Notes:
(i) This graph shows estimated propensity score densities resulting from the logit model presented in table 2.5.
(ii) *Source:* EU-SILC 07/08.

semi-parametric nature - not impose as strong linearity assumptions as logit and FILM are imposing. On the imposed common support we match the nearest - in terms of the propensity score - CATI-neighbour to every CAPI-household (1 to 1 matching) in order to balance the joint distribution of the covariates (controls).[18]

We use coarsened exact matching (CEM) as a further robustness check. Iacus, King, and Porro (2008) developed a method to temporarily coarse data based on ex-ante user choice and then run the analysis on the common support of the uncoarsened data.[19]

While PSM still uses a parametric model to match CAPI and CATI observations

[18]Note that propensity score matching is not exact matching. That means, that the common support imposed by the propensity score differs from the common support in the strict sense which would include only cells of covariate combinations which include both CAPI- and CATI-households (the later is a subset of the former.). As exact matching is not feasible given finite data and many (and including continous) covariates, the resulting model still extrapolates also for cells which do not include both CAPI- and CATI-households

[19]We use *cem*, a STATA program provided by Matthew Blackwell, Stefano Iacus, Gary King and Giuseppe Porro, see `http://ideas.repec.org/c/boc/bocode/s457127.html` [accessed on 19th of November 2012].

and therefore extrapolates outside the common support[20] in order to calculate treatment effects, CEM imposes a user input based non-parametric matching strategy to balance the joint distributions of covariates among CAPI- and CATI-observations. This reduces the necessary extrapolation outside the common support but of course comes with a decrease in sample-size.

As matching variables we use a subset of our household and personal level covariates. For categorical variables (household size and being female as the household head) we impose an exact matching strategy and for the two continuous variables we use a coarsening strategy for matching on certain parts of the distributions by imposing cut-points (household level: household disposable income 20000, 30000, 40000, 50000, 60000; and personal level: age 30, 40, 50, 60, 70). This matching strategy leads to a perfectly balanced dataset in terms of the joint distribution of the categorical variables. As household disposable income is allowed to be matched in approximately 10,000€ brackets (and below 20,000€ as well as over 70,000€) and age is allowed to be matched in about 10 year age brackets (and below 30 as well as older than 70) some imbalances remain with respect to those two variables. Out of the 343 covariate combinations defined we find 224 which define the common support, i.e. where at least one CAPI and one CATI observation can be found. In terms of the sample the total of 3,377 observations collapses to 3,190 observations which lie inside the common support. 63 CATI- and 124 CAPI-observations lie outside the common support. The weights for the matched observations are chosen in a way that CAPI and CATI observations are balanced for each covariate combination. To control for the remaining imbalances in the matched dataset we still use age and household disposable income as controls when we calculate the treatment effects.

[20] With relation to realizations of CAPI- and CATI-observations in cells defined by covariate combinations and not with relation to the common support of the propensity scores.

2.5 Results

2.5.1 Mode Effects on Income Item Non-response

Table 2.6 shows the estimated effects of CAPI on item non-response. All estimated effects resulting from the logit regressions are significant at a 1% significance level, but are not significantly different from each other indicating again that selection bias is rather small. The estimate using the largest set of controls, i.e. all household- and personal-level controls given in table 2.3, is −0.071 which indicates that the probability of item non-response when interviewed via CAPI is 7.1 percentage points (the average item non-response is 21.8%) lower than in the case of a CATI interview. This implies a reduction of item non-response of about 30% given the EU-SILC estimates.

Table 2.6: Interview Mode Effect on Income Item Non-Response

	I	*II*	*III*
Logit Model	-0.088	-0.075	-0.071
	(0.014)	(0.014)	(0.015)
Fully Interacted Model			-0.076
			(0.016)
Propensity Score Matching			-0.072
			(0.022)
Coarsened Exact Matching			-0.068
			(0.014)
N	3291	3290	
Household Controls		X	X
Personal Controls			X

Notes:
(i)This table shows average partial effects (APE) of being interviewed by CAPI on household and personal income item non-response. Results are reported from a logistic (using an item non-response dummy for household income [at least one item non-response in an income question] as dependent variable) as well as a fully interacted model and various matching techniques.
(ii) Standard errors are given in parentheses (for the standard errors of the marginal effects that delta method is applied).
(iii) *Source:* EU-SILC 07/08.

The estimates resulting from the FILM and PSM estimations closely resemble the logit estimates. The FILM and PSM estimates are −0.076 and −0.072 respectively and significant at the 5% level. The estimate using 3,190 CAPI households

after the coarsened exact matching and reweighting to balance the covariate distributions is –0.068 and significant at the 5% level. Thus these more flexible estimation approaches confirm the previous finding of the interview mode on the item non-response using the somewhat naive linear approach.

2.5.2 Mode Effects on the Income Level

Analogous to the estimates of the mode effects on item nonresponse we estimate the effect of the mode on household income again with increasing flexibility and less restrictive assumptions. All estimates shown in table 2.7 are produced analogously to the estimates in table 2.6, but with the logarithm of household income of 2008 being the dependent variable instead of HINR.

Table 2.7: Interview Mode Effect on Household Income

	I	*II*	*III*
Average effect of CAPI on log household income			
OLS-Regression	-0.205	-0.052	-0.040
	(0.022)	(0.016)	(0.015)
Fully Interacted Model			-0.043
			(0.016)
Propensity Score Matching			-0.017
			(0.032)
Coarsened Exact Matching			-0.060
			(0.017)
N	3377	3376	
Household Controls		X	X
Personal Controls			X

Notes:
(i) This table shows the effect (regression coefficient, as well as matching estimators) of being interviewed by CAPI on the logarithm of household disposable income.
(ii) Standard are given in parentheses.
(iii) *Source:* EU-SILC 07/08.

Overall we find a significant negative effect of the CAPI mode on average (log) household income. The estimate ranges between –0.02 (propensity score matching) and –0.2 (OLS without controls). This estimate (taking as an example 0.04) implies that interviewing via CAPI leads on average about 1,000€ lower reported income

values than interviewing via CATI. As the income distribution is very skewed and most households have income below the mean income this result already implies that CATI leads on average to income values closer to the mean income. Again the more flexible methods FILM, PSM and CEM confirm the negative effect, however the PSM estimate is not significant.

The average effect, however, is of no help if one wants to analyse the impact on the entire distribution and resulting inequality measures. In fact it might be quite misleading since the effect on inequality measures is positive instead of negative. Furthermore, even many small and insignificant effects over the distribution could accumulate to relevant effects on inequality measures.

2.5.3 Mode Effects on the Income Distribution

Differences

Starting from a simple idea in order to compare the whole distribution in more detail we calculate the difference of a given statistic for both sub-samples,[21] i.e. for households interviewed with CAPI and CATI respectively, and provide standard errors of this difference using a bootstrap with 500 replicates.[22] We apply this procedure to the Gini coefficient; the poverty rate, i.e. the proportion of population with lower income than 60% of the median; and the 90/10 percentile ratio. Additionally, we provide the estimates of the difference between the CAPI- and the CATI sample statistics (i) without any adjustments and (ii) on the common support resulting from the CEM-Procedure, which controls partly for selection bias but at the cost of a reduction of the sample.

Table 2.8 reports first the measures of the Gini coefficient, the Poverty Rate, and the 90/10 percentile ratio for both the CAPI and CATI sub-samples in EU-

[21] Appropriate weights are used in the calculation of the statistics, since we are now interested in the effect on the population and not, as before, in the sample.

[22] For the estimations the whole procedure is bootstrapped in the sense that a bootstrap subsample is drawn, and then the difference of the statistic calculated. With these 500 estimates of the differences we are able to estimate standard errors.

Table 2.8: Differences of Inequality Measures between Interview Modes

	Gini Coeff.		Poverty rate		90/10 Percentile Ratio	
	Ia	*IIa*	*Ib*	*IIb*	*Ic*	*IIc*
Inequality Measure CAPI	0.3341	0.3332	0.2280	0.2226	4.801	4.791
Inequality Measure CATI	0.3076	0.3096	0.1885	0.1920	4.523	4.527
Difference (Bootstrap)	0.026	0.024	0.039	0.031	0.278	0.263
Bootstrapped Std. Err.	(0.012)	(0.013)	(0.014)	(0.013)	(0.218)	(0.267)

Notes:
(i) This table shows the effect of the interview mode on aggregate measures of inequality. We report the inequality statistic (Gini coefficient, the Poverty Rate, and the Percentile Ratio), and the difference between the sub-samples.
(ii) Standard errors are reported using bootstrapping method.
(iii) Columns *Ia-c* show the results using the full (panel) sample, i.e. 3.377 observation; and columns *IIa-c* only use the matched observation from the above coarsened exact matching procedure.
(iv) *Source:* EU-SILC 07/08.

SILC 2008. We find that the difference of the Gini coefficient is 0.026 for the full sample and 0.024 using only the common support sample implying a 8.1% and 7.2% lower Gini coefficient using CATI instead of CAPI. While the estimate of the whole sample is significant at the 5% level, the difference for the balanced sample is only significant at the 10% level.[23] Table 2.8 additionally reports a significantly higher poverty rate for the CAPI sample (i.e. 3.9 percentage points in the whole and 3.1 percentage points in the balanced sample) and an (statistically) insignificant difference of the 90/10-percentile ratio between the two sub-samples. To control rigorously for selection we employ RIF-regressions to estimate the effect of interview mode on the distributional statistics of income.

Overall Effect

We regress the RIF of the Gini, the Poverty Rate, and the 90/10 percentile ratio of the income distribution of 2008 on the interview mode and (i) linear, squared and cubed income from 2007 as well as (ii) all our controls for households characteristics

[23]To check for the differences at the top we also used a General Entropy Class index with $\alpha = 2$ which shows huge differences (0.18 for the CATI versus 0.49 for the CAPI common support samples). This is due to the fact that it is very sensitive to top income and as we saw the highest income observations are all from CAPI interviews. This sensitivity, however, also renders a high variability of replicates in this bootstrapping procedure and thus generates high standard errors yielding insignificant results.

(including income). In both cases also all interactions of the control variables with the interview modes are included in the model. For the standard errors of the reported ATE we use the delta method.[24] Using CATI instead of CAPI as interview modes reduces the observed Gini coefficients significantly (see table 2.9). According to our estimates of 0.033 with the limited set of controls and 0.027 using the full set of controls, inequality measured by the Gini coefficient drops by around 10% if the interview mode is switched from CAPI to CATI. This is a severe effect on the most prevalent inequality measure. Furthermore, we see that the estimate on the poverty rate decreases greatly and looses its significance and the 90/10-percentile ratio switches signs but remains insignificant. The result suggests that using the latter two measures when comparing inequality over time or between countries might be more robust.

Table 2.9: Effect of Interview Mode an Inequality

	Gini Coeff.		Poverty rate		90/10 Percentile Ratio	
RIF-regression	0.0329	0.0274	0.0068	0.0019	-0.2344	-0.1829
	(0.0107)	(0.0101)	(0.0127)	(0.0134)	(0.2469)	(0.2571)
Controlling for						
past income	X	X	X	X	X	X
other covariates		X		X		X

Notes:
(i) This table shows the effect of the interview mode on aggregate measures of inequality. We report the inequality statistic (Gini coefficient, poverty rate, and the percentile ratio), the difference between the sub-samples and the effect of the interview mode using RIF-regressions.
(ii) Standard errors are reported using delta methods.
(iii) *Source:* EU-SILC 07/08.

Heterogeneity of the Effect

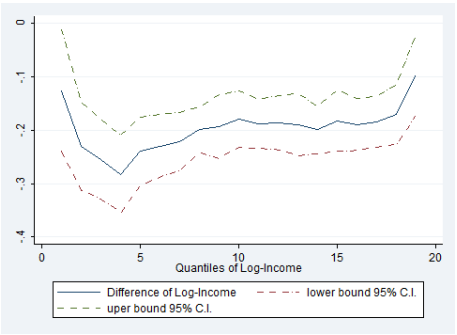
The estimates for the difference in percentiles are reported in figure 2.2. Panel (a) to (c) show that the effect of the interview mode on the percentiles[25] follows a u-shape.

[24] Thus the standard errors are based on the whole sample of the 2008 wave of EU-SILC and not only on the panel component. Bootstrapped standard errors using 500 replicates were also calculated as a robustness check and yield very similar results.

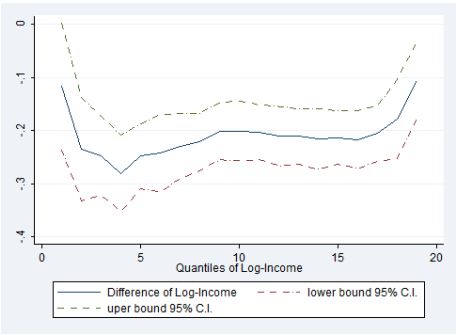
[25] The effect is estimated for 20 quantiles.

Figure 2.2: Average Treatment Effect on Quantiles

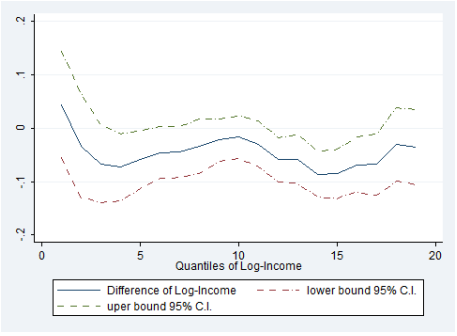
(a) Difference in percentiles



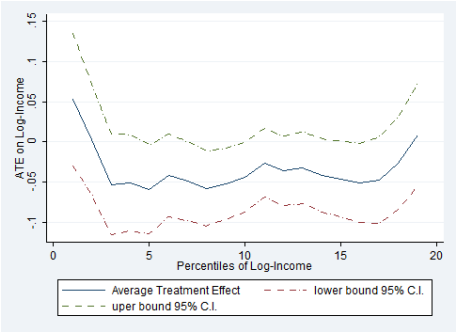
(b) Common support



(c) Re-weighting



(d) RIF-regression on percentiles



Notes:
(i) This graphs show the effect of the interview mode on the percentiles (20 percentiles were used) over the whole income distribution. Panel a) displays the simple difference of the percentiles, Panel b) shows the differences for the balanced sample using the coarsened exact matching technique explained above, Panel c) shows the results for the exact k to k matching within a bin of the matching procedure, and Panel d) shows the effect on the percentiles using RIF-regressions.
(ii) 95%-confidence intervals are provided using bootstrapping standard errors (Panel a) to c)) and the delta method (Panel d)).
(iii) *Source:* EU-SILC 07/08.

These graphs are based on the full sample in Panel a), the balanced (using the coarsened exact matching procedure outlined above) observations in Panel b), and the re-weighted k to k matching (same matching procedure as before) in Panel c). The u-shape means that the percentiles are lower for households interviewed with CAPI but less so (and in the more controlled version even higher) at the extremes of the distribution. This implies a higher spread of income in households interviewed with CAPI. Our understanding is, that there are two mechanism that play an important role here. First, it is easier to lie over the phone, hence biasing reported income at the extremes of the distribution towards the mean; and second, households at the tail of the distribution are especially hard to interview, and thus are mostly interviewed via CAPI (whilst in a CATI survey these households, however important they are for a correct estimate, do not take part at all).
Panel d) of figure 2.2 shows the ATE on the percentiles using RIF-regression estimates with the full set of control variables. We see again that it follows a u-shape. The effect of the interview mode on the percentiles is closer to zero, since the model controls for a wide range of characteristics, and is negative in most parts of the income distribution. In a distribution that is skewed to the right this coupled with the positive effect for the highest income percentiles translates once again to a more compressed income distribution for households interviewed with CATI. Note that the coverage effect, i.e. that households at the extremes of the distribution are only covered with CAPI interviews, is not accounted for in this estimation.

2.6 Conclusion

In this paper we exploited the rare situation of having panel data in combination with a change of the interview mode, which occurred in the 2008 wave of the Austrian EU-SILC data for some - but not all - panel households. The quasi-experimental nature of these data allowed us to estimate causal effects of this change of the inter-

view mode. We have studied the effect of the interview mode (CAPI versus CATI) on item non-response and the level as well as the distribution of household income. First, we find descriptive evidence that CATI compresses the income distribution by leading to less coverage in the final sample via higher unit-nonresponse especially at the tails of the distribution. Second, controlling for a rich set of covariates from the baseline survey, we find that the change from CAPI to CATI has increased item non-response significantly in statistical as well as real terms. This result is robust over all the parametric, semi-parametric and non-parametric methods - which allow for a huge degree of flexibility - we applied. Every researcher pursuing answers to economic questions with the evaluation of survey data should thus be concerned with the interviewing mode of data collection and the follow up imputations. One has to keep in mind that all missing values are usually imputed in various ways or, even more severe, dropped from the analysis altogether. Again CATI tends to compress the observed income distribution - in this case via item nonresponse especially at the tails. Third, we find that households which are interviewed by CATI on average report higher income values. This effect is stronger for lower income than for higher income. In general the income distribution is skewed implying that most part of the distribution lies below the mean. Compared to CAPI the effect of CATI therefore is a mean-reverting one, again compressing the income distribution. Ultimately we find that this level effect accumulates to a severe bias with regard to income inequality measured by the Gini coefficient. We conduct RIF-Regressions to compare the effect of the introduction of CATI on the unconditional distribution of income and find a highly significant effect which reduces the Gini coefficient by around 10%.

Given these results in terms of the effects of the mode of on income distribution we advocate the careful use of survey data taking into account different interview modes and being aware of the possibility that differences between countries and within countries over time might be due to yet another candidate, namely the inter-

view mode. Commonly used rankings of countries by Gini coefficients (see OECD (2011) or table 2.1) of the income distribution might - at least to some extent - be an artifact of different survey techniques such as the interview mode than true differences in income inequality.

2.A Technical Notes

Causality of the Interview Mode Effect

Let us assume N individuals which are indexed by $i = 1, ..., N$ and randomly drawn from a large population. Each individual can be characterized by a pair of potential responses, R_{0i} for the outcome without treatment and R_{1i} for the outcome under the treatment. In addition, each individual has a vector of characteristics, referred to as covariates, X_i. The problem at hand can be interpreted as a quasi-experiment,

$$\begin{aligned} \underbrace{\mathbb{E}\left[R_i|M_i = 1\right] - \mathbb{E}\left[R_i|M_i = 0\right]}_{\text{Difference in response}} &= \underbrace{\mathbb{E}\left[R_{1i}|M_i = 1\right] - \mathbb{E}\left[R_{0i}|M_i = 1\right]}_{\text{Average effect of CAPI on CAPI-Responses}} \\ &+ \underbrace{\mathbb{E}\left[R_{0i}|M_i = 1\right] - \mathbb{E}\left[R_{0i}|M_i = 0\right]}_{\text{Selection bias due to mode assignment}}, \end{aligned} \tag{2.A.1}$$

where R_i is the response behaviour, i.e. some outcome variable (in this analysis it is the item non-response behaviour and the reported income), and M_i is the mode the respondent is confronted with, i.e. CAPI or CATI. Equation 2.A.1 shows that the difference in the means of some outcome variable between the CAPI- and CATI-Respondents consists of an average treatment effect on the treated, i.e. the effect of CAPI (1 signifies CAPI in the subscripts whereas 0 signifies CATI) on those interviewed with mode CAPI. But what is added to the treatment effect is a selection bias resulting from the fact that respondents who are selected to CAPI might differ from those who are selected to CATI. If the mode would be assigned randomly, the selection bias would be zero. Formally we can interchange $\mathbb{E}\left[R_{0i}|M_i = 1\right]$ and $\mathbb{E}\left[R_{0i}|M_i = 0\right]$ and could rearrange equation 2.A.1 to

$$\mathbb{E}\left[R_i|M_i = 1\right] - \mathbb{E}\left[R_i|M_i = 0\right] = \mathbb{E}\left[R_{1i}|M_i = 1\right] - \mathbb{E}\left[R_{0i}|M_i = 1\right] = \mathbb{E}\left[R_{1i} - R_{0i}\right].$$

As we can not assume random assignment of the mode we need to control for

selection in order to to justify a causal interpretation of estimated effects. Under the Conditional Independence Assumption (CIA) also refered to as unconfoundedness assumption, given as

$$\{R_{0i}, R_{1i}\} \perp\!\!\!\perp M_i | X_i, \tag{2.A.2}$$

selection bias disappears. Under this assumption differences in averages across mode groups have a causal interpretation, because the assignment of the mode conditional on selection variables X_i is at random. The remaining question is how to select the vector of covariates X_i. This set of controls needs to be chosen in a certain way to reasonably ensure the CIA but without introducing bad control bias, which is another form of selection bias (see e.g. Angrist and Pischke (2009)). Good controls are in our case variables which are themselves not a (possible) outcome of the mode. A safe way to prevent the danger of using controls which could induce selection bias is to use controls which are themselves already fixed at the time of mode selection or strictly exogenous variables in relation to the mode (Imbens (2009)). In our case all characteristics of a certain household i gathered already before the interview (in 2008) are natural candidates (2007 wave variables).

Effect on the Income Distribution - Recentered Influence Functions

The above summarized approach, however, yields only a causal effect on the mean of the outcome variable. Especially, since we are interested in the effect on the income distribution, we turn to another approach in the literature. After evaluating in a naive way simply the differences of various statistics of the income distribution, we follow the approach suggested by Fortin, Lemieux, and Firpo (2009). To summarize, the Influence Function can be used in order to estimate the ATE of a a bivariate variable (i.e. CAPI vs. CATI) on any distributional statistic $\nu(F(Y))$, where $F(Y)$ is the unconditional distribution of household income. A linear approximation for

$\nu(F(Y))$ is given by

$$\nu(F) \approx \nu(F^0) + \int IF(Y;\nu,F^0)dF(Y),$$

where F^0 is the baseline distribution, which is in our case the distribution given CATI interviews. This approximation exists for "smooth" statistics ν, by Fréechet differentiation and the Riesz representation theorem. Given the CIA (see 2.A.2) and our treatment indicator M_i we have

$$E[IF(Y^i)|X] = E[IF(Y)|X, M_i = i],$$

which justifies estimation of the following linear model, with appropriate interactions and powers of the covariate vector X:

$$E[IF(Y)|X, M_i] = X \cdot \beta^0 + M_i \cdot X \cdot \beta^1 \tag{2.A.3}$$

The effect of the the mode M_i on the population ν is thus given by

$$\nu(F^1) - \nu(F^0) = E[X] \cdot \beta^1. \tag{2.A.4}$$

With this procedure we estimate the effect on the Gini coefficient, the $P90/P10$-percentile ratio as well as the poverty rate (two of the most widely published estimates from EU-SILC), i.e. the share of the household population with less than 60% of the median income. The Recentered Influence Function (RIF) for the Gini coefficient (see e.g. Firpo, Fortin, and Lemieux (2007) page 24f.) can be written as,

$$RIF(y,\nu^{GI},F^0) = 1 + \frac{2}{\mu^2}R(F_y)y - \frac{2}{\mu}\left[y(1-p(y)) + GL(p(y),F_y)\right],$$

where $R(F_y)$ is the integral of the Generalized Lorenz curve with the ordinates $GL(\bullet)$, and $p(y)$ is the density. Additionally the RIF for the poverty rate $\nu^{PR}(F) =$

$F\left(0.6 \cdot F^{-1}(0.5)\right)$, where $med := F^{-1}(0.5)$ is the median for the distribution given by F and $0.6 \cdot F^{-1}(0.5) = 0.6 \cdot med$ is the poverty ceiling, is given by

$$RIF(y; \nu^{PR}, F^0) = \mathbb{1}(y \leq 0.6 \cdot med) - \frac{0.6 \cdot f(0.6 \cdot med)}{f(med)} \cdot \mathbb{1}(y \leq med).$$

$\mathbb{1}$ is an indicator function that is one if a household is poor.

Finally, we can calculate the RIF of the the $P90/P10$-percentile ratio given as $\nu^{RA} = \frac{Q^{90}}{Q^{10}}$ where Q^i is the i's percentile of the income distribution by

$$RIF(y; \nu^{RA}, F^0) = \nu + \nu * \left[\frac{\mathbb{1}(y \leq Q^{10})}{Q^{10} * f(Q^{10})} - \frac{\mathbb{1}(y \leq Q^{90})}{Q^{90} * f(Q^{90})} \right].$$

Chapter 3

Net Wealth across the Euro Area - Why Household Structure Matters and How to Control for it[1]

[1]This chapter comprises a paper that is joint work with Pirmin Fessler and Esther Segalla.

3.1 Introduction

The most important stylized facts about the distribution of net wealth are well known and well documented. Net wealth is distributed more unequally than income. The distribution of inherited wealth is more unequally distributed than wealth in general; and the differences in wealth distribution across developed countries are large (Davies and Shorrocks (2000)). Despite the fact that net wealth and its distribution - resources and their allocation - is at the heart of economics, theoretical models struggle to reproduce the observed skewness of the distribution and empirical analyses suffer from the lack of comparable data across countries.

While consumption smoothing or intergenerational transmission are well covered by theoretical models, the strong differences across countries and the amount of wealth concentration are still a major obstacle for modelling (Cagetti and DeNardi (2005)). It is unclear how much of the empirically observed differences might be accountable (i) to differences in methodology of the underlying data production process, (ii) to institutional differences such as pension systems, taxation or welfare programs, (iii) to historical differences such as land reform or war, or (iv) to differences in the structure and behaviour of economic agents as households or individuals.

International comparisons of wealth distributions reliant on post-harmonized data and definitions originally collected through a broad variety of different methodologies are known to lead to differences in the observed wealth distributions and therefore might disguise or exaggerate the true differences. The Luxembourg Wealth Study is the main example for such an harmonization endeavour and the differences in estimates for the US net wealth distribution between the Panel Study of Income Dynamics (PSID) and Survey of Consumer Finances (SCF) illustrate the importance of methodological differences in data production (see Sierminska, Brandolini, and Smeeding (2006) or Bover (2005)).

The Household Finance and Consumption Survey (HFCS) of the Eurosystem is the first project to a priori harmonize the data production process of wealth surveys

across the 17 member countries[2] of the Eurosystem. Hence, it delivers the first large dataset allowing for reasonable cross-country comparison of net wealth among a large number of developed countries (ECB (2013a) and ECB (2013b)).

Acknowledging that there might remain some methodological differences - of importance especially at the tails of the wealth distribution - in this a priori harmonized cross country dataset. Also differences due to institutions, history and behaviour need to be elucidated in future research. In this study, however, we concentrate on differences in the wealth distribution due to variation in the form of the unit of observation - the household - and its different structure across countries. Upon publication of the HFCS data researchers as well as the public at large focussed on the comparability issues due to the unit of observation in the results and hence it is one of the major issues that have to be tackled. Net wealth is usually surveyed for households, not individuals, and cannot be partitioned to household members without further assumptions. This convention might be useful for different reasons. First, we might be interested in the possession of or access to resources instead of the legal ownership of an individual inside a household.[3] Second, the control over some assets inside a household might differ from the ownership structure.[4] Third, it might be impossible to allocate all assets inside a household to individuals. For an economic interpretation of cross-country comparisons, however, the differences in household structure are important. Imagine two countries A and B, both populated with three individuals each endowed with wealth 1 but in country A they form 3 and in country B only two households. In country A we would observe perfect wealth equality measured at the household level while in country B the distribution would be unequal. Additional to the number of household members also age profiles and

[2]The first wave of the HFCS includes only 15 of the 17 member countries, excluding Ireland and Estonia.

[3]Children who live in a house might benefit as much from the house as their parents who own the house and who even have some legal duty to give them shelter.

[4]As for example in the case of subsidized assets which are subsidized only for one individual and therefore often refer to children or elderly which might not be in control of the assets. However, the role of intra-household power structures can not be analyzed with our data.

gender composition might play an important role for the basic form of the household as unit of observation. Previous work on international comparisons either treated households as homogeneous across countries, applied some equalizing such as dividing net wealth by the number of household members or its square root, or compared conditional on certain age bands (Banks, Blundell, and Smith (2003)). However, given the large number of different countries and the heterogeneity of household structures observed in a dataset like the HFCS the topic of household structure deserves more attention. This is true for all household level variables but especially relevant for variables such as net wealth which are ultimately accumulated at the individual level and all main sources such as income and inheritance are phenomena at the individual level.

We estimate counterfactual distributions by imposing a common (the average) household structure on all countries. We do this by using non-parametric reweighting and examine to what extent differences in the unconditional distributions of net wealth between euro area countries are due to differences in household structure.[5] We further show the effect of these differences on the most common inequality measures and find that imposing a common household structure has strong effects on both, the full unconditional distributions as well as its mappings to different inequality measures. Our paper is closest to the recent papers by Bover (2010) who estimates counterfactual US net wealth distributions relying on the Spanish household structure and Peichl, Pestel, and Schneider (2012) who examine the effect of changes in the German household structure on the income distribution. It differs in using a large dataset of a priori harmonized cross country data and by conducting a more flexible non-parametric approach with regard to the definition of household types. We find that for the median 50% of the differences to the euro area estimate are explained for Austria, 15% for Germany, 25% for Italy, 14% for Spain and 38% for

[5] As a benchmark we divide household wealth and multiply the household weight by the number of household members to produce a measure of the individual wealth distribution under the assumption of equal intra household division of wealth. The results thereof are left out due to space constraints but can be obtained from the authors.

Malta. For others as Belgium, France, Greece, Luxembourg, Portugal, Slovenia and Slovakia household structure masks the differences to the euro area median. The impact on the mean and percentile ratios is similarly strong. This analysis also confirms the finding of Bover (2010) that the effect on the Gini coefficient is somewhat less pronounced, but might mask relevant information by being a net effect of different accumulated effects along the distribution. Country rankings based on almost all of these measures are severely affected by the household structure alluding to the need for cautious interpretation when dealing with such rankings. As household structure matters a lot we argue to use more flexible controls than the standard household size as well as age, age squared and gender of a more or less arbitrarily chosen representative member of the household. We provide such a set of controls to account for household type fixed effects which are based on the number of household members as well as possible combinations of age and gender of all of them.[6]

The rest of this paper is structured as follows. Section 3.2.1 introduces the dataset and its main properties and differences to existing data. In section 3.2.2 we discuss the implemented non-parametric estimation strategy to generate the relevant counterfactual distributions. The main part of the paper, section 3.3, discusses the results reached from our empirical exercise and is split in two parts. First we discuss the differences due to household structure on the full distributions, on wealth inequality measurement and on the ownership of certain assets in section 3.3.1. Second, in section 3.3.2 we argue for more flexible controls for household structure in regression analyses than are usually used and provide such controls based on our re-weighting approach. Section 3.4 discusses the results and concludes.

[6]These controls are a non-parametric and very flexible way to control for household structure and also can be used additionally to the information of a household reference person to further control inside the defined household types. Furthermore we provide the weights to standardize household structure across the euro area to further analyse the contribution of household structure of any distribution and any statistic with regard to the set of all HFCS variables using the approach proposed in this paper. The dataset containing dummies for the household types, the weights and variables to merge the dataset to the HFCS as well as Information on the household types will be made available online once the paper is published.

3.2 Data and Estimation Strategy

3.2.1 Data

We use the first wave of the HFCS, a euro area wide project to gather data on real and financial assets and liabilities of euro area households. In the first wave HFCS (2010) more than 62,000 households across the euro area were interviewed, leading to a micro dataset which is not only arguably representative at the euro area but also at the level of each member state. While the goal was maximized harmonization in terms of questionnaire, interview method, editing, multiple imputation and all other aspects of data production, national differences were accounted for by adapting to country specifics where necessary. Table 3.1 shows basic information such as fieldwork period, net sample size, response rates, oversampling of the wealthy as well as survey mode for all HFCS 2010 country surveys. Our main variable to

Table 3.1: General Information on the HFCS Wave 1

	Fieldwork period	Net sample Size	Response Rate (%)	Over-sampling	Survey Mode[i]
Austria	2010/2011	2,380	55.7	No	CAPI
Belgium	2010	2,364	21.8	Yes	CAPI
Cyprus	2010	1,237	31.4	Yes	CAPI(12%)/PAPI(88%)
Germany	2010/2011	3,565	18.7	Yes	CAPI
Spain	2008/2009	6,197	56.7	Yes	CAPI
Finland	2009/2010	10,989	82.2	Yes	CAPI(3%)/CATI(97%)
France	2009/2010	15,006	69.0	Yes	CAPI
Greece	2009	2,971	47.2	Yes	CAPI
Italy	2010	7,951	52.1	No	CAPI(85%)/PAPI(15%)
Luxembourg	2010/2011	950	20.0	Yes	CAPI
Malta	2010/2011	843	29.9	No	CAPI(81%)/PAPI(19%)
Netherlands	2010	1,301	57.5	No	CAWI
Portugal	2010	4,404	64.1	Yes	CAPI
Slovakia	2010	2,057	n.a.	No	CAPI
Slovenia	2010	343	36.4	No	CAPI

Notes:
(i) Computer-assisted personal interview (CAPI); paper based personal interview (PAPI); computer-assisted telephone interview (CATI); computer-assisted web interview (CAWI).
(ii) *Source:* Eurosystem HFCS 2010.

illustrate the importance of household structure is net wealth. While in general

household structure is important for all variables at the household level and might even be important for variables at the individual level it is crucial for net wealth. Net wealth is ultimately accumulated via savings from income or inheritance (and gifts), which are both (all) phenomena at the level of the individual. At the same time consumption and therefore savings as well as probabilities of inheritances received at the household level are highly dependent on the size of the household and the individuals age profile. Household net wealth is defined as real assets plus financial assets minus debt of a household. Real assets consist of the main residence, other real estate property, investments in self-employed businesses, vehicles and other valuables. Financial assets are current accounts, savings deposits, mutual funds, bonds, stocks, money owed to the household and other financial assets. Debt consists of collateralized debt as well as uncollateralized debt including credit card debt and overdrafts (see table 3.2).[7] See the ECB methodological report for detailed

Table 3.2: HFCS Household Balance Sheet

+ Real Assets	Real Estate	Household main Residence
		Other real estate property
	Business	Self-employed
	Other	Vehicles
		Valuables
+ Financial Assets	Deposits	Sight accounts
		Savings accounts
	Shares	
	Bonds	
	Mutual Funds	
	Business	Non-self-employed
	Money owed to households as private loans	
	Private Pension Plans	
	Other	Options, futures, royalities, etc.
- Debt	Collateralized	Household main Residence
		Other real estate property
	Non-Collateralized	Credit Cards/Overdraft
		Other loans
= Net Wealth		

Notes:
(i) *Source:* Eurosystem HFCS 2010.

[7] The HFCS does not collect information on the outstanding amount for a leasing contract, and hence these form of liability is not included in the debt level.

information on all HFCS variables (ECB (2013a)). Throughout our paper we use complex survey weights as well as the multiple imputations provided by the HFCS.

3.2.2 Estimation Strategy

Reweighting

We observe cross-sections with draws from the country-distribution functions P^c of the vector (W, H) consisting of net wealth W and household structure H. We want to identify and estimate differences in the distribution of wealth $P(W)$, or differences in statistics with regard to the distribution of wealth, $\nu(P(W))$, which are due to differences in household structure H between countries $c \in C$. Whereas in general many statistics ν might be of interest we focus on percentiles, certain percentile ratios, the Gini coefficient, as well as extensive and intensive margins of components of net wealth.

Let $P^{ea}(W, H)$ denote the overall distribution of (W, H) in the 14 countries surveyed in the first HFCS wave which include all variables necessary for this analysis[8] (henceforth called the euro area), and $P^c(W, H)$ the particular distributions for country $c \in C$. We then want to identify the counterfactual distribution $P^c_{ea}(W)$, in which the differences in the distribution of wealth W in a certain country c that are due to differences in the household structure H between the particular country and the euro area as a whole are eliminated. The differences between $P^c(W)$ and $P^c_{ea}(W)$, as well as differences between measures $\nu(P^c(W))$ and $\nu(P^c_{ea}(W))$ are the differences which are due to household structure. Formally we can write the counterfactual of interest of country c as,

$$P^c_{ea}(W) := \int_H P^c(W, H) dP^{ea}(H). \tag{3.2.1}$$

[8] We had to exclude Cyprus because it collects gender only for one household member, see ECB (2013a) page 83.

We can rewrite the counterfactual distribution in equation 3.2.1[9]

$$P^c_{ea}(W) := \int_H P^c(W,H)\Psi_H(H)dP^c(H), \tag{3.2.2}$$

where the re-weighting function Ψ_H is defined as

$$\Psi_H := \frac{P^{ea}(H)}{P^c(H)} \tag{3.2.3}$$

An overview of similar techniques which emerged after the contribution of DiNardo, Fortin, and Lemieux (1996) can be found in Fortin, Lemieux, and Firpo (2011). Instead of using a re-weighting approach we could also directly estimate the counterfactual distributions $P^c_{ea}(W)$ as proposed in Chernozhukov, Fernandez-Val, and Melly (2009). Another possibility recently proposed is the influence function regression approach by Fortin, Lemieux, and Firpo (2009) which is based on the first order approximation of ν, as a function of P^c around P^{ea}. Note that given the counterfactual distributions $P^c_{ea}(W)$, we can decompose the differences between any measure of the observed distributions $\nu(P^{ea}(W))$ and $\nu(P^c(W))$ in the following way:

$$\nu(P^{ea}(W)) - \nu(P^c(W)) = [\nu(P^{ea}(W)) - \nu(P^c_{ea}(W))] + [\nu(P^c_{ea}(W)) - \nu(P^c(W))], \tag{3.2.4}$$

where the first term reflects the differences remaining after controlling for differences in household structures across countries and the second covers the differences explained by differences in household structure.

We plot the respective quantile functions $Q^c(W)$, $Q^{ea}(W)$, the counterfactual quantile functions $Q^c_{ea}(W)$ as well as the resulting observed differences $d_{obs} := Q^{ea}(W) - Q^c(W)$ and the resulting differences after re-weighting $d_{rew} := Q^{ea}(W) - Q^c_{ea}(W)$. Note that analogous to equation 3.2.4 the relation $\frac{d_{obs}-d_{rew}}{d_{obs}}$ is a measure of the ob-

[9]Note that we could formulate the counterfactual distribution also as $P^c_{ea}(W \leq w) = E^1[\mathbb{1}(W \leq w)\Psi_H]$ (as e.g. in Bover (2010))

served differences which can be explained by differences in household structure at every quantile u.

Household Structure

We use the overall (or weighted average) household structure H^{ea} which refers to the union $\bigcup_{c \in C} H^c$ of the collection of country level household types $\{H^c : c \in C\}$ as a reference. First it includes by definition all household structures observed in all countries. Second it minimizes the overall need for re-weighting as it is the weighted average of country level household structures. However we have to assure that we choose a set of household structures that is large enough to flexibly control for the differences in household structures, i.e. helps us compare "apples to apples" but which is at the same time small enough to ensure enough overlap between the countries. In both extreme cases of a very small, where only one household structure is assumed, or very large number of household structure cells in which every type of household only exists in one certain country, re-weighting to the overall household structure would be without any effect.

We define household types by all possible combinations of 4 age categories and gender for each individual (member) up to 4 individuals in each household. We are not taking (i) a particular order of individuals or (ii) gender for individuals aged 15 or below into account. Households with 5 or more members are treated as 4 person households and sorted with regard to the first 4 members, the financially knowledgeable person (respondent) and the next 3 persons sorted by descending age. This results in 329 possible household types of which 249 are observed at least once in the euro area. A detailed description of the construction of these cells can be found in appendix 3.B. [10]

[10] We tried several possible definitions to construct household types based on gender as well as age of individuals living in a certain households. Results are robust across a great variety of combinations. Setting the limit at this 329 possible types is a compromise between ensuring enough flexibility to compare "apples to apples" as well as having enough common support between countries given a certain definition of household types to generate meaningful counterfactuals.

Table 3.3 shows the top 30 household types that occur most often in the euro area and include more than 90% of the observed euro area household population and between 82% and 95% in each individual country. It also shows the top 10 categories for each of the countries which all are subsets of the euro area top 30 household types and include between 48% and 72% of households in all countries. However, already large differences can be seen in the occurrence of the top types. The household type code describes the composition of the household. Two numbers for each individual in a household, where the first refers to age category $((1 = [-, 15]; 2 = [16, 34]; 3 = [35, 64]; 4 = [65, +]))$ and the second refers to gender for all individuals aged 16+ $(1 = male; 2 = female; 3 = below\ 16)$. The code is sorted by individual age. The most common household type 3132 is therefore a two person household (4 digits), consisting of a man aged between 35 and 64 (31) and a woman aged between 35 and 64 (32). This category consists of 10.2% of all euro area households. Around 9.5% of households are single households consisting of a woman aged 65+, which is the second most common household type. All other household type codes are to be read analogous. As one can see by the distributions of the top ten household categories in each country among the euro area top 30 categories, certain types which are rather relevant in southern or eastern countries (e.g. type 21223132 make up more than 4% of households in Spain, Greece, Italy, Malta and Slovakia) and are not even in the top ten in northern countries. Also the typical single households are rather different. Whereas middle age singles (types 31 and 32) are very typical for e.g. Austria, Germany, Finland or the Netherlands they are much less important in Spain, Greece or Portugal.

Table 3.3: Occurrence of Top 10 Country wise Household Types among the Euro Area Top 30 Household Types in % of the Respective Household Populations

Top 30 EA	HH Size	Categories	EA	AT	BE	DE	ES	FI	FR	GR	IT	LU	MT	NL	PT	SL	SK
1	2	3132	10.2	13.3	10.2	12.7	7.6	12.8	11.6	6.4	6.7	7.2	7.4	9.3	9.0	10.7	8.0
2	1	42	9.5	8.8	7.7	10.2	7.5	10.4	11.3	6.7	9.6	5.7	6.8	8.2	8.5	11.4	6.1
3	2	4142	9.1	7.3	9.4	10.1	8.1	7.6	7.0	8.8	11.2	7.9	8.4	9.1	9.3	5.2	4.9
4	1	31	7.0	7.7	7.9	10.0	3.7	9.2	6.7		4.1	9.8	4.6	12.3		8.8	3.5
5	1	32	5.7	10.0	6.9	5.7		8.0	6.6	3.6	5.2	5.6	3.8	8.1	2.8	3.8	5.2
6	4	13133132	5.6	4.0	4.6	4.7	6.1	4.9	6.0	7.4	6.8	7.6	8.0	5.8	5.1	7.0	
7	1	41	3.6	3.4	5.1	4.3		3.3	4.0		3.4	3.1		2.7			3.6
8	3	133132	3.4				5.2			3.3	4.7		4.3		4.7		
9	1	21	3.3	4.7	3.4	5.3		5.1	3.8								
10	2	2122	3.2	3.3	2.9		3.1	6.2	4.6			3.5		4.9			
11	3	213132	3.0				3.3			4.6	3.8	3.3	7.1		5.2	4.3	4.7
12	2	3241	2.8			3.4				4.0				2.7			
13	1	22	2.6	4.1		4.2		4.0				3.0		3.1			
14	4	21223132	2.4				4.3			4.1	4.3		5.2				5.5
15	3	223132	2.1				3.2						3.8		3.7	3.5	3.6
16	3	132122	2.0						2.8								
17	4	13132122	1.7		2.8												
18	4	13223132	1.6												2.7		
19	4	13213132	1.6												3.0		
20	4	13132231	1.4							3.4							
21	4	21213132	1.4													4.4	3.5
22	3	132231	1.1														
23	2	3142	1.0														
24	4	22223132	1.0													3.3	
25	2	2132	0.9														
26	2	2231	0.9														
27	2	1332	0.8														
28	2	2232	0.7														
29	3	131332	0.6														
30	2	3242	0.5														
Sum of Countrywise Top 10				66.6	60.9	70.5	52.3	71.5	64.4	52.3	59.8	56.7	59.6	66.1	54.0	62.3	48.7
Sum of Euroarea Top 30			90.6	91.7	90.2	94.1	83.7	94.7	92.6	84.0	89.2	89.8	87.1	92.8	82.9	85.6	82.7

Notes:

(i) The household type codes show the age an gender of each household member, i.e. the first character reveals age as 1 = $[-, 15]$; 2 = $[16, 34]$; 3 = $[35, 64]$; 4 = $[65, +]$ and the second refers to gender by 1 = $male$; 2 = $female$; 3 = $below$ 16. Only the Top 10 household types for each country are listed with their relative percentage share inside the country.

(ii) Common support exists for all euroarea top 30 types but household type 131332, which is not observed in Slovenia.

(iii) *Source:* Eurosystem HFCS 2010.

Figures 3.1, 3.2 and 3.3 show the occurrence of household types across countries and the distributions and re-weighted distributions of the euro area top 100 most populated household types respectively. The distributions are sorted by the occurrence of household types in the euro area. Figure 3.1 illustrates the differences in the variety of household types observed across countries. The larger the sample size, the higher the probability that also sparsely populated household types are drawn into the samples. Also certain sampling schemes or interviewer modes (such as CAWI[11] in the Netherlands) might lead to a smaller variety of household types. As can be seen in figure 3.2 all the top household types are relatively common across the euro area. The top 30 euro area types include at least about 83% in Portugal and at most about 95% in Finland and Germany. Figure 3.3 shows the distributions when the data is re-weighted to the euro area average as described in section 2.4. As can be seen the common support between countries is large but does not include all household types in all country samples. The country variation is completely eliminated for the first few types and in general strongly reduced but small variation remains (compare figures 3.2 and 3.3). This is because for some countries certain household types are not observed at all, implying that they can not be re-weighted which translates to the remaining extrapolation outside the common support (between countries) as discussed in section 2.4.

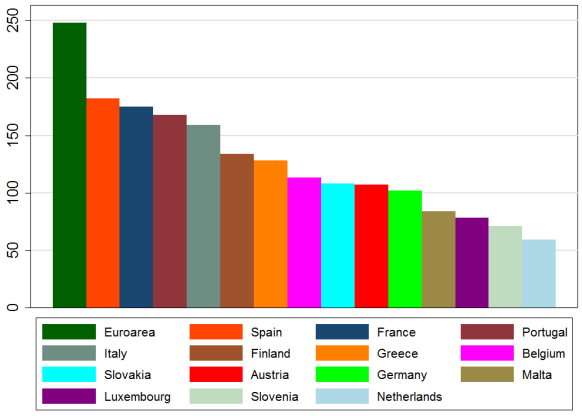
3.3 Results

3.3.1 Why Household Structure Matters

Table 3.4 includes all the main results of the paper. It shows estimates of the mean, selected percentiles as well as selected inequality measures estimated for the observed distributions, $P^{ea}(W)$ and $P^{c}(W)$ as well as estimated for the counterfactual distributions $P^{c}_{ea}(W)$, where euro area household structure is imposed using

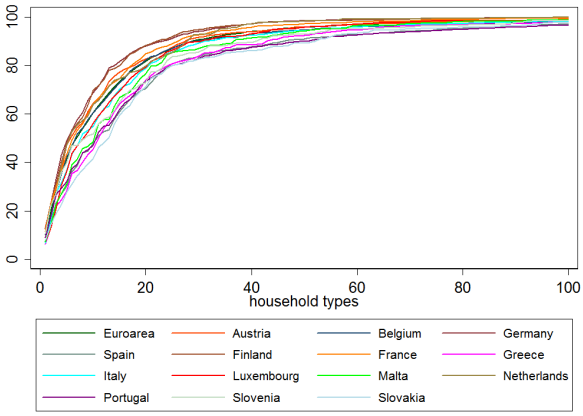
[11] Computer-assisted web interview.

Figure 3.1: Number of Household Types in the Euro Area



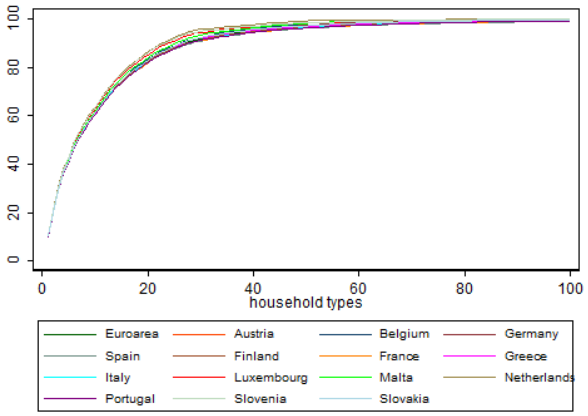
Notes:
(i) This graph shows the number of different household types observed as described in section 2.4.
(ii) *Source:* Eurosystem HFCS 2010.

Figure 3.2: Distribution of Household Types in the Euro Area



Notes:
(i) This graph shows the distributions of different household types observed as described in section 2.4.
(ii) *Source:* Eurosystem HFCS 2010.

Figure 3.3: Re-weighted Distribution of Household Types in the Euro Area



Notes:
(i) This graph shows the distributions of different household types observed as described in section 2.4 re-weighted to the euro area distribution.
(ii) *Source:* Eurosystem HFCS 2010.

non-parametric re-weighting as described in section 2.4. Differences between any statistic $\nu(P^c(W))$ and $\nu(P^c_{ea}(W))$ are the differences which can be explained by differences in household structure.

Mean

The euro area mean is 230 thousand Euro. The means of the euro area countries range from as low as 80 thousand Euro (Slovakia) to as high as 710 thousand Euro (Luxembourg). The impact of imposing the euro area household structure to all countries is large for many of them. Note that as the household structure of larger countries is more important for the euro area they are in general also more stable with regard to this type of re-weighting. There are different groups of countries. Austria, Belgium, France and Luxembourg already have an above euro area mean of net wealth but move even further away through re-weighting. Household structure in these countries is dampening means with regard to the average euro area household structure. Spain, Italy and Malta also have an above euro area mean but move closer to the euro area. In the case of Spain around 23% of the difference to the euro area

is explained only by household structure. For Italy this value is 47% and for Malta even 48%. Germany, Finland and the Netherlands have means below the euro area mean and move up towards the euro area mean. Around 43% of the difference to the euro area mean is explained for Germany, 39% also for Finland and about 32% for the Netherlands. Greece, Portugal, Slovenia and Slovakia all have below euro area means and their means decrease even more with re-weighting, implying that the household structure exaggerates their wealth in euro area comparisons.

Percentiles

The counterfactual percentiles in the second part of table 3.4 illustrates the variation of the importance of household structure between countries and along the net wealth distribution. In general differences between observed and counterfactual distributions are relatively stronger at the bottom than at the top, but effects are considerable all along the distribution. For example Finland's median is 86 thousand Euro and well below the euro area median of 109 thousand Euro whereas its re-weighted median 112 thousand Euro already lies above the euro area median. Also the Netherlands changes its position from below to above the median. For other countries large parts of the differences to euro area medians are explained by household structure (50% for Austria, 15% for Germany, 14% for Spain, 25% for Italy and 38% for Malta) others again move further away. Note especially that for the median it is not the same countries as for the mean where the gap between the euro area and their observed distribution is smaller or larger. Austria moves away for the mean but gets closer for the median. Belgium, France, Greece, Luxembourg, Portugal, Slovenia and Slovakia move away from the euro area for both. Germany, Spain, Italy and Malta get closer for both. Finland and the Netherlands even change position with regard to the median while both get closer to the euro area in case of the mean. Note, that the patterns differ along the distribution for many countries such as Austria, France, Slovenia and Slovakia which move closer

for some areas and further away for others. Finland, Greece, The Netherlands and Portugal even switch position with regard to the euro area in some areas. Belgium and Luxembourg always move away from the euro area whereas Germany and Italy are always getting closer.

Inequality Measures

The third part of table 3.4 shows the impact of household structure on selected percentile ratios as well as the Gini coefficient of net wealth. Again large parts of the differences to the euro area measure can be explained by household structure. For the most robust measure $P75/P25$, all countries but Belgium, Luxembourg (further away) and the Netherlands (switches position) get closer to the euro area measure. For Finland which has a $P75/P25$ ratio of 34 as opposed to only 17 for the euro area 95% of the difference can be explained by household structure. About 27% for Germany, 48% for France and 53% for Italy. Again effects are large for many countries, and again they are different for different measures. While $P75/P25$ gets closer to the euro area $P90/P10$ moves further away from the euro area for France even though in both cases french inequality is reduced by re-weighting. The Gini coefficient seems less sensitive to household structure and about half of the countries move closer to the euro area Gini and half of them move further away. However, as Bover (2010) mentioned, the Gini masks relevant information by being a net effect of different accumulated effects along the distribution.

The main driving force behind these results are the differences in household size. Values of the re-weighted net wealth distributions for southern and eastern European countries in general are lower than the observed values because of their above average household size and the ones of northern European countries are higher because of their below average household size. Furthermore, the size of the impact of imposing the common household structure along the distribution of net wealth also

Table 3.4: Effects of Household Structure Differences Across Countries (in thousand Euro)

Variable Names	EA	AT	BE	DE	ES	FI	FR	GR	IT	LU	MT	NL	PT	SI	SK
Mean	229.84	265.03	338.65	195.17	291.35	161.53	233.40	147.76	275.20	710.09	365.99	170.24	152.92	148.74	79.66
Counterfactual	.	280.50	343.62	210.00	277.05	188.16	255.68	134.46	254.11	729.93	300.59	189.11	141.73	132.02	73.90
P10	1.19	0.98	2.78	0.06	5.66	-0.57	1.58	2.00	5.00	5.04	16.11	-3.80	1.04	4.22	12.92
Counterfactual	.	1.12	2.97	0.16	3.49	0.11	1.87	0.92	4.00	6.51	7.88	-0.06	0.56	2.97	7.76
P25	15.47	10.31	40.24	6.60	77.87	6.38	9.80	30.00	34.24	59.24	88.54	14.10	18.37	40.84	36.45
Counterfactual	.	11.52	46.11	8.10	64.51	13.92	12.97	20.96	22.32	80.31	66.05	20.04	11.67	29.59	32.03
P50	108.85	76.44	206.25	51.36	182.72	85.75	115.80	101.93	173.50	397.84	215.93	103.56	75.21	100.66	61.18
Counterfactual	.	92.74	213.42	59.95	172.12	111.97	135.00	90.02	157.13	417.07	174.81	124.84	68.13	83.04	56.04
P75	268.35	250.47	417.36	209.82	330.98	220.22	279.10	193.27	321.43	738.13	394.09	259.10	160.13	212.09	98.66
Counterfactual	.	266.78	423.84	225.60	311.96	254.06	300.40	171.53	304.04	741.38	345.56	283.37	150.32	192.06	91.15
P90	504.89	542.16	705.14	442.32	607.68	397.32	511.58	331.78	577.13	1,375.37	693.08	427.64	297.23	317.18	151.86
Counterfactual	.	552.55	725.30	475.43	561.16	441.82	549.17	301.90	537.10	1,392.36	624.46	455.55	286.92	300.71	141.29
P75/P25	17.35	24.34	10.39	31.79	4.25	34.49	28.47	6.44	9.39	12.53	4.45	18.53	8.72	5.24	2.71
Counterfactual	.	23.18	9.23	27.86	4.84	18.25	23.17	8.18	13.62	9.24	5.24	14.33	12.90	6.61	2.85
P90/P50	4.64	7.13	3.42	8.62	3.33	4.63	4.42	3.25	3.33	3.46	3.21	4.13	3.95	3.15	2.48
Counterfactual	.	5.98	3.40	7.93	3.26	3.95	4.07	3.35	3.42	3.34	3.57	3.65	4.21	3.62	2.52
P90/P10	424.01	581.05	253.82	7,371.25	107.64	-692.19	323.64	165.89	115.43	274.28	43.33	-162.21	286.92	75.56	11.77
Counterfactual	.	519.28	244.58	3,155.16	160.65	3,909.91	293.80	329.98	134.28	217.16	79.24	#[iii]	513.40	101.61	18.22
Gini	0.68	0.76	0.61	0.76	0.58	0.66	0.68	0.56	0.61	0.66	0.60	0.65	0.67	0.53	0.45
Counterfactual	.	0.75	0.60	0.75	0.60	0.63	0.67	0.58	0.63	0.65	0.59	0.63	0.68	0.54	0.46

Notes:

(i) This table shows (in thousand Euro) the mean, percentiles, and distributional measures (percentile ratios and Gini coefficient) of net wealth in the euroarea. For each statistic, one can see the estimate based on original weights and the counterfactual estimates using the re-weighted households weights controlling for the differences of the household structure.

(ii) For the euroarea there is no counterfactual estimate by definition, thus cells are denoted with a dot.

(iii) In the Netherlands P10 is zero in implicate 1, hence the P90/P10 quantile ratio cannot be estimated.

(iv)*Source:* Eurosystem HFCS 2010.

depends heavily on the age and gender structures and occurrence of household types. That leads to the result that the relevance of household structure varies considerably along the distribution and has different patterns across different countries. In all (regression) analyses where controls for household structure (number of household members, age and gender) are desirable this strong variation of the importance of household structure with regard to different countries and along their net wealth distributions calls for very flexible controls.

More results of imposing common household structure are shown in the appendix 3.A. See table 3.6 and 3.7 for counterfactuals for the extensive and intensive margins of net wealth components. See also figures 3.4 to 3.17, for country wise comparisons of the countries' net wealth distributions to the euro area distribution, the countries re-weighted distribution, the differences between those as well as a comparison of a ad-hoc individual level net wealth distribution where net wealth is divided by and household weights are multiplied by household members for each household.

3.3.2 How to Control for Household Structure

Most empirical papers use the household size as well as age (age squared) and gender of a more or less arbitrarily selected so-called reference person or householder - such as the reference person according to the Canberra definition or highest income earner - to control for household structure. This approach implies strong functional assumptions about the relationship of household structure and the variable(s) of interest. Furthermore, it ignores age and gender of all other household members in two or more person household, which in all HFCS countries are the majority of households. However, for most household level variables such as net wealth, household income, participation rates in certain assets, transfers, inheritances and gifts, portfolio choice, and many more, age and gender of all household members will be relevant to the households realization of a certain variable. This fact is already

relevant when comparing households within countries but is especially important for cross country comparison if the patterns of household structure are different between countries. While a 3 person household with a reference person aged around 30 living still with her older parents might be relatively common in Spain it is not in Germany, where the three person household with a (female) reference person aged around 30 is more likely to be a couple with a child. Using only household size and a reference persons age and gender information will not differentiate between these household types. The more strongly the occurrence of such household types differs between countries the more explanatory power will be transferred to other variables with cross country differences or country fixed effects, possibly leading to large bias and therefore misleading results.

We argue for taking into account the most relevant characteristics, i.e. age and gender, and possible combinations of all household members when controlling for household structure is desirable. One way to do so is to add a household type fixed effect for most relevant household types (e.g. the top 30 for the euro area as shown in table 3.3).[12] If net wealth is regressed on standard household characteristic controls (a set of household size dummies, gender, age, and age squared of the reference person) and the top 30 set of household type fixed effects 22 of the 30 stay significant and the joint F-test on all of them having zero explanatory power is also rejected. Furthermore, one can regress each of the household type dummies on the standard household characteristics controls (as above) and then regress net wealth on those residuals, which represent the information in the household type dummies which is orthogonal to the standard household characteristics controls. Again the H_0 of the joint F-test that the explanatory power of the residuals (the information orthogonal to the standard controls) is zero is rejected.

[12] We provide such a set of dummy variables resulting from our non-parametric procedure to define a set of different household types as explained in section 2.4 online once the paper is published.

Therefore even with only the Top 30 populated household types in the euro area and including a quadratic age term in the standard controls additional explanatory power of the household type fixed effect can be shown. Naturally this result holds for any standard household structure control not including a quadratic age term or assuming linearity along household size. The explanatory power of course increases if all 249 household types are included. However, while such a large number of household type fixed effect are of no concern at the euro area level using the full HFCS dataset with more than 62,000 observations, it might be problematic for smaller subsets of the data. Note, however, that a small number - like the top 30 for which illustrated their significance as additional controls - might include already a large proportion of households and the subset of household fixed effects used can be chosen according to their occurrence in the data used. For the full HFCS sample all the 249 household types might be appropriate, whereas for the subset of southern countries another subset of household type fixed effects might be in order than for a northern subset of countries. Even though the explanatory power of using the top 30 household types alone is close to using the standard household characteristic controls we recommend using both together. Additionally. household type fixed effects or the re-weights might be used to check robustness of results when for some reasons standard household characteristic controls are preferred.

3.4 Concluding Remarks

In this paper we highlight the importance of household structure for household level analyses. We use non-parametric re-weighting to impose the euro area household structure on all observed euro area countries and examine the extent to which differences in the observed unconditional distributions of net wealth between countries are due to differences in the structure of the household as the unit of observation. We employ the euro area Household Finance and Consumption Survey, the first high

quality a-priori harmonized dataset which allows for such a cross country comparison across 14 euro area countries. We find that household structure plays a major role in explaining differences in net wealth distributions as well as their mappings to inequality measures across all countries. Additionally, country rankings are severely altered once controlling for household structure. At different parts of the net wealth distribution household structure either accounts for a large part of the differences to the euro area net wealth distribution or masks the extent of these differences. These patterns differ between countries with regard to direction and size of the effect of imposing a common household structure.

At the bottom differences between observed and counterfactual distributions are relatively stronger than at the top. For the median 50% of the differences are explained for Austria, 15% for Germany, 14% for Spain, 25% for Italy and 38% for Malta. For others as Belgium, France, Greece, Luxembourg, Portugal, Slovenia and Slovakia household structure masks the differences to the euro area median. Finland's median is 86 thousand Euro and well below the euro area median of 109 thousand Euro whereas its re-weighted median 112 thousand Euro already lies above the euro area median. The Netherlands also changes its position from below to above the euro area median. Patterns differ along the distribution for many countries such as Austria, France, Slovenia and Slovakia which move closer for some areas and further away for others. Finland, Greece, the Netherlands and Portugal even switch position with regard to the euro area in some areas. Belgium and Luxembourg always move away from the euro area whereas Germany and Italy are always getting closer. Beside the direction of the effects also their level changes considerably across countries as well as along their net wealth distributions.

The impact on percentile ratios is similarly strong. We can confirm the finding of Bover (2010) that the effect on the Gini is somewhat less pronounced, but might mask relevant information by being a net effect of different accumulated effects along the distribution.

Given those findings we argue for more flexible controls for household structure and illustrate that even a small subset of the top 30 populated of our non-parametrically defined household types adds explanatory power to the standard approach of using information on household size and age and gender of a reference person only. Together with the definition of our household types which can be used flexibly to control for household type fixed effects in regression we provide our re-weighting weights which allow to analyze the relevance of household structure for any HFCS variable.

3.A Cell Construction

We define household types by all possible combinations of 4 age categories and gender for each individual (member) up to 4 individuals in each household. We are not taking (i) a particular order of individuals or (ii) gender for individuals aged 15 or below into account. Households with 5 or more members are treated as 4 person households and sorted with regard to the first 4 members, the financially knowledgeable person (respondent) and the next 3 persons sorted by descending age. This results in 329 possible household types of which 249 are observed at least once in the euro area.

The following shows the construction of the different household types over all countries in the euro area.

1. We take the first four members of each household.
2. Each member belongs to one of four age groups:

    1: below 16 years

    2: between 16 and 34 years

    3: between 35 and 64 years

    4: above 64 years

3. Each household member belongs to one of three gender groups:

    1: male

    2: female

    3: children

4. Each household member belongs therefore to a unique age-gender cell.

    Examples:

    - A male, 30 year old household member belongs in the cell [21].

- A female, 68 year old household member belongs in the cell [42].

5. Each household consists of a unique combination of age-gender pairs of the first four household members. We refer to this combination as the household type code. The household type code describes the composition of the household. Two numbers for each individual in a household, where the first refers to age category ((1 = [-; 15]; 2 = [16; 34]; 3 = [35; 64]; 4 = [65;+])) and the second refers to gender for all individuals aged 16+ (1 = male; 2 = female; 3 = below 16). The code is sorted by individual age. The most common household type 3132 is therefore a two person household (4 digits), consisting of a man aged between 35 and 64 [31] and a woman aged between 35 and 64 [32].

   Examples:

   - A household with 2 household members consisting of one male, between 35 and 64 years old and one female, between 35 and 64 years old, belongs then to the unique household cell of [31,32].
   - A single household with a female, above 64 year old household member belongs to the unique household cell of [42].
   - A household with 2 household members consisting of one male, above 64 years and one female, above 64 years, belongs then to the unique household cell of [41,42].

   The above three examples represent the most common household types in the euro area.

In order to calculate the number of possible cell combination we start at the person level. One person can be identified by one of the following combinations of age-gender pairs, whereby the first digit reflects the age group and the second the gender of the person. This takes into account that a person below 16 is always a child and a child can never be older than 15 years old.

In table 3.5 below in the first rows we see the 7 different person types, that we will observe based on the possible combinations of age and gender groups [A,B,C,D,E,F,G]. Each person falls into one of the 7 categories. For a household size larger than one we may observe different combinations of persons living in such a household. Therefore the permutations for each household size can be determined as, as k (householdsize) combinations out of 7 elements (person characteristics).

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

The formula to determine all possible cell combinations for different household sizes is a permutation with repetition without taking the rank order into account. The maximum number of combinations is 329. It allows for combinations of persons living in a household that are very unlikely to be observed in the actual household data e.g. a single household consisting of a minor (below 16 years old). The actual number of household cells observed in the euro area are 249 different types.

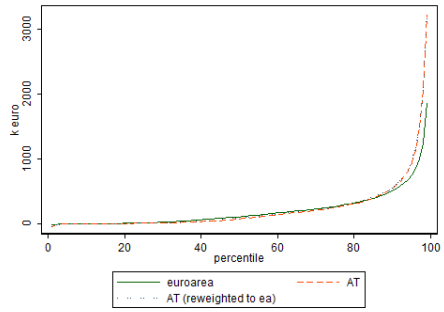
Table 3.5: Combinations of Household Cells

Size							
				7 Combinations			
1	[13]	[21]	[22]	[31]	[32]	[41]	[42]
	A	B	C	D	E	F	G
				28 Combinations			
2	[A,A]						
	[A,B]	[B,B]					
	[A,C]	[B,C]	[C,C]				
	[A,D]	[B,D]	[C,D]	[D,D]			
	[A,E]	[B,E]	[C,E]	[D,E]	[E,E]		
	[A,F]	[B,F]	[C,F]	[D,F]	[E,F]	[F,F]	
	[A,G]	[B,G]	[C,G]	[D,G]	[E,G]	[F,G]	[G,G]
				84 Combinations			
3	[AAA]	[ABB]	[ACC]	[ADD]	[AEE]	[AFF]	[AGG]
	[AAB]	[ABC]	[ACD]	[ADE]	[AEF]	[AFG]	[BGG]
	[AAC]	[ABD]	[ACE]	[ADF]	[AEG]	[BFF]	[BBD]
	[AAD]	[ABE]	[ACF]	[ADG]	[BEE]	[BFG]	[BBE]
	[AAE]	[ABF]	[ACG]	[BDD]	[BEF]	[BCE]	[BBF]
	[AAF]	[ABG]	[BCC]	[BDE]	[BBC]	[BDG]	[BDF]
	[AAG]	[BBB]	[BCD]	[BEG]	[BCF]	[BCG]	[BBG]
	[CCC]	[CDD]	[CEE]	[CFF]	[CGG]	[DEG]	[EFG]
	[CCD]	[CDE]	[CEF]	[CFG]	[DDD]	[EFF]	[EEE]
	[CCE]	[CDF]	[CEG]	[DFF]	[DDE]	[FGG]	[EEF]
	[CCF]	[CDG]	[DEF]	[DGG]	[DDF]	[FFG]	[EEG]
	[CCG]	[DEE]	[DFG]	[EGG]	[DDG]	[GGG]	[FFF]
				210 Combinations			
4	[AAAA]	[AABB]	[AACC]	[AADD]	[AAEE]	[AAFF]	[AAGG]
	[AAAB]	[AABC]	[AACD]	[AADE]	[AAEF]	[AAFG]	[ABGG]
	[AAAC]	[AABD]	[AACE]	[AADF]	[AAEG]	[ABFF]	[ACGG]
	[AAAD]	[AABE]	[AACF]	[AADG]	[ABEE]	[ABFG]	[ADGG]
	[AAAE]	[AABF]	[AACG]	[ABDD]	[ABEF]	[ACFF]	[AEGG]
	[AAAF]	[AABG]	[ABCC]	[ABDE]	[ABEG]	[ACFG]	[AFGG]
	[AAAG]	[ABBB]	[ABCD]	[ABDF]	[ACEE]	[ADFF]	[AGGG]
	[ABBC]	[ABCE]	[ABDG]	[ACEF]	[ADFG]	[BBGG]	[DDGG]
	[ABBD]	[ABCF]	[ACDD]	[ACEG]	[AEFF]	[BCGG]	[DEGG]
	[ABBE]	[ABCG]	[ACDE]	[ADEE]	[AEFG]	[BDGG]	[DFGG]
	[ABBF]	[ACCC]	[ACDF]	[ADEF]	[AFFF]	[BEGG]	[DGGG]
	[ABBG]	[ACCD]	[ACDG]	[ADEG]	[AFFG]	[BFGG]	[EEGG]
	[BBBB]	[ACCE]	[ADDD]	[AEEE]	[BBFF]	[BGGG]	[EFGG]
	[BBBC]	[ACCF]	[ADDE]	[AEEF]	[BBFG]	[CCGG]	[EGGG]
	[BBBD]	[ACCG]	[ADDF]	[AEEG]	[BCFF]	[CDGG]	[FFGG]
	[BBBE]	[BBCC]	[ADDG]	[BBEE]	[BCFG]	[CEGG]	[FGGG]
	[BBBF]	[BBCD]	[BBDD]	[BBEF]	[BDFF]	[CFGG]	[GGGG]
	[BBBG]	[BBCE]	[BBDE]	[BBEG]	[BDFG]	[CGGG]	[DDFF]
	[CCDF]	[CDEE]	[BBCF]	[BBDF]	[BCEE]	[BEFF]	[DDFG]
	[CCDG]	[CDEF]	[BBCG]	[BBDG]	[BCEF]	[BEFG]	[DEFF]
	[CDDD]	[CDEG]	[BCCC]	[BCDD]	[BCEG]	[BFFF]	[DEFG]
	[CDDE]	[CEEE]	[BCCD]	[BCDE]	[BDEE]	[BFFG]	[DFFF]
	[CDDF]	[CEEF]	[BCCE]	[BCDF]	[BDEF]	[CCFF]	[DFFG]
	[CDDG]	[CEEG]	[BCCF]	[BCDG]	[BDEG]	[CCFG]	[EEFF]
	[DDDD]	[DDEE]	[BCCG]	[BDDD]	[BEEE]	[CDFF]	[EEFG]
	[DDDE]	[DDEF]	[CCCC]	[BDDE]	[BEEF]	[CDFG]	[EFFF]
	[DDDF]	[DDEG]	[CCCD]	[BDDF]	[BEEG]	[CEFF]	[EFFG]
	[DDDG]	[DEEE]	[CCCE]	[BDDG]	[CCEE]	[CEFG]	[FFFF]
	[EEEE]	[DEEF]	[CCCF]	[CCDD]	[CCEF]	[CFFF]	[FFFG]
	[EEEF]	[DEEG]	[CCCG]	[CCDE]	[CCEG]	[CFFG]	[EEEG]

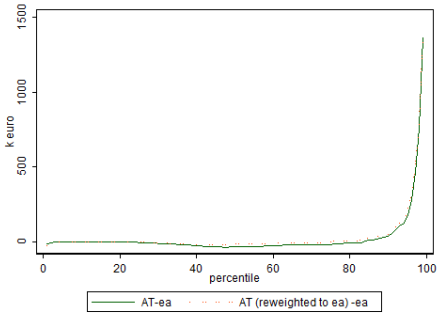
3.B Additional Results

Figure 3.4: Austrian Net Wealth Distribution with relation to the Euro Area

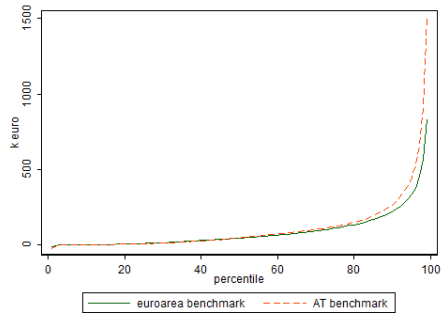
(a) (Reweighted) HH Dist.



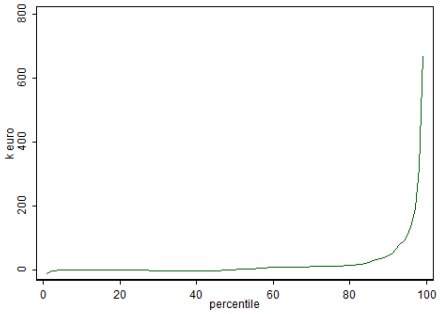
(b) Diff in (reweighted) HH Dist.



(c) Hyp. Personal Dist.



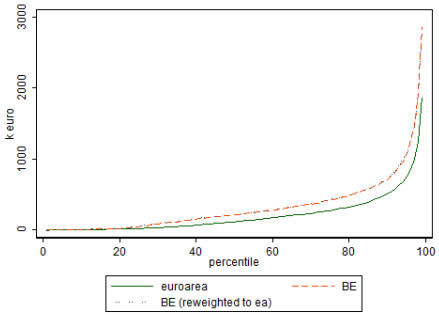
(d) Diff in Hyp. Pers. Dist.



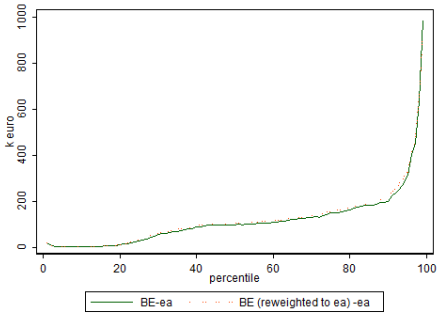
Notes:
(i) Graph (a) shows the net wealth distributions (99 percentiles) of Austria and the euro area as a whole, as well as a reweighted Austrian net wealth distribution to match euro area household structure (using non-parametric reweighting as described in section 2.4).
(ii) Graph (b) shows the differences between the net wealth distributions (99 percentiles) of Austria and the euro area as a whole as well as the differences of the reweighted Austrian distribution.
(iii) Graph (c) shows a hypothetical personal net wealth distributions (99 percentiles) of Austria and the euro area as a whole under the assumption that household wealth is equally shared among all household members.
(iv) Graph (d) shows the differences between the hypothetical personal distributions.
(v) *Source:* Eurosystem HFCS 2010.

Figure 3.5: Belgian Net Wealth Distribution with relation to the Euro Area

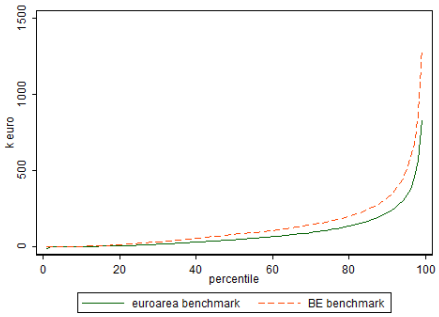
(a) (Reweighted) HH Dist.



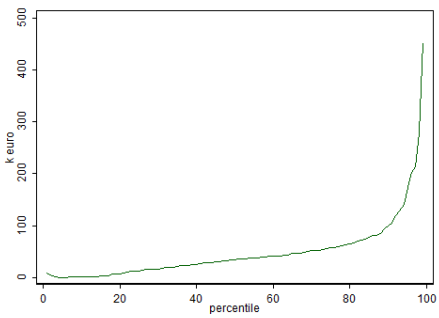
(b) Diff in (reweighted) HH Dist.



(c) Hyp. Personal Dist.



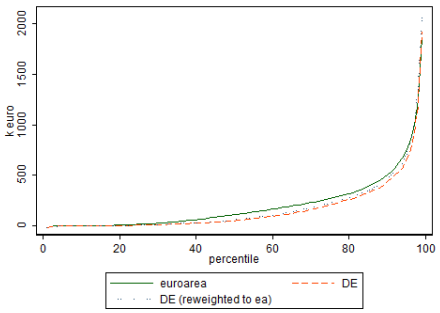
(d) Diff in Hyp. Pers. Dist.



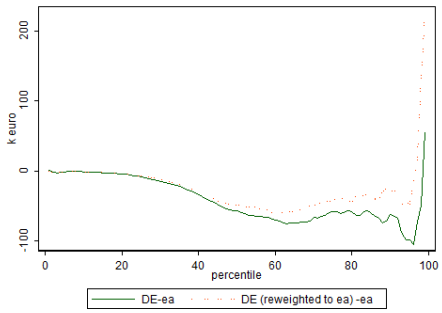
Notes:
(i) Graph (a) shows the net wealth distributions (99 percentiles) of Belgium and the euro area as a whole, as well as a reweighted Belgian net wealth distribution to match euro area household structure (using non-parametric reweighting as described in section 2.4).
(ii) Graph (b) shows the differences between the net wealth distributions (99 percentiles) of Belgium and the euro area as a whole as well as the differences of the reweighted Belgian distribution.
(iii) Graph (c) shows a hypothetical personal net wealth distributions (99 percentiles) of Belgium and the euro area as a whole under the assumption that household wealth is equally shared among all household members.
(iv) Graph (d) shows the differences between the hypothetical personal distributions.
(v) *Source:* Eurosystem HFCS 2010.

Figure 3.6: German Net Wealth Distribution with relation to the Euro Area

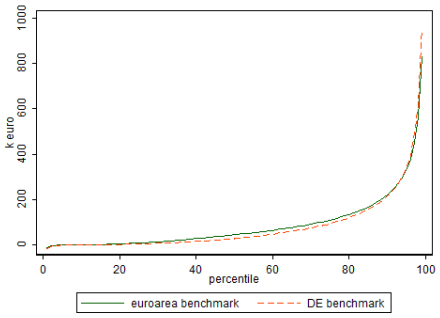
(a) (Reweighted) HH Dist.



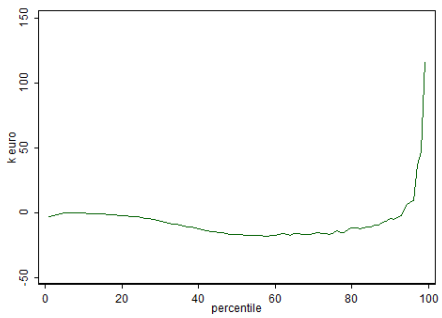
(b) Diff in (reweighted) HH Dist.



(c) Hyp. Personal Dist.

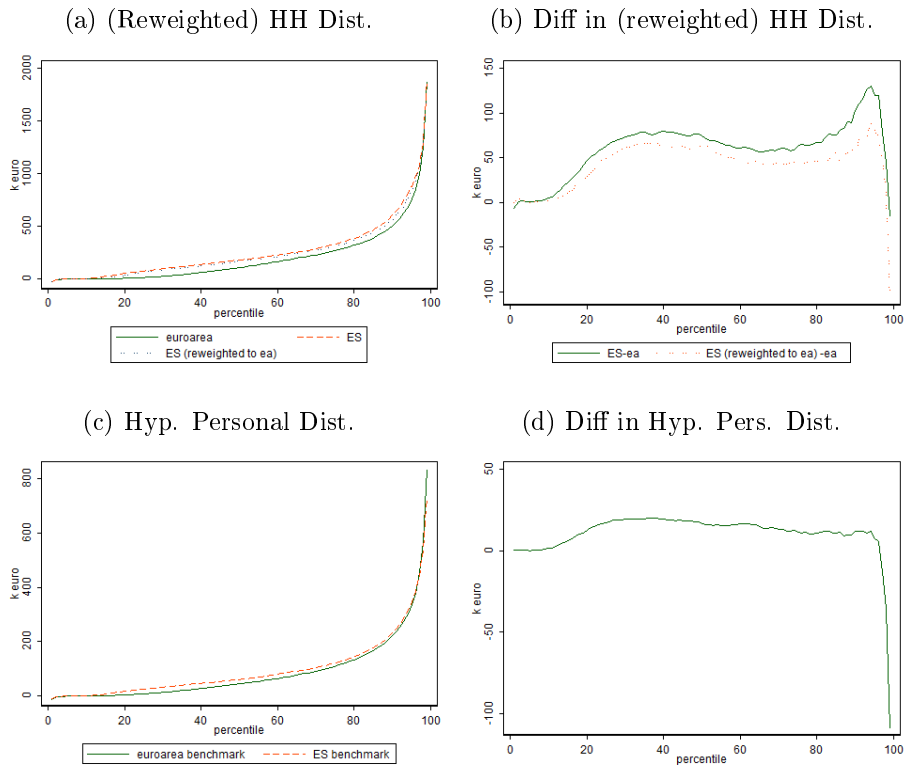


(d) Diff in Hyp. Pers. Dist.



Notes:
(i) Graph (a) shows the net wealth distributions (99 percentiles) of Germany and the euro area as a whole, as well as a reweighted German net wealth distribution to match euro area household structure (using non-parametric reweighting as described in section 2.4).
(ii) Graph (b) shows the differences between the net wealth distributions (99 percentiles) of Germany and the euro area as a whole as well as the differences of the reweighted German distribution.
(iii) Graph (c) shows a hypothetical personal net wealth distributions (99 percentiles) of Germany and the euro area as a whole under the assumption that household wealth is equally shared among all household members.
(iv) Graph (d) shows the differences between the hypothetical personal distributions.
(v) *Source:* Eurosystem HFCS 2010.

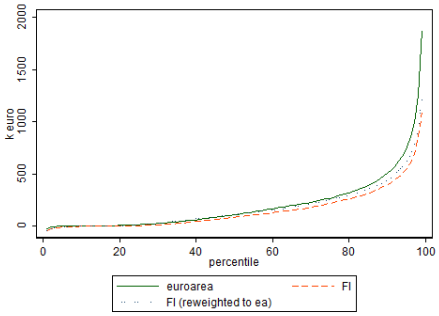
Figure 3.7: Spanish Net Wealth Distribution with relation to the Euro Area



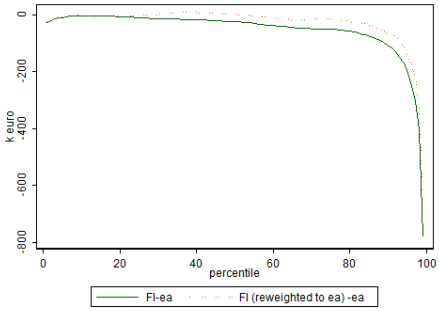
Notes:
(i) Graph (a) shows the net wealth distributions (99 percentiles) of Spain and the euro area as a whole, as well as a reweighted Spanish net wealth distribution to match euro area household structure (using non-parametric reweighting as described in section 2.4).
(ii) Graph (b) shows the differences between the net wealth distributions (99 percentiles) of Spain and the euro area as a whole as well as the differences of the reweighted Spanish distribution.
(iii) Graph (c) shows a hypothetical personal net wealth distributions (99 percentiles) of Spain and the euro area as a whole under the assumption that household wealth is equally shared among all household members.
(iv) Graph (d) shows the differences between the hypothetical personal distributions.
(v) *Source:* Eurosystem HFCS 2010.

Figure 3.8: Finnish Net Wealth Distribution with relation to the Euro Area

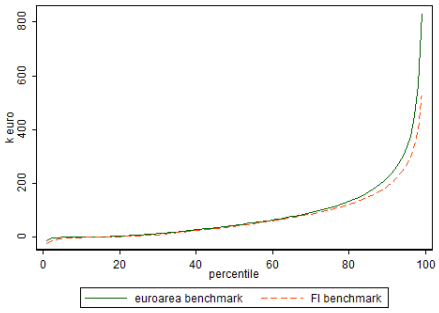
(a) (Reweighted) HH Dist.



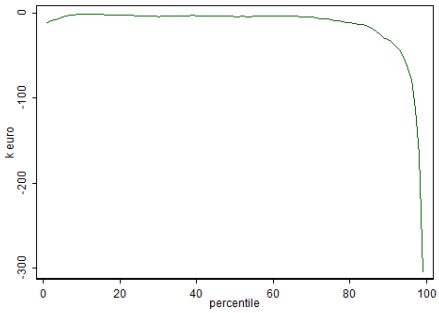
(b) Diff in (reweighted) HH Dist.



(c) Hyp. Personal Dist.



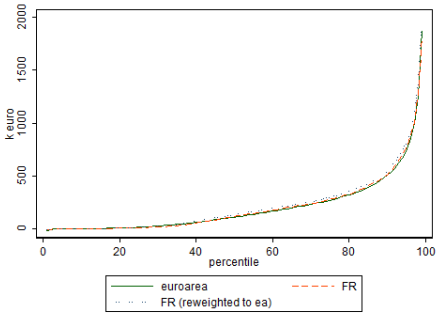
(d) Diff in Hyp. Pers. Dist.



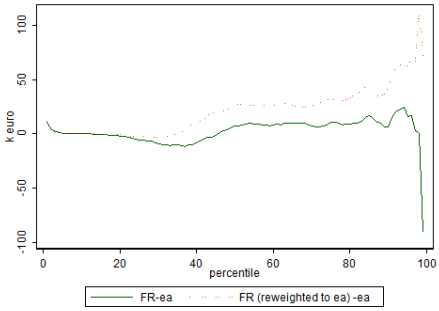
Notes:
(i) Graph (a) shows the net wealth distributions (99 percentiles) of Finland and the euro area as a whole, as well as a reweighted Finnish net wealth distribution to match euro area household structure (using non-parametric reweighting as described in section 2.4).
(ii) Graph (b) shows the differences between the net wealth distributions (99 percentiles) of Finland and the euro area as a whole as well as the differences of the reweighted Finnish distribution.
(iii) Graph (c) shows a hypothetical personal net wealth distributions (99 percentiles) of Finland and the euro area as a whole under the assumption that household wealth is equally shared among all household members.
(iv) Graph (d) shows the differences between the hypothetical personal distributions.
(v) *Source:* Eurosystem HFCS 2010.

Figure 3.9: French Net Wealth Distribution with relation to the Euro Area

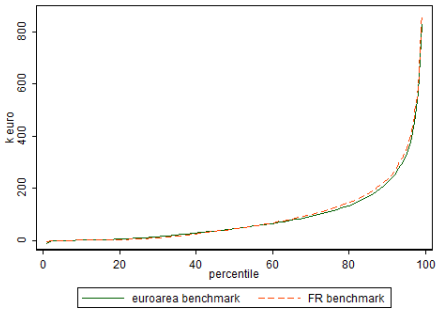
(a) (Reweighted) HH Dist.



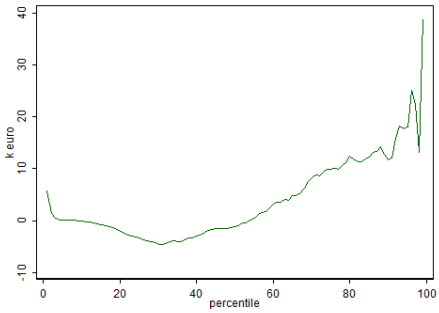
(b) Diff in (reweighted) HH Dist.



(c) Hyp. Personal Dist.



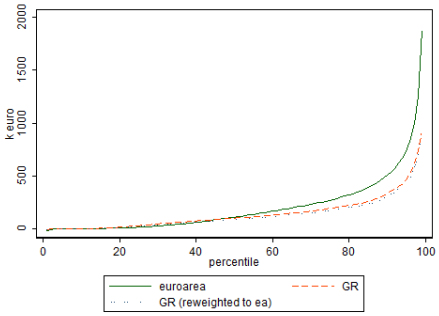
(d) Diff in Hyp. Pers. Dist.



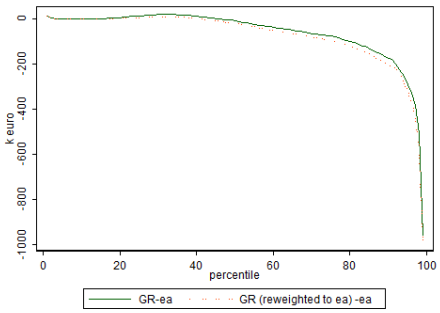
Notes:
(i) Graph (a) shows the net wealth distributions (99 percentiles) of France and the euro area as a whole, as well as a reweighted French net wealth distribution to match euro area household structure (using non-parametric reweighting as described in section 2.4).
(ii) Graph (b) shows the differences between the net wealth distributions (99 percentiles) of France and the euro area as a whole as well as the differences of the reweighted French distribution.
(iii) Graph (c) shows a hypothetical personal net wealth distributions (99 percentiles) of France and the euro area as a whole under the assumption that household wealth is equally shared among all household members.
(iv) Graph (d) shows the differences between the hypothetical personal distributions.
(v) *Source:* Eurosystem HFCS 2010.

Figure 3.10: Greek Net Wealth Distribution with relation to the Euro Area

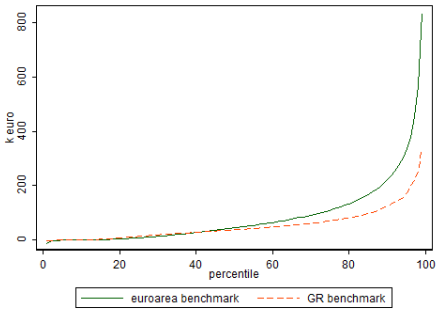
(a) (Reweighted) HH Dist.



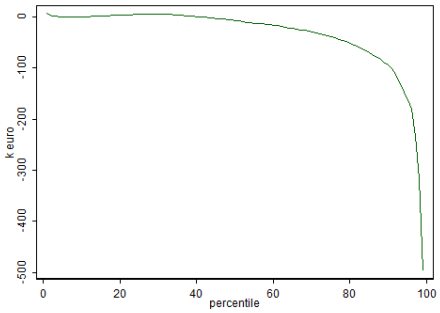
(b) Diff in (reweighted) HH Dist.



(c) Hyp. Personal Dist.



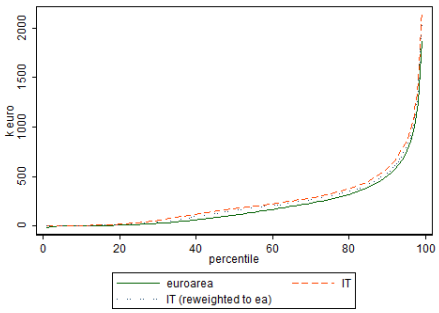
(d) Diff in Hyp. Pers. Dist.



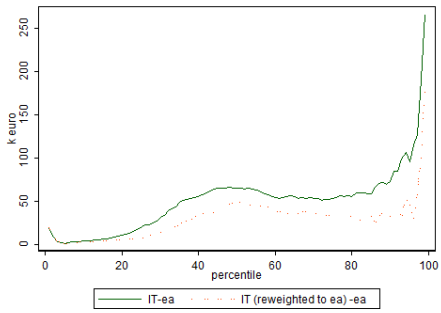
Notes:
(i) Graph (a) shows the net wealth distributions (99 percentiles) of Greece and the euro area as a whole, as well as a reweighted Greek net wealth distribution to match euro area household structure (using non-parametric reweighting as described in section 2.4).
(ii) Graph (b) shows the differences between the net wealth distributions (99 percentiles) of Greece and the euro area as a whole as well as the differences of the reweighted Greek distribution.
(iii) Graph (c) shows a hypothetical personal net wealth distributions (99 percentiles) of Greece and the euro area as a whole under the assumption that household wealth is equally shared among all household members.
(iv) Graph (d) shows the differences between the hypothetical personal distributions.
(v) *Source:* Eurosystem HFCS 2010.

Figure 3.11: Italian Net Wealth Distribution with relation to the Euro Area

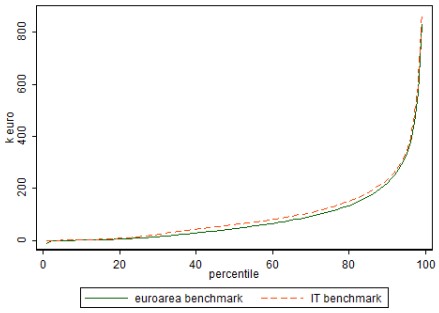
(a) (Reweighted) HH Dist.



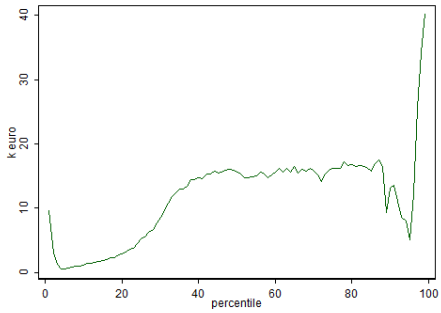
(b) Diff in (reweighted) HH Dist.



(c) Hyp. Personal Dist.



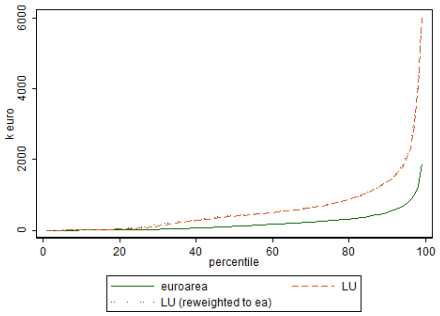
(d) Diff in Hyp. Pers. Dist.



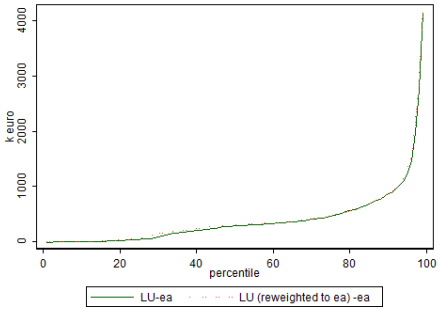
Notes:
(i) Graph (a) shows the net wealth distributions (99 percentiles) of Italy and the euro area as a whole, as well as a reweighted Italian net wealth distribution to match euro area household structure (using non-parametric reweighting as described in section 2.4).
(ii) Graph (b) shows the differences between the net wealth distributions (99 percentiles) of Italy and the euro area as a whole as well as the differences of the reweighted Italian distribution.
(iii) Graph (c) shows a hypothetical personal net wealth distributions (99 percentiles) of Italy and the euro area as a whole under the assumption that household wealth is equally shared among all household members.
(iv) Graph (d) shows the differences between the hypothetical personal distributions.
(v) *Source:* Eurosystem HFCS 2010.

Figure 3.12: Luxembourgian Net Wealth Distribution with relation to the Euro Area

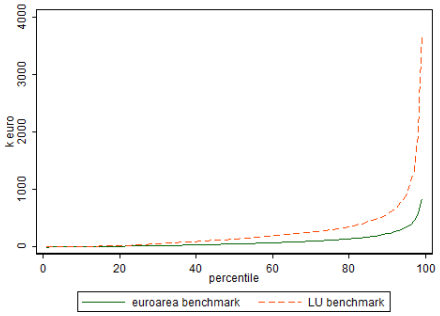
(a) (Reweighted) HH Dist.



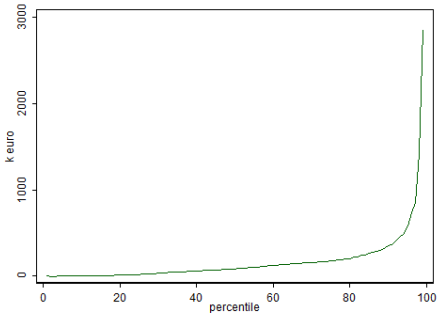
(b) Diff in (reweighted) HH Dist.



(c) Hyp. Personal Dist.



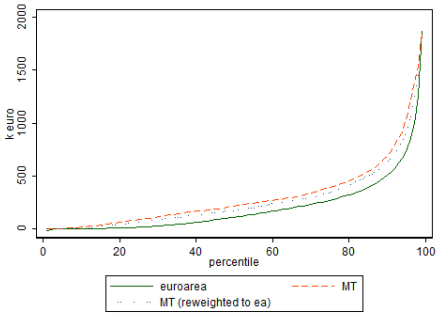
(d) Diff in Hyp. Pers. Dist.



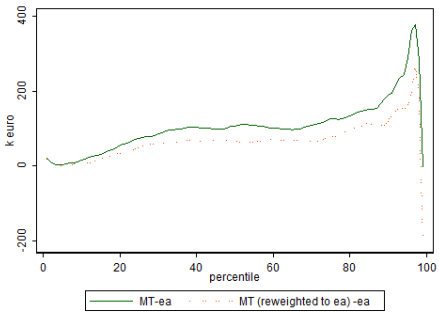
Notes:
(i) Graph (a) shows the net wealth distributions (99 percentiles) of Luxembourg and the euro area as a whole, as well as a reweighted Luxembourgish net wealth distribution to match euro area household structure (using non-parametric reweighting as described in section 2.4).
(ii) Graph (b) shows the differences between the net wealth distributions (99 percentiles) of Luxembourg and the euro area as a whole as well as the differences of the reweighted Luxembourgish distribution.
(iii) Graph (c) shows a hypothetical personal net wealth distributions (99 percentiles) of Luxembourg and the euro area as a whole under the assumption that household wealth is equally shared among all household members.
(iv) Graph (d) shows the differences between the hypothetical personal distributions.
(v) *Source:* Eurosystem HFCS 2010.

Figure 3.13: Maltese Net Wealth Distribution with relation to the Euro Area

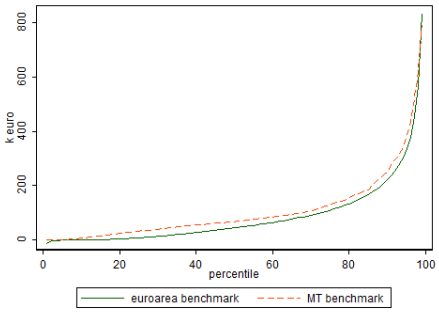
(a) (Reweighted) HH Dist.



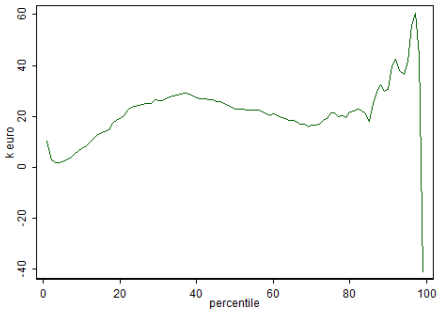
(b) Diff in (reweighted) HH Dist.



(c) Hyp. Personal Dist.



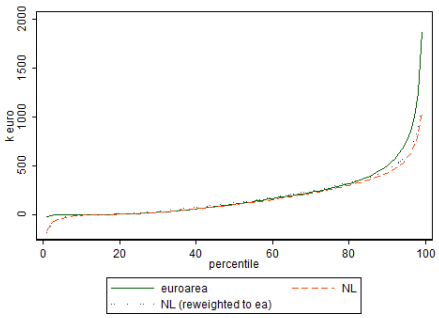
(d) Diff in Hyp. Pers. Dist.



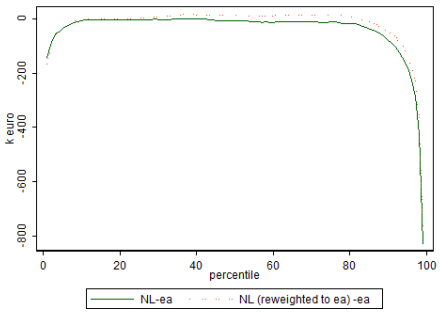
Notes:
(i) Graph (a) shows the net wealth distributions (99 percentiles) of Malta and the euro area as a whole, as well as a reweighted Maltese net wealth distribution to match euro area household structure (using non-parametric reweighting as described in section 2.4).
(ii) Graph (b) shows the differences between the net wealth distributions (99 percentiles) of Malta and the euro area as a whole as well as the differences of the reweighted Maltese distribution.
(iii) Graph (c) shows a hypothetical personal net wealth distributions (99 percentiles) of Malta and the euro area as a whole under the assumption that household wealth is equally shared among all household members.
(iv) Graph (d) shows the differences between the hypothetical personal distributions.
(v) *Source:* Eurosystem HFCS 2010.

Figure 3.14: Dutch Net Wealth Distribution with relation to the Euro Area

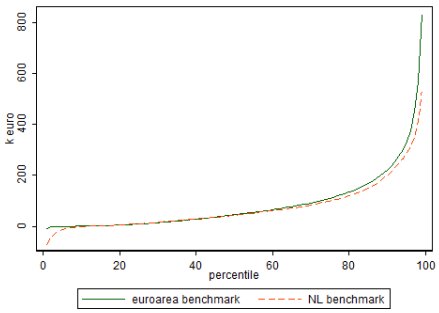
(a) (Reweighted) HH Dist.



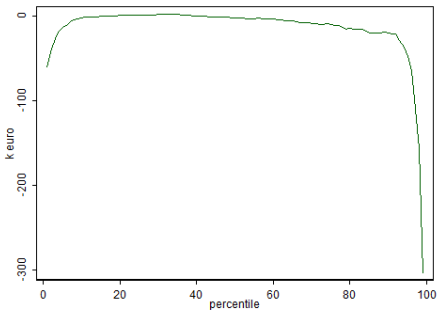
(b) Diff in (reweighted) HH Dist.



(c) Hyp. Personal Dist.



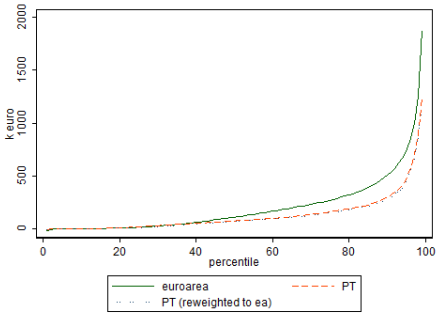
(d) Diff in Hyp. Pers. Dist.



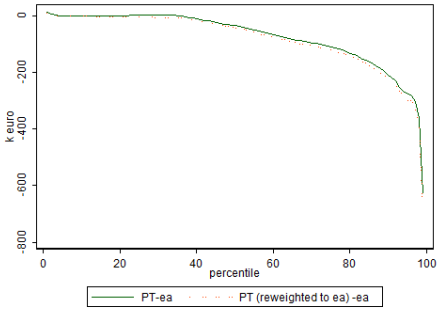
Notes:
(i) Graph (a) shows the net wealth distributions (99 percentiles) of the Netherlands and the euro area as a whole, as well as a reweighted Dutch net wealth distribution to match euro area household structure (using non-parametric reweighting as described in section 2.4).
(ii) Graph (b) shows the differences between the net wealth distributions (99 percentiles) of the Netherlands and the euro area as a whole as well as the differences of the reweighted Dutch distribution.
(iii) Graph (c) shows a hypothetical personal net wealth distributions (99 percentiles) of the Netherlands and the euro area as a whole under the assumption that household wealth is equally shared among all household members.
(iv) Graph (d) shows the differences between the hypothetical personal distributions.
(v) *Source:* Eurosystem HFCS 2010.

Figure 3.15: Portuguese Net Wealth Distribution with relation to the Euro Area

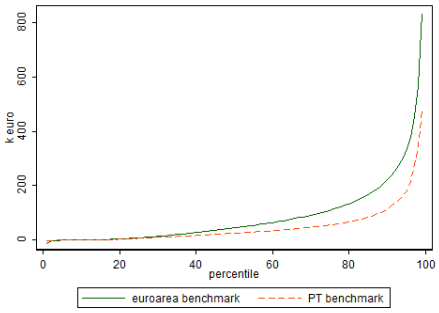
(a) (Reweighted) HH Dist.



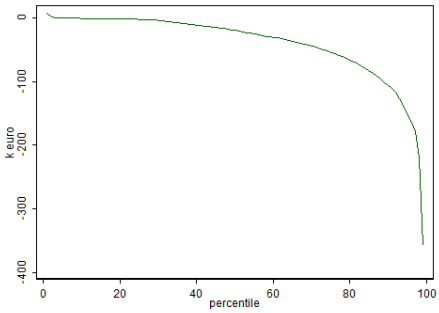
(b) Diff in (reweighted) HH Dist.



(c) Hyp. Personal Dist.



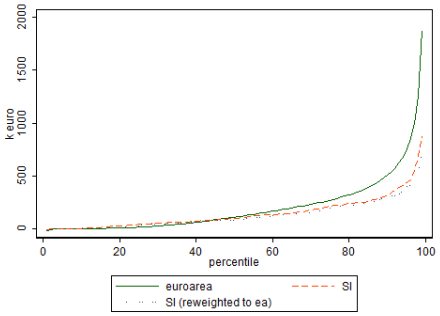
(d) Diff in Hyp. Pers. Dist.



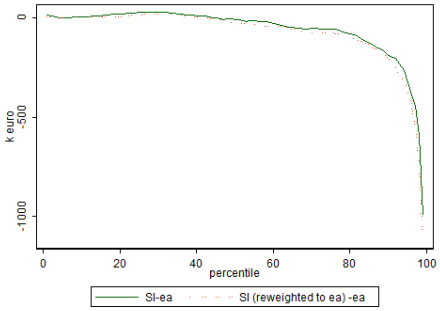
Notes:
(i) Graph (a) shows the net wealth distributions (99 percentiles) of Portugal and the euro area as a whole, as well as a reweighted Portuguese net wealth distribution to match euro area household structure (using non-parametric reweighting as described in section 2.4).
(ii) Graph (b) shows the differences between the net wealth distributions (99 percentiles) of Portugal and the euro area as a whole as well as the differences of the reweighted Portuguese distribution.
(iii) Graph (c) shows a hypothetical personal net wealth distributions (99 percentiles) of Portugal and the euro area as a whole under the assumption that household wealth is equally shared among all household members.
(iv) Graph (d) shows the differences between the hypothetical personal distributions.
(v) *Source:* Eurosystem HFCS 2010.

Figure 3.16: Slovenian Net Wealth Distribution with relation to the Euro Area

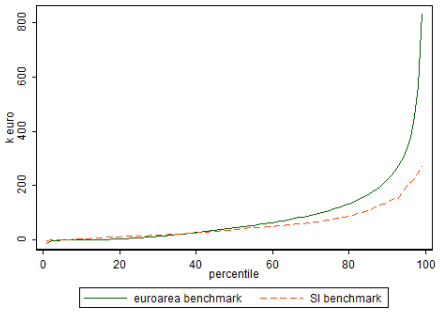
(a) (Reweighted) HH Dist.



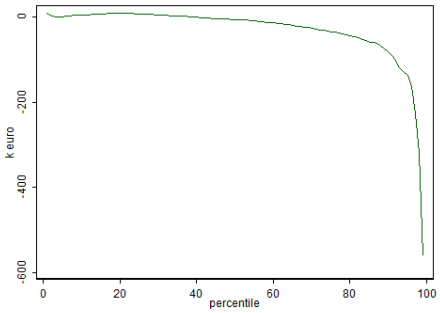
(b) Diff in (reweighted) HH Dist.



(c) Hyp. Personal Dist.



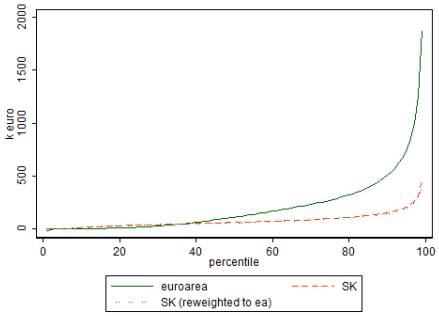
(d) Diff in Hyp. Pers. Dist.



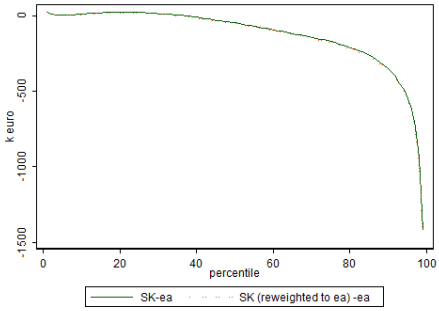
Notes:
(i) Graph (a) shows the net wealth distributions (99 percentiles) of Slovenia and the euro area as a whole, as well as a reweighted Slovenian net wealth distribution to match euro area household structure (using non-parametric reweighting as described in section 2.4).
(ii) Graph (b) shows the differences between the net wealth distributions (99 percentiles) of Slovenia and the euro area as a whole as well as the differences of the reweighted Slovenian distribution.
(iii) Graph (c) shows a hypothetical personal net wealth distributions (99 percentiles) of Slovenia and the euro area as a whole under the assumption that household wealth is equally shared among all household members.
(iv) Graph (d) shows the differences between the hypothetical personal distributions.
(v) *Source:* Eurosystem HFCS 2010.

Figure 3.17: Slovakian Net Wealth Distribution with relation to the Euro Area

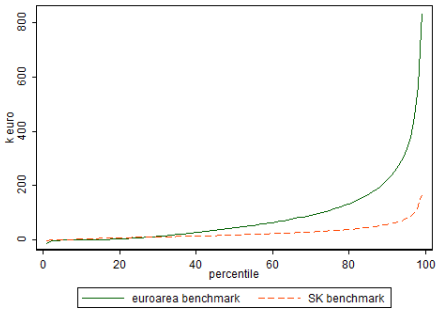
(a) (Reweighted) HH Dist.



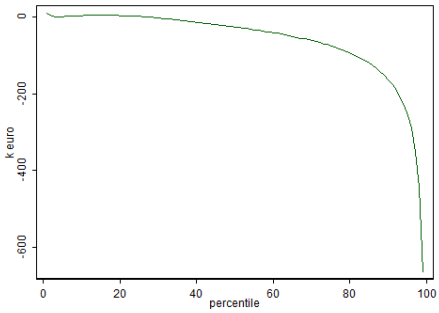
(b) Diff in (reweighted) HH Dist.



(c) Hyp. Personal Dist.



(d) Diff in Hyp. Pers. Dist.



Notes:
(i) Graph (a) shows the net wealth distributions (99 percentiles) of Slovakia and the euro area as a whole, as well as a reweighted Slovakian net wealth distribution to match euro area household structure (using non-parametric reweighting as described in section 2.4).
(ii) Graph (b) shows the differences between the net wealth distributions (99 percentiles) of Slovakia and the euro area as a whole as well as the differences of the reweighted Slovakian distribution.
(iii) Graph (c) shows a hypothetical personal net wealth distributions (99 percentiles) of Slovakia and the euro area as a whole under the assumption that household wealth is equally shared among all household members.
(iv) Graph (d) shows the differences between the hypothetical personal distributions.
(v) *Source:* Eurosystem HFCS 2010.

Table 3.6: Effects of Household Structure Differences Across Countries on Extensive Margin

Variable Names	Euro Area	AT	BE	DE	ES	FI	FR	GR	IT	LU	MT	NL	PT	SI	SK
Real Ass Part	0.911	0.848	0.898	0.802	0.953	0.848	1.000	0.922	0.977	0.936	0.948	0.898	0.901	0.962	0.960
Real Ass Count	.	0.863	0.898	0.824	0.937	0.876	1.000	0.907	0.974	0.941	0.924	0.911	0.875	0.953	0.951
HMR Part	0.601	0.477	0.696	0.442	0.827	0.692	0.553	0.724	0.687	0.671	0.777	0.571	0.715	0.818	0.899
HMR Count	.	0.502	0.704	0.468	0.800	0.740	0.582	0.695	0.651	0.695	0.741	0.623	0.674	0.774	0.883
FRE Part	0.238	0.134	0.164	0.178	0.362	0.298	0.285	0.379	0.249	0.282	0.314	0.061	0.271	0.232	0.153
FRE Count	.	0.140	0.165	0.193	0.335	0.339	0.301	0.344	0.225	0.281	0.285	0.076	0.256	0.205	0.133
BUS Part	0.110	0.094	0.066	0.091	0.142	0.138	0.089	0.098	0.180	0.052	0.115	0.048	0.077	0.116	0.107
BUS Count	.	0.090	0.069	0.100	0.122	0.157	0.093	0.083	0.179	0.048	0.093	0.043	0.071	0.097	0.099
Rest Real Part	0.841	0.799	0.802	0.732	0.799	0.679	1.000	0.735	0.951	0.884	0.875	0.826	0.741	0.804	0.685
Rest Real Count	.	0.815	0.803	0.756	0.753	0.717	1.000	0.678	0.946	0.875	0.833	0.834	0.694	0.806	0.632
Fin Ass Part	0.969	0.995	0.980	0.993	0.983	1.000	0.996	0.745	0.920	0.984	0.972	0.978	0.945	0.939	0.917
Fin Ass Count	.	0.995	0.978	0.995	0.984	1.000	0.996	0.735	0.917	0.984	0.959	0.977	0.935	0.937	0.888
Safe Fin Part	0.967	0.994	0.979	0.991	0.982	1.000	0.996	0.739	0.919	0.984	0.969	0.973	0.943	0.936	0.915
Safe Fin Count	.	0.994	0.977	0.992	0.983	1.000	0.996	0.729	0.916	0.984	0.958	0.975	0.932	0.935	0.885
Risky Fin Part	0.201	0.146	0.307	0.230	0.140	0.387	0.217	0.040	0.198	0.258	0.337	0.239	0.065	0.203	0.041
Risky Fin Count	.	0.152	0.318	0.232	0.138	0.408	0.222	0.034	0.189	0.259	0.291	0.243	0.070	0.179	0.043
Rest Fin Part	0.125	0.114	0.106	0.223	0.080	0.000	0.121	0.041	0.049	0.088	0.060	0.107	0.085	0.067	0.103
Rest Fin Count	.	0.106	0.107	0.222	0.078	0.000	0.121	0.039	0.050	0.084	0.056	0.101	0.085	0.062	0.097
Mortgage Part	0.231	0.184	0.305	0.215	0.325	0.328	0.244	0.175	0.108	0.388	0.156	0.447	0.267	0.141	0.096
Mortgage Count	.	0.203	0.309	0.239	0.305	0.345	0.245	0.154	0.109	0.357	0.158	0.492	0.256	0.119	0.102
Non-Mort Part	0.297	0.214	0.242	0.346	0.307	0.512	0.328	0.261	0.178	0.369	0.252	0.373	0.183	0.389	0.199
Non-Mort Count	.	0.212	0.238	0.353	0.280	0.523	0.325	0.241	0.176	0.337	0.224	0.364	0.171	0.345	0.164
Loan Tot Part	0.437	0.356	0.448	0.474	0.500	0.598	0.469	0.366	0.252	0.583	0.341	0.657	0.377	0.445	0.268
Loan Tot Count	.	0.369	0.445	0.494	0.467	0.608	0.467	0.338	0.251	0.541	0.327	0.676	0.358	0.393	0.240

Notes:

(i) This table shows the participations rates of the portfolio components of net wealth in the euro area. For each statistic, one can see the estimate based on original weights and the counterfactual estimates using the re-weighted households weights controlling for the differences of the household structure.

(ii) The household portfolio is separated in real assets [Real Ass] (comprising the households' main residence [HMR], further real estate [FRE], self-employment business participation [BUS], and other real assets [Rest Real]), financial assets [Fin Ass] (comprising safe financial assets [Safe Fin], risky financial assets [Risky Fin], and other financial assets [Rest Fin]), and liabilities [Loan Tot] (comprising mortgage [Mortgage] and non-mortgage [Non-Mort] loans).

(iii) For the euro area there is no counterfactual estimate by definition, thus cells are denoted with a dot.

(iv)*Source:* Eurosystem HFCS 2010.

Table 3.7: Effects of Household Structure Differences Across Countries on Intensive Margin

Variable Names	Euro Area	AT	BE	DE	ES	FI	FR	GR	IT	LU	MT	NL	PT	SI	SK
Real Ass Median	144.60	106.99	220.00	89.20	201.74	143.68	124.14	114.26	176.00	470.49	201.06	198.79	91.88	105.90	61.79
Real Ass Count	.	118.83	222.90	100.84	192.41	160.15	138.68	105.00	161.00	470.60	181.59	213.98	89.00	96.92	58.33
HMR Median	180.30	200.00	250.00	168.00	180.30	127.79	193.82	100.00	200.00	500.00	186.64	240.00	90.00	110.92	55.90
HMR Count	.	200.00	250.00	180.00	180.30	135.50	196.73	100.00	200.00	500.00	180.84	240.00	89.96	101.99	52.43
FRE Median	100.00	94.03	174.00	115.00	120.20	107.57	96.06	61.93	100.00	300.00	120.12	165.53	53.46	52.35	16.40
FRE Count	.	89.29	178.50	120.00	120.20	109.40	98.98	60.00	90.00	300.00	110.47	157.75	53.37	44.21	17.80
BUS Median	30.00	180.60	50.00	19.40	50.84	0.90	53.14	36.18	15.00	97.60	136.51	51.69	47.13	25.48	4.60
BUS Count	.	198.90	50.00	20.40	50.42	0.90	54.05	45.14	10.00	97.60	169.68	50.00	46.38	25.54	5.21
Rest Real Median	7.00	8.57	7.50	7.78	7.19	9.32	4.32	6.00	9.50	18.74	7.08	6.51	5.00	3.31	4.50
Rest Real Count	.	9.05	7.50	8.00	6.01	10.05	4.53	6.00	9.00	18.20	6.44	7.00	5.00	3.00	4.00
Fin Ass Median	11.35	13.47	26.48	17.11	6.00	7.36	10.67	4.37	10.00	27.91	26.23	34.72	4.27	1.72	2.54
Fin Ass Count	.	14.81	27.62	18.54	6.00	8.53	11.43	4.00	8.84	28.60	23.62	39.61	4.08	1.72	2.39
Safe Fin Median	9.14	11.88	20.68	13.16	5.06	5.66	8.91	3.95	7.40	23.06	17.71	30.38	3.83	1.13	2.33
Safe Fin Count	.	13.39	21.78	14.77	4.96	6.63	9.67	3.63	7.00	23.38	16.61	32.62	3.74	0.99	2.25
Risky Fin Median	12.21	12.25	20.14	12.08	12.00	3.72	8.14	7.34	22.44	28.48	21.65	8.22	8.85	3.42	1.11
Risky Fin Count	.	12.62	20.00	12.88	12.00	3.81	8.42	5.26	20.68	31.08	22.05	10.22	8.75	3.25	1.00
Rest Fin Median	3.25	3.39	4.40	2.63	7.00	.	4.00	3.00	10.00	5.00	8.75	2.23	4.30	4.00	1.29
Rest Fin Count	.	4.20	5.00	2.86	6.01	.	4.02	2.00	7.43	5.80	7.00	3.06	4.00	7.00	1.11
Mortgage Median	68.24	37.55	69.25	80.00	60.00	64.40	55.88	40.96	60.00	127.33	35.00	130.97	48.76	6.63	25.04
Mortgage Count	.	42.73	68.55	80.00	60.06	62.43	52.60	43.30	65.00	129.13	38.00	129.20	53.06	8.83	25.99
Non-Mort Median	4.98	3.02	5.16	3.17	7.19	6.78	5.19	4.32	5.70	10.02	4.00	13.71	3.31	3.07	1.05
Non-Mort Count	.	3.32	5.13	3.30	6.00	7.22	5.36	4.00	5.40	10.01	3.34	12.56	3.00	2.54	1.00
Loan Tot Median	21.37	13.78	39.30	12.62	36.00	29.45	18.38	14.57	15.00	73.44	15.67	89.13	31.69	4.34	3.20
Loan Tot Count	.	17.04	40.89	15.00	36.31	33.56	17.97	12.96	15.00	71.84	13.80	97.14	34.98	3.56	5.00

Notes:

(i) This table shows (in thousand Euro) the median of the portfolio components of net wealth in the euro area. For each statistic, one can see the estimate based on original weights and the counterfactual estimates using the re-weighted households weights controlling for the differences of the household structure.

(ii) The household portfolio is separated in real assets [Real Ass] (comprising the households' main residence [HMR], further real estate [FRE], self-employment business participation [BUS], and other real assets [Rest Real]), financial assets [Fin Ass] (comprising safe financial assets [Safe Fin], risky financial assets [Risky Fin], and other financial assets [Rest Fin]), and liabilities [Loan Tot] (comprising mortgage [Mortgage] and non-mortgage [Non-Mort] loans).

(iii) For the euro area there is no counterfactual estimate by definition, thus cells are denoted with a dot.

(iv)*Source:* Eurosystem HFCS 2010.

Chapter 4

Factor Decomposition of the Wealth Distribution in the Euro Area

4.1 Introduction

Economists and social scientists in general were always interested in the access to and control over resources in the society. Thus analysing and understanding the distribution of wealth and its foundations is at the heart of a large empirical literature. In order to investigate which component of the portfolio of households is contributing to what extent towards the overall distribution of wealth, we employ several decomposition methods according to factor components proposed by the literature. The research question that is underlying this analysis can be laid out as: What is the marginal contribution of a certain component to overall inequality? What proportion of the overall distribution of wealth can be attributed to which component of the portfolio of households? Doing so allows us to see where the differences in this respect lay across the euro area countries and also which of the components are most important for the differences in the overall wealth figures.

The motivation of the investigation at hand is based on two topics. On the one hand, economic policy is often focussed on or affects certain components of the household balance sheet. Consider for example the government assistance for construction or refurbishment of the primary residence of households mostly in form of relatively cheap financing. Also monetary policy in setting the interest rate affects different households differently. The recently very low nominal interest rate translates over the monetary policy transmission mechanism to (extremely) low interest rates in savings while financial markets and hence risky financial instruments of the portfolio are stabilized and thus favoured. These policies quite often are implemented without assessing the distributional wealth impacts or targeting the wealth distribution but clearly affect it.

On the other hand, there has been a huge public discussion about wealth figures and their cross country comparability since the publication of the first wave of the Household Finance and Consumption Survey (HFCS) from the European Central Bank. As it turns out the commonly held believe about the households which are

supposedly more affluent in northern and less affluent in southern European countries is a myth. Although this believe might hold for other sectors of the economy (e.g. the state sector) or to some extent also for comparison of countries concerning income the results for median wealth provided through the HFCS tell a different story. So is the median and mean household net wealth in Germany around 51 and 195 thousand € respectively (see table 4.9), while the according figures for Cyprus are 267 and 671 thousand €. Understanding where this differences come from will remain an important topic for applied research in the field of household finances. It necessitates a deep understanding of the structure and the explanatory factors of the wealth distribution, since comparisons of household wealth crucially depend on the distribution of this wealth within an economy. In the publication of the first results from the HFCS (ECB (2013b)) the ECB gives some explanation for why countries are different in terms of household wealth. These may be attributed to a wide array of potential possibilities ranging from historic developments over differences in the development of asset prices to differences in the data generating process. Fessler, Lindner, and Segalla (2013) take a closer look at the differences in the structure of the household across countries in order to see how much of the discrepancy can be attributed to the unit of collection. It is shown that a sizeable proportion of the differences between countries can be attributed to the household structure (household size, gender and age of the households members) using non-parametric reweighting techniques.

This paper comes from a different perspective and takes the level and distribution of household wealth as it is given by the results from the HFCS. In order to understand the importance and marginal contribution of various components of the household balance sheet we employ several decomposition methods according to components proposed by the literature and propose a new method based on the recentered influence function. Based thereon we can add to the understanding of the complexities of cross country comparisons in terms of the wealth distribution in the euro area and

discuss distributional effects of policy measures targeting specific components. The main contribution to the literature comes mainly in two respects. First, one can see how the recentered influence function which has recently been used widely but so far only for subgroup decompositions (such as unionised vs. non-unionised workers) can be applied in the factor decomposition according to components. And second, we report for the first time the results for this type of decomposition from an a-priori harmonized euro area wide data set. This is done not only for the whole euro area but also the 15 separate countries. Additionally, whereas most articles concentrate on one or the other decomposition method, we employ a variety of them in order to see the differences that come with the choice of the underlying method.

One finds in general that the overall wealth distribution is driven by the real asset side of the households' balance sheet explaining about 4/5 of the overall distribution, since this side contains by far the higher wealth share. However, the elasticity with respect to components shows differences over components as well as over countries. For example (taking the results form table 4.6) favouring safe financial assets in the euro area should tend to have an equalizing impact on the wealth distribution while favouring risky financial assets inhibits the reverse tendency. Such policies also might qualitatively as well as quantitatively display different impacts in different countries, as can be seen by the sign and size differences of these elasticities reported in tables 4.6 and 4.8.

In what follows the paper first shortly surveys the literature that is most closely related to the topic at hand in order to embed this work into the wider context (section 4.2). Section 4.3 gives an overview of the data that are used. This part also includes the definition of wealth and components thereof. For the convenience of the reader, section 4.4 revises the applied decomposition methods from the literature. The newly proposed method based in the recentered influence function is also introduced in this section. Section 4.5 discusses the overall statistics derived from the HFCS and then the results from the decomposition exercise. Concluding

remarks round up the paper.

4.2 Background

Certainly inequality and its decomposition[1] - mostly in terms of income - has been studied for a long time in economics. For example Shorrocks (1982) and Lambert and Aronson (1993) provide an early intuitive method using (strong) assumptions based on the covariance of total income and its components or subgroups of individuals. Lerman and Yitzhaki (1985) lays out an alternative including the derivation for the elasticity of inequality with respect to each specific component. In Silber (1989) one of the most widely used measure to characterize the income distribution, namely the Gini coefficient, is decomposed both according to components and subgroups based on the matrix algebraic computation of this index. A computationally different approach is taken by the so-called Shapley value decomposition, see e.g. Sastre and Trannoy (2002), Araar (2006), or Shorrocks (2013), where one estimates the contribution of each component by the weighted average of the differences of inequality over all possible combinations including and excluding a specific component. There are also, especially for sub-group decompositions however, regression based decomposition methods that can e.g. be found in a recent work by Cowell and Fiorio (2009). The regression approach is closely related to the natural decomposition proposed by Shorrocks (1982) for the factor components. More recently decomposition methods based on counterfactual analysis as in Chernozhukov, Fernandez-Val, and Melly (2009) and Fortin, Lemieux, and Firpo (2011) have been introduced. We make use of all these methods (except the regression based approach[2]) and lay out the details in section 4.4.

[1] See for example Kanbur (2006) for a relatively recent discussion on the importance of such decompositions exercises.

[2] This line of research makes use of the income generating function which has no equivalent for wealth.

Income

On the one hand there is a rich empirical literature on the various forms of decomposing income inequality. This strand covers a wide variety of countries (without any special ordering see for example Okamoto (2011), Chernozhukov, Fernandez-Val, and Melly (2009), Firpo, Fortin, and Lemieux (2007), Yun (2006), Fields (2003), Juhn, Murphy, and Pierce (1993) for the U.S.; Machado and Mata (2005) for Portugal; Paul (2004) for Australia; Celestin and Clovis (2012) for Cameroon; Fournier (2001) for Taiwan; Mookherjee and Shorrocks (1982) for the UK; Fiorio (2006) for Italy; Morduch and Sicular (2002) for China; Fräßdorf, Grabka, and Schwarze (2011) for a comparison of the UK, Germany and the US; Bourguignon, Ferreira, and Leite (2008) for a comparison of Brazil and the US; Gunatilaka and Chotikapanich (2009) for Sri Lanka; and López-Feldman, Mora, and Taylor (2006) for Mexico). Furthermore, there is an extensive conglomeration of research questions that were investigated with these decomposition approaches. For one, differences in income and its distribution are analysed with respect to differences in the socio-economic / work force characteristics (see e.g. Machado and Mata (2005), Morduch and Sicular (2002), Bourguignon, Ferreira, and Leite (2008) or Fields (2003)) and thus also in terms of returns to skill as in Juhn, Murphy, and Pierce (1993). Additionally, Gunatilaka and Chotikapanich (2009), Firpo, Fortin, and Lemieux (2007), Mookherjee and Shorrocks (1982), Fiorio (2006) as well as Yun (2006) use decomposition methods to analyse the change of the income distribution over time. What is more closely connected to the exercise below is the decomposition of income according to the different possible sources. Paul (2004) for example looks at the contribution income components exerts to the total distribution of income for Australia. Extending the decomposition procedure from Lerman and Yitzhaki (1985) it is shown in the study from Paul (2004) that in this country the overall distribution has a positive elasticity with respect to income from business activity, wages, the children's income and for some measures property income whilst the other forms show a negative elasticity.

Fräßdorf, Grabka, and Schwarze (2011) base their study about the development and contribution of capital income in the UK, USA, and Germany on the decomposition method proposed by Shorrocks (1982) and find that this form of income has a disproportionally high share of explanatory power for overall income inequality.

Wealth

The topic of the decomposition of the wealth distribution on the other hand seems to have gotten less attention, which might be due to the lack of data. This gap can now be filled with the information from the HFCS. One of the few examples is Sierminska and Doorly (2012) who looked at the differences of certain socio-economic characteristics in explaining the wealth portfolio of households in the United States, Germany, Italy, Luxembourg and Spain. This study finds that these socio-economic characteristics can - to a varying extent over the different portfolio components - explain differences in the participation as well as the level of wealth holdings of households across countries. Their study deviates from the exercise presented here in so far that it is not attempted to investigate the contribution of certain components to the overall wealth distribution. Accompanying is a paper about the decomposition of the gender wealth gap for Germany in Sierminska, Frick, and Grabka (2008). On a regression based approach[3] the wealth gap of the two subgroups can be attributed mostly to differences in individuals' characteristics. Even closer related to the investigation here is Azpitarte (2008) who, in his study of the household wealth distribution in Spain, includes a decomposition according to components, finding that in Spain for the year 2002 around 4/5 of the inequality is explained by real assets and only one fifth by financial assets.[4] There are some other attempts concerning the decomposition of wealth; e.g. Brandolini, Cannari, D'Alessio, and Faiella (2004) includes a decomposition with similar results as Azpitarte (2008) for

[3] Using the techniques explained and derived e.g. in Juhn, Murphy, and Pierce (1993) and Firpo, Fortin, and Lemieux (2007).

[4] Debt has a minor impact according to the calculation in Azpitarte (2008). However there seems to be no special treatment for the negativity of the debt components.

Italian data. Also Wolff (2002) uses a decomposition approach of the coefficient of variation to asses the impact of inheritances on the wealth distribution. The paper here, however, is mostly based on Lindner (2011) and further develops as well as extends the results in various ways. First, while the previous study only looked at Austria now we are able to look at 15 euro area countries due to the release of the HFCS data by the ECB. Secondly in addition to the decomposition of only a part of the household balance sheet, namely its focus on financial wealth[5] and income, we can also take real assets into account. Furthermore, this new analysis is not restricted to one particular decomposition method but estimates various potential ways to decompose wealth. So the overlap with the previous investigation is only minor.

4.3 Data

As is laid out in ECB (2013a) and ECB (2013b) the European Central Bank (ECB) coordinated the international effort to gather - under the name Household Finance and Consumption Survey (HFCS) - household level information on the whole balance sheet and the portfolio allocation in the euro area. In this project the central banks (in some countries in connection with the national statistical office) collected previously agreed information at the national level and provided the data over the Household Finance and Consumption Network (HFCN) to the research community. The following 15 countries conducted the first wave of the HFCS and all countries are included in this analysis: Austria (AT), Belgium (BE), Cyprus (CY), Germany (DE), Spain (ES), Finland (FI), Franc (FR), Greece (GR), Italy (IT), Luxembourg (LU), Malta (MT), Netherlands (NL), Portugal (PT), Slovenia (SI), and Slovakia (SK). There are also results for the overall euro area figures - excluding Ireland and Estonia that did not collect data in the first wave of the HFCS - provided. Although

[5]Although Lindner (2011) includes a decomposition of real assets in the appendix, the results on real and financial assets cannot be combined since they originate from two distinct data sources.

there remain some differences,[6] the ex-ante harmonisation not only of the collected information but also the various methods (sampling, weighting, interview mode, imputation, etc.) guarantees a cross country comparability to a very large extent.

All in all more then 62,000 households could be interviewed in this survey with voluntary participation.[7] The field period[8] of the HFCS spans from 2009 to 2011 and all the stock information used here refers to the time of the interview, i.e. the reference period for the information is 2009–2011. At the country level appropriate sampling procedures (mostly 2-stage stratified random sampling and in smaller countries a 1-stage stratified design) were used in the household selection process (see also ECB (2013a) chapter 4).[9] By the ECB provided are final household weights that are used throughout this paper. These weights are based on the sampling design and adjusted for unit non-response as well as post-stratified. The latter process is also country specific and hence varies between countries.[10] As in all voluntary surveys household not only do not participate at all, but also do not know or refuse to answer specific question. In the HFCS considerable effort was put on the correction of this so-called item non-response since the survey covers particularly sensitive topics such as the information on wealth holdings of households. As is proposed in the literature (e.g. see Rubin and Little (2002)) missing information is imputed using a Bayesian-based multiple imputation procedure.[11] Hence the HFCS data come in five multiple imputed implicates of the original data. In the analysis below all five implicates are used in the estimations below such that the estimation of a given statistic, say E, is first calculated in every implicate $a = 1, 2, ..., 5$ to get the statistic

[6] See for example the information concerning interview mode, which is computer assisted personal interview (CAPI) in all countries but two (FI and NL) as is documented in ECB (2013a).

[7] The number of successful participant households in each country ranges from just below 350 in SI to above 15,000 in FR and can be seen in table 4.2.

[8] The exception to this is ES where some households were interviewed already in 2008, see ECB (2013a).

[9] Notable exception are SK which employed a quota sampling technique.

[10] All the information about the country-specific weighting procedures in the HFCS can be found in ECB (2013a) in chapter 5.

[11] Italy has no imputed observations in the HFCS since there is no missing information.

in the given implicate E_a and then averaged over all five implicates, i.e.

$$E = \frac{1}{5} \sum_{a=1}^{5} E_a$$

In this study we use the information on the whole balance sheet in order to generate the (net) wealth estimate of each household. Table 4.1 shows the structure of net wealth based on gross wealth minus debt of a household. Gross wealth can be broadly split into real and financial assets, whereas the separation is along tangible assets on the one hand and all other claims on the other. Business assets, as long a one of the household members works in a leading position in the company is considered to be a real asset since a big proportion of worth is usually in tangible assets (e.g. real estate or vehicles, etc.).

Table 4.1: Composition of Wealth

Net Wealth			
Gross Wealth		**Debt**	
Real Assets	Financial Assets	Mortgage Debt	Non-Mortgage Debt
Primary Housing	Save Financial Assets	With Primary Housing Secured Debt	Overdraft / Credit Card
Further Real Estate	Risky Financial Assets	With Further Real Estate Secured Debt	Consumer Loans
Business Assets			
Other Real Assets	Other Financial Assets		

Notes:
(i) This table describes the components of net wealth as it can be derived from the Eurosystem HFCS 2010.

Real assets are further subdivided into assets held in the primary housing (HMR), further real estate (FRE), business assets (BUS), and other real assets (REST REAL). FRE may contain all forms of real estate such as holiday homes, landed

property, rented out property, but also forests or rented out office buildings. So it is comprised of all real estate owned by the household that is not in HMR (or potentially owned as BUS). BUS is recorded as the net value of business assets that are owned by the household and where at least one household member works in a leading position.[12] The remaining category (REST REAL) contains all other tangible assets; mainly there are all forms of vehicles, but in some cases also other valuables such as collections of pieces of art.

Financial assets are further subdivided into safe assets (SAFE), risky investments (RISKY), and other financial assets (REST FIN). The first component is comprised of sight and savings accounts (including savings in building societies accounts) as well as life insurances and voluntary private pension plans. These savings vehicles are classified as safe assets not because there is no risk attached to them at all, but because the risk attached should be considerable smaller than for other investments.[13] RISKY financial assets subsume shares, bonds, and mutual funds. Fewer households own these investments and they are considered to yield on average a higher return but come accordingly with a higher risk to it. The last category includes all other forms of financial assets such as silent partnerships, managed accounts, or money owed to a household. It potentially also is comprised of wealth hold in trusts. In the questionnaire at the end of both chapters for real and financial assets there is one question each to cover other forms of real or financial assets that may not be covered explicitly in other questions. The wealth recorded in this variable is included in REST REAL and REST FIN respectively.

In the liability side of the household balance sheet all debt can be split into mortgage and non-mortgage debt. The first class is subdivided into mortgages that use either the HMR or FRE as the collateral. Under non-mortgage debt, overdraft in sight accounts and credit card debt over and above the monthly repaid amount are one

[12] Silent partnerships are recorded as financial assets.

[13] To a great extent there is also another distinguishing feature in that these forms of assets are in general much more widely held than the other forms of financial assets.

component, whereas all other forms of consumer loans (including loans for business purposes as long as it is not taken out by the legal entity of the company itself) establish the remaining component.

4.4 Decomposition Methods

In order to describe the distribution of wealth we use several commonly applied measures. After descriptive statistics (minimum, median, mean, and maximum as well as participation rates) for overall wealth and all components, various percentile ratios (P90/P10, P75/P25, and P90/P50), the mean/median ratio, the Gini coefficient, the general entropy index[14] with sensitivity parameter $\alpha = 2$ and the shares of various groups (the top 5%, top 10% and bottom 50% of the wealth distribution) are estimated and discussed. Furthermore, we look at the share of each component in total wealth in order to see the importance. All inequality measures have strength and weaknesses; while a specific percentile ratio only looks at a specific position of the distribution, such as e.g. the P90/P10-ratio for which the distribution in between P11 and P89 is not considered at all, they are robust statistics in the sense that they have a bounded influence function. The chosen measure from the general entropy class is more sensitive to the top whereas the Gini coefficient is more sensitive to the middle of the wealth distribution.[15] Presenting a variety of them we try to capture various parts of the whole distribution.

In the further step of decomposing inequality according to its components we have to restrict ourself to gross wealth in order to include all households and assure that all values are non-negative. Already the inequality measures are generally developed for positive (non-negative) values only, as there are various difficulties in calculating distributional measures if one allows for negative values (e.g. see for example the special treatment of the Gini coefficient in Chen, Tsaur, and Rhai (1982), Berrebi

[14] This is equivalent to half the coefficient of variation.

[15] See e.g. Cowell and Flachaire (2002) and Cowell and Flachaire (2007).

and Silber (1985), and Stich (1996)). Since there is no agreement in the literature on how to adjust distributional measures for these values let alone decompose them, we restrict the decomposition analysis to gross wealth and present the descriptive results both for gross and net wealth and hence also some information for the debt side of the household balance sheet. Additionally, we restrict the decomposition to one index, namely the Gini coefficient,[16] since it is the most widely used measure. Note however, that the natural decomposition based on the assumptions in Shorrocks (1982) is unique and thus applies to all inequality measures. For the readers convenience, the remainder of this section reviews the various decomposition methods analytically. The last paragraph introduces a new application of the influence function, which is used for subgroup decomposition procedures in Fortin, Lemieux, and Firpo (2011), for decomposing the distribution according to components. Before we start, one needs to introduce some notation. We want to estimate the relative/proportional contribution s_k of each wealth component k, where w_i denotes gross wealth of household i and w_{ik} denotes wealth of household i in component k. So we have $w_k = \sum_i w_{ki}$ and $w_i = \sum_k w_{ki}$. Let us also denote total gross wealth by w, hence $w = \sum_k w_k = \sum_k \sum_i w_{ki} = \sum_i w_i$. Furthermore, for several decomposition methods required is the mean of wealth μ and the mean of component k denoted as μ_k. If one ranks the individuals i by their gross wealth holdings w_i the ranks of each is called j, so that both i and j are indexes that run from $1...N$ where the total number of all households is given by N, however, while in i there can be any ordering in j households are ranked in an ascending way according to gross wealth.

4.4.1 Natural Decomposition

Shorrocks (1982) provides this decomposition method which is based on an idea in Fei, Ranis, and Kuo (1978). It starts with the definition of the Gini coefficient

[16] This measure is commonly used and discussed extensively for example in Xu (2004).

$(GC(\bullet))$, which can be written as

$$GC(w) = \frac{2}{N^2\mu}\sum_{j}^{N}(j - \frac{N+1}{2})w_j.$$

This paper also looks at other measures of the income distribution such as variance or the coefficient of variation. Shorrocks (1982) shows based on six (strong) assumptions[17] that the unique decomposition rule is given by

$$s_k = \frac{cov(w_k, w)}{\sigma^2(w)}$$

where $\sigma^2(w)$ is the variance of wealth and thus the relative contribution is given by the coefficient of an OLS regression of each component on gross wealth. This natural decomposition method does in principle not depend on the distributional measure chosen, but needs the assumptions to be fulfilled for this uniqueness.

4.4.2 Additive Decomposition

Using the representation of the GC[18] based on the co-variance Lerman and Yitzhaki (1985) use the additive separability of the covariance in order to calculate the relative contribution of each component.[19] Take

$$GC(w) = \frac{2}{\mu}cov[w, F(w)]$$

[17] These assumption are (1) continuity and symmetry of the inequality measure as well as the requirement that it is zero if all individuals own an equal amount; (2) continuity and symmetry of the decomposition method; (3) independence of the aggregation level; (4) summation of all contributions equal the measure of inequality; (5) independence of numeration of individuals and contribution is zero if all individuals own an equal amount in a specific source; and (6) two factor components symmetry.

[18] See Yitzhaki (1998) for a variety of possibilities to calculate the Gini coefficient.

[19] Note that this is the exact same decomposition method that is also applied in Lindner (2011).

where $F(w)$ is the cumulative distribution of wealth and apply the additive separability, so that the Gini coefficient can be written as

$$GC = \frac{2}{\mu} \sum_{k=1}^{K} cov[w_k, F(w)].$$

This sum can be rearranged to be the sum over all components of the product of the relative contribution[20] of each component $s_k = \frac{cov(w_k, F(w))}{cov(w_k, F(w_k))}$, the GC of the component k denoted as gc_k, and the share of component k in total wealth $a_k = \frac{\mu_k}{\mu}$, so that the overall GC is

$$GC(w) = \sum_{k=1}^{K} \frac{cov[w_k, F(w)]}{cov[w_k, F(w_k)]} \frac{2cov[w_k, F(w_k)]}{\mu_k} \frac{\mu_k}{\mu}$$

$$GC(w) = \sum_{k=1}^{K} s_k gc_k a_k.$$

Furthermore, Lerman and Yitzhaki (1985) show that the local elasticity of the distribution with respect to component k, i.e. the relative partial derivative with respect to a small variation denoted as ε_k^* can be written as

$$\varepsilon_k^* := \frac{\frac{\partial GC(w)}{\partial \varepsilon_k}}{GC(w)} = \frac{a_k s_k gc_k}{GC(w)} - a_k.$$

One additional decomposition method based on GC is discussed in the early literature by Silber (1989). This approach, however, is - although it is often taken as a separate decomposition technique - empirically equivalent to the decomposition above and thus is only provided in appendix 4.A for completeness.

[20] In contrast to the natural decomposition from above here the covariance between a component and the distribution of total wealth instead of total wealth itself is taken and divided by the covariance of a component with its own distribution instead of the variance of total wealth.

4.4.3 Shapley Value Decomposition

The Shapley value decomposition stems from results in game theory on how to share a profit in a cooperative game (it goes back to Shapley (1953)). Basically the idea is that the players share the benefits from a coalition and each player gets the marginal contribution that this player makes to it, letting everybody potentially form new coalitions. In the decomposition analyses the application of this idea means nothing less than dropping one of the $k = 1, 2, ..., K$ components of wealth each at a time but taking into account all possible subsets of components of total wealth and weighting them appropriately. In our case. for example, not only the distribution of total wealth with and without the primary housing, but also the distribution of all possible subsets of the seven components that include primary housing with and without this component are taken into account. This approach has recently gotten new attention and is discussed in more detail in Shorrocks (2013) and Chantreuil and Trannoy (2013). Furthermore, one also has to mention an earlier work of one of the authors from these articles, namely Sastre and Trannoy (2002). The calculation is as follows.

From combinatoric we know that in total there are $Z = 2^K - 1$ subsets of wealth components, where the subtraction of 1 corresponds to the excluded empty set. This means that there are a total of $Z = 127$ potential subgroups in the empirical set-up of this paper with seven components of wealth. The contribution s_k of each component k is then given by the weighted sum of the differences of inequality in a subgroup including component k and this subgroup excluding k. The weighting factor captures how many potential subgroups there are of given size. So the contribution of component k can be stated as

$$s_k = \sum_{\substack{z \subseteq Z \\ k \in z}} \frac{(z^* - 1)!(K - z^*)!}{K!} [GC(w(z)) - GC(w(z - k))]$$

where the sum runs over all subsets z of possible subsets Z, that include compo-

nent k, which is denoted as $z \subseteq Z \; k \in z$. The number of components in the specific subset is denoted as z^* and K is the total number of components (i.e. seven in this analysis). $GC(w(z))$ is the GC of the subset z that includes component k whereas $GC(w(z-k))$ denotes exactly this subset z but excluding component k. Hence this decomposition provides a weighted (by the number of possibilities) average of the change when component k is excluded from all possible subsets of the components of gross wealth. The contributions may take on positive as well as negative values and sum up to the inequality measure under consideration.

4.4.4 Decomposition using the Influence Function

More recently the application of the influence function[21] has been popular also in studies investigating the decomposition of the income distribution. However, as in Fortin, Lemieux, and Firpo (2011) and Firpo, Fortin, and Lemieux (2007), this method was so far only used for sub-group decomposition and not as we want to investigate here the decomposition according to components.

Despite this lack of studies, it seems fairly straightforward to look at the marginal contribution of components to the overall distribution of wealth. Recall that gross wealth $w_i = \sum_{k=1}^{K} w_{ik}$, i.e. it is the sum over all components. One can then construct counterfactual distributions corresponding to a proportional increase of the various components k of total wealth w, such that

$$w(\alpha) := \sum_{k=1}^{K} (1 + \alpha_k) \cdot w_k,$$

and the corresponding inequality indices $\nu(w(\alpha))$. The marginal contribution of component k to inequality can then be defined as

$$s_k := \frac{\partial \nu(w(\alpha))}{\partial \alpha_k}.$$

[21] For this approach see Fortin, Lemieux, and Firpo (2011), Fortin, Lemieux, and Firpo (2009), and Firpo, Fortin, and Lemieux (2007).

If ν has a differentiable recentered influence function[22] (RIF), we can write this as

$$s_k = E[\frac{\partial RIF(\nu(w))}{\partial w} \cdot w_k].$$

An additional appealing fact is that if ν is scale invariant, i.e. homogeneous of degree zero, such that $\nu(\kappa \cdot \alpha) = \nu(\alpha)$ for any constant κ the marginal contributions sum up to zero, so that $\sum_k s_k = 0$. Applying this idea and taking the recentered influence function[23] of the GC from Essama-Nssah and Lambert (2011) given by

$$RIF(GC(w)) = -\frac{w}{\mu}GC + 1 - \frac{w}{\mu} + \frac{2}{\mu}\int_0^w F(x)dx$$

where all the parameters are as before and $F(x)$ is the cumulative distribution function of wealth. Taking the derivative with respect to w yields

$$\frac{\partial RIF(GC(w))}{\partial w} = -\frac{GC}{\mu} - \frac{1}{\mu} + \frac{2}{\mu}F(w).$$

Since gross wealth is the sum of k components the expectation of this derivative of the RIF plugging in each component gives us the contribution of the respective component. So it is given by

$$s_k = E[(-\frac{GC}{\mu} - \frac{1}{\mu} + \frac{2}{\mu}F(w)) \cdot w_k]$$

As we have survey data in the empirical part the expectation naturally is weighted. One advantage of this decomposition method compared to the natural and additive decomposition is that it does not rely an the six strong assumption in Shorrocks (1982) and is in principle applicable to many measures of inequality. In comparison to the Shapley value decomposition it is computationally less intensive.

[22] One could also use the influence function itself, since the recentring only adds a constant which disappears if one takes the derivative.

[23] This is the sum of the influence function and the statistic itself.

4.5 Results

4.5.1 Descriptive Statistics

In the HFCS 2010 of the Eurosystem gross wealth in the euro area ranges from zero to above €400 million (see table 4.2).[24] The overall median household owns investments just above €140 thousand while the average is close to €260 thousand.

Table 4.2: Descriptive Statistics of Gross Wealth in EA-Countries

	No. Obs.	Median (in k €)	Mean (in k €)	Minimum	Maximum (in k €)
EA	62,521	142	257	0	409,417
AT	2,380	93	282	0	21,948
BE	2,327	250	369	0	8,408
CY	1,237	332	742	0	51,627
DE	3,565	68	222	0	76,305
ES	6,197	210	324	0	409,417
FI	10,989	133	198	1	14,751
FR	15,006	150	258	81	152,506
GR	2,971	110	160	0	11,700
IT	7,951	188	287	0	26,127
LU	950	494	792	0	34,080
MT	843	227	378	0	30,851
NL	1,301	217	252	0	4,901
PT	4,404	93	170	0	27,096
SI	343	105	154	0	1,937
SK	2,057	64	83	0	1,223

Notes:
(i) This table shows for each country and the overall euro area the number of observations, median (in thousand €), mean (in thousand €), the minimum and maximum (the latter in thousand €) observation of gross wealth.
(ii) Country abbreviations: Euro area (EA), Austria (AT), Belgium (BE), Cyprus (CY), Germany (DE), Spain (ES), Finland (FI), France (FR), Greece (GR), Italy (IT), Luxembourg (LU), Malta (MT), Netherlands (NL), Portugal (PT), Slovenia (SI), and Slovakia (SK).
(iii) All estimates are weighted using appropriate household weights provided by the data producer and the multiple imputation structure is taken into account.
(iv) *Source:* Eurosystem HFCS 2010.

[24] See tables 4.9 in the appendix for the descriptive statistics of net wealth from the HFCS. Naturally, there are households with negative net wealth and the overall mean and median are substantially lower. More information on the debt side of the household balance sheet and the share of each component in term of net wealth rather then gross wealth as it is discussed in the paper can be found in the appendix in tables 4.11 and 4.10.

As was already mentioned in the introduction, there is considerable heterogeneity among countries. So, while the median gross household wealth in e.g. Germany and Slovakia is around €65 thousand the according figures for Cyprus and Luxembourg are around €330 and €490 thousand. Also one can see that in various countries the top of the wealth distribution is covered to a different extent.[25] As the exercise in this paper is focused on the decomposition of inequality table 4.2 gives a first idea about cross country differences in this respect if one looks at the mean to median gross wealth ratio (more on that below).
Gross wealth comprises several components, which can be seen in table 4.3. Figures 4.1a and 4.1b provide the share of gross wealth that is held in each portfolio component for real and financial assets respectively.[26]

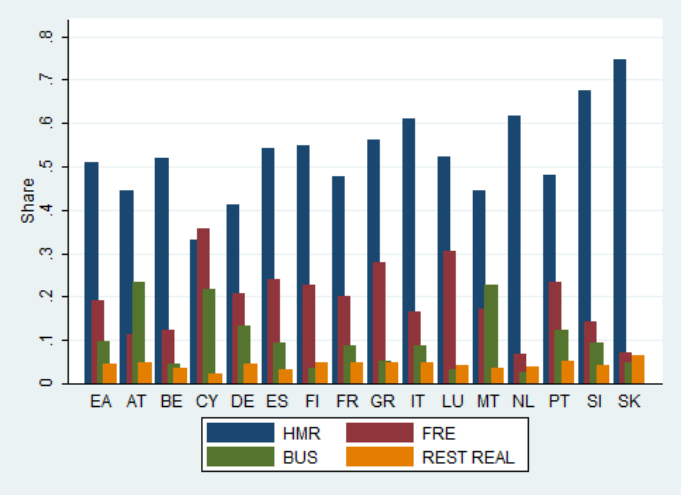
From figures 4.1a and 4.1b one can see that the highest share of wealth is held in form of the households main residence. In the euro area the figure is estimated to be 50.6.%, it ranges from a relatively low 32.8% in Cyprus to almost three quarter of the wealth in Slovakia. With the exception of the lowest share found (i.e. CY) it is in all countries the form of wealth where the highest share is invested. This finding already points towards a heavy impact of this component on the distribution of gross wealth. However, also participation rates are generally fairly high and the distribution of wealth hold in the HMR is generally less skewed than for other portfolio instruments, such as business assets, where - although participation of rather low - a fair amount of wealth is hold and the distribution is relatively more skewed (result not shown in a table).

Furthermore, an inspection of the cross country differences in participation rates for the HMR (see table 4.3) reveals that in most countries, i.e. all with the exception

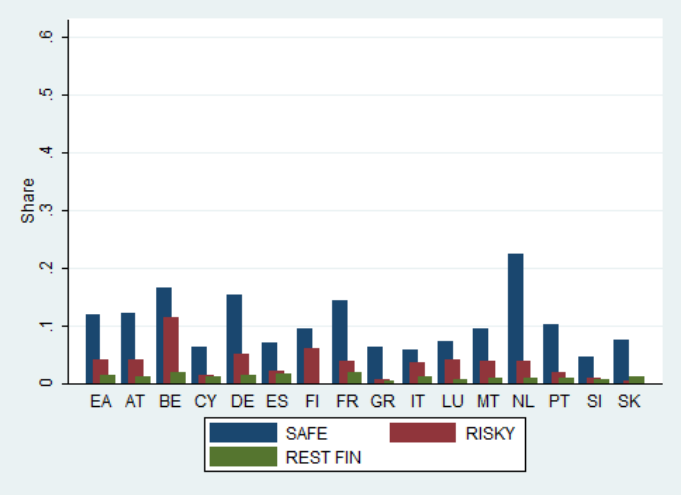
[25] Among other things, this might also be due to differences in sampling designs for each country. More details can be found in ECB (2013a).
[26] Again, additional results for the liability side and the share of components relative to net wealth can be found in the appendix in table 4.10 and 4.11 respectively.

Figure 4.1: Share of Totals of Asset Components in terms of Gross Wealth

(a) Real Asset Components



(b) Financial Asset Components



Notes:
(i) Graph (a) shows the share of the aggregate wealth held in each real asset component in terms of gross wealth.
(ii) Graph (b) shows the share of the aggregate wealth held in each financial asset component in terms of gross wealth.
(iii) Country abbreviations: Euro area (EA), Austria (AT), Belgium (BE), Cyprus (CY), Germany (DE), Spain (ES), Finland (FI), France (FR), Greece (GR), Italy (IT), Luxembourg (LU), Malta (MT), Netherlands (NL), Portugal (PT), Slovenia (SI), and Slovakia (SK).
(iv) All estimates are weighted using appropriate household weights provided by the data producer and the multiple imputation structure is taken into account.
(v) *Source:* Eurosystem HFCS 2010.

of DE and AT, is above 50% and hence the median households in terms of gross wealth in these countries own its HMR. For the two exceptions the median gross wealth level is relatively low (see table 4.2), but the conditional median of the owners of the HMR is in line with what is found for other countries (see table 4.3). In AT the median home-owner holds investments equal to €200 thousand in this asset form which is equal to the according estimate for IT and compares to the euro area figure of about €180 thousand and €168 thousand in DE. Safe financial assets are owned by almost all households in most of the countries and since it is so widespread the relatively low conditional median translates in a overall share of gross wealth of about 12%, which is about the same as the estimate for self-employed business assets. Especially in the Netherlands and Belgium (also DE and FR) a relative high share is invested in the former portfolio category (SAFE). On the other hand, risky financial assets as well as the other financial assets are confined to a very much lower proportion of households in all countries. There are, however, huge cross country differences. So e.g. only 4% of households in GR do take part in these kind of financial transactions whereas the estimate for Malta is about 33.7% or even higher for CY and FI with 36.3% and 38.7% respectively.

Table 4.3: Components of Gross Wealth in EA-Countries

	HMR		FRE		BUS		REST REAL		SAFE		RISKY		REST FIN	
	Part.	Median (in k €)	Part.	Median (in k €)	Part.	Median (in k €)	Part.	Median (in k €)	Part.	Median (in k €)	Part.	Median (in k €)	Part.	Median (in k €)
EA	0.601	180	0.231	103	0.111	30	0.841	7	0.967	9	0.202	12	0.125	3
AT	0.477	200	0.134	94	0.094	181	0.799	9	0.994	12	0.146	12	0.114	3
BE	0.696	250	0.164	174	0.066	50	0.802	8	0.979	21	0.307	20	0.106	4
CY	0.767	240	0.516	202	0.195	99	0.895	10	0.859	18	0.363	2	0.101	9
DE	0.442	168	0.178	115	0.091	19	0.732	8	0.991	13	0.230	12	0.223	3
ES	0.827	180	0.362	120	0.142	51	0.799	7	0.982	5	0.140	12	0.080	7
FI	0.692	128	0.298	108	0.138	1	0.679	9	1.000	6	0.387	4	0.000	.
FR	0.553	194	0.247	116	0.089	53	1.000	4	0.996	9	0.217	8	0.121	4
GR	0.724	100	0.379	62	0.098	36	0.735	6	0.739	4	0.040	7	0.041	3
IT	0.687	200	0.249	100	0.180	15	0.951	10	0.919	7	0.198	22	0.049	10
LU	0.671	500	0.282	300	0.052	98	0.884	19	0.984	23	0.258	28	0.088	5
MT	0.777	187	0.314	120	0.115	137	0.875	7	0.969	18	0.337	22	0.060	9
NL	0.571	240	0.061	166	0.048	52	0.826	7	0.973	30	0.239	8	0.107	2
PT	0.715	90	0.271	53	0.077	47	0.741	5	0.943	4	0.065	9	0.085	4
SI	0.818	111	0.232	52	0.116	25	0.804	3	0.936	1	0.203	3	0.067	4
SK	0.899	56	0.153	16	0.107	5	0.685	5	0.915	2	0.041	1	0.103	1

Notes:
(i) This table shows for each asset component (main residence [HMR], further real estate [FRE], self-employed business assets [BUS], and rest real assets [REST REAL], safe financial assets [SAFE], risky financial assets [RISKY], and rest financial assets [REST FIN]) participation rate (Part.) and conditional median in the respective component.
(ii) Country abbreviations: Euro area (EA), Austria (AT), Belgium (BE), Cyprus (CY), Germany (DE), Spain (ES), Finland (FI), France (FR), Greece (GR), Italy (IT), Luxembourg (LU), Malta (MT), Netherlands (NL), Portugal (PT), Slovenia (SI), and Slovakia (SK).
(iii) All estimates are weighted using appropriate household weights provided by the data producer and the multiple imputation structure is taken into account.
(iv) *Source:* Eurosystem HFCS 2010.

Table 4.4 reports various indicators for the euro area and country specific distribution of gross wealth. As is well known, table 4.4 shows that the distribution

Table 4.4: Inequality Measures of Gross Wealth in EA-Countries

	Ratios				Ineq. Measures		Shares		
	P90/ P10	P75/ P25	P90/ P50	Mean/ Median	Gini	GE2	Top 5%	Top 10%	Bottom 50%
EA	183.0	15.2	3.9	1.8	0.645	4.289	0.348	0.476	0.078
AT	233.7	22.4	6.2	3.0	0.734	4.014	0.455	0.588	0.039
BE	206.5	7.7	3.1	1.5	0.573	1.139	0.293	0.415	0.126
CY	107.8	5.2	4.9	2.2	0.662	2.652	0.397	0.534	0.092
DE	449.1	25.7	7.3	3.3	0.725	4.719	0.424	0.557	0.038
ES	233.7	22.4	6.2	1.5	0.542	6.795	0.290	0.410	0.155
FI	221.5	15.0	3.4	1.5	0.583	1.371	0.268	0.398	0.107
FR	233.7	22.4	6.2	1.7	0.651	5.446	0.345	0.476	0.069
GR	94.1	5.7	3.2	1.4	0.539	0.730	0.242	0.373	0.138
IT	233.7	22.4	6.2	1.5	0.599	1.722	0.314	0.441	0.111
LU	145.2	6.9	3.0	1.6	0.614	2.757	0.371	0.479	0.118
MT	39.0	4.5	3.2	1.7	0.592	5.788	0.344	0.456	0.129
NL	96.4	9.6	2.4	1.2	0.514	0.569	0.206	0.324	0.137
PT	164.0	6.7	3.6	1.8	0.635	5.775	0.375	0.494	0.102
SI	62.6	4.6	3.1	1.5	0.521	0.647	0.214	0.352	0.148
SK	10.4	2.5	2.4	1.3	0.436	0.526	0.213	0.322	0.215

Notes:
(i) This table shows for each country and the overall euro area various frequently used measures of the distribution of gross wealth.
(ii) Country abbreviations: Euro area (EA), Austria (AT), Belgium (BE), Cyprus (CY), Germany (DE), Spain (ES), Finland (FI), France (FR), Greece (GR), Italy (IT), Luxembourg (LU), Malta (MT), Netherlands (NL), Portugal (PT), Slovenia (SI), and Slovakia (SK).
(iii) All estimates are weighted using appropriate household weights provided by the data producer and the multiple imputation structure is taken into account.
(iv) *Source:* Eurosystem HFCS 2010.

of gross wealth is more unequal than what is known from income distributions. This holds not only for the euro area but also in the 15 countries under investigation.[27] Considering as a measure of the skewness of the wealth distribution the mean/median ratio, one can see that most countries display a relatively similar pattern with the statistic ranging from 1.5 to 1.8. NL and SK seem to be somewhat more equal while AT and DE (to some extent also CY) are the countries that are

[27]SK could be argued to be an exemption to this finding with a GC of 0.44 which is close to what is found for the income distribution in some countries.

most unequal with estimates of 3.0 and 3.3 respectively. This refers back to whether the median household in terms of gross wealth owns its HMR, which is not the case in the latter two countries mentioned. In other parts of the distribution, however, cross country differences are found to be more pronounced. The interquartile ratio (P75/P25) is about 15 for the euro area and ranges from below 5 in some countries to more than 25 in DE. The discrepancies between countries increase if we move towards the tails of the distribution (looking e.g. at the P90/P10 ratio) where estimates range from 10 to almost 450. Although there are differences between the various measures of the distribution the general pattern - namely that there are cross country differences and that some countries exhibit a relatively high level of inequality (namely AT and DE) whilst others a relatively low level (e.g. SI and SK but also in some parts of the distribution ES) - can be found with (almost) any measure. The following part tries to give some explanation where this differences in the distribution of gross wealth stem from by decomposing the overall distribution according to its components.

4.5.2 Natural Decomposition

As we have seen in section 4.4.1, one of the ideas to decompose the distribution according to its components is based on the covariance of a component and overall wealth. Table 4.5 displays the results of this exercise. Since the components sum up to gross wealth the estimates in table 4.5 sum up to 1. Overall in the euro area it can be seen that the distribution of wealth is coming mainly from real assets while the financial side accounts for only about 13%. This might be due to the fact that the relative share of financial assets is rather low and even in the most widely used asset component - save financial assets - the distribution is relatively more even. Within the real asset side the highest proportion is estimated for self-employed business assets followed by further real estate, since relatively few household own these asset types but if it is owned a substantial amount is invested, i.e. the distribution of

Table 4.5: Natural Decomposition of Gross Wealth

	HMR	FRE	BUS	REST REAL	SAFE	RISKY	REST FIN
EA	0.113	0.303	0.444	0.011	0.044	0.043	0.041
AT	0.211	0.172	0.530	0.014	0.032	0.031	0.009
BE	0.228	0.171	0.119	0.017	0.187	0.252	0.026
CY	0.076	0.299	0.580	0.005	0.021	0.011	0.008
DE	0.121	0.285	0.500	0.011	0.038	0.029	0.016
ES	0.051	0.415	0.452	0.006	0.017	0.021	0.038
FI	0.245	0.234	0.205	0.020	0.046	0.250	.
FR	0.076	0.248	0.426	0.014	0.076	0.070	0.090
GR	0.317	0.487	0.086	0.025	0.070	0.014	0.002
IT	0.314	0.235	0.337	0.018	0.029	0.044	0.023
LU	0.247	0.639	0.054	0.015	0.017	0.026	0.002
MT	0.051	0.088	0.844	0.007	0.005	0.003	0.001
NL	0.479	0.224	0.025	0.020	0.167	0.074	0.011
PT	0.057	0.182	0.644	0.008	0.065	0.043	0.001
SI	0.381	0.276	0.266	0.024	0.034	0.007	0.010
SK	0.550	0.147	0.180	0.051	0.044	0.003	0.024

Notes:
(i) This table shows for each country and the overall euro area the decomposition of gross wealth according to its components (main residence [HMR], further real estate [FRE], self-employed business assets [BUS], rest real assets [REST REAL], safe financial assets [SAFE], risky financial assets [RISKY], and rest financial assets [REST FIN]) following the approach of Shorrocks (1982).
(ii) Country abbreviations: Euro area (EA), Austria (AT), Belgium (BE), Cyprus (CY), Germany (DE), Spain (ES), Finland (FI), France (FR), Greece (GR), Italy (IT), Luxembourg (LU), Malta (MT), Netherlands (NL), Portugal (PT), Slovenia (SI), and Slovakia (SK).
(iii) Statistics denoted with a "." depicts countries, i.e. FI, where certain components are not collected.
(iv) All estimates are weighted using appropriate household weights provided by the data producer and the multiple imputation structure is taken into account.
(v) *Source:* Eurosystem HFCS 2010.

these components is highly skewed.

In each country the overall picture that emerges is that only BE and NL exhibit a higher proportion in the estimates for SAFE and RISKY (the latter only for the former country) financial assets. Accordingly their contributions of the real asset components are somewhat lower. Within the real asset side there seems to be a high variability of the contribution of BUS, so while in MT this asset type explains 84% of the overall distribution it is only about 3% in NL. However, also the HMR displays different levels; Looking at countries with high ownership rates of the HMR like ES and CY where almost 83% and 77% respectively, the estimate is only 5% and 8%, whilst higher levels are found in other countries. Still ownership rates are not the only factor driving the results as can be seen in the case of SK where ownership is high at almost 90% and the contribution estimate for HMR is 55%. So the estimation is based on the participation in and the within distribution of a certain component as well as the correlation with overall wealth.

4.5.3 Additive Decomposition

Table 4.6 displays the results from the method laid out in section 4.4.2. The contribution s_k as well as the elasticity of the distribution with respect to a small change of one component ε_k^* are shown. The contributions sum up to overall inequality measured by the GC and are bounded from below at zero.[28]

The separation of the contributions of real and financial asset components in this decomposition procedure is similar to what was found in the natural decomposition, that is the contribution of the financial components is relatively low. The exceptions again are BE and NL with estimates of 0.09 and 0.08 (which translates to about 16% of total inequality as can be seen in the relative contributions in table 4.13 in the appendix). On the specific components of the real asset side, however, the

[28]The appendix provides in table 4.13 the relative contributions that sum up to 1.

Table 4.6: Decomposition of Gross Wealth following Lerman and Yitzhaki (1985)

	HMR		FRE		BUS		REST REAL		SAFE		RISKY		REST FIN	
	s_k	ε_k^*	s_k	ε_k^*	s_k	ε_k^*	s_k	ε_k^*	s_k	ε_k^*	s_k	ε_k^*	s_k	ε_k^*
EA	0.287	-0.061	0.153	0.048	0.087	0.039	0.018	-0.014	0.060	-0.022	0.030	0.008	0.010	0.002
AT	0.297	-0.037	0.093	0.017	0.221	0.066	0.022	-0.017	0.063	-0.034	0.032	0.006	0.006	-0.001
BE	0.231	-0.113	0.092	0.042	0.038	0.024	0.012	-0.009	0.092	-0.002	0.094	0.053	0.014	0.006
CY	0.148	-0.105	0.266	0.047	0.196	0.081	0.007	-0.010	0.031	-0.014	0.008	0.000	0.007	0.001
DE	0.277	-0.027	0.179	0.041	0.125	0.042	0.020	-0.014	0.080	-0.040	0.035	0.001	0.008	-0.002
ES	0.211	-0.150	0.176	0.088	0.078	0.054	0.011	-0.010	0.038	0.002	0.015	0.008	0.013	0.008
FI	0.277	-0.071	0.166	0.060	0.030	0.019	0.019	-0.014	0.045	-0.015	0.046	0.021	.	.
FR	0.265	-0.068	0.160	0.046	0.076	0.032	0.023	-0.011	0.084	-0.011	0.030	0.009	0.013	0.004
GR	0.250	-0.097	0.201	0.095	0.034	0.015	0.016	-0.017	0.033	0.000	0.005	0.003	0.001	0.000
IT	0.327	-0.062	0.129	0.052	0.071	0.035	0.015	-0.021	0.024	-0.015	0.025	0.007	0.008	0.004
LU	0.256	-0.101	0.253	0.109	0.026	0.013	0.016	-0.011	0.031	-0.020	0.029	0.010	0.002	0.000
MT	0.185	-0.129	0.125	0.044	0.212	0.133	0.013	-0.010	0.031	-0.038	0.020	-0.002	0.006	0.003
NL	0.319	0.008	0.056	0.044	0.012	0.002	0.011	-0.013	0.083	-0.059	0.027	0.015	0.005	0.002
PT	0.233	-0.110	0.193	0.074	0.112	0.057	0.022	-0.014	0.057	-0.011	0.013	0.005	0.004	-0.001
SI	0.291	-0.114	0.103	0.058	0.083	0.067	0.015	-0.011	0.022	-0.001	0.004	0.000	0.003	0.001
SK	0.289	-0.081	0.048	0.043	0.040	0.046	0.024	-0.005	0.027	-0.010	0.002	0.001	0.006	0.006

Notes:
(i) This table shows for each country and the overall euro area the financial asset side of the decomposition of gross wealth according to its components (main residence [HMR], further real estate [FRE], self-employed business assets [BUS], rest real assets [REST REAL], safe financial assets [SAFE], risky financial assets [RISKY], and rest financial assets [REST FIN]) following the approach of Lerman and Yitzhaki (1985).
(ii) For each component the first column reports the contribution (s_k) of this component to inequality and the second column reports the elasticity (ε_k^*) of the Gini coefficient with respect to the respective component.
(iii) Country abbreviations: Euro area (EA), Austria (AT), Belgium (BE), Cyprus (CY), Germany (DE), Spain (ES), Finland (FI), France (FR), Greece (GR), Italy (IT), Luxembourg (LU), Malta (MT), Netherlands (NL), Portugal (PT), Slovenia (SI), and Slovakia (SK).
(iv) Statistics denoted with a "." depicts countries, i.e. FI, where certain components are not collected.
(v) All estimates are weighted using appropriate household weights provided by the data producer and the multiple imputation structure is taken into account.
(vi) *Source:* Eurosystem HFCS 2010.

results present a different picture. Whereas the main component in the natural decomposition was BUS, it is the HMR following this approach. So much so that from a GC of gross wealth of 0.645 the HMR contributes 0.287 or almost half of the total measure. The reason for this discrepancy is due to the computation. In the natural decomposition the driving factor is the covariance between a component and total gross wealth while here it is the covariance of the component with the *distribution* of gross wealth. So, since the household hold the biggest share of wealth in the HMR the contribution of this factor is higher and hence this procedure follows relatively closely the relative share that is held in a certain component. This can also be seen by the two extreme cases. On the one hand households in CY hold only about a third of total gross wealth in the HMR the (relative) contribution is (22%) 0.148 and on the other hand in SK households hold almost three quarters of total wealth in the HMR and the corresponding (relative) contribution is (66%) 0.289.

Turning towards the estimates for the elasticity of the distribution with respect to its components (columns headed by ε_k^*), one can see that two distinct sets of asset components emerge. On the one hand HMR, REST REAL and SAFE exhibit a negative elasticity and on the other hand the remaining (FRE, BUS, RISKY and REST FIN) show a positive estimate. This means that for the first set a cet. par. 10% increase in wealth hold in a specific component would reduce inequality measured by the GC by the respective estimate, e.g. for the component of safe financial assets in the euro area by 0.22%. This indicates the impact on the distribution of changes in the interest rate - mentioned in the introduction - over the last years. Real interest rates recently have been negative resulting in a reduction of real wealth in safe financial assets, for which the result here points towards an increasing tendency in the overall inequality of the distribution of wealth. The highest equalizing effect is found for the HMR while on the other side of the spectrum in the euro area FRE displays the highest positive estimate for this elasticity. Also there are huge cross

country differences. For the HMR component, the results range from a positive 0.008 in the NL to -0,15 in ES. Also for BUS one can see a high elasticity of 0.13 in MT and relatively low figures such as 0.015 in GR. For the other components cross country differences are less pronounced. The calculations were additionally done with the decomposition described in appendix 4.A and - as described - the results are identical to table 4.6 in the columns containing the relative contributions and headed by s_k and thus are not repeated.

4.5.4 Shapley Value Decomposition

In contrast to the decomposition procedures from above that are ultimately based on the breakdown of the covariance between the components and total wealth, this subsection provides the results from the approach based on the Shapley value. Here the marginal contribution is based - as was laid out in section 4.4.3 - on the average of all possible combination of potential subsets of components and the difference that each specific component contributes there. Table 4.7 reports the absolute values from the calculation, the relative estimates giving the percentage of the distribution explained by a component can be found in table 4.14 in the appendix. This implies the sum of the contributions add up to the GC and - as mentioned above - the estimates can take on negative as well as positive values.

The pattern that emerges from this exercise is somewhat different to what is found above. Overall in the euro area the lowest contributions can be found in the components that held most widely by the households. In the case of the HMR the contribution is actually negative (-0.067) meaning that without this component the distribution of wealth would be less equal. The components SAFE and REST REAL also display a relatively small but positive contribution. Especially for the latter one this seems to go back to the widely distributed usage of cars and thus the high ownership rates. RISKY and REST FIN financial assets contribute, in contrast to

Table 4.7: Shapley Value Decomposition of Gross Wealth (absolute contribution)

	HMR	FRE	BUS	REST REAL	SAFE	RISKY	REST FIN
EA	-0.067	0.148	0.172	0.058	0.053	0.140	0.142
AT	0.006	0.151	0.200	0.058	0.037	0.144	0.139
BE	-0.146	0.144	0.155	0.070	0.050	0.157	0.144
CY	-0.078	0.121	0.210	0.052	0.077	0.136	0.144
DE	0.021	0.162	0.185	0.068	0.023	0.133	0.133
ES	-0.192	0.133	0.171	0.058	0.083	0.142	0.148
FI	-0.072	0.152	0.182	0.074	0.077	0.169	.
FR	-0.067	0.140	0.164	0.064	0.066	0.142	0.144
GR	-0.145	0.128	0.144	0.041	0.088	0.142	0.141
IT	-0.095	0.158	0.171	0.024	0.058	0.136	0.147
LU	-0.113	0.193	0.156	0.051	0.046	0.140	0.140
MT	-0.129	0.140	0.236	0.067	0.022	0.112	0.144
NL	-0.083	0.164	0.141	0.045	-0.030	0.138	0.140
PT	-0.129	0.160	0.183	0.066	0.073	0.145	0.137
SI	-0.209	0.141	0.173	0.058	0.094	0.126	0.138
SK	-0.244	0.146	0.162	0.044	0.046	0.141	0.142

Notes:
(i) This table shows for each country and the overall euro area the decomposition of gross wealth according to its components (main residence [HMR], further real estate [FRE], self-employed business assets [BUS], rest real assets [REST REAL], safe financial assets [SAFE], risky financial assets [RISKY], and rest financial assets [REST FIN]) following the approach of Shorrocks (2013) and Sastre and Trannoy (2002).
(ii) Country abbreviations: Euro area (EA), Austria (AT), Belgium (BE), Cyprus (CY), Germany (DE), Spain (ES), Finland (FI), France (FR), Greece (GR), Italy (IT), Luxembourg (LU), Malta (MT), Netherlands (NL), Portugal (PT), Slovenia (SI), and Slovakia (SK).
(iii) Statistics denoted with a "." depicts countries, i.e. FI, where certain components are not collected.
(iv) All estimates are weighted using appropriate household weights provided by the data producer and the multiple imputation structure is taken into account.
(v) *Source:* Eurosystem HFCS 2010.

the results above, almost as much to inequality as FRE (each component for the euro area has an estimate of about 0.14 to 0.15) and only the component BUS, which is held by only a small minority of households, contributes marginally relatively more. Apart from the HMR cross country differences are relatively small in absolute terms. Since GCs are different in different countries (see table 4.4) the differences in relative marginal contribution are more pronounced. In the case of the HMR differences seem to reflect the differences in ownership rates. There are positive estimates for AT and DE - the two countries with the lowest ownership rates - and all the remaining countries exhibit a negative contribution with the highest values found for SK, SI and ES which are also the countries with the highest participation rates for the HMR.

4.5.5 Decomposition using the Influence Function

The results of the newly proposed method - based on the recentered influence function - of decomposing a distribution according to components and hence investigate the marginal contribution is presented in table 4.8.

We find that the (qualitative) results very closely match what is found in table 4.6. So that in all countries but NL the wealth held in form of the HMR inhibits a negative marginal contribution. Also REST REAL and SAFE financial assets are found to have negative sign, but a much smaller point estimate than the HMR. All the other other components with the exception of RISKY financial assets in LU as well as the REST FIN component in AT, DE, and PT show a positive result. In general estimates for financial components are far smaller than their real asset counterparts. This again reflects the importance of real assets in the household balance sheet. Considering one of the motivating examples in the introduction, one can see that in the euro area a policy that fosters the existing stock of the households main residence in general should cushion the inequality of wealth. Households in different countries, however, are affected to a varying degree. So would the impact

Table 4.8: RIF-based Decomposition of Gross Wealth

	HMR	FRE	BUS	REST REAL	SAFE	RISKY	REST FIN
EA	-0.039	0.031	0.025	-0.009	-0.014	0.005	0.001
AT	-0.027	0.012	0.049	-0.012	-0.025	0.004	-0.001
BE	-0.065	0.024	0.014	-0.005	-0.001	0.031	0.004
CY	-0.069	0.031	0.054	-0.007	-0.009	0.000	0.001
DE	-0.019	0.030	0.030	-0.010	-0.029	0.001	-0.001
ES	-0.081	0.048	0.029	-0.006	0.001	0.004	0.005
FI	-0.041	0.035	0.011	-0.008	-0.009	0.012	.
FR	-0.044	0.030	0.021	-0.007	-0.007	0.006	0.003
GR	-0.051	0.052	0.008	-0.009	0.000	0.001	0.000
IT	-0.037	0.031	0.021	-0.012	-0.009	0.004	0.003
LU	-0.062	0.067	0.008	-0.007	-0.012	0.006	0.000
MT	-0.076	0.027	0.079	-0.006	-0.023	-0.001	0.002
NL	0.005	0.023	0.001	-0.007	-0.030	0.008	0.001
PT	-0.070	0.047	0.036	-0.009	-0.007	0.003	-0.000
SI	-0.055	0.031	0.036	-0.005	-0.000	0.000	0.000
SK	-0.034	0.019	0.020	-0.002	-0.004	0.001	0.003

Notes:
(i) This table shows for each country and the overall euro area the decomposition of gross wealth according to its components (main residence [HMR], further real estate [FRE], self-employed business assets [BUS], rest real assets [REST REAL], safe financial assets [SAFE], risky financial assets [RISKY], and rest financial assets [REST FIN]) following the new approach laid out in section 4.4.4.
(ii) Country abbreviations: Euro area (EA), Austria (AT), Belgium (BE), Cyprus (CY), Germany (DE), Spain (ES), Finland (FI), France (FR), Greece (GR), Italy (IT), Luxembourg (LU), Malta (MT), Netherlands (NL), Portugal (PT), Slovenia (SI), and Slovakia (SK).
(iii) Statistics denoted with a "." depicts countries, i.e. FI, where certain components are not collected.
(iv) All estimates are weighted using appropriate household weights provided by the data producer and the multiple imputation structure is taken into account.
(v) *Source:* Eurosystem HFCS 2010.

of such a policy be relatively small in DE with estimates of -0.019 and even reversed in NL with an estimate of 0.005, whereas one would expect a larger impact in ES and MT where the according estimates are -0.081 and -0.076 respectively.

4.6 Conclusion

Since the publication of the first wave of the Household Finance and Consumption Survey (HFCS) from the ECB one of the most widely discussed topic is the inequality of wealth between but also within euro area countries. This paper investigates the composition of the wealth distribution using various decompositions procedures proposed in the literature. In doing so one can see which component of wealth contributes more or less of the overall distribution. Furthermore, one can inspect cross country differences and thus get a better understanding of the differences that go beyond simple comparisons of specific estimates such as medians or means.

Although one finds that the basic structure of the contributions to inequality is relatively stable across countries (e.g. the separation of real and financial assets and its according contributions), there are notable exceptions such as self-employed business assets are far more important in explaining the distribution of wealth in CY and AT than in others like NL or SK. In general, based on the breakdown of the covariance the contribution of real assets is much higher than that of the financial assets. This might be due to the relative share of the respective components. If one takes, however, a different perspective with the Shapley value findings differ. Thus one should be careful with firm conclusions based on only a specific decomposition procedure. Since there are very few factor decomposition with respect to components of wealth in the literature so far, it is difficult to compare the results found here. Azpitarte (2008) reports remarkable similar results based on a survey from 2002 in Spain, although the debt side of the balance sheet is included without any special treatment, and Lindner (2011) at least does not find contradicting evidence,

although very incomplete data is used.

To see the implication of these results, consider the estimate for the elasticity from table 4.6. A back of the envelop calculation looking at the negative real interest rate on safe financial assets of about 2%, which households faced in recent years (at least in some parts of the euro area) and is driven by the interest rate set by the ECB, yields an increase in wealth inequality measured by the GC in the euro area of about 0.05% per period. This comes additional to the stabilisation of the financial markets for which the component displays a positive elasticity, i.e. the GC increases if wealth in this component gains. This effect is more pronounced in some countries, such as Germany or the Netherlands, whereas it might have an opposite effect in Spain or a very much lower effect such as in France or Italy. Also, policy favouring wealth in the existing primary residence - such as support of energy efficiency/green technology of this real estate- might have the tendency to even the wealth distribution in all countries but the Netherlands, the impact, however, is very much lower for example in Austria than it is in Malta. Overall one can see that even policies that are generally thought not to impact or at least not target the wealth distribution, e.g. monetary policy, still affect households differently and hence yield different outcomes.

In a separate, but closely related field of research the distribution - usually of income - is decomposed by subgroups. Such an exercise with a decomposition of the wealth distribution in the euro area by subgroups of households, such as countries or household characteristics, is beyond the scope of the paper and left for future research. Additionally, it would be interesting to investigate whether the decomposition according the components is influenced heavily if one controls for country specific characteristics.

4.A Additional Decomposition Method

Apart from the decomposition procedures discussed in section 4.4 there is one more approach in the literature. Silber (1989) basically starts of with a matrix form representation of the Gini coefficient. One possible way to calculate the GC is

$$GC(w) = \sum_{j=1}^{N}(\frac{N-j}{N} - \frac{j-1}{j})p_j$$

where individuals are ranked according to increasing wealth and the rank j is taken. p_j denotes the wealth share of individual i in rank j, hence it is given by $p_j = \frac{w_j}{w}$. This can also be written as

$$GC(w) = \sum_{i}^{N}[\sum_{j\geq i}(\frac{1}{N}) - \sum_{j\leq i}(\frac{1}{N})]p_j = e'XP$$

where the last step is a matrix multiplication with the $1 \times N$-vector

$$e = \begin{pmatrix} \frac{1}{N} \\ \frac{1}{N} \\ \vdots \\ \frac{1}{N} \end{pmatrix},$$

$N \times N$-matrix

$$X = \begin{pmatrix} 0 & -1 & -1 & \dots & & -1 \\ 1 & 0 & -1 & & & \vdots \\ 1 & 1 & 0 & & & \\ \vdots & & & \ddots & & \\ & & & & 0 & -1 \\ 1 & \dots & & & 1 & 0 \end{pmatrix},$$

and $1 \times N$-vector

$$P = \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_N \end{pmatrix}.$$

Note that the matrix X here is for simplicity written as if we had data from the whole population, i.e. there are no survey weights that have to be applied. If one needs to account for the survey structure of the data the (negative) ones are replaced with weights, such that each row j is multiplied with the appropriate household weight of household j. As individual household wealth share is based on the sum of components this households owns, i.e. $w_i = \sum_k w_{ki}$ one can rewrite the vector P so that it represents each component for each individual separately. Thus, denoting the share of component k of household i as $p_{ik} = \frac{w_{ik}}{w_i}$ and including an additional vector P as the first column, we can write the matrix P^* as

$$P^* = \begin{pmatrix} p_1 & p_{11} & p_{12} & p_{13} & \cdots & p_{1K} \\ p_2 & p_{21} & p_{22} & p_{23} & \cdots & p_{2K} \\ p_3 & p_{31} & p_{32} & p_{33} & \cdots & p_{2K} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ p_N & p_{N1} & p_{N3} & p_{N3} & \cdots & p_{NK} \end{pmatrix}$$

Combining the matrix P^* with the calculation of the GC from above gives us a vector s that has as first element the GC and following elements the relative contribution of component k, i.e. s_k. Since individuals are still ranked according to household wealth and not a certain component of wealth these contributions are sometimes also called pseudo-Gini.

$$s = \begin{pmatrix} Gini & s_{k=1} & s_{k=2} & s_{k=3} & \cdots & s_{k=K} \end{pmatrix} = e'XP^*$$

Also since individuals are still ranked in an ascending order according to total wealth, this decomposition is (at least) empirically equivalent to the additive decomposition discussed in the text.

4.B Additional Results

Table 4.9: Descriptive Statistics of Net Wealth in EA-Countries

	No. Obs.	Median (in k €)	Mean (in k €)	Minimum	Maximum (in k €)
EA	62,521	109	231	-1,371	409,417
AT	2,380	76	265	-1,371	21,948
BE	2,327	206	339	-421	8,408
CY	1,237	267	671	-831	51,607
DE	3,565	51	195	-359	76,305
ES	6,197	183	291	-1,227	409,417
FI	10,989	86	162	-634	14,723
FR	15,006	116	233	-464	152,506
GR	2,971	102	148	-91	11,700
IT	7,951	174	275	-45	26,127
LU	950	398	710	-392	34,080
MT	843	216	366	-33	30,675
NL	1,301	104	170	-440	4,901
PT	4,404	75	153	-88	27,096
SI	343	101	149	-28	1,937
SK	2,057	61	80	-41	1,181

Notes:
(i) This table shows for each country and the overall euro area the number of observations, median, mean, the minimum and maximum (all but the first in thousand €) observation of net wealth.
(ii) Country abbreviations: Euro area (EA), Austria (AT), Belgium (BE), Cyprus (CY), Germany (DE), Spain (ES), Finland (FI), France (FR), Greece (GR), Italy (IT), Luxembourg (LU), Malta (MT), Netherlands (NL), Portugal (PT), Slovenia (SI), and Slovakia (SK).
(iii) All estimates are weighted using appropriate household weights provided by the data producer and the multiple imputation structure is taken into account.
(iv) *Source:* Eurosystem HFCS 2010.

Table 4.10: Components of Gross Wealth in EA-Countries (Liabilities)

	Mortgage HMR			Mortgage FRE			Overdraft			Consumption Loan		
	Part.	Median (in k €)	Share	Part.	Median (in k €)	Share	Part.	Median (in k €)	Share	Part.	Median (in k €)	Share
EA	0.194	65	0.065	0.055	57	0.020	0.125	1	0.002	0.224	6	0.016
AT	0.166	37	0.043	0.024	36	0.007	0.146	1	0.001	0.111	8	0.009
BE	0.285	67	0.066	0.032	57	0.008	0.106	1	0.001	0.179	7	0.008
CY	0.350	85	0.053	0.154	66	0.029	0.338	4	0.004	0.293	13	0.010
DE	0.180	67	0.071	0.060	82	0.035	0.215	2	0.003	0.217	5	0.012
ES	0.268	54	0.061	0.073	80	0.026	0.078	1	0.001	0.272	8	0.013
FI	0.328	64	0.132	.	.	.	.	.	.	0.512	7	0.051
FR	0.169	61	0.050	0.101	22	0.023	0.070	1	0.001	0.287	6	0.023
GR	0.139	40	0.045	0.039	42	0.013	0.168	3	0.006	0.126	6	0.010
IT	0.096	65	0.027	0.016	25	0.003	0.048	2	0.001	0.153	7	0.010
LU	0.328	122	0.070	0.084	116	0.023	0.118	1	0.001	0.308	12	0.009
MT	0.121	34	0.016	0.045	37	0.008	0.163	1	0.002	0.137	8	0.006
NL	0.439	130	0.250	0.025	103	0.020	0.230	2	0.006	0.246	26	0.050
PT	0.245	46	0.082	0.033	58	0.012	0.075	1	0.001	0.133	5	0.007
SI	0.125	7	0.014	0.016	4	0.001	0.255	1	0.002	0.271	5	0.018
SK	0.093	25	0.031	0.006	13	0.002	0.107	0	0.001	0.126	2	0.006

Notes:
(i) This table shows for each liability component (mortgage using main residence as collateral [Mortgage HMR], mortgage using further real estate as collateral [Mortgage FRE], overdraft and credit card debt [Overdraft], and consumption loan) participation rate (Part.), conditional median (in thousand €), and the share relative to gross wealth hold in the respective component.
(ii) Country abbreviations: Euro area (EA), Austria (AT), Belgium (BE), Cyprus (CY), Germany (DE), Spain (ES), Finland (FI), France (FR), Greece (GR), Italy (IT), Luxembourg (LU), Malta (MT), Netherlands (NL), Portugal (PT), Slovenia (SI), and Slovakia (SK).
(iii) Statistics denoted with a "." depicts countries, i.e. FI, where certain components are not collected.
(iv) All estimates are weighted using appropriate household weights provided by the data producer and the multiple imputation structure is taken into account.
(v) *Source:* Eurosystem HFCS 2010.

Table 4.11: Share of Asset Components relative to Net Wealth in EA-Countries

	HMR	FRE	BUS	REST REAL	SAFE	RISKY	REST FIN	MORT. HMR	MORT. FRE	OVER-DRAFT	CONS. LOAN
EA	0.564	0.211	0.107	0.046	0.129	0.043	0.014	0.073	0.023	0.002	0.018
AT	0.471	0.117	0.246	0.049	0.128	0.042	0.010	0.046	0.007	0.001	0.009
BE	0.562	0.130	0.047	0.034	0.177	0.121	0.019	0.071	0.009	0.001	0.009
CY	0.363	0.393	0.237	0.023	0.067	0.013	0.010	0.059	0.032	0.004	0.011
DE	0.466	0.234	0.148	0.049	0.172	0.054	0.015	0.081	0.040	0.003	0.014
ES	0.599	0.264	0.100	0.034	0.076	0.021	0.017	0.068	0.029	0.001	0.014
FI	0.669	0.275	0.040	0.057	0.113	0.071	.	0.162	.	.	0.063
FR	0.526	0.221	0.093	0.052	0.156	0.040	0.018	0.056	0.025	0.001	0.025
GR	0.605	0.300	0.051	0.050	0.066	0.006	0.002	0.049	0.015	0.007	0.010
IT	0.634	0.171	0.086	0.047	0.058	0.036	0.010	0.028	0.003	0.001	0.011
LU	0.578	0.337	0.033	0.042	0.078	0.042	0.004	0.078	0.026	0.001	0.010
MT	0.456	0.173	0.232	0.033	0.094	0.037	0.008	0.017	0.008	0.002	0.006
NL	0.907	0.095	0.033	0.052	0.327	0.056	0.010	0.370	0.029	0.008	0.074
PT	0.532	0.256	0.133	0.054	0.112	0.018	0.009	0.091	0.014	0.001	0.007
SI	0.697	0.145	0.095	0.041	0.045	0.007	0.006	0.014	0.001	0.002	0.019
SK	0.775	0.070	0.047	0.063	0.075	0.003	0.009	0.032	0.002	0.001	0.007

Notes:
(i) This table shows for each asset component (main residence [HMR], further real estate [FRE], self-employed business assets [BUS], rest real assets [REST REAL], safe financial assets [SAFE], risky financial assets [RISKY], rest financial assets [REST FIN], mortgage using main residence as collateral [MORT. HMR], mortgage using further real estate as collateral [MORT. FRE], overdraft and credit card debt [OVERDRAFT], and consumption loan [CONS. LOAN]) the share relative to net wealth hold in the respective component.
(ii) Country abbreviations: Euro area (EA), Austria (AT), Belgium (BE), Cyprus (CY), Germany (DE), Spain (ES), Finland (FI), France (FR), Greece (GR), Italy (IT), Luxembourg (LU), Malta (MT), Netherlands (NL), Portugal (PT), Slovenia (SI), and Slovakia (SK).
(iii) Statistics denoted with a "." depicts countries, i.e. FI, where certain components are not collected.
(iv) All estimates are weighted using appropriate household weights provided by the data producer and the multiple imputation structure is taken into account.
(v) *Source:* Eurosystem HFCS 2010.

Table 4.12: Inequality Measures of Net Wealth in EA-Countries

	Ratios				Ineq. Measures		Shares		
	P90/ P10	P75/ P25	P90/ P50	Mean/ Median	Gini	GE2	Top 5%	Top 10%	Bottom 50%
EA	423.8	17.3	4.6	2.1	0.678	5.170	0.372	0.504	0.060
AT	581.1	24.3	7.1	3.5	0.762	4.468	0.476	0.611	0.028
BE	253.8	10.4	3.4	1.6	0.608	1.328	0.313	0.440	0.101
CY	200.8	6.8	5.5	2.5	0.698	3.070	0.425	0.567	0.074
DE	7,371.3	31.8	8.6	3.8	0.758	5.760	0.456	0.592	0.028
ES	581.1	24.3	7.1	1.6	0.580	8.277	0.309	0.434	0.130
FI	-692.2	34.5	4.6	1.9	0.664	1.837	0.306	0.450	0.052
FR	581.1	24.3	7.1	2.0	0.679	6.505	0.365	0.500	0.054
GR	165.9	6.4	3.3	1.4	0.561	0.818	0.254	0.388	0.124
IT	581.1	24.3	7.1	1.6	0.609	1.827	0.321	0.448	0.102
LU	274.3	12.5	3.5	1.8	0.661	3.311	0.399	0.513	0.086
MT	43.3	4.5	3.2	1.7	0.600	6.104	0.353	0.468	0.125
NL	-162.2	18.5	4.1	1.6	0.654	0.977	0.257	0.401	0.049
PT	286.9	8.7	4.0	2.0	0.670	7.104	0.407	0.527	0.083
SI	75.6	5.2	3.2	1.5	0.534	0.685	0.224	0.357	0.138
SK	11.8	2.7	2.5	1.3	0.448	0.556	0.218	0.328	0.206

Notes:
(i) This table shows for each country and the overall euro area various frequently used measures of the distribution of net wealth.
(ii) Country abbreviations: Euro area (EA), Austria (AT), Belgium (BE), Cyprus (CY), Germany (DE), Spain (ES), Finland (FI), France (FR), Greece (GR), Italy (IT), Luxembourg (LU), Malta (MT), Netherlands (NL), Portugal (PT), Slovenia (SI), and Slovakia (SK).
(iii) All estimates are weighted using appropriate household weights provided by the data producer and the multiple imputation structure is taken into account.
(iv) *Source:* Eurosystem HFCS 2010.

Table 4.13: Decomposition of Gross Wealth following Lerman and Yitzhaki (1985) (relative contribution)

	HMR	FRE	BUS	REST REAL	SAFE	RISKY	REST FIN
EA	0.445	0.237	0.135	0.027	0.094	0.047	0.015
AT	0.406	0.127	0.297	0.030	0.087	0.045	0.009
BE	0.402	0.161	0.067	0.022	0.160	0.164	0.024
CY	0.223	0.402	0.295	0.010	0.046	0.012	0.010
DE	0.382	0.246	0.172	0.028	0.111	0.049	0.012
ES	0.389	0.325	0.144	0.020	0.070	0.028	0.024
FI	0.475	0.284	0.052	0.033	0.076	0.079	.
FR	0.407	0.245	0.116	0.036	0.130	0.046	0.021
GR	0.463	0.373	0.063	0.029	0.062	0.009	0.002
IT	0.546	0.216	0.118	0.025	0.041	0.041	0.014
LU	0.417	0.411	0.043	0.027	0.051	0.047	0.004
MT	0.312	0.212	0.358	0.022	0.053	0.033	0.010
NL	0.621	0.109	0.024	0.022	0.162	0.053	0.009
PT	0.367	0.304	0.176	0.035	0.090	0.021	0.007
SI	0.559	0.198	0.159	0.028	0.043	0.007	0.006
SK	0.663	0.110	0.091	0.056	0.062	0.004	0.015

Notes:
(i) This table shows for each country and the overall euro area the decomposition of gross wealth according to its components (main residence [HMR], further real estate [FRE], self-employed business assets [BUS], rest real assets [REST REAL], safe financial assets [SAFE], risky financial assets [RISKY], and rest financial assets [REST FIN]) following the approach of Lerman and Yitzhaki (1985) (relative contribution).
(ii) Country abbreviations: Euro area (EA), Austria (AT), Belgium (BE), Cyprus (CY), Germany (DE), Spain (ES), Finland (FI), France (FR), Greece (GR), Italy (IT), Luxembourg (LU), Malta (MT), Netherlands (NL), Portugal (PT), Slovenia (SI), and Slovakia (SK).
(iii) Statistics denoted with a "." depicts countries, i.e. FI, where certain components are not collected.
(iv) All estimates are weighted using appropriate household weights provided by the data producer and the multiple imputation structure is taken into account.
(v) *Source:* Eurosystem HFCS 2010.

Table 4.14: Shapely Value Decomposition of Gross Wealth (relative contribution)

	HMR	FRE	BUS	REST REAL	SAFE	RISKY	REST FIN
EA	-0.104	0.230	0.267	0.089	0.082	0.216	0.219
AT	0.008	0.206	0.272	0.080	0.050	0.195	0.189
BE	-0.254	0.251	0.270	0.121	0.088	0.274	0.250
CY	-0.118	0.183	0.317	0.079	0.117	0.206	0.217
DE	0.029	0.223	0.255	0.094	0.032	0.184	0.184
ES	-0.354	0.245	0.315	0.107	0.153	0.262	0.272
FI	-0.124	0.261	0.313	0.128	0.132	0.290	.
FR	-0.103	0.214	0.251	0.099	0.101	0.218	0.220
GR	-0.268	0.237	0.267	0.077	0.163	0.264	0.261
IT	-0.158	0.263	0.286	0.041	0.097	0.227	0.245
LU	-0.183	0.314	0.255	0.083	0.074	0.228	0.229
MT	-0.218	0.236	0.398	0.113	0.038	0.189	0.244
NL	-0.162	0.318	0.274	0.088	-0.059	0.269	0.273
PT	-0.203	0.251	0.288	0.105	0.115	0.227	0.216
SI	-0.401	0.271	0.333	0.111	0.180	0.242	0.265
SK	-0.558	0.334	0.372	0.100	0.105	0.322	0.325

Notes:
(i) This table shows for each country and the overall euro area the decomposition of gross wealth according to its components (main residence [HMR], further real estate [FRE], self-employed business assets [BUS], rest real assets [REST REAL], safe financial assets [SAFE], risky financial assets [RISKY], and rest financial assets [REST FIN]) following the approach of Shorrocks (2013) and Sastre and Trannoy (2002).
(ii) Country abbreviations: Euro area (EA), Austria (AT), Belgium (BE), Cyprus (CY), Germany (DE), Spain (ES), Finland (FI), France (FR), Greece (GR), Italy (IT), Luxembourg (LU), Malta (MT), Netherlands (NL), Portugal (PT), Slovenia (SI), and Slovakia (SK).
(iii) Statistics denoted with a "." depicts countries, i.e. FI, where certain components are not collected.
(iv) All estimates are weighted using appropriate household weights provided by the data producer and the multiple imputation structure is taken into account.
(v) *Source:* Eurosystem HFCS 2010.

Chapter 5

Bibliography

Abbring, J. H., and J. J. Heckman (2007): "Econometric Evaluation of Social Programs, Part III: Distributional Treatment Effects, Dynamic Treatment Effects, Dynamic Discrete Choice, and General Equilibrium Policy Evaluation," in *Handbook of Econometrics*, ed. by J. J. Heckman, and E. Leamer, vol. 6B, chap. 72. Elsevier, 1 edn.

Albacete, N. (2011): "Multiple Imputations in the Household Survey on Housing Wealth," OeNB Working Paper Series 176, Oesterreichische Nationalbank.

Andreasch, M., P. Fessler, and M. Schürz (2009): "Austrian Households' Equity Capital - Evidence from Micro Data," *Monetary Policy and the Economy*, Q4/09, 66–84.

Angrist, J., and A. B. Krueger (1998): "Empirical Strategies in Labor Economics," Working papers 98-7, Massachusetts Institute of Technology (MIT), Department of Economics.

Angrist, J., and J.-S. Pischke (2009): *Mostly harmless econometrics*. Princeton University Press, Princeton, New Jersey.

Araar, A. (2006): "On the Decomposition of the Gini Coefficient: an Exact Ap-

proach, with an Illustration Using Cameroonian Data," Working Paper 0602, CIRPEE.

Azpitarte, F. (2008): "The Household Wealth Distribution in Spain: The Role of Housing and Financial Wealth," ECINEQ Working Paper Series 83.

Banks, J., R. Blundell, and J. P. Smith (2003): "Understanding Differences in Household Financial Wealth between the United States and Great Britain," *Journal of Human Resources*, 38(2), 241–279.

Beer, C., P. Mooslechner, M. Schürz, and K. Wagner (2006): "Austrian Households' Financial Wealth: An Analysis Based on Microeconomic Data," *Monetary Policy and the Economy*, Q2/06, 94–110.

Berrebi, Z. M., and J. Silber (1985): "The Gini Coefficient and Negative Income: A Comment," *Oxford Economic Papers*, 37(3), 525–26.

Blundell, R., and M. C. Dias (2002): "Alternative approaches to evaluation in empirical microeconomics," CeMMAP Working Papers CWP10/02, Centre for Microdata Methods and Practice, Institute for Fiscal Studies.

Bourguignon, F., F. Ferreira, and P. Leite (2008): "Beyond Oaxaca–Blinder: Accounting for differences in household income distributions," *Journal of Economic Inequality*, 6(2), 117–148.

Bover, O. (2005): "The Wealth of Spanish Households: A Microeconomic Comparison with the United States, Italy and the United Kingdom," *Economic Bulletin*, Banco D'España, 1–23.

——— (2010): "Wealth inequality and household structure: US vs. Spain," *The Review of Income and Wealth*, 56(2), 259–290.

Brandolini, A., L. Cannari, G. D'Alessio, and I. Faiella (2004): "Household Wealth Distribution in Italy in the 1990s," Temi di discussione (Economic Working Papers) 530, Bank of Italy.

Buhmann, B., L. Rainwater, G. Schmaus, and T. M. Smeeding (1988): "Equivalence Scales, Well-Being, Inequality, and Poverty: Sensitivity Estimates across ten Countries using the Luxembourgh Income Study (LIS) Database," *Review of Income and Wealth*, 34(2), 115–142.

Cagetti, M., and M. DeNardi (2005): "Wealth inequality: data and models," Working Paper Series 05-10, Federal Reserve Bank of Chicago.

Celestin, C. N., and M. W. Clovis (2012): "Inequality of Cameroonian Households: An Analysis Based on the Shapley-Shorrocks Decomposition," *International Journal of Economics and Finance*, 4(6), 149–156.

Chantreuil, F., and A. Trannoy (2013): "Inequality decomposition values: the trade-off between marginality and efficiency," *Journal of Economic Inequality*, 11(1), 83–98.

Chen, C.-N., T.-W. Tsaur, and T.-S. Rhai (1982): "The Gini Coefficient and Negative Income," *Oxford Economic Papers*, 34(3), 473–478.

Chernozhukov, V., I. Fernandez-Val, and B. Melly (2009): "Inference on counterfactual distributions," CeMMAP Working Papers CWP09/09, Centre for Microdata Methods and Practice, Institute for Fiscal Studies.

Conti, G., and S. Pudney (2011): "Survey design and the analysis of satisfaction," *The Review of Economics and Statistics*, 93(3), 1087–1093.

Cowell, F. A., and C. V. Fiorio (2009): "Inequality Decomposition - A Reconciliation," *LSE STICERD Research Paper*, (DARP 100), 1–32.

Cowell, F. A., and E. Flachaire (2002): "Sensitivity of Inequality Measures to Extreme Values," STICERD - Distributional Analysis Research Programme Papers 60, Suntory and Toyota International Centres for Economics and Related Disciplines, LSE.

Cowell, F. A., and E. Flachaire (2007): "Income distribution and inequality measurement: The problem of extreme values," *Journal of Econometrics*, 141(2), 1044–1072.

Cowell, F. A., and M.-P. Victoria-Feser (1996): "Robustness properties of inequality measures," *Econometrica*, 64(1), 77–101.

Davies, J. B., S. Sandström, A. B. Shorrocks, and E. N. Wolff (2009): "The Level and Distribution of Global Household Wealth," NBER Working Paper 15508.

Davies, J. B., and A. F. Shorrocks (2000): "The Distribtion of Wealth," in *Handbook of Income Distribution*, ed. by A. Atkinson, and F. Bourguignon, chap. 11, pp. 605–675. Elsevier.

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

——— (2008): "The Effect of Computer-Assisted Interviewing on Data Quality: A Review of the Evidence," Unpublished manuscript, Department of Methodology and Statistics, Utrecht University.

De-Leeuw, E. D., J. J. Hox, and D. A. Dillman (2008): *International Handbook of Survey Methodology*. European Association of Methodology.

Dell, F., T. Piketty, and E. Saez (2005): "Income and Wealth Concentration in Switzerland Over the 20th Century," CEPR Discussion Paper 5090.

DiNardo, J., N. Fortin, and T. Lemieux (1996): "Labor market institutions and the distribution of wages, 1973-1992: A semiparametric approach," *Econometrica*, 64(5), 1001–1044.

Dorfmann, R. (1979): "A Formula for the Gini Coefficient," *The Review of Economics and Statistics*, 61(1), 146–149.

Duflo, E., R. Glennerster, and M. Kremer (2008): "Using Randomization in Development Economics Research: A Toolkit," *Handbook of Development Economics*, 4, 3895–3962.

Ebert, U., and P. Moyes (2003): "Equivalence Scales Reconsidered," *Econometrica*, 71(1), 319–343.

ECB (2013a): "The Eurosystem Household Finance and Consumption Survey Methodological Report for the First Wave," Statistics Paper Series 1, European Central Bank.

——— (2013b): "The Eurosystem Household Finance and Consumption Survey Results from the first wave," Statistics Paper Series 2, European Central Bank.

Essama-Nssah, B., and P. J. Lambert (2011): "Influence functions for distributional statistics," Working Papers 236, ECINEQ, Society for the Study of Economic Inequality.

Fei, J. C. H., G. Ranis, and S. W. Y. Kuo (1978): "Growth and the Family Distribution of Income by Factor Components," *The Quarterly Journal of Economics*, 92(1), 17–53.

Fessler, P., P. Lindner, and E. Segalla (2013): "Household Structure and the Distribution of Net Wealth - Evidence from the Euroarea," *Mimeo*.

Fessler, P., P. Mooslechner, M. Schürz, and K. Wagner (2009): "Housing

Wealth of Austrian Households," *Monetary Policy & The Economy*, Q2/09, 104–124.

Fessler, P., and M. Schürz (2008): "Stock Holdings in Austria," *Monetary Policy & the Economy*, Q2/08, 83–100.

Fields, G. S. (2003): "Accounting for income inequality and its changes: a new method, with application to the distribution of earnings in the United States," *Research in Labor Economics*, 22, 1–38.

Fiorio, C. V. (2006): "Understanding Inequality Trends:Microsimulation Decomposition for Italy," STICERD - Distributional Analysis Research Programme Papers 78, Suntory and Toyota International Centres for Economics and Related Disciplines, LSE.

Firpo, S., N. Fortin, and T. Lemieux (2007): "Decomposing Wage Distributions using Recentered Influence Function Regressions," Unpublished manuscript, University of British Columbia.

Fortin, N. M., T. Lemieux, and S. Firpo (2009): "Unconditional Quantile Regression," *Econometrica*, 77(3), 953–973.

——— (2011): "Decomposition methods in economics," *Handbook of Labor Economics*, 4, 1–102.

Fournier, M. (2001): "Inequality decomposition by factor component: a "rank-correlation" approach illustrated on the Taiwanese case," *Recherches Economiques de Louvain - Louvain Economic Review*, 67(4), 381–304.

Fräßdorf, A., M. Grabka, and J. Schwarze (2011): "The impact of household capital income on income inequality—a factor decomposition analysis for the UK, Germany and the USA," *Journal of Economic Inequality*, 9(1), 35–56.

Frosini, B. V. (2012): "Approximation and decomposition of Gini, Pietra–Ricci and Theil inequality measures," *Empirical Economics*, 43, 175–197.

Gastwirth, J. L. (1970): "A General Definition of the Lorenz Curve," *Econometrica*, 39(6), 1037–1039.

——— (1972): "The estimation of the Lorenz curve and the Gini index," *The Review of Economics and Statistics*, 54, 306–316.

Gastwirth, J. L., and M. Glauberman (1976): "The Interpolation of the Lorenz Curve and the Gini Index from Grouped Data," *Econometrica*, 44(3), 479–483.

Gunatilaka, R., and D. Chotikapanich (2009): "Accounting for Sri Lanka's expenditure inequality 1980-2002: Regression-based decomposition approaches," *Review of Income and Wealth*, 55(4), 882–906.

Huber, P. J. (2003): *Robust Statistics*, Wiley Series in Probability and Statistics. John Wiley & Sons, Hoboken, New Jersey.

Iacus, S., G. King, and G. Porro (2008): "Matching for Causal Inference Without Balance Checking," Unpublished manuscript, iqss, Harvard University.

Imbens, G. (2009): "Nonparametric Estimation of Average Treatment Effects under Exogeneity: A Review," *The Review of Economics and Statistics*, 86(1), 4–29.

Imbens, G., and J. Wooldridge (2009): "Recent Developments in the Econometrics of Program Evaluation," *Journal of Economic Literature*, 47(1), 5–86.

Jenkins, S. P., and M. Jäntti (2005): "Methods for Summarizing and Comparing Wealth Distribution," Discussion Paper 2005-05.

Juhn, C., K. M. Murphy, and B. Pierce (1993): "Wage inequality and the Rise in Returns to Skill," *Journal of Political Economy*, 101(3), 410–442.

Kanbur, R. (2006): "The policy significance of inequality decompositions," *Journal of Economic Inequality*, 4(3), 367–374.

Kendall, M. G., and A. Stuart (1963): *The Advanced Theory of Statistics*. Charles Griffen, London, second edn.

Koulovatianos, C., C. Schröder, and U. Schmidt (2005): "On the income dependence of equivalence scales," *Journal of Public Economics*, 89, 967–996.

Lambert, P. J., and R. J. Aronson (1993): "Inequality decomposition analysis and the Gini coefficient revisited," *The Economic Journal*, 103(420), 1221–1227.

Lerman, R. I., and S. Yitzhaki (1985): "Income inequality effects by income source: a new approach and applications to the United States," *The Review of Economics and Statistics*, 67(1), 151–156.

Lindner, P. (2011): "Decomposition of Wealth and Income using Micro Data from Austria," OeNB Working Papers 173, Oesterreichische Nationalbank.

List, J. A., and I. Rasul (2011): "Field Experiments in Labor Economics," *Handbook of Labor Economics*, 4, 103–228.

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), 37–54.

López-Feldman, A. (2006): "Decomposing inequality and obtaining marginal effects," *The STATA Journal*, 6(1), 106–111.

López-Feldman, A., J. Mora, and E. J. Taylor (2006): "Does natural resource extraction mitigate poverty and inequality? Evidence from rural Mexico and a Lacandona rainforest community," Agriculture and Resource Economics Working Papers 005.

Machado, J. A. F., and J. Mata (2005): "Counterfactual decomposition of changes in wage distributions using quantile regression," *Journal of Applied Econometrics*, 20(4), 445–465.

Mookherjee, D., and A. F. Shorrocks (1982): "A Decomposition Analysis of the Trend in UK Income Inequality," *Economic Journal*, 92(368), 886–902.

Morduch, J., and T. Sicular (2002): "Rethinking inequality decomposition, with evidence from China," *The Economic Journal*, 112(476), 93–106.

Morgan, S. L., and C. Winship (2007): *Counterfactuals and Causal Inference*. Cambridge, Cambridge University Press.

OECD (2011): "Divided We Stand - Why Inequality Keeps Rising," Discussion paper, OECD.

Okamoto, M. (2011): "Source decomposition of changes in income inequality: the integral-based approach and its approximation by the chained Shapley-value approach," *Journal of Economic Inequality*, 9(2), 145–181.

Paul, S. (2004): "Income sources effects on inequality," *Journal of Development Economics*, 73(1), 435–451.

Peichl, A., N. Pestel, and H. Schneider (2012): "Does Size Matter? The Impact Of Changes In Household Structure On Income Distribution In Germany," *Review of Income and Wealth*, 58(1), 118–141.

Peichl, A., T. Schäfer, and C. Scheicher (2010): "Measuring Richness and Poverty: A Micro Data Apllication to Europa and Germany," *Review of Income and Wealth*, 56(3), 597–619.

Rubin, D. B., and R. J. A. Little (2002): *Statistical Analysis with Missing Data*. John Wiley and Sons, Inc. New York, NY, USA, second edn.

Sastre, M., and A. Trannoy (2002): "Shapley Inequality Decomposition by Factor Components: Some Methodological Issues," *Journal of Economics*, Suppl. 9, 51–89.

Shapley, L. F. (1953): "A value for n-person games," in *Contribution to the Theory of Games*, ed. by H. Kuhn, and A. W. Tucker, vol. 2. Princeton University Press.

Shorrocks, A. F. (1982): "Inequality decomposition by factor components," *Econometrica*, 50(1), 193–211.

——— (2013): "Decomposition procedures for distributional analysis: a unified framework based on the Shapley value," *Journal of Economic Inequality*, 11(1), 99–126.

Sierminska, E., A. Brandolini, and T. Smeeding (2006): "The Luxembourg Wealth Study - A cross-country comparable database for household wealth research," *Journal of Economic Inequality*, 4(3), 375–383.

Sierminska, E., and K. Doorly (2012): "Decomposing household wealth portfolios across countries: An age-old question?," CEPS/INSTEAD Working Paper Series 32, CEPS/INSTEAD.

Sierminska, E., J. R. Frick, and M. M. Grabka (2008): "Examining the Gender Wealth Gap in Germany," IZA Discussion Papers 3573, Institute for the Study of Labor (IZA).

Sierminska, E., and T. Smeeding (2006): "Measurement Issues: Equivalence Scales, Accounting Framework, and Reference Unit," *Paper prepared for the LWS Workshop: Construction and Usage of Comparable Microdata on Wealth: the LWS, Banca de Italia, Perugia, 27-29 January.*

Silber, J. (1989): "Factor Components, Population Subgroups and the Computa-

tion of the Gini Index of Inequality," *Review of Economics and Statistics*, 71(1), 107–115.

Smeeding, T. M., and D. H. Weinberg (2001): "Toward a Uniform Definition of Household Income," *Review of Income and Wealth*, 47(1), 1–24.

Statistik, A. (2010): "Standard Dokumentation Metainformationen - Definitionen, Erläuterungen, Methoden, Qualität," Documentation, Statistik Austria.

Stich, A. (1996): "Inequality and negative income," *Journal of the Italian Statistical Society*, 5(3), 297–305.

Taylor, J., J. Mora, R. Adams, and A. López-Feldman (2005): "Remittances, inequality and poverty: Evidence from rural Mexico," Agriculture and Resource Economics Working Papers 003.

Wolff, E. N. (2002): "Inheritances and Wealth Inequality, 1989–1998," *American Economic Review*, 92(2), 260–264.

Wolff, E. N., and A. Zacharias (2009): "Household wealth and the measurement of economic well-being in the United States," *Journal of Economic Inequality*, 7, 83–115.

Xu, K. (2004): "How has the Literature on Gini's Index Evolved in the Past 80 Years?," Economics working paper, Dalhousie University.

Yitzhaki, S. (1998): "More than a Dozen Alternative Ways of Spelling Gini," *Research on Economic Inequality*, 8, 13–30.

Yun, M.-S. (2006): "Earnings inequality in USA, 1969-99: Comparing Inequality using Earnings equation," *Review of Income and Wealth*, 52(1), 127–144.

Appendix A

Abstracts

A.1 English Abstracts

Abstract of "Decomposition of Wealth and Income using Micro Data from Austria"[1]

This paper analyses Austrian data on (financial) wealth and income. The main focus lies on the distribution of these indicators. Using a decomposition procedure of the Gini Index first proposed by Lerman and Yitzhaki (1985), it is possible to recover the effect of specific forms of investments of assets and of certain sources of income on the overall distribution (in terms of an elasticity) of wealth and income in Austria. For the first time Austrian wealth and income data are used to decompose the total distribution into various categories. Additionally, there are, due to the lack of available data on household wealth, internationally only very few attempts to compare wealth and income using decomposition methods. The analysis shows that specific forms of assets (mainly more sophisticated ones) as well as income from sources that are concentrated on a small group tend to increase inequality whereas the others have an equalizing effect.

Key Words: Wealth Distribution, Income Distribution, Decomposition by Components / Factors

JEL Classifications: D3

[1]This chapter comprises a single authored paper. However, I want to thank Nicolás Albacete, Pirmin Fessler, Markus Knell, Karin Wagner and Siegried Zottel for very helpful comments and guidance. Additional to the usual disclaimer, the opinions expressed in this work are those of the author and do not necessarily reflect the ones of the OeNB.

Abstract of "Survey Mode Effects on Item Non-response and Income Inequality Measurement"[2]

Exploiting a quasi-experiment we use non-parametric re-weighting and regression approaches to estimate the causal effect of the interview method on item-nonresponse and the unconditional observed income distribution. The minor change of interviewing households via the telephone instead of personally leads to major changes in item non-response, which increases by roughly 20% – 30%. The observed distribution of income is compressed translating to a decrease in the Gini coefficient by about 10%. Commonly used rankings of countries by Gini coefficients of the income distribution might largely be an artifact of different survey techniques as the interview mode than true differences in income inequality.

Key Words: EU - Statistics on Income and Living Conditions, Computer Assisted Personal Interviewing, Computer Assisted Telephone Interviewing, Survey Methodology, Unconditional Distribution, Non-Parametric Re-Weighting

JEL Classifications: B41, C01, C14, C64, C81, C83, D31

Abstract of "Net Wealth across the Euro Area - Why Household Structure Matters and How to Control for it"[3]

We study the link between household structure and cross country differences in the wealth distribution using a recently compiled household dataset for the euro area (HFCS). We estimate counterfactual distributions using non-parametric reweighting to examine the extent to which differences in the unconditional distribu-

[2]This chapter comprises a paper joint with Pirmin Fessler and Maximilian Kasy. We'd like to thank Markus Knell for valuable comments and discussion. Additional to the usual disclaimer, the opinions expressed in this work are those of the authors and do not necessarily reflect the ones of the OeNB.

[3]This chapter comprises a paper joint with Pirmin Fessler and Esther Segalla. Additional to the usual disclaimer, the opinions expressed in this work are those of the authors and do not necessarily reflect the ones of the OeNB.

tions of wealth across euro area countries can be explained by differences in household structure. We find that imposing a common household structure has strong effects on both the full unconditional distributions as well as its mappings to different inequality measures. For the median 50% of the differences are explained for Austria, 15% for Germany, 25% for Italy, 14% for Spain and 38% for Malta. For others as Belgium, France, Greece, Luxembourg, Portugal, Slovenia and Slovakia household structure masks the differences to the euro area median and Finland and the Netherlands change their position from below to above the euro area median. The impact on the mean and percentile ratios is similarly strong and varies with regard to direction and level across countries and their distributions. We can confirm the finding of Bover (2010) that the effect on the Gini coefficient is somewhat less pronounced, but might mask relevant information by being a net effect of different accumulated effects along the distribution. Country rankings based on almost all of these measures are severely affected alluding to the need for cautious interpretation when dealing with such rankings. Furthermore, the explanatory power of household structure changes along the net wealth distribution. Therefore we argue for more flexible controls for household structure. We provide such a set of controls to account for household type fixed effects which are based on the number of household members as well as possible combinations of age categories and gender.

Key Words: Wealth Distribution, Household Structure, Survey Methodology, Unconditional Distribution, Non-Parametric Re-Weighting, Counterfactuals

JEL Classifications: D30, D31

Abstract of "Factor Decomposition of the Wealth Distribution in the Euro Area"[4]

Economic policy often targets specific components of the household balance sheet such as policies supporting the financing of real estate or setting the interest rate. In order to understand the potential impact of these policy measures on the wealth distribution this paper investigates the link between components and the overall wealth distribution using factor component decomposition methods. Applying methods from the literature and proposing a new decomposition approach based on the recentered influence function, it uses the newly published household survey data from the Household Finance and Consumption Survey. This paper studies the (marginal) contributions of and the elasticity with respect to components in the euro area as well as in 15 euro area countries separately. We find that some factors (such as wealth hold in the households' main residence) contribute a higher proportion to inequality while others (such as risky financial assets) exhibit only a relatively weak contribution. Elasticities reveal additionally that components have diverging impacts on the distribution. Furthermore, the relative contribution as well as the point elasticities for the components of total wealth show differences between the countries in the euro area pointing towards regionally distinctive impacts of the policy measures.

Key Words: Decomposition, Distribution, Wealth, Shapley Value, RIF, HFCS

JEL Classifications: D31, D63

[4]This chapter comprises a single authored paper. However, I want to thank Esther Segalla, Maximilian Kasy, Helmut Elsinger, Pirmin Fessler, and Nicolás Albacete for very helpful discussions and comments. Additional to the usual disclaimer, the opinions expressed in this work are those of the authors and do not necessarily reflect the ones of the OeNB.

A.2 German Abstracts

Abstract zum Kapitel "Decomposition of Wealth and Income using Micro Data from Austria"

Der Artikel, der diesem Kapitel zu Grunde liegt, analysiert die österreichische Volkswirtschaft anhand von Daten zu (finanziellem) Vermögen und Einkommen. Der Fokus liegt auf der Verteilung dieser Indikatoren. Anhand einer Dekompositionsmethode, die erstmals von Lehrman und Yitzhaki (1985) in der Literatur vorgeschlagen wurde, ist es möglich den Beitrag einzelner Vermögenskomponenten beziehungsweise Einkommensarten auf die gesamte Verteilung von Vermögen und Einkommen in Bezug auf deren Elastizität zu bestimmen. Zum ersten mal werden Daten aus Österreich zu Vermögen und Einkommen in einer Analyse gemeinsam für diese Dekomposition nach Vermögens- und Einkommenskomponenten herangezogen. Zusätzlich gibt es international insbesondere aufgrund fehlender Daten wenige Ansätze der Dekomposition von Haushaltsvermögen und dessen Vergleich mit den Ergebnissen zum Einkommen. Diese Analyse kann aufzeigen, dass die gesamte Verteilung tendenziell eine positive Elastizität nach manchen Anlagemöglichkeiten (meist komplexere Finanzanlageprodukte) und Einkommensformen (meist Formen, die von relativ wenigen Haushalten bezogen werden) aufgrund der Nutzung und Verteilung der jeweiligen Komponenten vorweist. Für andere Komponenten liegt ein gegenteiliges Ergebnis, also eine negative Elastizität der gesamten Verteilung nach diesen Komponenten, in einer ceteris paribus Analyse vor.

Abstract zum Kapitel "Survey Mode Effects on Item Nonresponse and Income Inequality Measurement"

In diesem Papier wird ein annähernd natürliches Experiment genutzt, um mit nichtparametrischen und regressions-bassierten Analysemethoden der kausale Effekt der

Interviewmethode auf die vorgefundene Item Non-Response (Antwortverweigerung bei einzelnen Fragen einer Erhebung) und die resultierende Verteilung des Haushaltseinkommens zu eruieren. Es ist möglich zu zeigen, dass eine oft als geringfügig eingeschätzte Veränderungen der Interviewmethode von einem computer-gestützten persönlichen Interview hin zu einem computer-gestützten Telefoninterview eine schwerwiegende Auswirkung auf das Antwortverhalten der Respondenten insofern ausübt, dass die Item Non-Response-Rate um rund 20% – 30% erhöht ist. Ebenfalls ist die resultierende Einkommensverteilung durch eine auf der Telefonmethode basierten Erhebung komprimiert, was durch eine Reduktion der Verteilung gemessen mit dem Gini Index um rund 10% gesehen werden kann. Somit könnten aufgrund der internationalen Unterschiede in der Verwendung von Interviewmethoden die meist verwendeten Rangordnungen und international Vergleiche der Einkommensverteilungen anhand des Gini Koeffizienten lediglich diese Unterschiede der Erhebungsmethode aufzeigen anstatt vergleichbare Ergebnisse zu liefern.

Abstract zum Kapitel "Net Wealth across the Euro Area - Why Household Structure Matters and How to Control for it"

In diesem Kapitel wird der Zusammenhang zwischen der Haushaltsstruktur und den Länderunterschieden in der Vermögensverteilung auf Basis eines neu veröffentlichten Datensatzes (Household Finance and Consumption Survey - HFCS) zum Euroraum untersucht. Anhand von nicht-parametrischen Methoden zur Umgewichtung der Haushalte können sogenannte counterfactual Verteilungen, also jene Verteilungen, die zwar nicht beobachtbar sind, jedoch vorzufinden wären, wenn die Haushaltsstruktur in allen Ländern identisch wär, geschätzt werden, um für die Unterschiede in der Haushaltsstruktur zu kontrollieren. Ein signifikanter Teil der Unterschiede in den länderspezifischen Verteilungen kann so auf die unterschiedlichen Haushaltsstrukturen zurückgeführt werden. Die Haushaltsstruktur erklärt zum Beispiel für den Unterschied zwischen dem Euroraum und dem jeweiligen Land beim Median 49%

für Österreich, 16% für Deutschland, 25% für Italien, 14% für Spanien und 38% für Malta. In anderen Ländern (Belgien, Frankreich, Griechenland, Luxemburg, Portugal, Slowenien, und Slowakei) hingegen, verschleiert die Haushaltsstruktur größere Unterschiede zwischen dem Euroraum und dem jeweiligen Land im Median des Vermögens. Für die Perzentile der Verteilung können ähnlich große Einflüsse der Haushaltsstruktur erkannt werden, wohingegen das Ergebnis von Bover (2010) in Bezug auf einen schwächeren Einfluss auf den Gini Koeffizienten bestätigt werden kann. Diese Analyse zeigt, dass internationale Rangordnungen, die die Haushaltsstruktur der einzelnen Länder außer acht lässt, mit Vorsicht zu interpretieren sind. Darüber hinaus wird für eine flexible Kontrolle der Haushaltsstruktur, die über eine Beachtung eines notwendigerweise arbiträr gewählten Haushaltsvorstandes hinaus geht, argumentiert, sofern für die Unterschiede in der Erhebungseinheit kontrolliert werden will. Das Papier gibt die Möglichkeit dies anhand von Dummy-Variablen in einfacher Weise zu bewerkstelligen.

Abstract zum Kapitel "Factor Decomposition of the Wealth Distribution in the Euro Area"

Politikmaßnahmen in wirtschaftlichem Umfeld, wie zum Beispiel die öffentliche Unterstützung zur Schaffung von privatem Wohnraum oder die Zinsentscheidungen der Geldpolitik, zielen oft auf bestimmte Komponenten der Haushaltsbilanz ab. Um die Auswirkungen dieser Entscheidungen auf die Verteilung von Haushaltvermögen analysieren zu können, wird in diesem Kapitel die Verbindung zwischen den einzelnen Vermögenskomponenten und der gesamten Verteilung des Haushaltsvermögens untersucht. Dies basiert auf sogenannten Dekompositionsmethoden. Für die Analyse anhand der Daten des Household Finance and Consumption Survey der Europäischen Zentralbank (EZB) werden die in der Literatur vorgeschlagen Methoden sowie eine neu vorgeschlagene Art und Weise der Dekomposition nach Komponenten auf Basis der 'Recentered Influence Function' verwendet. Dadurch

können in diesem Kapitel sowohl die (marginalen) Beiträge der einzelnen Komponenten zur gesamten Verteilung als auch die Elastizität dieser nach den Komponenten im Euroraum und den 15 vorhandenen Ländern separat dargestellt und diskutiert werden. Die Ergebnisse zeigen, dass manche Komponenten (wie zum Beispiel Vermögen, das in Form des Hauptwohnsitzes gehalten wird) einen größeren Teil der Ungleichheit des gesamten Vermögens erklären, wohingegen andere (wie zum Beispiel risikoreiche Finanzanlagen) relativ gesehen weniger beitragen. Auch die Elastizitäten zeigen, dass unterschiedliche Komponenten divergierende Verbindung zur Verteilung des gesamten Vermögens aufweisen. Zusätzlich treten auch länderspezifische Unterschiede im Euroraum zu Tage, die auf regionale Unterschiede von identen Politikmaßnahmen hinweisen.

Appendix B

Curriculum Vitae

Peter R. Lindner

WORKING ADDRESS
Oesterreichische Nationalbank
Economic Analysis Devision
Otto-Wagner-Platz 3
1090 Vienna, Austria
Phone: +43 1 40420 7425
E-Mail: Peter.Lindner@oenb.at

PRIVATE ADDRESS
Brunnengasse 41 Top 6-9
1160 Vienna, Austria
Phone: +43 699 17135614

EDUCATION

2009– ***PhD Candidate in Economics***
University of Vienna, Department of Economics, Austria
Proposed dissertation title: "Microeconometric analysis of households regarding the income, consumption behaviour and investment decisions using survey data"
Supervisors: Robert Kunst; Christine Zulehner
Expected completion data: Spring 2014

2007–2008 ***Master of Science in Economics***
London School of Economics and Political Science, United Kingdom
Thesis title: "How to Fight Corruption: An Empirical Analysis"
Supervisor: Ronny Razin

2002–2006 ***Magister in Economics (Master equivalent)***
University of Vienna, Department of Economics, Austria
Diploma thesis title: "Trade and Labour Markets: A comparison between the Efficiency Wage and Search Matching Model"
Supervisor: Julien Prat

EMPLOYMENT

2010– ***Economist***
Oesterreichische Nationalbank, Austria

2009 ***Referent (Senior Consulting Officer)***
Austrian Council for Research and Technological Development, Austria

2009 ***Research Assistant***
IIASA International Institute for Applied Systems Analysis, Austria

2008 ***Teacher in Economics***
Cranleigh School, United Kingdom

2006–2007 ***Research Associate***
Austrian Institute of Economic Research (WIFO), Austria

2005–2006 ***Teaching Assistant***
Technical University of Vienna, Austria

2001–2002 ***Assistant Hospital Nurse***
Inn-Salzach-Klinikum, Social Service, Germany

RESEARCH INTERESTS

Decomposition Methods, Portfolio Composition of Households, Wealth and Income Inequality, Survey Methodology, Microeconometrics

PUBLICATIONS

2013 ***Household Vulnerability in Austria - A Microeconomic Analysis Based on the Household Finance and Consumption Survey***
joint with Nicolás Albacete; Financial Stability Report 25, 2013

Verknüpfung der Informationen zum Finanzvermögen der Haushalte aus Mikro- und Makrodaten anhand des HFCS und der GFR für Österreich
joint with Michael Andreasch and Pirmin Fessler; Statistiken - Daten & Analysen, Sonderheft Sektorale VGR in Österreich 2012 [in German], 2013

2012 ***Eurosystem Household Finance and Consumption Survey 2010 Methodological Notes for Austria***
joint with Nicolás Albacete, Karin Wagner, and Siegfrid Zottel; Monetary Policy & the Economy Q3/12 - Addendum, 2012

2011 ***Decomposition of Wealth and Income using Micro Data from Austria***
OeNB Working Paper No. 173, 2011

2009 ***Social norms, third-party observation and third-party reward***
joint with Daniela Platsch and Matthias Sutter; Working Paper No. 2009-8, Faculty of Economics and Statistics, University of Innsbruck

WORK IN PROGRESS

Survey mode effects on item non response and income inequality measurement
joint with Pirmin Fessler and Maximilian Kasy

Decomposing Wealth in Europe

Household structure and the distribution of net wealth - Evidence from the Euro area
joint with Pirmin Fessler and Esther Segalla

Income Decomposition using the Income Generating Function

Micro- and Macro data: A comparision of the HFCS and the National Accounts for Austria
joint with Michael Andreasch

PRESENTATIONS

2013 ***Survey Mode Effects on Income Inequality Measurement***
Joint Statistical Meeting, Montréal, Québec

Household Assets and Households Portfolios
Annual Conference of the Royal Economic Society, Royal Holloway, University of London

2012 ***The HFCS in Austria: Methodology and First Results***
Austrian Institute of Economic Research, WIFO Extern, Vienna

Survey Mode Effects on Item Non Response and Income Inequality Measurement
HFCN Meeting ECB, Frankfurt

	Survey Mode Effects on Income Inequality Measurement 32nd IARIW General Conference, Boston
2010	***The HFCS in Austria*** Bank of Greece, Athens

SUMMER SCHOOLS

2012	***Winter School on Inequality and Social Welfare Theory*** Organized by the University of Verona
2011	***Survey Statistics*** Taught by Stephanie Eckman and Brady T. West, organized by the Bundesbank Deutschland
2011	***Barcelona GSE Summer School in Microeconometrics*** Taught by Badi H. Baltagi and Stephan Litschig
2010–2011	***Stata Programming Courses NC151 and NC152*** StataCorp LP
2009	***OeNB Summer School in Microeconometrics*** Taught by Colin Cameron

EXTRA-CURRICULAR ACTIVITIES

2009–	***Referee for the EAP, Australia*** Referee for the journal Economic Analysis and Policy (EAP) a scientific publication of the Economic Society of Australia (Queensland) Inc.
2004–2005	***Study Abroad*** Victoria University Wellington, New Zealand
1998–2000	***Elected Student Representative, Germany*** Vice chairman of the student representative board and treasurer of the students financial holdings.

COMPUTER SKILLS

STATA, Eviews, Microsoft Office, Latex, Scientific Workplace, TYPO3, z-tree, Windows Operating System, Ubuntu, Adobe Photoshop

LANGUAGE SKILLS

German (native), English (fluent), Spanish and Arabic (beginner)

References are available upon request. Last updated: October 2013.

Appendix C

Acknowledgements

First of all, I want to thank my supervisors Robert Kunst and Christine Zulehner for the critical reviews and helpful recommendations at various stages of this thesis. Obviously I am deeply indebted and grateful to my co-authors Pirmin Fessler, Esther Segalla, and Maximilian Kasy. It was and still is always a grate experience to work and discuss ideas with them as well as learn from them.

As the thesis was written while I was working for the Oesterreichsiche Nationalbank (OeNB) in the team for the first wave the the Household Finance and Consumption Survey (HFCS), it could not have been done without the thoughtful insights and critical reflection of other team members. So I am also indebted to Martin Schürz and Nicolás Albacete.

Last but not least, the whole endeavour of studying at the University of Vienna, the Victoria University of Wellington and the London School of Economics and Political Science is based on the support and background help from my family and friends to which I count a lot of great and wonderful people around the world. My deep gratefulness goes out to all of them. In the journey that eventually ended in this work there were many instances where one or more of them made it possible for me to go on and enjoy what I am doing.