Exploring the possibilities and boundaries of survey data for the analysis of wealth and wealth transfers

INAUGURAL-DISSERTATION

zur Erlangung des akademischen Grades
eines Doktors der Wirtschaftswissenschaft
des Fachbereichs Wirtschaftswissenschaft
der Freien Universität Berlin

vorgelegt von

Christian Westermeier, M.Sc.

geboren in Landshut

Berlin, 2017

Gedruckt mit der Genehmigung des Fachbereichs Wirtschaftswissenschaft der Freien Universität Berlin.

Dekan:
Professor Dr. Dr. Andreas Löffler

Erstgutachter:
Professor Dr. Dr. Giacomo Corneo

Zweitgutachter:
Professor Dr. Carsten Schröder

Tag der Disputation: 19.05.2017

Erklärung zu Vorveröffentlichungen und Zusammenarbeit mit Koautoren

Die vorliegende Dissertation umfasst eine Einleitung (Kapitel 1), sowie vier Forschungspapiere (Kapitel 2 bis 5). Kapitel 3 ist ohne Koautoren entstanden. Teile von Kapitel 3 und 5 wurden überarbeitet, bevor sie Eingang in die vorliegende Dissertation fanden.

Kapitel 2: Longitudinal wealth data and multiple imputation – An evaluation study

Dieses Kapitel ist in Zusammenarbeit mit Dr. Markus M. Grabka entstanden. Es erschien zum einen als Arbeitspapier in der Serie *SOEP Papers* am DIW Berlin, darüber hinaus wurde es im Journal *Survey Research Methods* veröffentlicht.

- SOEPpapers Nr. 790, Deutsches Institut für Wirtschaftsforschung
- *Survey Research Methods* 10(2016), S. 327-352
  DOI: http://dx.doi.org/10.18148/srm/2016.v10i3.6387

Kapitel 3: Estimating top wealth shares using survey data – An empiricist's guide

- Diskussionspapiere des Fachbereichs Wirtschaftswissenschaften der Freien Universität Berlin, Nr. 21/2016

Kapitel 4: Breaking down Germany's private wealth into inheritance and personal efforts – A distributional analysis

Dieses Kapitel ist in deutscher Sprache unter dem Titel „Erbschaft und Eigenleistung im Vermögen der Deutschen – Eine Verteilungsanalyse“ in Zusammenarbeit mit Prof. Dr. Dr. Giacomo Corneo sowie Jun.-Prof. Dr. Timm Bönke entstanden. Es erschien zum einen als Arbeitspapier, darüber hinaus wurde es in der Zeitschrift *Perspektiven der Wirtschaftspolitik* veröffentlicht.

- Diskussionspapiere des Fachbereichs Wirtschaftswissenschaften der Freien Universität Berlin, Nr. 10/2015
- *Perspektiven der Wirtschaftspolitik* 17(2016), S. 35-53
  DOI: https://doi.org/10.1515/pwp-2016-0003

Kapitel 5: Comparing the joint distribution of intergenerational transfers, income and wealth across the Euro area

Dieses Kapitel ist in zusammenarbeit mit Anita Tiefensee entstanden. Es wurde als Arbeitspapier veröffentlicht.

- Diskussionspapiere des Fachbereichs Wirtschaftswissenschaften der Freien Universität Berlin, Nr. 4/2016
- Diskussionspapiere des Deutschen Instituts für Wirtschaftsforschung, Nr. 1556

Acknowledgments

First of all, I would like to express my sincere gratitude to Giacomo Corneo for supervising my thesis, providing excellent advice, and giving me the opportunity to co-author an article, a project through which I learnt a lot, and that led to a lot of interesting discussions. I would also like to thank Carsten Schröder for co-supervising and being a good advisor on and off topic.

Special thanks go to Markus M. Grabka, whose door was always open for me, and who provided continuous guidance, support and encouragement. I cannot possibly thank him enough. I am also greatly indebted to my co-author and colleague Anita Tiefensee, who not only set up the research project that led to the topic of this dissertation, but was there for me on countless occasions for inspiring discussions, and who always provided a helping hand.

I need to thank Timm Bönke for our fruitful team work and inspiring discussions. I also would like to offer my special thanks to Charlotte Bartels for competently managing our Ph.D. program. Speaking of which, I would like to thank all the other members of the Ph.D. program for suffering through some of my less flowery talks, providing helpful comments, and making the long trips to Dahlem so much more bearable.

I would also like to thank the Hans Böckler Foundation for making both our research project and this dissertation possible.

To my sister, to my brother, as well as to all my friends for always supporting me through the highs and lows of life, and in particular, over the past few years.

To my father, who left too early – but I know he would really be proud of me now.

Lastly, since words cannot express my gratitude to my mother, I dedicate this thesis to her.

Contents

1 General introduction

The economic well-being of individuals and households is determined by their core resources income and wealth. Private wealth fulfills a multitude of functions for individuals and households, that go well beyond the utility derived from consuming parts or all of the regular income. Davies and Shorrocks (2000) name a few: private wealth is useful for consumption smoothing, households prepare themselves for shocks and expected low income periods (precautionary savings, old-age protection); it is also accumulated to make bequests to the descendants; large fortunes are associated with both economic and political power. Hauser (2007) refines the purposes of wealth even further: individuals generate capital income from their investments directly for consumption; parts of the wealth portfolio, such as real property, directly benefit their holders through their usage; wealth signals status and wealth indicates a household's position, its upward mobility; and lastly, it also benefits the socialization of children through better education. It is telling, then, that private wealth is much more unequally distributed than disposable incomes (Figure 1.1).

According to the *OECD Wealth Distribution Database* and *Income Distribution Database*, which includes 18 OECD countries, the average share of household disposable income in the top decile is roughly 25 %, whereas the richest wealth decile holds more than 50 % of all assets (see also OECD, 2015). Data suggest that these estimates are conservative with respect to many countries such as the United States (Bricker et al., 2014) or Germany (Westermeier and Grabka, 2015). Evidence about the distribution of incomes in Germany seems to be readily available from survey data and published on a regular basis (see Goebel et al., 2015), and even top incomes appear to be represented well in surveys

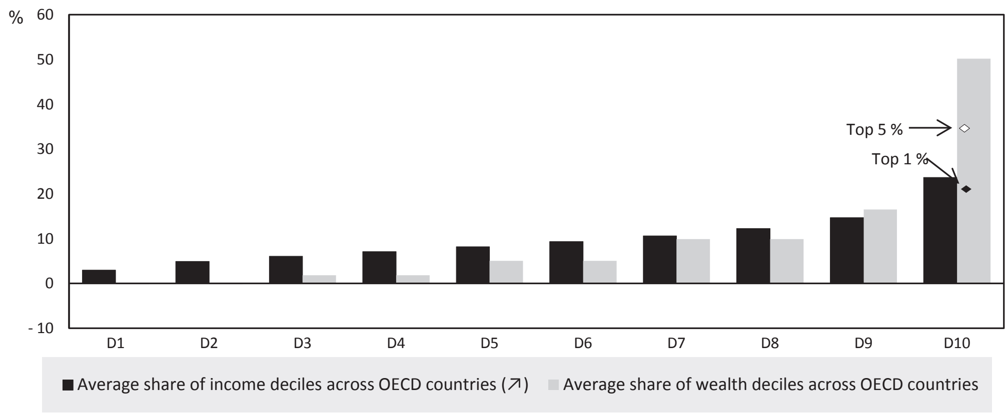


Figure 1.1: Distributions of household disposable income and net worth across deciles. Average over 18 OECD countries. Source: OECD Wealth Distribution Database and OECD Income distribution data base, OECD (2015).

(Bartels and Schröder, 2016). It is considerably more difficult to adequately map the asset holdings of households due to the skewness of its distribution. However, evidence about the distribution of wealth and its development in the long run are supposed to be core criteria for decision makers in social policy.

There are alternatives to survey data, such as register data (Waldenström, 2016) and evidence from wealth or inheritance tax statistics (Henrekson and Waldenström, 2016). Researchers also draw conclusions from capitalized incomes reported by taxpayers (Saez and Zucman, 2016). However, in Germany the wealth tax was abandoned in 1997, after the German Federal Constitutional Court (*Bundesverfassungsgericht*) rightfully criticized its inconsistent taxation of real property and other assets in 1995. In effect, even for the time period up to 1997, the tax data are hardly useful for researchers, as deviating definitions and assessment rules do not match with research agendas and would paint a biased picture (Bartels and Bönke, 2015). Alternatively, an official statistics of inheritances and gifts is available, as wealth transfers still are subject to taxation in Germany (Bach et al., 2014b; Bach and Thiemann, 2016). However, a closer look reveals that the tax statistics only records the aggregate of taxable wealth (*Reinnachlass*). The taxable wealth accounts for a mere fraction of the actually transferred assets, as only transfers exceeding a certain amount–corresponding to generous high tax allowances–are subject to taxation in the first place.

The launch of an income and consumption sample (*Einkommens- und Verbrauchstichprobe* EVS) by the Federal Statistical Office would have presented researchers with an opportunity to access micro data on private wealth from 1978 onwards. However, a major drawback is that until 1993 real property was valuated with a uniform price (*Einheitswert*) instead of market values resulting in limited comparability (Statistisches Bundesamt, 2014). Furthermore, if the income exceeds a certain threshold, households are excluded from the sample, severely limiting the usability of the sample for the analysis of highly concentrated net worth. A time series on aggregate private wealth published jointly by the German Federal Statistical Office and the German Federal Bank (Statistisches Bundesamt, 2013) is available from 1991 onward. However, it also includes the private non-profit sector and real property is evaluated with replacement values, which frequently deviate from the market values (Grabka and Westermeier, 2015). Once decidedly favorable official data sources such as register data or wealth tax data are dried up, household survey data remains as an empiricist's last resort.

This thesis consists of four research papers, two of which are setting out to improve the survey data situation, and two of which consider the joint distribution of wealth and wealth transfers, in order to assess the possibilities and boundaries of inheritance data collected in surveys. In Chapter 2, it is shown, by means of a simulation project similar to Watson and Starick (2011), how the process of multiply imputing for item non-response might be adjusted to account for very unequally distributed wealth assets. Adjusting for item non-response, however, is not enough to compensate for missing high-net-worth-individuals in the data. Once wealthy households are missing from the data, due to either non-observation bias (Eckerstorfer et al., 2015) or systematic unit non-response (Kennickell, 2007, 2009; Bover et al., 2014; Vermeulen, 2014), top wealth shares and aggregates calculated from survey data are biased downward. Chapter 3 sets out to guide empiricists through several cures that have been proposed to counter the absence of the wealthiest households. In Chapter 4 the joint distribution of wealth and wealth transfers, as recorded by a new

German household survey, is exploited to assess the role of inheritances and gifts for the current distribution of household net worth. In the fifth chapter of this doctoral thesis, the scope is widened to assess the role of inheritances and gifts for households' financial situation in a European cross-country comparison and focusing on the correlation between wealth transfers and both disposable income and education.

1.1 Wealth

Statistical analysis in surveys is generally facing missing data. The second chapter, entitled 'Longitudinal wealth data and multiple imputation – An evaluation study', is dedicated to the successful imputation of missing items in survey wealth data. As in longitudinal studies for some missing values there might be past or future data points available, the question arises how to successfully transform this advantage into improved imputation strategies. In a simulation study, six combinations of cross-sectional and longitudinal imputation strategies for wealth panel data are compared. The imputation quality is assessed using wealth data collected for the German Socio-Economic Panel study (SOEP) from waves 2002, 2007 and 2012 (Frick et al., 2007, 2010b; Grabka and Westermeier, 2014). The simulation data sets are generated by blanking out observed data points: item non-response is induced by a missing at random (MAR) and two differential non-response (DNR) mechanisms. Three imputation methods are considered: a state-of-the-art multiple imputation using chained equations (MICE, see for instance Royston, 2004 and van Buuren et al., 2006), an imputation procedure for panel data known as the row-and-column method (Little and Su, 1989) and a regression prediction with correction for sample selection. The regression and MICE approaches serve as fallback methods if only cross-sectional data is available.

The contribution of this chapter to the literature is manifold. First, single imputation proves to have undesired properties, because the uncertainty reflected by the respective parameters based on a single stochastic imputation is likely to be biased downwards, since the estimators treat the imputed values as if they were actually observed ones (Rubin, 1986, 1987). Yet, many surveys still address missing values with single imputation methods (e.g. wealth in the Panel Study of Income Dynamics, 2011; income variables in the SOEP, see Frick and Grabka, 2005). The drawbacks of case-wise deletion strategies have been well documented (Little and Rubin, 1987). Multiple imputation addresses this issue. In many ways this work is a follow-up study to the evaluation study of single imputation methods for income panel data conducted by Watson and Starick (2011). However, apart from their focus on income variables there are quite a few more differences: they only consider the MAR assumption as a non-response generating mechanism, an issue that is addressed in this study. Furthermore, they focus on single imputation methods and leave it to other researchers to evaluate the performance of multiple imputation methods.

Despite a lack of theoretical justification, the US Survey of Consumer Finances (Kennickell, 1991) and its European counterparts in Spain or France apply imputation procedures similar to MICE for the imputation of cross-sectional wealth variables (see Bover, 2004). Since the initiative for a harmonized European panel survey on household finances started (see European Central Bank, 2013a,b), the question of how to impute for missing wealth items in panel data is of renewed interest.

As Chapter 2 shows, the univariate row-and-column method by Little and Rubin (1987) performs surprisingly well considering the cross-sectional evaluation criteria. For trend estimates and the measurement of inequality, combining MICE with the row-and-column technique regularly improves the results based on our catalogue of six evaluation criteria including three separate inequality indices. As for wealth mobility, two additional criteria show that a model based approach, such as MICE, might be the preferable choice. Overall the results show that if the panel variables, which ought to be imputed, are highly skewed, the row-and-column technique should not be dismissed beforehand.

However, once a cure for item non-response is found, survey wealth data quickly shows its next chronic disease: survey data tends to be biased towards the middle class.[1] In the

1 The term 'middle class bias' is typically associated with income variables as documented in Riphahn and Serfling (2005) or Frick and Grabka (2005).

third chapter, entitled 'Estimating top wealth shares using survey data – An empiricist's guide', we take into consideration that survey data often fails to adequately cover the highly relevant group of multi-millionaires and billionaires, which in turn results in biased estimates for both aggregate wealth and top wealth shares, and yields large corridors of uncertainty (see Westermeier and Grabka, 2015). In order to overcome the under-coverage and obtain more reliable measurements of wealth inequality, researchers are simulating the tail of wealth distributions using Pareto distributions both with and without information on high-net-worth-individuals from rich lists (see Bach et al., 2016 and Vermeulen, 2016 for recent examples). In a series of Monte Carlo experiments, this chapter assesses the determining factors for such an exercise to yield reliable results. The contribution of this chapter is to shed light on some aspects of enhancing survey data using Pareto simulated tails that previously have been neglected.

First, aggregate private wealth and top wealth shares are estimated under conditions typically encountered by empiricists. It is shown that wealth data, which is plagued by differential non-response, as opposed to a non-observation bias, might not be treated with a simple maximum likelihood estimation of the top tail based on survey data alone (as in Eckerstorfer et al., 2015), as estimates are still inherently biased downward. Including rich list data and switching to a regression estimation (as in Bach et al., 2014a, 2016 or Vermeulen, 2014, 2016) impacts top wealth shares, but the aggregate wealth remains biased downward. In the last step of the simulation, it is show what potential effects are to be expected, if publishers of rich lists data systematically overestimate the top fortunes as suggested by Raub et al. (2010). Overall, all empirically encountered estimations of the aggregate wealth and top wealth shares using corrected data yield inherently biased results, once the survey weights are uninformed and no additional data is available for calibration. In an application using German survey wealth data, it is shown that re-weighting the provided frequency weights based on exogenous information possibly affects the estimates more severely than choosing the right parameters of the Pareto distribution. However, three separate empirically derived functional forms of non-response yield wildly different estimates. The validity of exogenous data–and the rich list data–remains a matter of trust on the part of the empiricist.

1.2 Inherited wealth

In Germany, taxes on inheritances and gifts are virtually regressive due to its comprehensive exemptions on large assets (Bach and Thiemann, 2016). Research suggests that saving rates from income and intergenerational wealth transfers (inheritances and gifts) are two key determinants of wealth held by private households (Davies and Shorrocks, 2000); accordingly, we observe a surge in research on inherited wealth over the last few years (Semyonov and Lewin-Epstein, 2013; Arrondel et al., 2014; Mathä et al., 2014; Fessler and Schürz, 2015). A key point since the 1980s is the debate over which of the two determinants contributes more to the current net worth of private households (Modigliani, 1986, 1988; Kotlikoff and Summers, 1981; Kotlikoff, 1988). Recent research stresses that intergenerational transfers are a dominant factor for households' positions in the distribution of wealth (Piketty, 2011, 2014; Piketty and Zucman, 2015), thus fueling a discussion about the legitimacy of wealth without effort. Figure 1.2 depicts the inheritance flow in France, Germany and UK as reported in Piketty and Zucman (2015), himself drawing from works by Schinke (2013) for Germany and Atkinson (2013) for Britain. The inheritance flow as a percent of national income suggests that after two economic shocks due to World War I and II the inheritance flows are rebounding to their former levels since the 1980s.

The fourth chapter 'Breaking down Germany's private wealth into inheritance and personal efforts – A distributional analysis' investigates the role of inheritance in the distribution of wealth in Germany. Recently collected survey data from the German Panel on Household Finances (PHF)[2] allows to compute inheritance-wealth ratios for various quantiles based on several assumptions concerning the capitalization of past bequests and

2 For an overview of PHF waves 1 and 2 see Deutsche Bundesbank (2013, 2016).

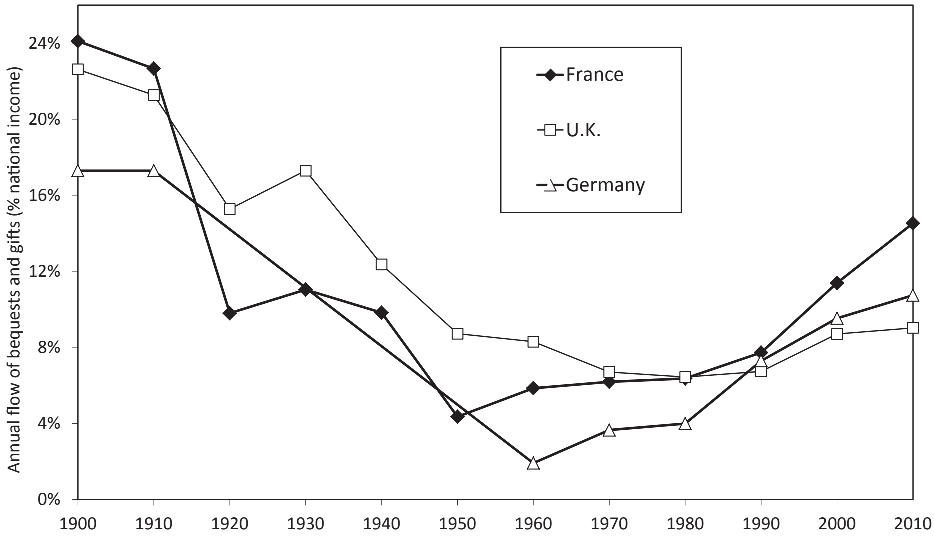


Figure 1.2: Inheritance flow in France, Germany and UK, 1900–2010.
Source: Piketty and Zucman (2015, p. 1339).

gifts. Traditionally, the classic methodology introduced by Kotlikoff and Summers (1981) and Kotlikoff (1988) was opposed by Modigliani (1986, 1988). The former proposed to capitalize past inheritances and gifts and compute the inheritance-wealth ratio. The latter proposed to merely adjust past inheritances and gifts for inflation before computing the inheritance-wealth ratio. However, none of the two concepts provided researchers with a satisfying formula to compute inherited wealth as a percent of net worth–and their application yields wildly different results. Thus, in Chapter 4 the analysis relies on a new approach by Piketty et al. (2014), which we deem superior to the prior approaches, as it divides households into 'rentiers' and 'savers'. As *rentiers* consumed more than they would have been able to from their labor income alone, their inheritance-wealth ratio is 100 %, which cannot be exceeded by any household. This is the major advantage of the method as compared to both Kotlikoff and Summers (1981)/Kotlikoff (1988) and Modigliani (1986, 1988), as these approaches allow single households, or whole (sub-)populations to have inherited more than 100 % of their current net worth.

However, as shown in Chapter 3, the results based on survey data might be severely biased to the middle class. The PHF markedly improves the data situation with regard to wealth and inheritances in Germany, but it still applies that participation is voluntary and false information is not penalized. It is reasonable to assume that particularly the richest one percent of the population concentrates large fortunes and inherits fundamentally different portfolios of assets (primarily valuable business assets). Therefore, in another exercise, we assume that the results for the bottom 99 % of the households are intact and combine them with exogenous sources for wealth (Vermeulen, 2014) and inheritance (Piketty and Zucman, 2015) for the overall population, thus, assessing the role of inherited wealth for the top percentile anew.

Our results indicate that wealth inequality below the top-1% is hardly affected by inheritances: the share of inheritances in wealth is about one third on average and it does not change much across quantiles of the wealth distribution. We also find that retirees exhibit similar inheritance-wealth ratios to the whole population. Additionally, including pension wealth reduces the significance of past wealth transfers for the poorer wealth deciles in particular, whereas the pension wealth of the *upper middle class* or the *upper class* reduces the ratio only by a few percentage points. The findings for the *upper class* are not the result of the low interest rate: we modify this assumption and assume that the top-1% might invest into risky stock markets with higher returns. We find that not until an equity risk premium of 9 % would their inherited wealth share align with the value observed for the middle class. However, the combination of PHF data with alternative sources arrives at a much higher inheritance-wealth ratio for the wealthiest: more than 80 % of their wealth might be inherited.

The fifth chapter 'Comparing the joint distribution of intergenerational transfers, income and wealth across the Euro area' widens the scope of Chapter 4. We investigate the current role of wealth transfers in the Euro-area (Austria, Belgium, France, (West) Germany, Cyprus, Greece, Portugal, and Spain). Whereas harmonized data on household incomes has been available and is used by researchers (Bönke and Schröder, 2014), the availability of harmonized wealth data was limited before the Eurosystem Household Finance and

Consumption Survey (HFCS)[3] was released, this is the first time that cross-country comparisons focusing on Europe are possible. We also contribute to the literature by giving an overview of the inheritance and gift taxation in each country. The chapter describes the distribution of intergenerational transfers in the Euro-area in absolute terms and analyzes the socio-demographic characteristic of heirs applying several regression analyses via logit and OLS. Additionally, we analyze the role of past intergenerational transfers for current net worth using recently established methods by Wolff and Gittleman (2014) and Piketty et al. (2014) as well as fractional logit models that explain the relative importance of transfers received, while controlling for several socio-demographic characteristics simultaneously.

The joint distribution of income and transfers reveals that the relationship between income and the propensity to receive an inheritance or gift is higher in core Europe, indicating less intergenerational mobility, whereas the correlation between income positions and the capitalized present values of those transfers is high across the board. A series of country-specific multivariate regressions confirms these findings and suggests that higher education also goes hand in hand with higher absolute transfer values. As expected, the present transfer values monotonously increase with household net worth. However, when analyzing the capitalized present value as a percent of current net worth on the household level we see some of the results reversed, as apparently the relative importance of intergenerational transfers does not increase with the level of income or wealth. Using a fractional logit regression we find that for higher income quintiles the ratio of current net worth attributable to past intergenerational transfers tends to be decreasing.

1.3 Concluding remarks

In summary, Chapters 2 and 3 explore possibilities to improve the measurement of wealth distributions with data collected in household surveys. The results concerning item non-response in wealth surveys are encouraging. A different set of imputation methods might

3 For an overview of the first wave of HFCS data, harmonization and participating countries see European Central Bank (2013a,b).

improve the imputation quality for highly skewed variables such as wealth assets and liabilities. The results concerning unit non-response–and differential non-response–are decidedly less encouraging. The Monte Carlo experiments suggest that the methods proposed so far, for correcting for the missing rich in household surveys, fail to yield reliable results under empirically encountered circumstances. Top wealth shares and aggregate wealth remain biased if the survey weights are uninformed about the actual response probabilities. The findings emphasize the need to use exogenous information in sample design. Moreover, they underline the present lack of exhaustive data sources and constitute an appeal to tax authorities to cooperate more closely with researchers and survey data providers.

This also applies to chapters 4 and 5 of this thesis. While the results for 99 % of the households are firmly rooted in the empirical framework, there is reason to believe that the top tail of the German wealth distribution is insufficiently covered by surveys such as the SOEP or the HFCS, which in turn inhibits the computation of inheritance-wealth ratios for the wealthiest households. HFCS data providers from France and Spain greatly benefit from the stratification using tax registers, and are enabled to release superior data. Chapters 4 and 5 suggest that in order to limit the tax burden to inherited wealth an inheritance and gift tax is preferable to a wealth tax. For research purposes, only data derived from the latter would yield sufficient evidence about the distribution of wealth for the whole population. Wealth-related tax registers for scholarly use do not necessarily increase the tax burden, however, as a 0% tax-rate would suffice.

2 Longitudinal wealth data and multiple imputation – An evaluation study

2.1 Introduction

Large-scale surveys are usually facing missing data, which poses problems for researchers and research infrastructure providers alike. In longitudinal studies for some missing values there might be past or future data points available. The question arises how to successfully transform this advantage into improved imputation strategies. Single imputation proves to have undesired properties, because the uncertainty reflected by the respective parameters based on one single stochastic imputation is likely to be biased downwards, since the estimators treat the imputed values as if they were actually observed ones (Rubin, 1987, 1986).[4] Multiple imputation addresses this issue. Our study examines the performance of several multiple imputation methods for the adjustment for item-non response (INR) in wealth panel data. Wealth is considered a sensitive information that is usually collected with rather high non-response rates compared to less sensitive questions such as demographic variables like age, sex, migration status (e.g. Riphahn and Serfling, 2005; Frick et al., 2010b). In addition, there is a rather high state-dependency in terms of ownership status of wealth components, which facilitates the consideration of longitudinal information in the imputation process.

In many ways this work is a follow-up study to the evaluation study of single imputation

4 The drawbacks of case-wise deletion strategies have been well documented (Little and Rubin, 1987).

methods for income panel data conducted by Watson and Starick (2011). They conclude their study with a few remarks: future research should test the performance of imputation methods under different assumptions concerning the non-response mechanism, an issue that we are trying to address in this study. Furthermore, they focus on single imputation methods and leave it to other researchers to evaluate the performance of multiple imputation methods. Again, this is something we are tackling with this study. In our simulation study we compare six combinations of cross-sectional and longitudinal imputation strategies for German wealth panel data collected for the German Socio-economic Panel Study (SOEP) in 2002, 2007 and 2012. We create simulation data sets by setting observed data points to missing based on three separate non-response generating mechanisms. We examine the performance of imputation models assuming the mechanisms are missing at random (MAR) or the data suffers by differential non-response (DNR). We test the performance of multiple imputation by chained equations (MICE, named after one of the first popular implementations, see Royston, 2004). We test a univariate imputation procedure for panel data known as the row-and-column method introduced by Little and Su (1989). Additionally, we test a regression specification with correction for sample selection including a stochastic error term, which was the standard imputation method for the SOEP wealth data in survey waves 2002 and 2007.

The paper is organized as follows: Section 2.2 gives an overview of wealth surveys and their imputation strategies and of item non-response in the SOEP wealth data, Section 2.3 describes how we generate simulation data sets with missing values from observed cases. Section 2.4 explains the evaluation set-up in detail and the criteria we are choosing to compare the imputation methods. In Section 2.5 we summarize the imputation methods and discuss their strengths and weaknesses. Section 2.6 details the performance of these methods using our simulated wealth data derived from the SOEP. Section 2.7 concludes.

2.2 Wealth surveys and incidence of item non-response in SOEP wealth data

Household panel surveys typically provide their users with imputed information. However, such surveys differ with respect to the imputation strategies applied to address item non-response and also in the way how available longitudinal information is incorporated. In the following we present panel surveys, which collect wealth information, and their imputation strategies. Their consideration might give useful clues for the imputation of wealth data in this study.

The recently established Eurosystem Household Finance and Consumption Survey (HFCS) is a household survey conducted in 15 euro area countries and organized by the European Central Bank (ECB) (see European Central Bank, 2013b). This survey uses an iterative and sequential regression design for the imputation of missing data, similar to the sequential approach we evaluate in this paper (see Section 2.5.1). The method used by the HFCS is adopted from similar surveys by the Federal Reserve Board and Banco de España (see Kennickell, 1991, 1998; Barcelo, 2006). The number of implicates provided by the HFCS is five, which seems to be the generally agreed on number of imputations provided with survey data.[5] In most of the participating countries the HFCS will be continued as a panel study (see European Central Bank, 2013a). However, the sequential approach the data providers are using has only been tried and tested in cross-sectional surveys thus far. We argue that the evaluation of multiple imputation strategies for longitudinal wealth data will increase in relevance in the future.

The Survey of Health, Aging and Retirement in Europe (SHARE) is a cross-national panel survey including more than 85,000 individuals from 20 European countries aged 50 and older. SHARE also imputes data using a method that is similar to MICE (see Christelis, 2011).

The Household, Income and Labour Dynamics in Australia Survey (HILDA) is a

5 The same number of implicates is also provided by e.g. the SCF, the SOEP, and SHARE.

household-based panel study which collects information about economic and subjective well-being, labor market dynamics and family dynamics in Australia (see Watson and Mark, 2002). HILDA uses a combination of nearest neighbor regression imputation and the row-and-column imputation, depending on the availability of longitudinal information from other waves of the survey (Hayes and Watson, 2009). The US panel study of income dynamics (PSID) is the longest running household panel survey, it started in 1968. The PSID asks about nine broad wealth categories; INR is imputed using a single hot-deck imputation technique, home equity is imputed using a simple carry-forward method (see Panel Study of Income Dynamics, 2011).

The German Socio-economic Panel Study (SOEP)–the survey used for this study–is a longitudinal representative survey collecting socio-economic information on private households in Germany (Wagner et al., 2007). In contrast to other wealth surveys that interview only one household representative, the SOEP collected wealth information separately for all household members (with age 17 or older) in 2002, 2007 and 2012. This survey strategy seems to be advantageous compared to collecting wealth information by one reference person per household only, given that accuracy and comparability to official statistics seem to perform better (Uhrig et al., 2012). One major drawback of this strategy is inconsistency on the household level. Given that asset values held by several household members can deviate from each other and may result in an even higher share of INR. The major disadvantage of surveys collecting the data solely interviewing one reference person is that the risk to overlook wealth, assets or debts of other household members increases. However, the methods we test in this evaluation study can be easily applied to wealth data collected at the household level, we do not expect the results to be significantly different in such a set-up. The first wave of SOEP data was collected prior to the German reunification in 1984 with 12,245 respondents. The original sample was eventually supplemented by 10 additional samples to sustain a satisfactory number of observations and to control for panel effects. In 2002, an additional sample of high-income earners was implemented (2,671 individuals), which is particularly relevant for the representation of high net worth individuals in the sample given that income and wealth is rather highly correlated. In 2012, more than 21,000 individuals were interviewed.

The SOEP wealth module collects 10 different types assets and debts: value of owner-occupied and other property (and their respective mortgages), private insurances, building loan contracts, financial assets (such as savings accounts, bonds, shares), business assets, tangibles and consumer credits.

A filter question is asked whether a certain asset is held by the respondent, then the market value is collected and finally information about the personal share of property is requested (determining whether the respondent is the sole owner or, if the asset is shared, the individual share).

The imputation of wealth data consists of three steps (for more information see Frick et al., 2007, 2010b): First, the filter imputation determines whether an individual has a certain asset type in his or her portfolio. These variables are imputed using logit regression models. Second, the metric asset values are imputed. And third, a personal share is imputed with logit regressions. In this simulation study we concentrate on item non-response (INR) for the metric asset values.[6]

In Table 2.1 we summarize the observed INR incidences for the SOEP wealth data 2002, 2007 and 2012 for the metric values and the filter variables. The respective share of INR varies between about zero for debts on other property and about 14 percent for private insurances.

6 (Partial) unit non-response and wave non-response–persons or households dropping out of the sample for a limited time or permanently–do not receive any imputation treatment in the person-level SOEP wealth data. Unit non-response generally is addressed by survey weighting procedures (see Kalton, 1998).

Table 2.1: Item non-response rates in SOEP Wealth Questions.

Wave	Type of wealth question		missing filter information	share of missing filter in %	missing (metric) values*	share of missing values* in %
	gross	home market value	83	0.48	1,104	4.60
	wealth	other property	227	0.79	453	1.90
		financial assets	418	1.89	1,822	7.63
		building-loan contract	(in 2002 together with private insurances)			
2002		private insurances	333	1.53	3,308	13.85
(n=23,892)		business assets	243	1.15	350	1.46
		tangible assets	373	1.70	592	2.48
	gross	debts owner-occupied property	-	-	63	0.26
	debt	debts other property	-	-	6	0.00
		consumer credits	251	1.19	366	1.53
	gross	home market value	139	0.67	1,093	5.23
	wealth	other property	178	0.85	364	1.74
		financial assets	239	1.14	1,931	9.25
		building-loan contract	187	0.90	921	4.41
2007		private insurances	221	1.06	2,781	13.32
(n=20,886)		business assets	177	0.85	290	1.39
		tangible assets	199	0.85	214	1.02
	gross	debts owner-occupied property	-	-	179	0.86
	debt	debts other property	-	-	40	0.19
		consumer credits	180	0.86	212	1.02
	gross	home market value	308	1.68	958	5.22
	wealth	other property	350	1.91	341	1.81
		financial assets	470	2.56	1,469	8.00
		building-loan contract	349	1.90	812	4.42
2012		private insurances	390	2.12	2,385	12.99
(n = 18,361)		business assets	344	1.87	270	1.47
		tangible assets	402	2.19	196	1.07
	gross	debts owner-occupied property	-	-	276	1.50
	debt	debts other property	-	-	53	0.29
		consumer credits	395	2.15	219	1.19

(*) Note that the absolute number of missing metric values, as well as the share, is determined by the sample members who did report that they are holding a certain asset type and could not or refuse to provide a value, it excludes all members who did not report filter information, which has yet to be determined in a separate pre-value imputation. That is why for some variables with a low incidence (such as business assets) the filter information is missing for more individuals than the metric value.

2.3 Simulating non-response

The first step in every imputation procedure that accounts for INR in a data set is to make an assumption concerning the non-response mechanism, which may be either explicitly formulated or implicitly derived from the imputation framework. The commonly used framework for missing data inference traces back to Rubin (1976), who differentiates the response mechanism for three assumptions: Missing Completely At Random (MCAR), Missing At Random (MAR) and Missing Not At Random (MNAR). If the observation is assumed to be MCAR the probability of an observation being missing does not depend on any observed or unobserved variables. With MCAR, excluding all observations with missing values yields unbiased estimators, but also results in a loss of efficiency. Under MAR, given the observed data, the missing values do not depend on unobserved variables. That is, two units with the same observed values share the same statistical behavior on other variables, whether observed or not. If neither of the two assumptions holds, the data is assumed to be MNAR: the response status is dependent on the value of unobserved variables (e.g. the missing value itself) and cannot be accounted for by conditioning on observed variables.

The most commonly used assumption about the non-response mechanism is MAR. However, ‘as with other statistical assumptions, [...] the missing at random assumption may be a useful approximation even if it is believed to be false’ Allison (1987, p. 77). Thus, we focus on the evaluation of the imputation methods described in Section 2.5 only assuming MAR and two variants of MNAR. We focus on three components of the asset portfolio covered by the SOEP: home market value, financial assets and consumer credits. Home market value is easily the most important component in the average wealth portfolio in Germany. Financial assets are subject to both comparatively high non-response rates and rather high incidences. Additionally, regression models for the home market value tend to yield a good model fit, whereas models for financial assets tend to have a relatively poor model fit (Frick et al., 2007). We choose consumer credits as the third component to cover in this study, because it exhibits rather low incidences and modeling for both response and asset value tends to fare mediocre; the reason being that the imputation cannot rely on a high number of sound co-variates given that the SOEP does not collect additional information about this type of liability in comparison to other assets.

A large pool of fully observed observations remains after blanking out all INR cases, which turns out to be useful for the creation of simulation data sets. Depending on component and wave there are between 2 291 and 8 103 nonzero asset values (see the sum of ‘Number to be imputed’ and ‘Nonzero observations’ in Table 2.1). Since it is not possible to compare imputed values with the true ones in our imputation set-up, we need to go one step back and create a simulation data set. Basically, we estimate a set of logit regression models for the non-response mechanism based on the full data set including all observations with empirically missing data.

The variables included in the non-response models are the employment status und the total personal income, the interview mode, a set of socio-demographic variables (e.g. gender, age, number of children, years of schooling, region) and a rather small set of supplemental economic indicators (e.g. financial support received). Additionally, a set of dummies indicate non-response in other wealth components in the same survey wave and a lag (or lead) dummy variable indicates non-response of the same variable in one of the other waves as state-dependency matters for INR in subsequent waves (Frick and Grabka, 2005). The simulation data sets, then, are generated by taking all complete cases of one wealth variable and one wave and predicting missingness based on the non-response models and conditional on non-response in other wealth variables and in other waves as already predicted. In order to fully generate the same patterns of missing values, depending on missingness in other variables and waves in the simulation data set, we need to update the prediction in a second sequence.

However, since then the predicted probability that the value of a certain wealth component is missing is highly dependent on whether the value has been observed in any of the two other waves, the share of observations in our simulation data sets with non-response in every wave was too high compared to the original dataset, as the information on the response status in other waves is the most important predictor. Therefore we added a small stochastic component to the predictions to incorporate uncertainty. After the addition of this random error terms the share of observations for which information from the other two waves is available for longitudinal imputation is approximately the same as in the original datasets.[7]

Table 2.2 displays the McFadden R^2 for the non-response models under MAR, the number of observations with missing values and the number of nonzero observations for the simulation assets and waves. Note that the number to be imputed is fixed at around 10 percent of all valid nonzero observations, which is a rather high non-response incidence for home market value and consumer credits. The share of missing values for questions concerning the financial assets tends to be higher than 10 percent. However, the majority of our performance criteria are not affected by the share, as the focus is on the differences between imputed and observed data sets using only the respective imputed cases.

However, to assume the (non-)response mechanism is fully explained once we conditioned on observed variables may be putting things too simple. Thus, we simulate two additional response mechanisms under the assumption of differential non-response: in two different set-ups we assume that the probability to provide the value of a certain asset depends on the value itself. The empirically observed relationship between non-response incidence and the corresponding values tends to be U-shaped, which is better documented for income questions than it is for wealth questions: In fact, Frick and Grabka (2005) state that the incidence for non-response of a component of the post-government income for the lowest and highest income deciles is between 28 and 60 percent higher than for the fifth and sixth income deciles. Additionally, characteristics that are typically observed for low income and low wealth households, such as level of schooling and part time employment, have significant explanatory power in non-response models (Riphahn and Serfling, 2005). As Kennickell and Woodburn (1997) conclude with U.S. wealth data, the higher the household wealth is, the higher the probability that the household refuses to participate.[8]

7 Sequentially inducing non-response across several waves, assets and NR assumptions is a lengthy and complex exercise; the code for this section as well as Sections 2.4 and 2.6 is available to researchers, we urge our readers to not hesitate to contact us, if anything is unclear. The code covering the data preparation and imputation is based on the imputations of waves 2002 and 2007 and even lengthier; as it would be a massive undertaking to provide it with decent commentary, it is available from the authors upon request.

8 Vermeulen (2014) gives a comprehensive overview of the potential effects of differential non-response for high-net-worth-individuals on the measurement of inequality in the European HFCS survey data.

Table 2.2: Descriptive statistics for observed and simulated data (#1).

INR assumption	Wave		McFadden R^2	Mean in €	Number to be imputed	Nonzero observations	Coefficient of Variation
Observed	2002	Home market value	-	243,769	-	7075	0.731
		Financial assets	-	39,798	-	8103	3.209
		Consumer Credits	-	26,544	-	2088	4.792
	2007	Home market value	-	237,508	-	6775	0.762
		Financial assets	-	40,114	-	8377	3.651
		Consumer Credits	-	17,935	-	2978	2.850
	2012	Home market value	-	230,613	-	6164	0.726
		Financial assets	-	44,740	-	7377	2.901
		Consumer Credits	-	16,866	-	2552	4.911
MAR	2002	Home market value	0.595	225,724	707	6368	0.773
		Financial assets	0.410	44,921	810	7293	2.026
		Consumer Credits	0.524	26,475	208	1880	1.733
	2007	Home market value	0.518	214,858	677	6098	0.746
		Financial assets	0.391	54,026	837	7540	6.060
		Consumer Credits	0.618	16,191	297	2681	2.048
	2012	Home market value	0.540	202,057	637	5527	0.789
		Financial assets	0.406	59,015	737	6640	3.010
		Consumer Credits	0.597	18,689	255	2297	1.871
DNRI	2002	Home market value	-	204,609	716	6359	0.634
		Financial assets	-	15,762	808	7295	1.894
		Consumer Credits	-	10,168	176	1912	1.801
	2007	Home market value	-	190,218	692	6083	0.756
		Financial assets	-	11,242	809	7568	2.917
		Consumer Credits	-	6,190	301	2677	2.304
	2012	Home market value	-	195,064	636	5528	0.873
		Financial assets	-	11,287	773	6604	2.306
		Consumer Credits	-	6,682	256	2296	1.871
DNRII	2002	Home market value	-	283,085	760	6315	0.705
		Financial assets	-	73,853	805	7298	2.253
		Consumer Credits	-	39,505	209	1879	1.748
	2007	Home market value	-	284,654	637	6138	0.800
		Financial assets	-	75,950	858	7519	2.690
		Consumer Credits	-	41,856	309	2669	2.334
	2012	Home market value	-	301,754	626	5538	0.924
		Financial assets	-	84,956	763	6614	2.629
		Consumer Credits	-	36,835	261	2291	6.917

Source: SOEP v29, the number of observations to be imputed in the simulated data sets vary slightly around 10 percent of the nonzero observations in the observed data sets, as the exact number of missing values in each data set depends on a stochastic components under both MAR and DNR. Likewise, the exemplary results here for MAR, DNR1 and DNR2 are from #1 of the randomly generated data sets.

Under the assumption that wealth components share a similar non-response behavior, we assume in the DNR1 data sets that the probability that a value is missing is the higher, the lower the true value is (i.e. differential non-response at the bottom of the distribution). In the DNR2 data sets, we assume the contrary, the higher the actual value of the asset the higher is the probability that the value is missing. Table 2.2 compares the effects on the mean and the coefficient of variation of one of the respective generated simulation data sets. Consequently, the means for the observations to be imputed in the DNR1 data sets are substantially lower, whereas in the DNR2 data sets they are substantially higher than in the data sets containing all observed cases.

As all non-response generating mechanisms have a stochastic component, we can easily repeat the steps involved for each assumption to generate 1 000 simulation data sets per item non-response assumption. Those 1 000 data sets are imputed separately using each of the six imputation methods presented in Section 5, yielding in total 3 x 6 x 1 000 imputation procedures.

2.4 Evaluation criteria

Our evaluation criteria differ from those of Watson and Starick (2011); we focus on a set of 8 instead of 11 criteria applied by the authors. We divide the main applications of wealth data into three sections. (I) Cross-sectional analyses focus on point estimates, trend and distributional analyses. (II) Inequality measurement focuses on the computation of the GINI coefficient and other inequality indices. (III) Longitudinal analyses focus on wealth mobility. (I) and (II) are rather closely related and should be adequately replicated by the imputation procedure. (III) is an additional focus, which we tackle in a separate evaluation. We divide the criteria into two subsets to account for the comparatively higher importance of wave-specific trend and inequality analyses (six criteria in Section 2.4.1) compared to rare analyses that specifically make use of the panel structure of the data (two additional longitudinal criteria in Section 2.4.2). Ultimately, an ideal imputation model would account for cross-sectional, longitudinal and inequality accuracy.

Generally, multiple imputation is supposed to yield valid inference as, in comparison to single imputation, the parameters calculated using imputed data do not exhibit biased standard errors. Thus, in the last step of this evaluation we assess the impact of the imputation methods on statistical inference. We compute the **relative bias of standard errors** (2.1) and compare the results by non-response assumption, method and asset.

$$\hat{SE}(\hat{\theta}) = \sum_{j=1}^{1000} \left(\frac{\hat{SE}(\hat{\theta})_j - SE(\hat{\theta})}{SE(\hat{\theta})} \right) \tag{2.1}$$

$SE(\hat{\theta})$ is the empirical standard error of the mean calculated using the originally observed data, $\hat{SE}(\hat{\theta})_j$ is the standard error of the mean calculated using the j-th replication of imputed data. Hoogland and Boomsma (1998) suggest that the bias shall not exceed 5%.

2.4.1 Wave-specific evaluation criteria

Finding suitable evaluation criteria for multiple imputation is challenging. Most criteria applied by Watson and Starick (2011) are not applicable to the task at hand, as they would be heavily biased in favor of a replication of the observed value; for instance, an evaluation of the correlation between observed and imputed value does neglect the fact, that it is not the goal of multiple imputation to create a valid value for an individual missing item, but rather create a valid data set that takes the uncertainty of the imputation procedure into account. Hence, multiple imputation is best understood as simulating values for valid inference. In this study, we chose to evaluate trend, distributional and inequality accuracy jointly in a set of six evaluation criteria that take the overall data set into account instead of the replications of single values.

Chambers (2001) notes the imputation results should reproduce the lower order moments of the distribution of the true values. Given that we can directly compare the lower order moments between imputed and observed data sets, we chose to include the **absolute relative difference in means** (2.2) for the assessment of trend accuracy and the **absolute**

difference in the coefficient of variation (2.3) as an indicator of distributional and inequality accuracy. Generally, the dot symbol indicates imputed values, whereas symbols without dots indicate observed values.

$$CR(1) = \left| \frac{(\bar{y} - \bar{\dot{y}})}{\bar{\dot{y}}} \right| \tag{2.2}$$

$$CR(2) = \left| \frac{\sigma}{\bar{y}} - \frac{\sigma}{\bar{\dot{y}}} \right| \tag{2.3}$$

Additionally, distributional accuracy is achieved when the distributional properties of the original data set is replicated by the imputed data sets. The **Kolmogorov-Smirnov distance** (2.4) is the higher the more the two tested empirical distributions of the imputed and the true values deviate from each other. Thus, the smaller the Kolmogorov-Smirnov distance is, the more accurate the imputation method.

$$d_{KS} = \max_j \left(\left| \frac{1}{n} \sum_{i=1}^{n} I(y_i \leq x_j) - \frac{1}{n} \sum_{i=1}^{n} I(\dot{y}_i \leq x_j) \right) \right| \tag{2.4}$$

For the assessment of inequality we include three additional criteria. The **Gini coefficient** is especially sensitive against changes in the center of the distribution. The **mean log deviation** is sensitive for shifts at the bottom of the distribution. Those two criteria are complemented by an inequality measure for the top tail of the distribution, by using the **99/50 ratio of percentiles**.[9]

9 This indicator is not responsive to outliers–a relevant phenomenon in wealth analyses–compared to e.g. the half squared coefficient of variation (HSCV).

2.4.2 Longitudinal evaluation criteria

We apply two additional evaluation criteria that help to examine the effects of the imputation on wealth mobility. The first criterion assesses the distributional accuracy of **wealth mobility between waves** (2.5) for specific components and includes all observations with a positive value for the specific wealth type in two waves simultaneously. Here, wealth mobility is defined by the change in wealth decile group membership in 2002 vs. 2007, 2007 vs. 2012 and 2002 vs. 2012. A standard Chi-square test for fit of the distributions is performed, where the imputed cell frequencies are the observed ones and the expected cell frequencies are the true cell frequencies.

$$\chi^2 = \sum_{j=1}^{10} \sum_{i=1}^{10} \frac{(\dot{n}_{ij} - n_{ij})^2}{n_{ij}} \tag{2.5}$$

Thus, the higher the Chi-square test statistic (2.5) the worse the imputation method can replicate the observed mobility for the wealth component in consideration.

The second longitudinal criterion is the **cross-wave correlation** (2.6) for each wealth type separately: before and after the imputation procedure the differences of the correlations between each wealth type are compared and should be close to zero. The higher the deviation from zero the worse the performance of the imputation method.[10]

$$\begin{aligned} r_{y_1 y_2} - r_{\dot{y}_1 \dot{y}_2} = & \frac{\sum_{i=1}^{n}(y_{i1} - \bar{y}_1)(y_{i2} - \bar{y}_2)}{\sqrt{\sum_{i=1}^{n}(y_{i1} - \bar{y}_1)^2 \sum_{i=1}^{n}(y_{i2} - \bar{y}_2)^2}} \\ & - \frac{\sum_{i=1}^{n}(\dot{y}_{i1} - \bar{\dot{y}}_1)(\dot{y}_{i2} - \bar{\dot{y}}_2)}{\sqrt{\sum_{i=1}^{n}(\dot{y}_{i1} - \bar{\dot{y}}_1)^2 \sum_{i=1}^{n}(\dot{y}_{i2} - \bar{\dot{y}}_2)^2}} \end{aligned} \tag{2.6}$$

10 For comparison's sake we need to mention that we opt to not include four criteria applied by Watson and Starick (2011) that we find do not add another dimension to the evaluation at hand and, thus, are redundant. This includes the preservation of skewness and kurtosis, since the replication of the shape of the distribution is covered by the Kolmogorov-Smirnow distance (2.4). Furthermore, unlike Watson and Starick (2011) we do not include Pearson correlations between two wealth types. There is not enough covariation for this criterion to be applied for the asset types we choose for this study.

2.5 Imputation methods

The imputation methods which can be considered in our simulation study are limited by the fact that we are interested to use multiple imputation techniques. We have to rule out all single imputation techniques beforehand. This includes all carryover methods, which use valid values observed in the last or next wave of the survey (and variations thereof, which have been applied in the PSID for home equity). This also excludes, more generally, all imputation methods without a stochastic component. The methods we choose to examine are commonly used by other important wealth surveys (see Section 2.2).

We also refrain from considering (longitudinal) hotdeck imputation given that Watson and Starick (2011, p. 711) already present evidence in a simulation study that the hotdeck imputation method does 'not perform particularly well on either cross-sectional or longitudinal accuracy'.

2.5.1 Multiple imputation by chained equations (MICE)

MICE is an iterative and sequential regression approach that grew popular among researchers, because it demands very little technical preparation and is easy to use. We present the basic set-up for imputations using chained equations in this chapter, but for more detailed information we refer to van Buuren et al. (1999), Royston (2004), and van Buuren et al. (2006), among others. Multiple imputation by chained equations (MICE) is not an imputation model by itself, it is rather the expectation that by sequentially imputing the variables using separate univariate imputation models there will be convergence between the imputed variables after a certain number of iterations. For each prediction equation all but the variable for which missing values ought to be imputed are included, that is, each prediction equation exhibits a fully conditional specification. It is necessary for the chained equations to be set up as an iterative process, because the estimated parameters of the model are possibly dependent on the imputed values. Formally, we have p wealth components $Y_1, Y_2, \dots, Y_p$ and a set of predictors (without missing values) Z, then for iterations $n = 0, 1, \dots N$, and with ϕ_j as the corresponding model parameters with uniform prior probability distribution, the missing values are drawn from

$$
\begin{aligned}
Y_1^{(n+1)} &\sim g_1(Y_1|Y_2^{(n)},\ldots,Y_p^{(n)},Z,\phi_1)\\
Y_2^{(n+1)} &\sim g_2(Y_2|Y_1^{(n+1)},Y_3^{(n)},\ldots,Y_p^{(n)},Z,\phi_2)\\
&\vdots\\
Y_p^{(n+1)} &\sim g_p(Y_p|Y_1^{(n+1)},Y_2^{(n+1)},\ldots,Y_{p-1}^{(n+1)},Z,\phi_p)
\end{aligned}
\tag{2.7}
$$

until convergence at $n = N$ is achieved. That is, in iteration $(n + 1)$ the dependent variables of each imputation model $g_j(.)$ are updated with the corresponding imputed values of the last iteration n (or the ongoing iteration, if the dependent variable already has been imputed). The MICE imputation converges, once the distributions of Y_1,$Y_2,\ldots$,Y_p all have become stationary conditional on the observed data and the other imputed wealth variables. One of the main advantages is that the univariate imputation models $g_j(.)$ may be chosen separately for each imputation variable, which is also why even though MICE lacks any theoretical justification, it is widely used by researchers and practitioners. We did not make use of this specific feature at the project at hand, as all wealth variables exhibit similar statistical and distributional characteristics. However, we barely adjusted the set of additional independent variables Z_j for each imputation variable Y_j. The most important variables among Z_j are the lag and lead variables of the respective asset value, which are drawn from the other waves. Additionally, and in line with the experiences of other countries and surveys for the imputation of wealth data, the independent variables Z_j we choose are in line with the framework laid out in Barcelo (2006). We present a detailed overview and further explanations in Appendix 2.C.

We specified the imputation models $g_j(.)$ in (2.7) using predictive mean matching (PMM) to account for the restricted range of the imputation variables and to circumvent the assumption that the normality of the underlying models holds true. Predictive mean matching (PMM) was introduced by Little (1988) and is a nearest-neighbor matching technique used in imputation models to replace the outcome of the imputation model for every missing value (a linear prediction) with an observed value. The set of observed values, from which the imputed value is randomly drawn, consists of (non-missing) values derived from one randomly drawn out of the five nearest neighbors which are closest to the linear prediction.

2.5.2 Regression with Heckman correction for sample selection

For the first two waves of wealth information in the SOEP, the researchers opted for a regression design with Heckman correction for sample selection for the imputation of the missing asset values (Frick et al., 2007, 2010b). The first step involved a cross-sectional imputation of missing values for 2002. The data were then used for a longitudinal imputation of the 2007 data using the lagged wealth data from 2002 as covariates. The third step was a re-imputation of 2002 wealth data using the now-completed longitudinal information from 2007, and starting a cycle of regression models with longitudinal info until convergence between 2002 and 2007 was achieved. In total, Frick et al. (2010b) repeat this cycle five times; as this study aims to replicate their approach, we conduct the same number of iterations. The stochastic component in each step, which is necessary to generate multiple implicates, is added through the assignment of randomly drawn residuals derived from the respective regression models. With the 2012 wealth data and three available waves, the pool of available longitudinal information grows considerably. We add the regression models for 2012 after convergence between 2002 and 2007 has been achieved, with 2007 now serving as the base year. Consequently, longitudinal information from the survey wave 2007 is used for the imputation of missing values in 2002 and 2012 alike.

The variables included in those models are similar to the set of covariates used in the MICE approach (see Appendix 2.C). As in Frick et al. (2007, 2010b) we use ‘life satisfaction’ and a dummy for civil servants as selection instruments. However, generally in the Heckman regression the prediction equation does not include the metric values of the other wealth types, as they are not imputed yet.[11] All imputation models are specified separately. Additionally, in comparison to the MICE procedure, this regression model imputation does not include draws of the model parameters–the stochastic component is generated by draws from the residuals–, the uncertainty in the model estimation is not propagated in the imputations.

2.5.3 Row-and-column imputation technique

Little and Su (1989) proposed the row-and-column imputation technique (RC) as a procedure for item non-response adjustment in panel surveys. It takes advantage of available cross-sectional as well as individual longitudinal information. It combines data available from the entire panel duration for every unit (row) and cross-sectional trend information (column) and adds a residual derived from a nearest neighbor matching, thereby attaching a stochastic component to an otherwise deterministic approach.

Since we have three waves of wealth data, the column effects (for any wealth asset) are given by

$$c_t = \frac{(3 \cdot \bar{y}_t)}{\sum_k \bar{y}_k} \tag{2.8}$$

and are calculated for each wave separately. $\bar{y}_t$ is the sample mean wealth asset for $t = 2002, 2007, 2012$. The row effects are given by

$$r_i = \frac{1}{m_i} \cdot \sum_j \frac{y_{it}}{c_j} \tag{2.9}$$

and are calculated for each member of the sample. y_{it} it is the value of the wealth asset

11 There are a few exceptions: The regression model for home value (other property values) additionally includes the home debt (other property debt). The imputations for both these values are generated in an iterative process in itself, since both values have very high explanatory power in the respective models.

for individual i in wave t. m_i is the number of recorded waves in which the asset value of individual i has been observed.

Originally, the row-and-column-method was designed as a single imputation method. However, the last step–assigning the residual term from the nearest neighbor–may be modified in such a way that for every individual unit and wave multiple imputed values can be derived. After sorting the units by their row effects r_i, the residual effect of the nearest complete unit l in year j is used to calculate the imputed value for unit i:

$$\dot{y}_{it} = r_i \cdot c_t \cdot \overbrace{\frac{y_{lt}}{r_j \cdot c_t}}^{\text{residual term}} \quad . \tag{2.10}$$

$\dot{y}_{it}$ is the single imputed value using the residual effect from the nearest neighbor l. To generate multiple imputations we only need two additional steps. Instead of only assigning the residual of the nearest neighbor in (2.10), we assign the residuals of the k nearest neighbors. Then terms (2.8) and (2.9) are identical for every computation and the residual terms are used to generate k imputed values for every unit i and every year t. Since there is a tradeoff between the number of imputations and the distance to the 'farthest' nearest neighbor, we reasoned that the generally agreed on number of five imputations would present a reasonable balance (see e.g. the HFCS, other SOEP-variables, the Survey of Consumer Finances (SCF)). However, this decision is merely based on our expectations and has not been subject to an empirical analysis. It is also noteworthy, that the residual terms of the five nearest-neighbors have been randomly assigned to imputed values independently for every unit i in order to avoid any systematic differences of imputation accuracy in the five imputation data sets.

2.5.4 Row-and-column imputation with age classes

When using the row-and column imputation the donor of the residual term (and the distance between donor and recipient) in (2.10) is solely depending on the sorting of the units by their row effects r_i. Additionally, the trend component (2.8) is calculated using the complete sample. At the same time, as Watson and Starick (2011) state, recipients and the respective donors should have similar characteristics, and those characteristics should be associated with the variable being imputed. They introduce an addition to the basic row-and-column imputation; the method is extended to take into account basic characteristics of the donors and recipients. For a comparison between the standard row-and-column imputation and an imputation with age classes (RCA, see Table 2.3) we match donors and recipients within longitudinal imputation classes defined by the following age classes (at the time, the survey was conducted) in the respective wave: $17-19, 20-24, 25-34, 35-44, 45-54, 55-64, 65$ and older. Thereby it is guaranteed that donors share their residual with recipients from the same age range. The column term (2.8) is calculated using observations from the respective age classes.

A restriction of the row-and-column imputation is that it cannot be applied if no longitudinal information on the person level is available, thus we need a fallback method (e.g. the first wave of a respondent, or a specific wealth component is collected for the first time). As for the evaluation, we need a set-up that determines the superior combination of basic and fallback imputation methods simultaneously (see Table 2.3). The results of the evaluation should provide answers to several questions: (1) If a row-and-column imputation is used for observations that have valid information in other waves, does the addition of age classes improve the performance when compared to the standard row-and-column imputation? (2) Which combination of basic and fallback methods yields the best results? Basic imputation method means the technique that is used for observations with missing

Table 2.3: Basic and fallback imputation methods, and evaluation set-up.

acronym used in Section 5	**Basic** –for observations with missing values, information from other waves is available–	**Fallback** –for some observations with missing values, only cross-sectional information and variables are available–
MICE-RC	Standard row-and-column imputation	Multiple imputation by chained equations
REG-RC	Standard row-and-column imputation	Regression model with Heckmann correction
MICE-RCA	Row-and-column imputation plus age classes	Multiple imputation by chained equations
REG-RCA	Row-and-column imputation plus age classes	Regression model with Heckmann correction
MICE	Multiple imputation by chained equations	
REG	Regression model with Heckmann correction	

values and values from other waves of that same individual have been observed. Fallback imputation method means that for an observation with missing values only cross-sectional information and variables are available and, therefore, only either of the two model based approaches can be applied. Hence, in addition to the combinations using model based and row-and-column imputations, we test the performance of using a multiple imputation by chained equations as both basic and fallback method (MICE), and we proceed similarly with the regression with Heckman correction (REG).

2.6 Results

As we illustrated in Table 2.3, we compare the performance of the six combinations of prevalent imputation methods using the eight evaluation criteria we discussed in Section 2.4. As we wanted to compare the performance of the methods on a metric scale, we refrain from any ranking of the results. Second, we favor the property that the punishment for large deviations is larger than for smaller deviations, which should depend on the overall variance of the outcomes considering the individual evaluation criteria. That means, if the overall variance is small, outliers are punished harder, and deviations that are close to each other are punished similarly. Again, this is a property that is not fulfilled by any ranking of the results. It is, however, fulfilled, if we choose a distance measure that shows the distance between a well-defined optimum and the respective values calculated with imputed data. The optimum is simple to define, as all criteria are either calculated in a way that zero is representing no deviations from the original data or may be transformed to have this respective property. As for the distance measure, using the Euclidian distance would either require a normative decision on a weighting matrix or, alternatively, all criteria would contribute similarly (after normalizing). In order to avoid normative weighting we choose the Mahalanobis distance measure, as it additionally accounts for the observed covariance structure (Mahalanobis, 1936), and thereby is removing any redundancy in our evaluation criteria.

Our evaluation shows the distance between the ideal imputation (all values are zero for all criteria) and the deviation of the imputed values from this ideal point after using the respective imputation method (all tables in this section). Furthermore, this evaluation set-up allows us to compare the distances directly and interpret them on a metric scale, as the respective outcomes for the different methods are independent from each other (but depending on the overall variation and covariation of the evaluation criteria).

As already mentioned, we show the results for the three wealth items, the three years, and the three assumed non-response mechanisms separately and compare the outcomes for the imputation methods. The evaluation criteria (equations (2.2) – (2.4), plus inequality indices) are used for the trend, distributional and inequality evaluations. The longitudinal criteria (2.5) and (2.6) are additional criteria, which can solely be computed using the joint results of two waves (2002/07, 2007/12 and 2002/12) as reported in Section 2.6.2. In Section 2.6.3 we present the results for the relative bias of standard errors.

2.6.1 Evaluation of trend, distributional and inequality accuracy

If we would have solely considered the home market value in this study (Table 2.4), we would conclude that the combination of MICE and the RC imputation yield better results than a pure MICE imputation: Only taking into account the average distances for the trend evaluation reveals that in all cases the MICE imputation performs worse than the combinations with the RC imputation with and without age classes. Looking at the performance for all single waves, in all cases the addition of the RC technique as basic imputation improves the performance of MICE. Combining REG with the RC imputation on the other hand does not regularly improve the results. What is even more surprising, even though the combination of MICE and RC technique seems to perform best overall, the pure MICE approach rarely performs better than the pure REG approach. A possible explanation for these findings is that the home market values tend to be an asset type with a rather high state-dependency. The RC approach as univariate imputation technique, which solely considers future and past observed values and an overall trend effect, is closer to the trend and inequality estimates based on the observed data sets than both model-based approaches that may incorporate the uncertainty of the imputation procedure. Note that these outcomes are basically independent of the non-response mechanism that is assumed.

Generally, financial assets exhibit less state-dependency than home market values and regression models for both the imputation of the metric values and the non-response mechanism are mediocre compared to other asset types (Table 2.5). Thus, there is comparatively more uncertainty to consider by the imputation method, and the lag or lead variables have, in theory, considerably less explanatory power. However, if the mechanism of missingness is MAR, combining MICE with the RC method, again, yields the best results. If the missing mechanism is differential non-response at the bottom of the distribution, MICE-RCA seems to yield the best results as well. Only if differential non-response at the top is assumed, it is equally viable to choose between any RC method including age classes. Interestingly, including age classes does not improve the results for the RC technique, the differences between RC and RCA seem to be random.

Interestingly, for the evaluation criteria that are considered in this study and for financial assets, it seems to be more viable to choose a pure REG approach over a pure MICE approach if there is differential non-response at the top. Combining REG and RC improves the results under MAR. However, it is notable that all combinations of MICE with the RC method again regularly perform better than both pure model based approaches under any non-response assumption, but considerably less so under DNR2.

Consumer credits have the lowest state-dependency of the three wealth types we consider in this study. Note that the SOEP wealth data is collected in five-year intervals and credit periods for consumer credits are typically shorter. Following the same argumentation we already laid out for home market values und financial assets, we expect that the RC imputation performs rather weak. The results of the evaluation prove us mostly wrong. As shown in Table 2.6, both RC methods perform oftentimes better if MAR, DNR1 or DNR2 is assumed. Additionally, RCA has an advantage as compared to RC. One possible explanation is that even if the overall state-dependency is much lower for consumer credits,

Table 2.4: Performance of imputation methods: home market value.

	Wave 2002	Wave 2007	Wave 2012	Overall average distance
Missing at random (MAR)				
REG	4.93	5.64	5.46	5.34
REG-RC	5.23	5.86	5.82	5.64
REG-RCA	5.32	5.81	5.93	5.69
MICE	6.05	7.02	6.76	6.61
MICE-RC	**4.12**	4.94	**4.73**	**4.60**
MICE-RCA	4.16	**4.91**	**4.73**	**4.60**
Differential non-response 1 (DNRI)				
REG	5.79	6.32	5.77	5.96
REG-RC	6.50	6.25	6.46	6.40
REG-RCA	6.47	6.49	6.65	6.54
MICE	6.98	7.24	6.91	7.04
MICE-RC	5.61	**5.52**	**5.57**	**5.57**
MICE-RCA	**5.53**	5.72	5.71	5.65
Differential non-response 2 (DNRII)				
REG	6.45	5.91	6.06	6.14
REG-RC	6.34	4.94	5.38	5.55
REG-RCA	5.76	4.45	5.08	5.10
MICE	5.96	5.91	5.80	5.89
MICE-RC	5.59	4.42	4.68	4.90
MICE-RCA	**5.02**	**3.96**	**4.42**	**4.47**

Note: Bold figures indicate the smallest average distance among the six imputation variants.

the state-dependency at the bottom of the distribution may still be considerably high and the RC imputation might still yield more accurate imputed data sets in this case. Incorporating age classes seems to improve the results, because consumer credits are more prevalent among younger age classes, who are paying off their debts as they get older.

Comparing the distributions of the distances to the optimal imputations separately for the MAR assumption for all three waves and all assets jointly, confirms the conclusions we draw above (Figure 2.1).

Including the RC imputation does improve the performance of MICE considerably and significantly. The distance between the optimal imputation and MICE versus both MICE-RC and MICE-RCA is roughly 1.3 units higher, the respective means and standard deviations are shown in Figure 2.1 together with the boxplots of the distributions.

Table 2.5: Performance of imputation methods: financial assets.

	Wave 2002	Wave 2007	Wave 2012	Overall average distance
Missing at random (MAR)				
REG	5.82	6.37	5.46	5.88
REG-RC	**5.41**	5.78	5.19	5.46
REG-RCA	5.43	5.86	5.15	5.48
MICE	6.60	5.81	5.18	5.86
MICE-RC	5.49	**4.81**	4.89	**5.06**
MICE-RCA	5.55	4.91	**4.85**	5.10
Differential non-response 1 (DNRI)				
REG	6.28	6.89	6.07	6.41
REG-RC	5.80	6.84	6.09	6.24
REG-RCA	**5.68**	6.73	6.11	6.17
MICE	6.82	6.53	6.03	6.46
MICE-RC	6.17	6.18	**5.69**	6.01
MICE-RCA	6.12	**6.09**	5.70	**5.97**
Differential non-response 2 (DNRII)				
REG	**7.09**	6.51	6.59	6.73
REG-RC	7.38	6.26	**6.31**	**6.65**
REG-RCA	7.49	**6.24**	6.37	6.70
MICE	8.38	7.72	7.54	7.88
MICE-RC	7.22	6.49	6.44	6.72
MICE-RCA	7.35	6.44	6.40	6.73

Note: Bold figures indicate the smallest average distance among the six imputation variants.

Considering the performance of REG versus REG-RC and REG-RCA the differences are miniscule. Moreover, results for REG exhibit considerably more variance over the 1000 simulation data sets. Similar figures for DNR1 and DNR2 are presented in Appendix 2.B.

Additionally, we observe that the incorporation of age classes in the RC imputation does not improve the overall imputation results. Watson and Starick (2011) report an advantage for the performance of the RC imputation with age classes for the imputation of income items. One possible explanation, why we do not identify a similar advantage, is that there are less regular trends of increase and spend-down of asset values over the life cycle for home market value und financial assets as compared to income variables.

Table 2.6: Performance of imputation methods: consumer credits.

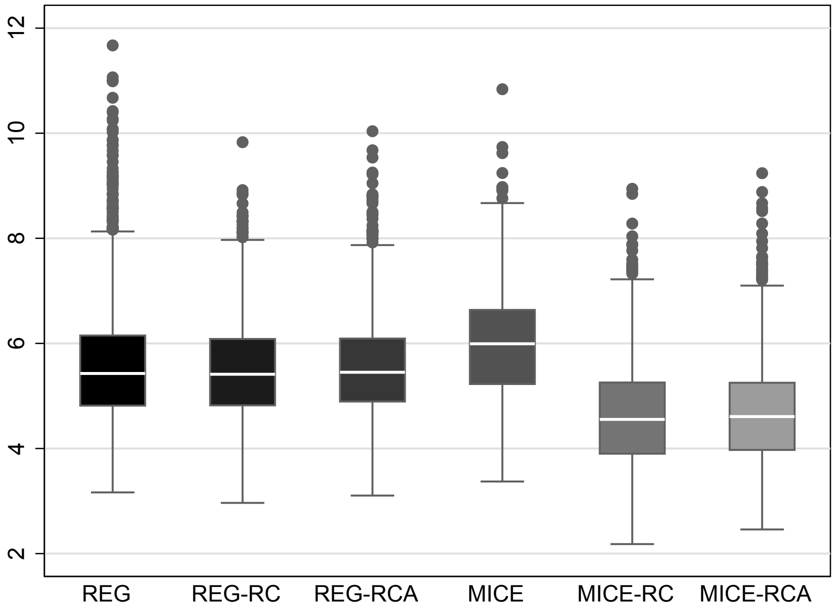
	Wave 2002	Wave 2007	Wave 2012	Overall average distance
Missing at random (MAR)				
REG	4.25	4.51	3.83	4.20
REG-RC	4.79	4.62	2.59	4.00
REG-RCA	**4.09**	4.31	**2.25**	**3.55**
MICE	5.35	4.65	4.63	4.88
MICE-RC	4.44	3.70	4.10	4.08
MICE-RCA	4.34	**3.48**	4.24	4.02
Differential non-response 1 (DNRI)				
REG	4.97	4.36	4.48	4.60
REG-RC	5.52	**3.90**	**3.44**	4.29
REG-RCA	4.39	3.95	3.84	**4.06**
MICE	5.30	5.26	4.97	5.18
MICE-RC	4.55	4.50	4.38	4.48
MICE-RCA	**4.22**	4.38	4.51	4.37
Differential non-response 2 (DNRII)				
REG	4.96	**4.56**	5.77	5.10
REG-RC	4.77	5.16	4.51	4.81
REG-RCA	4.85	4.86	**4.39**	4.70
MICE	5.07	4.85	4.63	4.85
MICE-RC	5.09	4.89	4.80	4.93
MICE-RCA	**4.41**	4.71	4.74	**4.62**

Note: Bold figures indicate the smallest average distance among the six imputation variants.

2.6.2 Evaluation of wealth mobility

As for the two additional longitudinal criteria, which focus on the changes in the observed mobility structures before and after imputations, the overall average distances include all pair-wise comparisons (2002/2007, 2007/2012, and 2002/2012) and are presented in Table 2.7. We expected that using RC imputations would overestimate the state-dependency for the wealth assets and undermine the actually observed mobility structures. This expectation gets confirmed to a certain extent.

Under MAR the pure MICE approach seems to perform better than the pure REG approach and all combinations with the RC method (at least for home market value and financial assets). This is to be expected, as the mobility seems to be severely reduced, once



	REG	REG-RC	REG-RCA	MICE	MICE-RC	MICE-RCA
Mean	5.60	5.49	5.54	5.97	4.61	4.66
Std. dev	0.47	0.31	0.31	0.34	0.34	0.33

Figure 2.1: Boxplots for the distances to optimal imputations by imputation methods under Missing at Random (MAR).

the only included variable is the lag or lead variable of the respective variable that is to be imputed. What is more surprising is that the REG approach performs considerably worse than MICE. One possible statistical explanation could be that the regression set-up is not taking into account one source of uncertainty, which the MICE procedure does take into account: the drawing of the respective model parameters. The only stochastic component in the REG approach is the drawing of a residual from the observed residuals, whereas MICE imputes values after drawing of the respective model parameters. Here, REG might underestimate the uncertainty of the imputation procedure and produce too less variation in the imputed values, thereby as well reducing mobility.

For differential non-response and especially consumer credits the results are less clear. MICE seems to reproduce the observed mobility structures slightly better than REG, in many cases the combination of MICE and the RC imputations yield satisfying results too, but generally distances to an optimal imputation seem to increase. We conclude that (1) a researcher interested in mobility structures would probably prefer the model based MICE

Table 2.7: Average performance on longitudinal evaluation criteria, all assets.

	Home market value	Financial assets	Consumer credits	Overall average distance
Missing at random (MAR)				
REG	1.51	2.11	**2.47**	2.03
RC-REG	2.53	2.25	2.56	2.45
RCA-REG	2.56	2.10	2.52	2.39
MICE	**0.71**	**0.94**	3.05	**1.57**
RC-MICE	1.79	1.37	3.00	2.05
RCA-MICE	1.77	1.33	2.94	2.01
Differential non-response 1 (DNRI)				
REG	1.77	2.77	**3.08**	2.54
RC-REG	2.77	2.52	3.19	2.83
RCA-REG	2.81	2.50	3.15	2.82
MICE	**1.14**	2.57	3.15	**2.29**
RC-MICE	2.30	**2.36**	3.13	2.60
RCA-MICE	2.35	2.37	3.12	2.61
Differential non-response 2 (DNRII)				
REG	**1.27**	2.35	**3.26**	2.29
RC-REG	2.23	2.31	3.53	2.69
RCA-REG	2.21	2.31	3.48	2.67
MICE	1.46	**0.77**	3.29	**1.84**
RC-MICE	1.72	1.40	3.62	2.25
RCA-MICE	1.63	1.42	3.60	2.22

Note: Bold figures indicate the smallest average distance among the six imputation variants.

approach to an univariate imputation procedure such as the RC method, and (2) even though REG yields imputed values using model prediction equations as well, the REG imputation performs worse than the MICE approach.

2.6.3 Evaluation of standard errors

Interestingly, inference seems to be affected differently for assets, and less for non-response assumption or imputation method (Table 2.8). Overall, the relative bias of standard errors is smallest for the imputation of home market values, it is slightly higher for financial assets under MAR and DNR1, and it is the highest for any of the imputation of consumer credits. The negative impact on standard errors by RC or RCA is not alarming in any of the cases analyzed here. Hoogland and Boomsma (1998) suggest that the relative bias of standard errors should not exceed 5 percent; here, only standard errors of the imputed values of consumer credits are showing worrisome results, but they do not indicate that a specific method yields considerably worse results. As for the intuition, why consumer credits are impacted the most, apparently once small liability values are missing (DNR1), imputation data sets tend to overstate the standard errors, and vice versa (DNR2 and MAR, see Table 2.2). This appears to be the result of generally poor imputation models, as the set of covariates with high explanatory power is smaller than for other assets in the SOEP as well as considerably less observations to rely on. Our experience with SOEP data shows that it is substantially more challenging to impute for missing liability values, which is reflected by the results of this study.

Table 2.8: Relative bias of standard errors. Note: Bold figures indicate that the relative bias exceeds 5 percent.

	Home market value	Financial assets	Consumer credits	Overall bias
Missing at random (MAR)				
REG	-1.80	**5.88**	**-9.94**	-1.95
RC-REG	-0.85	0.14	**-10.98**	-3.90
RCA-REG	-1.00	0.27	**-10.98**	-3.90
MICE	3.04	**5.49**	**-6.10**	0.81
RC-MICE	1.22	1.52	**-7.60**	-1.62
RCA-MICE	1.13	1.67	**-7.73**	-1.64
Differential non-response 1 (DNRI)				
REG	-1.33	**9.00**	**19.01**	**8.89**
RC-REG	-1.17	3.66	**11.19**	4.56
RCA-REG	-1.32	3.77	**11.21**	4.55
MICE	1.72	2.16	4.28	2.72
RC-MICE	-0.16	-2.18	1.51	-0.28
RCA-MICE	-0.31	-2.20	1.50	-0.34
Differential non-response 2 (DNRII)				
REG	-0.74	-0.34	**-7.38**	-2.82
RC-REG	-0.22	-0.03	**-9.20**	-3.15
RCA-REG	-0.08	-0.01	**-9.04**	-3.04
MICE	1.24	0.59	**-7.56**	-1.91
RC-MICE	0.88	0.19	**-8.92**	-2.62
RCA-MICE	1.05	0.22	**-8.71**	-2.48

Note: Bold figures indicate the smallest average distance among the six imputation variants.

2.7 Conclusion

In an assessment of the performance of several imputation methods for longitudinal wealth data we use a set of eight evaluation criteria and three assumptions for the non-response mechanism. The overall result does not yield that a single imputation method performs consistently better for all wealth types in a cross-sectional and longitudinal analysis. We compare the row-and-column imputation (with or without age classes) for observations with available longitudinal data with two methods that rely on the prediction equations of regression models. In our analyses of the performance of the imputation methods we identified several effects the researcher has to consider for studies using multiple imputation and imputed data.

As for the trend and inequality evaluation, if the missing data are truly missing at random (MAR), for all three assets we consider the combination of MICE and row-and-column imputation is at least among the best performing methods. Unexpectedly, this holds true independently of the level of state-dependency prevalent in the items. If the missing data are missing not at random and instead are the result of differential non-response (DNR1 and DNR2) the combination of the row-and-column imputation with MICE does improve the performance in our evaluation study as well. This is the core outcome of this study: If the missing at random assumption is violated, the row-and-column imputation technique yields less biased overall imputation results for trend and inequality estimates. We like to stress that–based on this study and our experience with data imputation–this conclusion holds only true for variables that are highly skewed (such as assets, net worth or income variables). The imputation technique itself–and thus an improvement of the performance–is applicable to panel data only.

Furthermore, we find that adding age classes to the standard row-and-column imputation as introduced by Little and Su (1989) does not regularly improve the performance based on our criteria and the input data. However, there is an advantage for the imputation of consumer credits.

As for the wealth mobility criteria, the conclusions are less clear. Generally, MICE seems to reproduce the observed mobility structures better than the regression approach, in many cases the combinations of MICE and the row-and-column imputations yield satisfying results, too. However, it is clearly noticeable that for most assets and non-response assumptions the mobility is reduced, once the row-and-column imputation is applied. Hence, a data provider needs to weigh the options: for the SOEP we decided that the method of choice depends on data usage; as the data are mainly used for trend and inequality analyses and much less for mobility analyses, we opt for the combination of MICE and the row-and-column imputation.

One thing that remains to be addressed is that we refrained from including partial unit non-response (PUNR) in this simulation, e.g. individuals within households that choose not to respond, whereas the rest of the household did. The reason is that analyses with the SOEP wealth data focus on the individual level observation and PUNR observations would only affect household wealth estimators. However, we do not expect the results to be significantly different, had we considered PUNR observations. Potential extensions to this study could be the inclusion of additional wealth types, examining the effects of imputation methods on the total net worth and the aggregate net worth, and additional imputation methods we did not consider for now.

2.A Predictive mean matching versus standard regression design

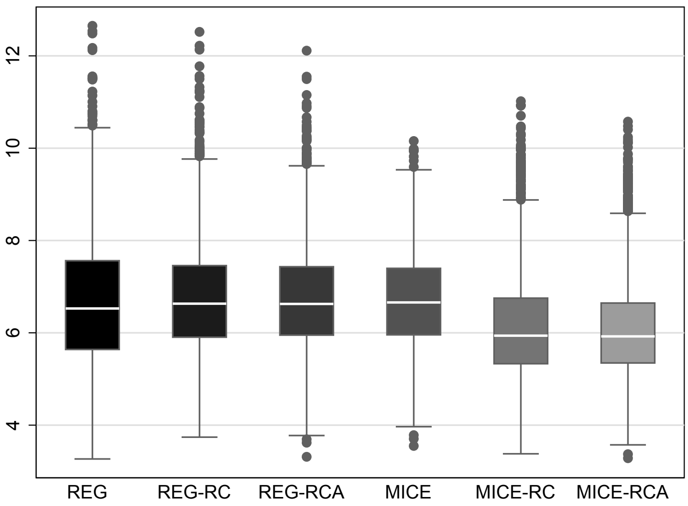
As described in Section 2.5, for the basic MICE and MICE with row-and-column imputation we chose to impute values using predictive mean matching. However, it might be the case that predictive mean matching performs worse than a standard regression design, if the missing at random assumption is violated, as potential donors with observed values similar to the missing ones might be rare in the upper tail of the wealth distribution (differential non-response at the top). Therefore, for this robustness check we repeat the multiple imputation using MICE, MICE-RC and MICE-RCA assuming DNR2 and choosing a standard regression instead of predictive mean matching. The results have been computed identically to Section 2.6. Overall, for all assets the results are very similar to the results using predictive mean matching. We conclude that in our simulation set-up and under differential non-response at the top, the results for MICE do not improve, if a standard regression imputation is used.

Table 2.9: Performance of imputation methods under DNR2, for MICE with standard regression design instead of predictive mean matching.

	Wave 2002	Wave 2007	Wave 2012	Overall average distance
Variable: Home Market Value				
REG	5.79	5.92	5.71	5.81
REG-RC	5.40	4.88	5.00	5.09
REG-RCA	5.06	4.43	4.83	4.77
MICE	5.78	5.56	5.54	5.63
MICE-RC	5.00	4.55	4.53	4.69
MICE-RCA	**4.61**	**4.08**	**4.36**	**4.35**
Variable: Financial Assets				
REG	11.04	6.16	6.73	7.98
REG-RC	10.50	6.05	6.68	7.74
REG-RCA	10.59	5.87	6.72	7.73
MICE	10.90	6.24	6.74	7.96
MICE-RC	**10.47**	4.89	5.67	7.01
MICE-RCA	10.63	**4.68**	**5.59**	**6.97**
Variable: Consumer Credits				
REG	7.28	6.93	6.48	6.90
REG-RC	7.68	6.67	**5.96**	**6.77**
REG-RCA	7.80	**6.63**	6.05	6.83
MICE	8.31	7.55	7.03	7.63
MICE-RC	**7.61**	7.00	6.29	6.97
MICE-RCA	7.76	6.94	6.30	7.00

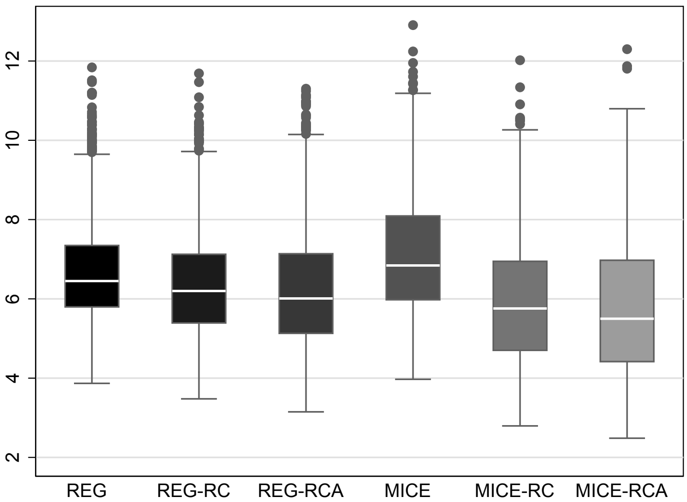
Note: Bold figures indicate the smallest average distance among the six imputation variants.

2.B Boxplots for the distances to optimal imputations



	REG	REG-RC	REG-RCA	MICE	MICE-RC	MICE-RCA
Mean	6.75	6.78	6.76	6.69	6.15	6.11
Std. dev	0.63	0.44	0.42	0.36	0.41	0.39

Figure 2.2: Boxplots for the distances to optimal imputations by imputation methods under Differential Non-Response 1 (DNRI).



	REG	REG-RC	REG-RCA	MICE	MICE-RC	MICE-RCA
Mean	6.67	6.36	6.25	7.12	5.89	5.77
Std. dev	0.46	0.46	0.52	0.50	0.51	0.57

Figure 2.3: Boxplots for the distances to optimal imputations by imputation methods under Differential Non-Response 1 (DNRII).

2.C List of covariates

Below is a list of all covariates used in the imputations using regressions with Heckman correction for sample selection and MICE. For the regressions the choice of variables is based on Frick et al. (2007, 2010), for MICE all of the variables below have been included as it is one large imputation that handles missing values for all assets simultaneously. The variables include (1) a set of covariates determining the non-response (variables of the non-response model under the MAR assumption mentioned in Section 2.4), (2) covariates that are considered good predictors for the variable we want to impute (3) economic variables that are possibly related to the outcome variable (according to economic theory) and (4) variables that are good predictors of the covariates included in the rest the groups of variables. However, the last group is especially important in the first iterations and the more association between the imputation variables is expected. Generally, we hereby follow the guidelines laid out in Barceló (2006) for the independent variables in the prediction equations. We additionally include dummies indicating non-response in other assets and other waves (for the respective asset).

As in the regressions using Heckman correction, in MICE we include lagged and/or lead variables of the assets we impute. Theoretically, for MICE we could build one giant model covering all assets in all waves. While all would be imputed in one step, we chose to code a sequence of MICE procedures, which impute all assets (in one chain) and draw the respective lagged and/or lead variables from the results of the other waves, as it is easier to check the properties of models in between imputations. To set this up, we first cross-sectionally impute all of the 2007 variables in one step drawing lag and lead variables from the 2002 and 2012 variables (unless they are missing as well). This completes the set of lead variables for the imputation of the 2002 variables. After this, we run the 2007 imputations again and may use the partially imputed 2002 variables as covariates. The procedure for the 2012 assets is carried out in a similar manner, drawing from the already imputed 2002 and 2007 variables.

Variable name	Description	Home market value	Financial assets	Consumer credits
How dwelling was acquired	3 dummies: acquired by purchase, inheritance or newly built?	X		
Age Age of the respondent	Missing values were estimated on the basis of the age of other household members and the relationship to the head of household	X[2]	X[2]	X[2]
Age of house	7 dummies: Year of construction: before 1918, 1918-1948, 1949-1971, 1972-1980, 1981-1990, 1991-2000, 2001 and later	X		
Savings account	Dummy: Household holds savings account (yes=1)			X
Capia02	4 dummies for the categorical CAPI-information on the market value of owner-occupied property (<150.000, <200.000, <400.000, >400.000)	X		
Capic02	4 dummies for the categorical CAPI-information on the value of financial assets (<5.000, <20.000, <100.000, >100.000)		X	
Capig02	4 dummies for the categorical CAPI-information on the value of consumer credits (<5.000, <10.000, <50.000, >50.000)			X[2]
Children	Dummy: children younger than sixteen in the household (yes=1)		X	X
Civil servant	Dummy: civil servant (yes=1)	X[1]	X[1]	X[1]
Condition of house	2 dummies. Dwelling is in a good condition (yes=1); Dwelling needs major refurbishment (yes=1)	X		
Credit	Dummy: household raised a consumer credit (yes=1)			X
Credit value	Monthly amount of loan repayment			X
Debts owner-occupied property	Debts related to owner-occupied property (edited/imputed; the first of the imputed versions is taken)	X[2]		
Dishwasher	Dummy: Dishwasher in the household (yes=1)	X		
District type	10 dummies on categorical information of the district's size	X		
Dividend	Dividend income in the household, metric information are logarithmized, categorical information from are recoded into 6 dummies, one for each category (<250, <1.000, <2.500, <5.000, <10.000, >10.000)		X	
Dwelling satisfaction	Satisfaction with the dwelling. For those without a valid info. the mean of all other household members was used or (if no household member gave a valid information to this satisfaction question) a random number between 0 and 10 was taken	X		
Education	Years of education. Those who are still in education are assigned the minimum of seven years.	X	X	
Enterprise	Dummy: household owns a commercial enterprise (yes=1)	X		X

continued on next page

Variable name	Description	Home market value	Financial assets	Consumer credits
Equipment	2 dummies: household with garden / balcony	X		
Estimated rent	Estimation of monthly rent by owners if they had to rent their dwelling	X[2]		
Financial worries	Dummy: At least some concerns about finances (yes=1)		X	
Household income	Annual post-government household income in eruos	X[2]	X[2]	
Inheritance	Dummy: Household received inheritance/other windfall profits in the previous year (yes=1)		X	
Dissatisfaction with life	Dummy: Individual is unhappy with his/her life (life satisfaction<6)	X[1]	X[1]	X[1]
Satisfaction with life	Dummy: Individual is happy with his/her life (life satisfaction>=9)	X[1]	X[1]	X[1]
Missing	Dummies for all those variables where missing values exist: missing or valid information	X	X	X
Monthly savings	Dummy: Household has monthly savings (yes=1)			X
New car	Dummy: Purchase of a new car in the last 12 months (yes=1)			X
No debt owner-occupied property	Dummy: Debts for owner-occupied property (no debt=1)	X	X	
No partner	Dummy: no partner within the household (yes=1)			X
No paym. to others	Dummy: no payments/support to persons outside the household (yes=1)		X	
Occupancy Year	moved into dwelling			X[2]
Owner	Dummy: Does the person have own property (yes=1)		X	X
Partner's filter	Dummy: Does the partner possess the respective wealth component (yes=1)		X	X
Partner's value	Partner's value of the respective wealth component		X[2]	X[2]
Paym. dwelling (metric)	Monthly loan payments for owner-occupied property in €		X[2]	
Old-age provisions	5 dummies: Interest in building-up private old-age provision (very strong/strong/medium/less/not at all)		X	
Public sector	Dummy: Individual works in the public sector (yes=1)			X
Region	97 dummies: Raumordnungsregion	X		
Rent income	Dummy: Household receives income from renting leasing (yes =1)	X		
Rent income (metric)	Household income from renting leasing in the previous year in €			X[2]
Rent level	6 dummies identifying regional level of rent	X		
Residential area	3 dummies on type of residential area: predominantly old houses / predominantly new houses / other	X		
Self-employed	Dummy: individual is self-employed (yes=1)			X

continued on next page

Variable name	Description	Home market value	Financial assets	Consumer credits
Sex	Dummy: female = 1		X	
Size housing unit	Size of the housing unit in square meters. For missing values, the mean of those with the same number of rooms resp. the same number of household members (if the information on the amount of rooms was also missing) was imputed	X^2		X^2
Size of household	3 dummies for size of household (one person / two or three / 4+ persons)		X	X
Fixed interest securities	Dummy: Household owns stocks (yes=1)		X	
Type of house	8 dummies: Type of house (farm house, one- or two-family house, one- or two-family row house, 3-4 unit building, 5-8 unit building, 9- or more unit building, other)	X		
Value owner-occupied property	Market value of owner-occupied property (edited/imputed; the first of the imputed versions is taken)			X^2
West	Dummy: West Germany (yes=1)		X	

Table 2.10: List of all covariates used in the imputations using regressions with Heckman correction for sample selection and MICE.

2.D Results for individual evaluation criteria

Assumption: Missing at Random

Financial Assets						
Wave 2002	Mean	Coeff. of Var.	KS-distance	Gini coeff.	MLD	p99/p50
REG	-0.048	-0.234	0.100	-0.0269	-0.0839	-7.33
REG-RC	-0.028	-0.142	0.095	-0.0303	-0.1048	-6.75
REC-RCA	-0.043	-0.237	0.093	-0.0314	-0.1120	-6.08
MICE	-0.064	-0.274	0.105	-0.0274	-0.1351	-4.29
MICE-RC	-0.035	-0.140	0.082	-0.0320	-0.1400	-5.76
MICE-RCA	-0.050	-0.237	0.081	-0.0330	-0.1470	-5.31
Wave 2007						
REG	0.021	-0.383	0.144	-0.0583	-0.2924	-23.86
REG-RC	0.146	0.012	0.141	-0.0247	-0.1540	-9.83
REC-RCA	0.132	0.059	0.140	-0.0230	-0.1265	-9.54
MICE	-0.103	-0.188	0.081	-0.0153	-0.0652	-5.35
MICE-RC	0.035	-0.018	0.083	-0.0132	-0.0806	-6.02
MICE-RCA	0.020	0.026	0.081	-0.0113	-0.0510	-4.89
Wave 2012						
REG	-0.002	-0.781	0.131	-0.0632	-0.3297	-22.14
REG-RC	0.127	-0.055	0.128	-0.0270	-0.1718	-8.68
REC-RCA	0.131	-0.054	0.130	-0.0254	-0.1568	-8.62
MICE	-0.083	0.251	0.078	-0.0108	-0.0506	-3.90
MICE-RC	0.014	0.275	0.063	-0.0073	-0.0596	-1.71
MICE-RCA	0.018	0.276	0.065	-0.0060	-0.0452	-1.50
Home Market Value						
Wave 2002	Mean	Coeff. of Var.	KS-distance	Gini coeff.	MLD	p99/p50
REG	0.018	0.044	0.088	0.0089	0.0227	0.02
REG-RC	0.036	0.007	0.101	0.0032	0.0104	0.13
REC-RCA	0.035	0.014	0.103	0.0051	0.0115	0.14
MICE	0.009	-0.062	0.110	-0.0044	-0.0042	0.03
MICE-RC	0.027	-0.055	0.071	-0.0079	-0.0094	0.15
MICE-RCA	0.027	-0.048	0.073	-0.0060	-0.0084	0.14
Wave 2007						
REG	0.016	0.014	0.084	0.0038	0.0163	-0.23
REG-RC	0.021	0.016	0.082	0.0021	0.0047	0.07
REC-RCA	0.022	0.015	0.083	0.0027	0.0066	0.04
MICE	0.001	-0.061	0.094	-0.0067	-0.0094	-0.21
MICE-RC	0.010	-0.017	0.058	-0.0065	-0.0133	0.12
MICE-RCA	0.010	-0.018	0.060	-0.0059	-0.0113	0.08
Wave 2012						
REG	0.007	-0.022	0.073	0.0041	0.0135	-0.13
REG-RC	0.021	0.005	0.074	0.0095	0.0150	0.32
REC-RCA	0.019	-0.005	0.075	0.0097	0.0159	0.18
MICE	-0.015	-0.032	0.080	-0.0055	-0.0087	0.08
MICE-RC	0.003	-0.026	0.051	-0.0021	-0.0074	0.23
MICE-RCA	0.000	-0.034	0.051	-0.0020	-0.0065	0.11
Consumer Credits						
Wave 2002	Mean	Coeff. of Var.	KS-distance	Gini coeff.	MLD	p99/p50
REG	-0.076	0.094	0.110	-0.0012	-0.0137	-1.53
REG-RC	-0.051	0.277	0.100	-0.0015	-0.0368	0.23

continued on next page

Consumer Credits						
Wave 2002	Mean	Coeff. of Var.	KS-distance	Gini coeff.	MLD	p99/p50
REC-RCA	-0.044	0.283	0.110	-0.0021	-0.0417	0.48
MICE	-0.190	-0.338	0.130	-0.0559	-0.2546	-6.95
MICE-RC	-0.129	-0.178	0.110	-0.0440	-0.2166	-4.46
MICE-RCA	-0.122	-0.179	0.110	-0.0447	-0.2222	-4.19
Wave 2007						
REG	-0.711	-1.043	0.111	-0.1121	-0.3591	-24.40
REG-RC	-0.408	-0.831	0.096	-0.0965	-0.3079	-17.34
REC-RCA	-0.396	-0.854	0.096	-0.0985	-0.3149	-18.17
MICE	-0.463	-0.631	0.104	-0.0873	-0.2578	-16.05
MICE-RC	-0.266	-0.575	0.076	-0.0776	-0.2434	-12.47
MICE-RCA	-0.254	-0.598	0.077	-0.0794	-0.2497	-13.11
Wave 2012						
REG	-0.177	-0.462	0.095	-0.0281	-0.0566	-6.50
REG-RC	-0.142	-0.686	0.094	-0.0529	-0.2179	-10.54
REC-RCA	-0.141	-0.730	0.094	-0.0525	-0.2037	-11.75
MICE	0.021	-0.410	0.088	0.0065	0.0237	1.46
MICE-RC	0.016	-0.657	0.087	-0.0255	-0.1404	-2.59
MICE-RCA	0.016	-0.704	0.087	-0.0249	-0.1258	-3.62

Longitudinal criteria

Financial Assets		**Home Market Value**		**Consumer Credits**	
Chi-square test stat.	Cross-wave correlation	Chi-square test stat.	Cross-wave correlation	Chi-square test stat.	Cross-wave correlation
336.57	0.139	902.44	-0.121	222.00	0.020
370.21	-0.227	901.46	-0.281	172.60	-0.176
333.63	-0.235	918.06	-0.284	168.77	-0.175
102.52	0.117	341.59	-0.018	277.46	0.028
143.00	-0.256	570.30	-0.192	191.79	-0.248
126.25	-0.263	501.67	-0.194	186.67	-0.246

Table 2.11: Mean results all evaluation criteria, assumption: missing at random (MAR).

Assumption: Differential non-response I

Financial Assets						
Wave 2002	Mean	Coeff. of Var.	KS-distance	Gini coeff.	MLD	p99/p50
REG	0.216	0.160	0.131	0.0112	0.0852	-9.45
REG-RC	0.211	-0.270	0.131	-0.0168	-0.0110	-9.21
REC-RCA	0.207	-0.340	0.130	-0.0148	-0.0075	-8.21
MICE	0.222	-0.273	0.119	-0.0085	-0.0166	-3.84
MICE-RC	0.220	-0.316	0.118	-0.0216	-0.0587	-6.08
MICE-RCA	0.216	-0.378	0.116	-0.0194	-0.0545	-5.44
Wave 2007						
REG	-0.416	-0.316	0.136	-0.0224	-0.1161	-55.74
REG-RC	0.113	-0.648	0.132	-0.0377	-0.1347	-15.61
REC-RCA	0.108	-0.609	0.129	-0.0345	-0.0993	-14.61
MICE	0.180	-0.254	0.105	-0.0122	0.0199	-2.39
MICE-RC	0.216	-0.130	0.117	-0.0219	-0.0644	-6.56
MICE-RCA	0.211	-0.089	0.114	-0.0185	-0.0290	-5.29

continued on next page

Financial Assets						
Wave 2012	Mean	Coeff. of Var.	KS-distance	Gini coeff.	MLD	p99/p50
REG	-0.486	-1.216	0.109	-0.0496	-0.2343	-51.85
REG-RC	0.039	-0.765	0.108	-0.0345	-0.1272	-15.31
REC-RCA	0.041	-0.755	0.108	-0.0337	-0.1150	-14.93
MICE	0.202	-0.158	0.102	-0.0027	0.0672	-0.67
MICE-RC	0.203	-0.071	0.095	-0.0085	-0.0052	-1.19
MICE-RCA	0.204	-0.066	0.095	-0.0079	0.0066	-1.03
Home Market Value						
Wave 2002	Mean	Coeff. of Var.	KS-distance	Gini coeff.	MLD	p99/p50
REG	0.100	0.154	0.111	0.0289	0.0380	0.10
REG-RC	0.094	0.089	0.125	0.0160	0.0182	-0.22
REC-RCA	0.092	0.088	0.125	0.0182	0.0205	-0.19
MICE	0.089	0.039	0.139	0.0190	0.0206	-0.05
MICE-RC	0.080	0.036	0.100	0.0067	0.0053	-0.32
MICE-RCA	0.078	0.035	0.100	0.0088	0.0076	-0.31
Wave 2007						
REG	0.109	0.164	0.114	0.0317	0.0469	0.85
REG-RC	0.103	0.133	0.113	0.0283	0.0345	0.77
REC-RCA	0.106	0.142	0.119	0.0282	0.0364	0.74
MICE	0.105	0.107	0.133	0.0303	0.0359	0.76
MICE-RC	0.095	0.115	0.095	0.0245	0.0274	0.77
MICE-RCA	0.098	0.123	0.101	0.0244	0.0293	0.74
Wave 2012						
REG	0.105	0.097	0.104	0.0305	0.0424	0.71
REG-RC	0.106	0.061	0.114	0.0273	0.0323	0.77
REC-RCA	0.107	0.045	0.119	0.0268	0.0331	0.71
MICE	0.099	0.083	0.124	0.0281	0.0325	0.64
MICE-RC	0.093	0.043	0.091	0.0206	0.0211	0.64
MICE-RCA	0.094	0.028	0.096	0.0200	0.0219	0.58
Consumer Credits						
Wave 2002	Mean	Coeff. of Var.	KS-distance	Gini coeff.	MLD	p99/p50
REG	0.313	0.449	0.131	0.0422	0.2105	0.32
REG-RC	0.306	0.510	0.128	0.0310	0.1536	0.47
REC-RCA	0.312	0.514	0.127	0.0318	0.1498	0.08
MICE	0.374	-0.065	0.169	0.0154	0.0326	-1.19
MICE-RC	0.356	0.057	0.154	0.0089	0.0186	-1.18
MICE-RCA	0.361	0.056	0.154	0.0091	0.0144	-1.17
Wave 2007						
REG	-0.122	-2.203	0.145	-0.0766	-0.2927	-41.76
REG-RC	0.107	-1.504	0.157	-0.0522	-0.1940	-23.13
REC-RCA	0.105	-1.492	0.154	-0.0512	-0.1874	-22.46
MICE	0.361	-0.200	0.155	0.0087	0.0962	-0.37
MICE-RC	0.368	-0.136	0.149	0.0046	0.0517	-0.87
MICE-RCA	0.366	-0.122	0.147	0.0060	0.0580	-0.47
Wave 2012						
REG	0.118	0.034	0.136	-0.0295	-0.0354	-15.98
REG-RC	0.186	0.169	0.144	-0.0224	-0.0728	-11.58
REC-RCA	0.186	0.160	0.145	-0.0219	-0.0597	-12.14
MICE	0.222	0.075	0.134	0.0120	0.0783	-3.44
MICE-RC	0.240	0.168	0.132	0.0040	-0.0026	-3.73
MICE-RCA	0.240	0.150	0.133	0.0046	0.0103	-3.96

continued on next page

Longitudinal criteria

Financial Assets		**Home Market Value**		**Consumer Credits**	
Chi-square test stat.	Cross-wave correlation	Chi-square test stat.	Cross-wave correlation	Chi-square test stat.	Cross-wave correlation
418.87	0.213	1139.16	-0.176	197.20	-0.054
447.85	-0.216	1152.72	-0.360	194.76	-0.183
443.89	-0.216	1161.64	-0.365	192.60	-0.177
398.90	0.181	785.63	-0.066	205.39	-0.023
426.31	-0.287	961.66	-0.291	172.28	-0.343
426.77	-0.286	1028.53	-0.295	172.46	-0.337

Table 2.12: Mean results all evaluation criteria, assumption: differential non-response at the bottom (DNRI).

Assumption: Differential non-response II

Financial Assets						
Wave 2002	Mean	Coeff. of Var.	KS-distance	Gini coeff.	MLD	p99/p50
REG	-0.462	0.887	0.319	0.0467	0.0748	3.86
REG-RC	-0.515	0.475	0.248	-0.0121	-0.0524	0.05
REC-RCA	-0.546	0.407	0.253	-0.0117	-0.0567	0.13
MICE	-0.403	0.123	0.232	-0.0570	-0.2091	-3.71
MICE-RC	-0.413	0.160	0.135	-0.0622	-0.2028	-3.79
MICE-RCA	-0.443	0.102	0.140	-0.0619	-0.2081	-3.91
Wave 2007						
REG	-0.690	1.240	0.297	0.0850	0.3024	6.04
REG-RC	-0.647	0.801	0.243	0.0442	0.1427	2.78
REC-RCA	-0.657	0.795	0.241	0.0436	0.1672	2.71
MICE	-0.544	0.340	0.189	-0.0077	-0.0068	-5.20
MICE-RC	-0.471	0.453	0.133	-0.0051	-0.0339	-4.88
MICE-RCA	-0.481	0.447	0.131	-0.0055	-0.0102	-5.18
Wave 2012						
REG	-0.603	1.117	0.275	0.0825	0.3098	7.37
REG-RC	-0.661	0.544	0.230	0.0355	0.1420	3.15
REC-RCA	-0.662	0.595	0.232	0.0363	0.1540	3.05
MICE	-0.594	0.441	0.217	0.0026	0.0285	-0.63
MICE-RC	-0.560	0.228	0.144	-0.0082	-0.0335	-2.27
MICE-RCA	-0.560	0.290	0.144	-0.0073	-0.0212	-2.87
Home Market Value						
Wave 2002	Mean	Coeff. of Var.	KS-distance	Gini coeff.	MLD	p99/p50
REG	-0.052	0.107	0.108	0.0245	0.0388	0.26
REG-RC	-0.050	0.047	0.091	0.0077	0.0120	0.14
REC-RCA	-0.047	0.054	0.087	0.0116	0.0160	0.22
MICE	-0.073	-0.015	0.087	0.0013	0.0066	0.01
MICE-RC	-0.061	-0.011	0.069	-0.0072	-0.0069	-0.04
MICE-RCA	-0.058	-0.006	0.065	-0.0035	-0.0029	0.02
Wave 2007						
REG	-0.064	0.117	0.110	0.0280	0.0464	0.45
REG-RC	-0.046	0.062	0.081	0.0131	0.0185	0.38
REC-RCA	-0.042	0.071	0.076	0.0159	0.0234	0.36
MICE	-0.078	0.007	0.096	0.0033	0.0087	0.08

continued on next page

Home Market Value						
Wave 2002	Mean	Coeff. of Var.	KS-distance	Gini coeff.	MLD	p99/p50
MICE-RC	-0.049	0.025	0.065	0.0020	0.0032	0.21
MICE-RCA	-0.045	0.034	0.060	0.0047	0.0081	0.17
Wave 2012						
REG	-0.046	0.085	0.095	0.0252	0.0391	0.45
REG-RC	-0.039	0.024	0.078	0.0163	0.0207	0.45
REC-RCA	-0.042	0.003	0.076	0.0168	0.0226	0.40
MICE	-0.068	0.013	0.082	0.0031	0.0071	0.13
MICE-RC	-0.047	-0.013	0.061	0.0023	0.0012	0.19
MICE-RCA	-0.050	-0.034	0.059	0.0027	0.0028	0.15
Consumer Credits						
Wave 2002	Mean	Coeff. of Var.	KS-distance	Gini coeff.	MLD	p99/p50
REG	-0.418	0.006	0.199	-0.0470	-0.0980	-0.44
REG-RC	-0.540	-0.083	0.203	-0.0776	-0.1766	-3.07
REC-RCA	-0.538	-0.050	0.204	-0.0789	-0.1846	-3.09
MICE	-0.612	-0.320	0.201	-0.1340	-0.3744	-8.57
MICE-RC	-0.625	-0.295	0.172	-0.1287	-0.3512	-7.85
MICE-RCA	-0.622	-0.267	0.174	-0.1301	-0.3596	-7.69
Wave 2007						
REG	-0.735	-0.090	0.177	-0.0166	0.0136	-3.38
REG-RC	-0.629	0.099	0.171	-0.0083	0.0037	-2.23
REC-RCA	-0.625	0.112	0.170	-0.0089	0.0063	-2.22
MICE	-0.925	-0.133	0.213	-0.0319	-0.0588	-6.11
MICE-RC	-0.759	0.014	0.177	-0.0197	-0.0520	-3.76
MICE-RCA	-0.755	0.025	0.175	-0.0204	-0.0499	-3.63
Wave 2012						
REG	-0.562	0.040	0.156	-0.0155	0.0335	-4.73
REG-RC	-0.485	0.132	0.141	-0.0169	-0.0291	-4.01
REC-RCA	-0.487	0.118	0.145	-0.0146	-0.0118	-3.91
MICE	-0.681	0.027	0.186	-0.0092	0.0101	-3.31
MICE-RC	-0.581	0.053	0.149	-0.0146	-0.0559	-2.97
MICE-RCA	-0.583	0.051	0.151	-0.0120	-0.0379	-2.92

Longitudinal criteria

Financial Assets		**Home Market Value**		**Consumer Credits**	
Chi-square test stat.	Cross-wave correlation	Chi-square test stat.	Cross-wave correlation	Chi-square test stat.	Cross-wave correlation
2009.71	0.063	283.6 9	-0.108	304.44	0.022
1436.30	-0.328	221.85	-0.294	295.77	-0.259
1425.01	-0.334	203.01	-0.296	290.35	-0.266
469.17	0.024	489.29	0.008	311.75	0.089
272.81	-0.330	252.18	-0.206	307.64	-0.243
273.08	-0.336	207.75	-0.206	305.22	-0.251

Table 2.13: Mean results all evaluation criteria, assumption: differential non-response at the top (DNRII).

2.E Results for relative bias of standard errors

Table 2.14: Relative bias of standard errors: home market value.

	Wave 2002	Wave 2007	Wave 2012	Overall bias
Missing at random (MAR)				
REG	-0.17	-4.75	-0.48	-1.80
REG-RC	0.30	-3.09	0.24	-0.85
REG-RCA	0.03	-3.02	0.00	-1.00
MICE	3.97	2.11	3.03	3.04
MICE-RC	2.47	-0.41	1.59	1.22
MICE-RCA	2.20	-0.19	1.39	1.13
Differential non-response 1 (DNRI)				
REG	0.55	-4.49	-0.06	-1.33
REG-RC	0.33	-3.97	0.14	-1.17
REG-RCA	0.34	-4.35	0.05	-1.32
MICE	3.45	-0.65	2.36	1.72
MICE-RC	1.43	-2.70	0.78	-0.16
MICE-RCA	1.47	-3.11	0.70	-0.31
Differential non-response 2 (DNRII)				
REG	0.95	-3.13	-0.02	-0.74
REG-RC	0.70	-2.18	0.82	-0.22
REG-RCA	0.81	-2.07	1.02	-0.08
MICE	2.61	-1.02	2.13	1.24
MICE-RC	1.92	-0.94	1.65	0.88
MICE-RCA	1.98	-0.78	1.95	1.05

Note: Bold figures indicate that the relative bias exceeds 5 percent.

Table 2.15: Relative bias of standard errors: financial assets.

	Wave 2002	Wave 2007	Wave 2012	Overall bias
Missing at random (MAR)				
REG	3.13	3.19	**11.32**	**5.88**
REG-RC	1.74	-2.49	1.17	0.14
REG-RCA	2.37	-2.58	1.01	0.27
MICE	**5.07**	**6.26**	**5.14**	**5.49**
MICE-RC	2.65	0.65	1.25	1.52
MICE-RCA	3.27	0.65	1.09	1.67
Differential non-response 1 (DNRI)				
REG	-3.97	**7.07**	**23.91**	**9.00**
REG-RC	-1.08	1.62	**10.43**	3.66
REG-RCA	-0.65	1.49	**10.47**	3.77
MICE	1.72	2.02	2.73	2.16
MICE-RC	0.07	**-5.27**	-1.33	-2.18
MICE-RCA	0.09	**-5.41**	-1.29	-2.20
Differential non-response 2 (DNRII)				
REG	-0.44	-0.55	-0.04	-0.34
REG-RC	-0.05	-0.43	0.40	-0.03
REG-RCA	0.15	-0.45	0.26	-0.01
MICE	0.47	0.44	0.87	0.59
MICE-RC	0.19	-0.27	0.64	0.19
MICE-RCA	0.40	-0.28	0.53	0.22

Note: Bold figures indicate that the relative bias exceeds 5 percent.

Table 2.16: Relative bias of standard errors: consumer credits.

	Wave 2002	Wave 2007	Wave 2012	Overall bias
Missing at random (MAR)				
REG	-4.96	**-24.64**	-0.23	**-9.94**
REG-RC	**-6.39**	**-26.15**	-0.41	**-10.98**
REG-RCA	**-6.39**	**-26.15**	-0.41	**-10.98**
MICE	0.07	**-19.10**	0.73	**-6.10**
MICE-RC	-1.80	**-21.16**	0.15	**-7.60**
MICE-RCA	-1.84	**-21.48**	0.11	**-7.73**
Differential non-response 1 (DNRI)				
REG	**16.52**	**35.72**	4.79	**19.01**
REG-RC	**10.21**	**20.94**	2.41	**11.19**
REG-RCA	**10.15**	**20.98**	2.50	**11.21**
MICE	**9.29**	-2.02	**5.57**	4.28
MICE-RC	**6.04**	**-5.04**	3.53	1.51
MICE-RCA	**6.00**	**-5.00**	3.51	1.50
Differential non-response 2 (DNRII)				
REG	1.35	**-19.96**	-3.53	**-7.38**
REG-RC	0.99	**-24.52**	-4.06	**-9.20**
REG-RCA	1.22	**-24.29**	-4.04	**-9.04**
MICE	3.38	**-24.67**	-1.40	**-7.56**
MICE-RC	2.07	**-26.82**	-2.02	**-8.92**
MICE-RCA	2.40	**-26.48**	-2.05	**-8.71**

Note: Bold figures indicate that the relative bias exceeds 5 percent.

3 Estimating top wealth shares using survey data – An empiricist's guide

3.1 Introduction

In the wake of a renewed discussion on inequality, distributive justice, and social cohesion, the distributions of income and wealth are, again, in the focus of science, media, and policy makers (Piketty, 2014). While research on income and its distribution in Europe and across the world is widely available, research on wealth is comparatively scarce. One reason might be that until recently the availability of comparable data for Europe was severely limited, which changed after the Eurosystem's Household Finance and Consumption Survey (HFCS) became accessible to researchers (European Central Bank, 2013a,b). Although the latest report by the OECD (2015) thoroughly analyzes income inequality, there is only one chapter dedicated to research on wealth inequality. It addresses two major problems: finding comparable data sources, and the fact that information on the long-term developments is even harder to come by.

One issue is that researchers need to rely on survey data, if tax return data is not available.[12] However, survey data typically has the problem of a middle-class bias, it lacks a sufficient number of observations for the margins of the distribution. Due to the pronouncedly skewed distribution of net worth, the upper tail of the wealth distribution is of utmost relevance when analyzing wealth inequality. Some wealth surveys try to overcome

12 Even if wealth tax data is available, the information does not typically cover the whole population, as only taxable wealth components are recorded.

this problem by oversampling rich households.[13] However, even with an oversampling of affluent households, there is the tendency that the truly rich households–in particular multi-millionaires and billionaires–are still not adequately represented in such surveys (Westermeier and Grabka, 2015). In order to overcome the under coverage of high-net-worth-individuals and -households in wealth survey data, researchers started simulating the top tail of wealth distributions using Pareto distributions both with and without information on high-net-worth-individuals from rich lists.

The aim of this study is to shed light on some aspects of enhancing lacking survey data using Pareto simulated top tails that are previously neglected. Using Monte Carlo experiments, we show that wealth data, which is plagued by differential non-response, as opposed to a non-observation bias, might not be treated with a simple maximum likelihood estimation of the top tail, as estimates are still inherently biased downwards. Including rich list data and switching to regression estimation impacts top wealth shares, but the total net worth is still biased downwards. In the last step of the simulation, I show what potential effects are to be expected, if publishers of rich lists data systematically overestimate the top fortunes. Overall, all empirically encountered estimations of the aggregate wealth and top wealth shares using corrected data yield inherently biased results, once survey weights are uninformed and no additional data is available for calibration. As shown in an application using German survey data, if survey weights are re-calibrated to carry information on the distribution of households from exogenous sources, the estimates change tremendously. U.S. and Spanish survey data, which include sampling via wealth strata, are the best guesses as to how response behavior and wealth may be related.

In one stream of the existing literature, the bulk of studies explore the consistency of 'rich list' data from magazines such as Forbes with the power law distribution (Klass et al., 2007; Brzezinski, 2014). Generally, these studies fall under the label econophysics (see

13 'Relatively wealthy households account for a disproportionate share of the total wealth, and existing evidence suggests that the likelihood that they will not complete interviews when included in a sample is disproportionately high. Thus, there are potentially both bias and variance implications stemming from the treatment of wealthy households. Standard designs used when measuring income or expenditure might not be adequate for measuring wealth.' (OECD, 2015, p. 147).

Chatterjee et al., 2005) and are concerned with questions of exact statistics and alternative models describing the distribution of wealth (Clauset et al., 2009). Some of the most recent works concentrate on the question of whether rich-list data, such as the yearly-published list of billionaires by the American Forbes magazine, can be better described using other distributional assumptions (Brzezinski, 2014; Capehart, 2014). While the questions studied among the researchers in the econophysics camp are valid questions to study, the more relevant questions for public finance and policy makers are, (1) do power law distributions approximate the reality well enough; and (2) can we draw conclusions for the estimates of wealth distribution and top wealth shares? The statistical properties of rich-list data alone are of limited use, once a researcher needs to impute for missing observations at the upper tail of the distribution *between* rich-list data and survey data.

The second stream of literature is decidedly more empirically oriented and studies whether Pareto distributions are a useful complement to survey data. For instance, Vermeulen (2014) shows in a Monte Carlo experiment that the inclusion of rich lists' entries (such as the Forbes magazine) increases the precision of estimators for both the Pareto index and, as a result, the key figures of the entire wealth distribution, if compared to survey estimates without top-net-worth-holders from rich lists. Using data from the HFCS, he presents results for adjusted wealth distributions based on arbitrarily chosen minimum values for the Pareto distribution. Bach et al. (2014a) carry out a similar exercise using survey wealth data and rich list data from Germany. Eckerstorfer et al. (2015) rely only on survey data and present a method for the identification of a Pareto distribution's minimum value using statistical hypothesis testing and Austrian data from the HFCS. They assume that, due to the skewness of wealth distributions, there is a non-observation bias at the top in survey data, as very rich households are randomly missing from the sample. Vermeulen (2014), on the other hand, based his simulation on the assumption that the under coverage at the top does not solely happen by chance, based on several reports on response rates in the US Survey of Consumer Finances (SCF), instead concluding that the response rates decrease due to differential non-response (as reported by Kennickell and

Woodburn, 1997; Kennickell and McManus, 1993). The term encapsulates the observation that the non-response rate is increasing, the higher the net worth value of a household is, i.e. the richer a household, the lower the probability that it is included in a survey sample. Survey data alone then yields severely downward biased results, even more so without a dedicated oversample for very rich households, as included in the SCF (Kennickell, 2007, 2009) and some countries that are part of the HFCS sample (European Central Bank, 2013a). The 2011 Spanish subsample of the HFCS includes such an oversample for very wealthy households; it is based on individual wealth tax file information from the 2007 wealth tax.[14] All people with taxable wealth over €108,000 in Spain were subject to this tax. The wealth strata were chosen based on the percentile distribution of households filing a wealth tax return. The resulting cooperation rates show that non-participation is much less likely for the middle class (wealth below €500,000) than for the upper class (wealth greater than €6 million). The former strata had cooperation rates exceeding 50 percent, the latter below 26.5 percent (Bover et al., 2014, p. 27). Findings from the 2002, 2005, and 2008 Encuesta Financiera de las Familias conducted in Spain confirm the progressively increasing non-response rates (Bover, 2004, 2008, 2011). As for the mechanisms that might cause differential non-response, they remain largely unexplored, but it is straightforward to think of a series of unknown household characteristics that might correlate with both net worth and response probability (e.g. availability of the household's head and time use, the value of opportunity costs), none of which are observed by survey providers or used in the post-stratification of the survey weights, thus yielding weights that do not reflect the distribution of households in the sampling population.

In Section 3.2, a series of Monte Carlo experiments is conducted, adjusting several of the assumptions and testing various methods used by researchers to correct for the under coverage of high-net-worth-individuals and -households in survey studies. In Section 3.3 the findings are applied to to German survey wealth data, showing how re-calibrating the survey weights might affect top wealth shares, and compares the results. Section 3.4

14 The wealth tax in Spain was discontinued afterwards, but re-established in 2011.

concludes.

3.2 A simulation study

The correction for the missing rich in survey data, using a Pareto distribution as an approximation for the upper tail, involves several steps. First, the parameters of the Pareto distribution must be identified. In Section 3.2.1, non-observation bias is illustrated using a similar simulation set-up to Eckerstorfer et al. (2015). It is shown that both sample estimates and Pareto-corrected estimates become more precise as more observations are sampled from the tail of a wealth distribution (Specification 1). Next, the assumption of non-observation bias as the motivating factor is changed to differential non-response and it is shown that non-informative survey weights, ceteris paribus, result in downward biased estimates, even though a Pareto-correction is applied (Specification 2).

3.2.1 Non-observation bias versus differential non-response

It is assumed that the top net worth population of a fictive country consists of 600,000 households with a net worth greater than 1 million; they are distributed following a Pareto distribution

$$f_p(w) = \begin{cases} 0 & w < w_m \\ \frac{\alpha w_m^\alpha}{w^{\alpha+1}} & w \geq w_m \end{cases}, \tag{3.1}$$

w_m is the threshold parameter of a pareto distribution, also called minimum value. All data exceeding this threshold follow a Pareto distribution. The parameter α is known as the Pareto index or the scaling parameter, which determines the shape of the distribution–the lower α the higher the inequality of the wealth distribution in the upper tail exceeding the threshold parameter w_m.

In the first Monte Carlo experiment, α is set to 1.3 and w_m equals 1 million (cf.

Eckerstorfer et al., 2015). To estimate the Pareto index from the data, the maximum likelihood estimator is used, as it is the preferable estimator compared to the regression estimator (Clauset et al., 2009). The units of the sample are denoted by $i = 1, \ldots, n$, hence, w_i equals the net worth of household i. Then, the maximum likelihood estimator for Pareto index α is given by

$$a_{ml} = 1 + n \left[\sum_{i=1}^{n} \ln \frac{w_i}{w_m} \right]^{-1}. \tag{3.2}$$

Samples with varying sample sizes are drawn from this population: the number of Pareto distributed households n varies between 100 and 2,000 observations in steps of 100.[15] As each step involves 1,000 samples, in total 20,000 samples are drawn and the respective Pareto index is calculated using equation (3.2).

In contrast to Eckerstorfer et al. (2015), the x-axis does not show the sample size as a percent of the population, it is shown in absolute numbers. The precision of the estimation depends on the absolute sample size rather than the relative sample size, i.e. the overall size of the population is only relevant for the extrapolation of the aggregate wealth. As the sampled total net worth in the top panel of Figure 3.1 depicts, the median aggregate wealth is likely biased downwards before correction. However, estimates of the Pareto index are unbiased (against the median) and the precision expectedly gets higher as the sample size increases. On the lower right panel in Figure 3.1 the total net worth is recalculated, based on the extrapolation of the Pareto estimates. As it is known how many units exceed w_m from the overall population, it is straightforward to calculate the resulting total net worth

15 Based on my own calculations, the 2012 sample of the German Socio-economic Panel Study (SOEP) has 270 households that have a net worth exceeding €1 million, in the German subsample of the HFCS data there are 246 households with a net worth greater than €1 million (means over 5 implicates, see Table 3.2). In the Austrian subsample only 113 households exceed 1 million, while in Belgium, Finland, and Italy the number varies between 200 and 300 households. In Spain and France the number is well above 1000 households, which in turn seems to greatly affect variance between separate implicates of the multiply imputed data (see Appendix 3.B).

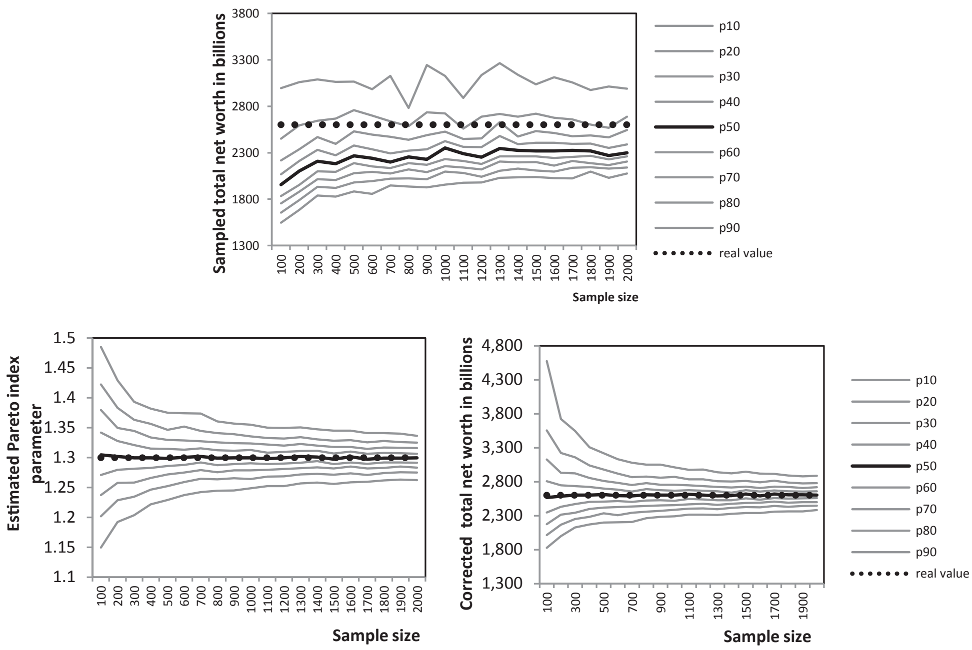


Figure 3.1: Specification 1. Deciles of the estimated a_{ml} and total net worth for sample sizes between 100 and 2,000 of a Pareto distributed population with an actual $\alpha = 1.3$, $w_m = 1{,}000{,}000$ and population size $N = 600{,}000$.

via the expected mean:

$$E(W) = \begin{cases} \infty & \alpha \leq 1 \\ \frac{\alpha w_m}{\alpha - 1} & \alpha > 1 \end{cases}. \tag{3.3}$$

The corrected estimates for the total net worth are unbiased, the precision increases with the sample size. In this case, any downward bias would be the result of a non-observation bias. Sample selection randomly excludes very rich households and the resulting estimated totals are too low.

However, surveys in Spain and the US show that the non-response rate is increasing with the level of wealth (Kennickell and Woodburn, 1997; Bover et al., 2014, p. 27). Hence, differential non-response might be a more viable explanation for the lack of statistical power at the top of wealth distributions and biased top wealth shares in survey samples. Eckerstorfer et al. concentrate their study on a statistically sound method to determine the correct minimum value w_m and apply the method to HFCS data from Austria, assuming only a non-observation bias. They note that their method could also be applied if the data suffer from differential non-response. As shown next, differential non-response yields biased estimates of Pareto index α and, thus, biased totals, even if both the threshold parameter w_m and the population size exceeding w_m are known.

In Specification 2, the same Monte Carlo experiment as in Specification 1 is repeated, assuming that the probability to refuse to participate in the survey increases with the level of wealth. For this simulation, any assumption of the mechanism would suffice, as long as it progressively increases the probability of non-response with the level of net worth. Vermeulen (2014) calculates non-response probabilities for the 1992 sample of the U.S. Survey of Consumer Finances, as reported in Kennickell and Woodburn (1997): the mechanism is (approximately) described by $\Pr(\text{non-response}) = 0.1 + 0.04\ln(w)$. This means a person with a net worth value of 1 million would refuse with a probability 65.3% and a person with a value of 10 million with a probability of 74.47%.[16] However, it is assumed that the survey provider increases the gross sample size by a factor of 3 in order to ensure that the net sample size stays roughly the same as in the first simulation. Implementing the same mechanism in our data and drawing random gross samples of, again, increasingly large gross sample sizes between 300 and 6,000 units yields net samples of roughly the size as in Specification 1.

In Figure 3.2 we estimate Pareto index α, again using the maximum likelihood estimator a_{ml}. However, the true mechanism of non-response is unknown to the researcher; hence, the empiricist assumes a random sample. This assumption is empirically warranted if no external data for calibration is available. Additionally, it is mathematically identical to an estimation of the Pareto index α from survey data with complex sampling without using

16 Low response rates are not unique to surveys in the US. The response rate for the first wave of the German HFCS subsample was as low as 18.7%; overall the rate was below 50% in about half of the countries included in the HFCS (European Central Bank, 2013a, p. 41).

weights as in Eckerstorfer et al. (2015).[17]

The total net worth calculated from the sample is severely biased downward (top panel in Figure 3.2). Additionally, estimating Pareto index α using the survey data with an unknown non-response mechanism results in an overestimation of the Pareto index (bottom left panel in Figure 3.2). In the Monte Carlo experiment the estimates are about 0.14 units too high. Hence, the first result is that if the survey data are plagued by increasing non-response rates, then a_{ml} cannot be consistently estimated. The effect of an overestimation of α leads to a corrected total net worth that is consistently too low (Figure 3.2, bottom right panel), in this case, roughly 550 Billion or 20% less than the real

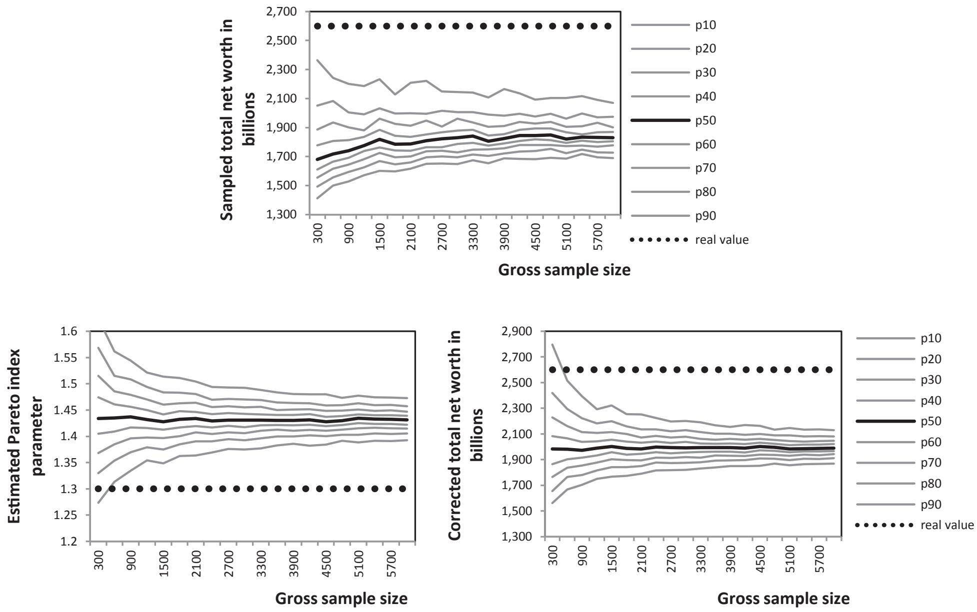


Figure 3.2: Specification 2. The impact of differential non-response on the estimation of α. Deciles of the estimated a_{ml} and total net worth for gross sample sizes between 300 and 6,000 of a Pareto distributed population with an actual $\alpha = 1.3$, $w_m = 1{,}000{,}000$ and population size $N = 600{,}000$.

17 In practice, the mechanism of non-response is never completely unknown to survey providers and unit non-response is addressed when calculating the survey weights, which combine sampling probabilities with additional paradata, that are possibly correlated with the wealth level of households. Thus, including survey weights in any estimation of Pareto index α is advisable, especially if an oversample of rich households is included in the sample. This was the case in Germany, Austria and several other countries included in the HFCS (European Central Bank, 2013a). Estimation without survey weights will surely yield biased results stemming from the complex survey sampling.

value. Note that it is still assumed that both the value of the threshold parameter w_m and the number of households exceeding it are known, both of which are unknown in practice.

Survey statisticians differentiate between *sampling error* and *sampling bias*. The former case describes a situation where a survey sample's calculated value differs from a population's aggregate private wealth purely by chance, as the value from a sample is never identical to the actual value. In this case, due to the skewness of the distribution, some super rich individuals or household are excluded from the sample. Their inclusion would drive up the mean and, subsequently, the aggregate private wealth. Their absence from the sample causes a *sampling error*.

A *sampling bias*, however, is to be expected if some members of the intended population are less likely included in the sample than others: in this case, differential non-response with respect to wealth. For the sake of simplicity, let us assume that private wealth W is independently identically Pareto distributed following equation (3.1). A sampling error is less of a problem, because any sampled value is still drawn randomly from the underlying distribution, thus the expected value equals the population's mean (is unbiased) as estimated by formula (3.3).

If the survey sample suffers by differential non-response, the probability to be included in the sample depends on w. In the example in Specification 2 it was $\Pr(i \in s|w_i) = 0.9 - 0.04\ln(w_i)$; more generally, one could describe such a response function as $\Pr(i \in s|w_i) = a - b\ln(w_i)$. In this case, the probabilities to be sampled depends on the value w itself and, thus, differs at different points of the wealth distribution. With a fixed at 0.9, and w_m supposed to be known, the empirically observed sample distribution exhibits much less inequality depending on b. As it is difficult to show analytically, one typically resorts to Monte Carlo experiments. However, for Specification 2 it is possible to show that the sample distribution differs from $f_p(w)$.

For some probability density functions (pdf) $f_p(y_i|\theta)$, depending on parameters θ, it is possible to derive the sample pdf of Y_i, defined as $f_s(y_i|i \in s)$, where S denotes the selected sample. It is obtained by application of the Bayes theorem (Peffermann et al., 1998):

$$f_s(y_i|\theta^*) = f_s(y_i|i \in s) = \Pr(i \in s|y_i)f_p(y_i|\theta)/\Pr(i \in s). \tag{3.4}$$

The parameters θ^* are a function of θ and the parameters indexing $\Pr(i \in s|y_i)$. It is important to note that, as $\Pr(i \in s|y_i) \neq \Pr(i \in s)$ for all y_i, the sample and population probability density functions are different and survey weights derived from the sampling become informative (if available). In most cases empiricists resort to Monte Carlo methods to show the impact of various non-response mechanisms on the population pdf. This paper's assumptions on the differential non-response affecting a Pareto distributed wealth tail generally yield sample pdfs that are not Pareto distributed any more, which becomes visible when examining the resulting shapes. One of the properties of a Pareto distribution is that the conditional probability distribution of a Pareto distributed random variable, given that it is greater or equals a particular value w_1 exceeding the threshold value w_m, is again a Pareto distribution with unchanged Pareto index α but minimum value w_1 instead of w_m. This property does not hold, if the sampling suffers by differential non-response, indicating that the resulting sample pdf $f_s(w_i|\alpha, w_m, \theta^*)$, with θ^* a function of parameters indexing the biased sampling procedure $\Pr(i \in s|w_i)$, is not Pareto distributed (see also counterexamples in Appendix 3.A and 3.C). With population distribution $f_p(w_i|\alpha, w_m) = \alpha w_m^{\alpha} * w_i^{-\alpha-1}$ and the probability to be sampled conditional on household wealth given as $\Pr(i \in s|w_i) = a - b\ln(w_i)$ the sample distribution can be written as

$$f_s(w|\alpha,w_m, a, b) = (a - b\ln(w))\alpha w_m^{\alpha} w^{-1-\alpha} / \int_{w_m}^{\infty} (a - b\ln(w))\alpha w_m^{\alpha} w^{-1-\alpha} \mathrm{d}w. \tag{3.5}$$

As in Specification 2 the wealth distribution was Pareto with an index of 1.3 and a threshold value of 1000000

$$W_i \sim Pareto(\alpha = 1.3, w_m = 1000000),$$

the sample distribution reduces to

$$f_s(w|a,b) = \frac{8.20245 * 10^7 * (a - b\ln(w))}{w^{2.3} * (a - 14.5847 * b)}. \tag{3.6}$$

For comparison's sake one might now plot both population distribution and the sample distribution from Specification 2 (with $a = 0.9$ and $b = 0.04$) to visualize that the sample distribution clearly exhibits less inequality and smaller mean (Figure 3.3).

To prove that the sample pdf necessarily has a smaller mean than the population pdf it is possible to compute the expected means. For a Pareto distributed random variable the expected value is given by (3.3). Thus, in Specification 2 the expected value of the population pdf equals $4\frac{1}{3}$ million. For the sample population the expected value is given

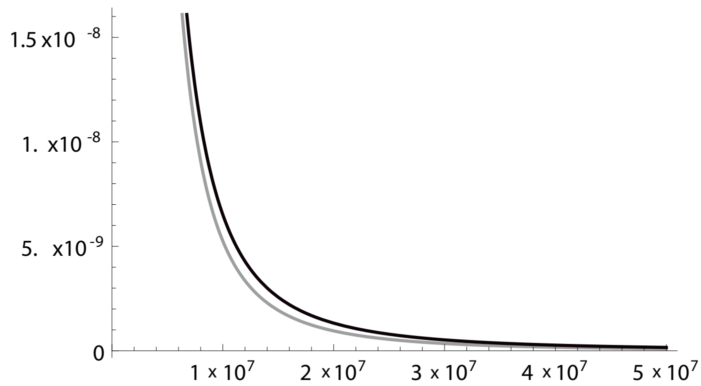


Figure 3.3: Population pdf (black) and sample pdf (grey) in Specification 2.

by

$$
\begin{aligned}
E_s(W) &= \int_{w_m}^{\infty} w f_s(w|a,b)\mathrm{d}w \\
&= \int_{w_m}^{\infty} w \frac{8.20245 * 10^7 * (a - b\ln(w))}{w^{2.3} * (a - 14.5847 * b)}\mathrm{d}w \\
&= \frac{4\frac{1}{3} * 10^6 a - 7.43 * 10^7 b}{a - 14.5847b}.
\end{aligned} \tag{3.7}
$$

In order to illustrate by how much the expected value differs between sample and population pdf depending on different values of parameters a and b, Figure 3.4 shows $E_s(W)/E_p(W)$ with a fixed at 0.9 (left panel) and b fixed at 0.04 (right panel).

If $b = 0.00$ the probability to be sampled is independent of the level of wealth, and, thus, there is no sampling bias (Figure 3.4, left panel). The greater b is, the steeper the non-response function, the more the sampled mean is biased downwards. The higher the parameter a is, the higher the overall survey response, the smaller the sampling bias (Figure 3.4, right panel). In Specification 2 the expected downward bias between sampled expected value and the population's expected value is 32.39%. Via extrapolation this directly translates to biased estimates of the population's aggregate private wealth.

Taking the results of this section further, in Specification 3 it is assessed how the ML estimation of the Pareto index as a function of the threshold parameter, which is

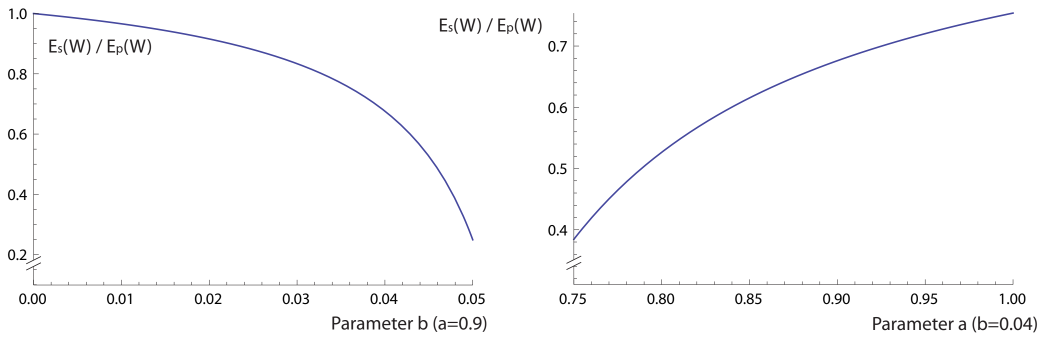


Figure 3.4: Expected value of sample pdf as a function of non-response parameters a and b.

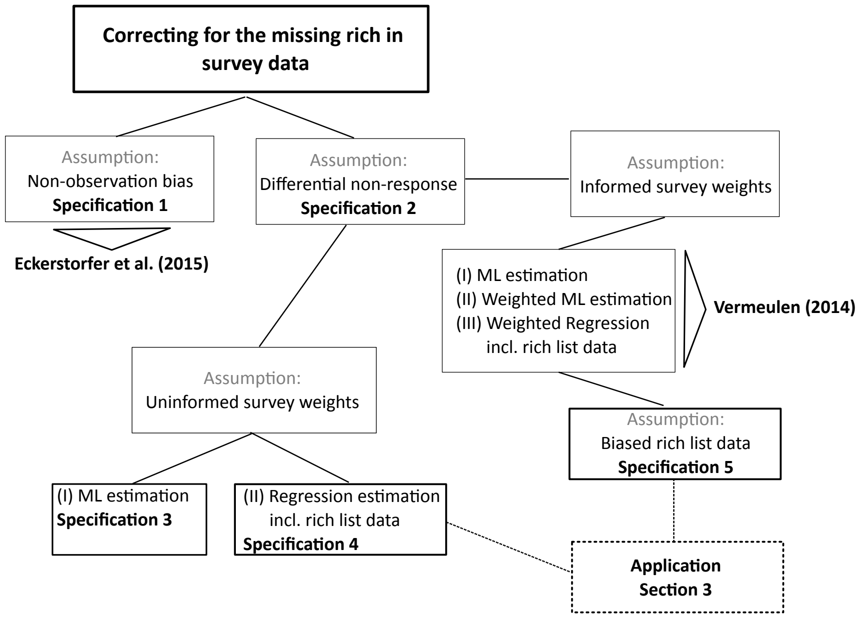


Figure 3.5: Overview. Correcting for the missing rich in survey data: assumptions, specifications and literature.

unknown in practice, behaves, assuming the sample suffers from differential non-response. In Specification 4 the effect of adding rich list data is shown, in which case an empiricist needs to switch to a weighted regression estimator. In addition, Specification 5 tests the impact of biased rich list data on the estimates assuming both informed and uninformed survey weights. The theoretical results of the simulations will then be compared to German wealth data in an application in Section 3.3; it is also discussed how a re-calibration of survey weights based on assumptions concerning the relationship between response and wealth levels might be conducted.

3.2.2 Maximum likelihood estimation of Pareto index α as function of the threshold parameter w_m

In a survey environment the threshold parameter w_m is crucial for the correct estimation of Pareto index α but unknown. If we set w_m too low, we include data in the estimation of α that do not follow a Pareto distribution and, thus, will end up with biased results.

Eckerstorfer et al. (2015) note, the inclusion of observations below the true minimum value of w_m yields downward biased estimates of α, the exclusion of data above w_m yields upward biased estimates of α. In the source cited, Clauset et al. (2009, p. 10), estimation behavior of α as a function of w_m is simulated, but the data below the threshold parameter follow an arbitrarily chosen exponential distribution. In fact, it is easy to prove another case using a different representation below the threshold parameter. Empirically, the expected value of α does not follow a clear pattern, as the shape of the plot depend on the empirical distribution of the data below the threshold parameter w_m (see Appendix 3.A and 3.B for simulated and empirical results, respectively).

In Specifications 3, 4, and 5, a population of 30 million households is assumed and net worth is distributed following a lognormal distribution below a certain threshold w_m and following a Pareto distribution above w_m. This means the wealth distribution is given by

$$f_p(w) = \begin{cases} 0 & w \leq w_m \\ \frac{1}{w\sigma\sqrt{2\pi}} e^{-\frac{(\ln w - \mu)^2}{2\sigma^2}} & 0 < w < w_m \\ \frac{\alpha w_m^{\alpha}}{w^{\alpha+1}} & w \geq w_m \end{cases} . \tag{3.8}$$

In effect, this population is characterized by a strictly positive net worth, which does not diminish the results of the simulation, as one may solely look at the upper parts of the distribution exceeding w_m, the Pareto distributed tail. The parameters and sample sizes are chosen to roughly resemble the West German population from Socio-Economic Panel study–with increased wealth concentration at the top–, so that the effect of non-response can be illustrated: With parameters $\mu = 11$ and $\sigma = 1.8$, and the Pareto distribution characterized by w_m=1,000,000 and $\alpha = 1.4$, the result is a population with an aggregate wealth of about 9.9 trillion, which translates to a mean net worth of roughly 330,000, the top percentile holds a share of 37.7% and the top 0.1% a share of 19.5% of the total net worth. It is assumed that an individual (or household) responds to the survey with a probability of P(Response) $= 0.62 + 0.02 * \ln(w_i) - 0.0024 * \ln(w_i)^2$. This relationship between wealth and survey response is directly derived from the 2012 Spanish EFF strata, as documented in Bover et al. (2014, p. 27). Figure 3.6 depicts the prototypical densities of the (log) wealth distribution before and after taking the response probabilities into account, as well as the response probabilities as a function of (log) wealth.

What is the resulting shape if the Pareto index α is estimated as a function of the threshold parameter, which is unknown in practice? In this simulation the ML estimator is used, while the survey weights are uninformed about the non-response mechanism. As the gross sample size in this simulation is 30,000 households, the non-response results in a net sample size of roughly half the size. Figure 3.7 shows that the resulting estimates of the Pareto index using maximum likelihood estimation are, in this case, against the median, about 0.15 units too high at the threshold value $w_m = 1$ million. In addition, the higher the assumed threshold value is set, the less precise the estimation of Pareto index α is. An upward biased Pareto index directly translates to downward biased estimates of the inequality of the Pareto distributed wealth. Had a researcher only used the raw sample, he would estimate a total net worth of 6.8 trillion (see Figure 3.7, right panel).

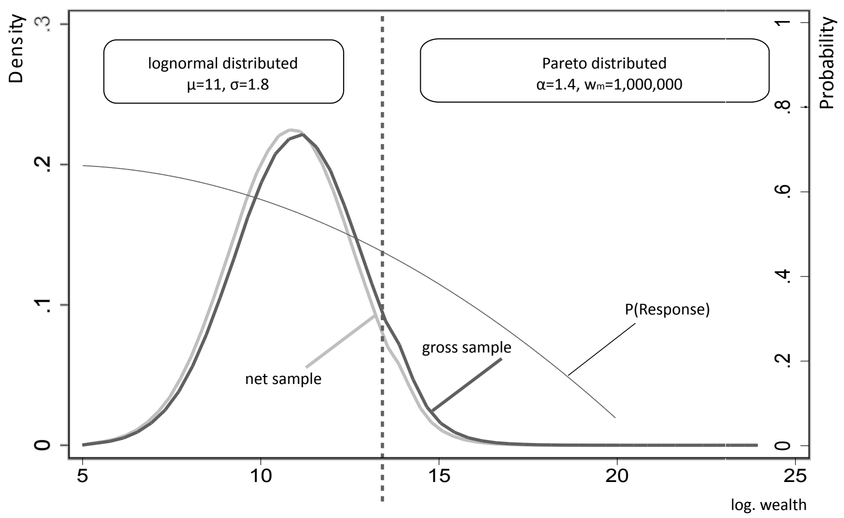


Figure 3.6: Simulation set-up. Assumptions in the Monte Carlo experiments in Specifications 3, 4 and 5.

The ML estimation of the Pareto index and a simulation of the tail based on the results barely improve on lacking survey data, if non-informative survey weights are included.

By the raw sample, the top-0.1% of the population holds 10.1% of the net worth; the target value would be 19.5%. The re-assessment slightly improves on the raw sample as Figure 3.8 shows: after applying a Pareto correction the top wealth shares are somewhat higher, as the wealth is redistributed from households in the 90th to 99th percentiles, whose wealth is overestimated before correction, to the top percentile. However, while the estimates are slightly improved, they are still far from acceptable. Generally, if the survey weights are uninformed, the number of Pareto distributed households is too low, and their distribution too equal, both before and after correction. For comparison's sake, the same simulation with informed survey weights using a weighted ML estimator is repeated in Appendix 3.C.

3.2.3 The regression method including rich list data

In a next step, it is assumed that information on the net worth of the top 50 net worth holders is available to the researcher from an external source. This is where the so-called rich lists come into play. It is evaluated, whether external data may be used in order

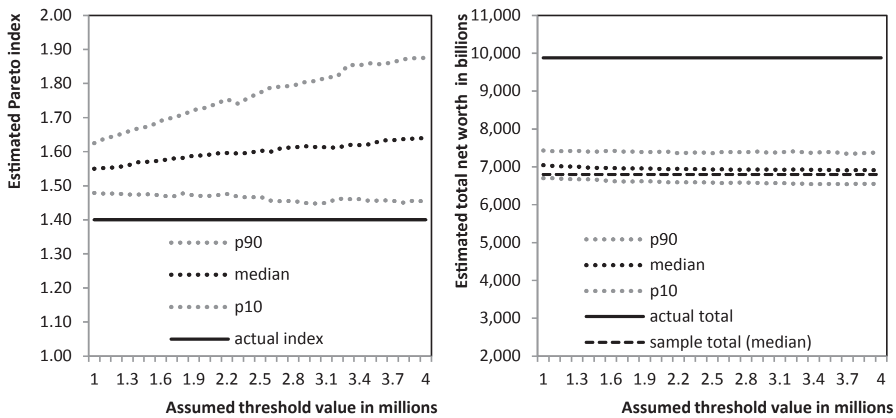


Figure 3.7: Specification 3. Impact of differential non-response on the maximum likelihood estimates for the Pareto index α and total net worth, plotted as a function of the value assumed for w_m. 1,000 samples each, drawn from test distributions, Eq. 3.8 , see also Fig. 3.6.

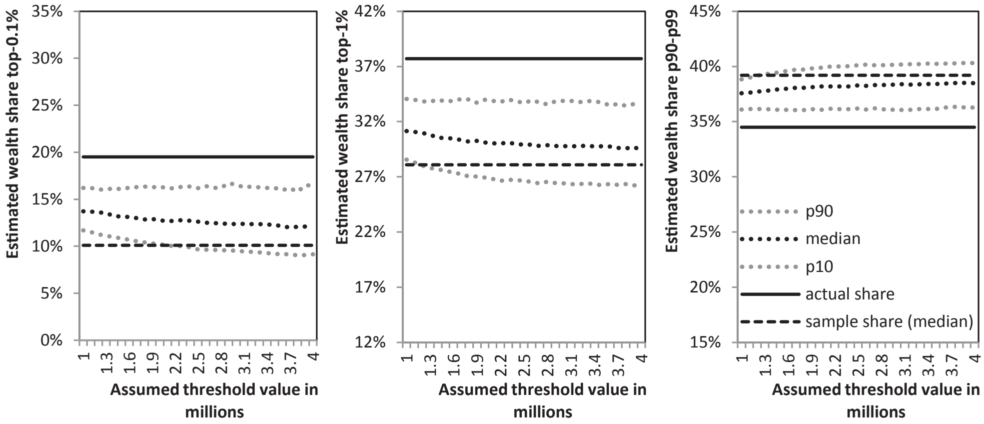


Figure 3.8: Specification 3. Impact of differential non-response on the top wealth shares before and after Pareto correction, plotted as a function of the value assumed for w_m. 1,000 samples each, drawn from test distributions, Eq. 3.8, see also Fig. 3.6.

to obtain unbiased estimates, if a survey sample suffers from differential non-response and the very rich are missing as a result. As in Specification 3, we draw a net sample of about 15,000 households. Then, we add the 50 wealthiest households taken from the base population for the estimation (each carrying a survey weight of one). In this case, using a sample combined of two sources, the maximum likelihood estimator would yield biased results, hence leaving us with the regression method as the only option. As above, N_i denotes the frequency weight of household i; $N_{w>w_i}$ is the sum of the frequency weights exceeding the net worth of household i; thus, it corresponds to the rank of a household, if survey weights are included. The regression estimator then is the estimated parameter from a regression of the log of the net worth on the log of the rank of all households holding a net worth of w_m and higher:

$$\ln(N_{w>w_i}) = c - a_{reg} \ln(w_i). \tag{3.9}$$

It is assumed that the weights are uninformed, i.e. the frequency weight for any household is the inverse of the sampling probability. The same specification is, again, repeated with informed survey weights in Appendix 3.C.

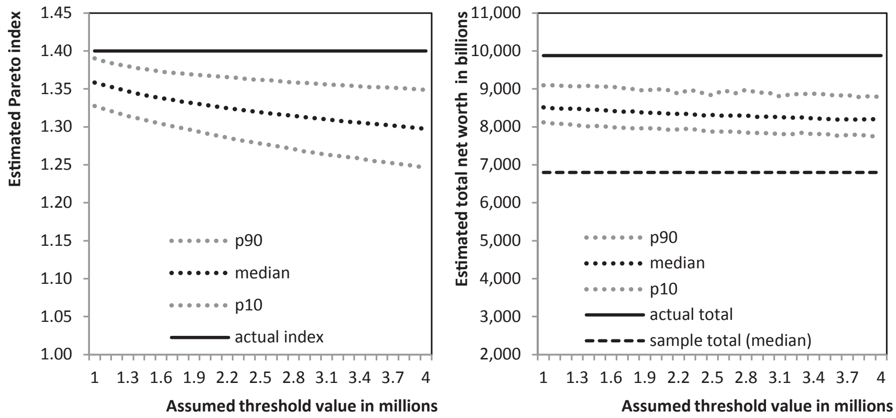


Figure 3.9: Specification 4. Including rich list data in the regression estimation. Regression estimates for the Pareto index α and corrected total net worth, plotted as a function of the value assumed for w_m. Uninformed survey weights. 1,000 samples drawn from the test distribution, Eq. 3.8, see also Fig. 3.6.

In Figure 3.9 the results on the left hand side depict the median estimates and selected percentiles of the estimation of Pareto index α. If the non-response mechanism is unknown, α is underestimated at the threshold $w_m = 1000000$ by 0.04 units, and steadily decreases thereafter as less and less households are included in the estimation. Apparently, if the survey sample is affected by differential non-response, including rich list data will result in an underestimation of the Pareto index, the inequality in the top tail is too high. The target estimate of an unbiased total net worth would be 9.9 trillion. An obvious question arises: if the Pareto index α is too low, why are estimates of the aggregate wealth too low? The adjustment for the missing rich here is plagued by two separate biases with countering effects: on the one hand, the Pareto index is too low, resulting in an overestimation of wealth concentration for the Pareto distributed part of the wealth distribution. On the other hand, the number of observations above the Pareto threshold w_m is too low. In effect, too few observations are distributed too unequally in the re-assessed tail of the wealth distribution, and the latter effect dominates the computation of the aggregate wealth. Only if the survey weights are informed about the exact mechanism of non-response, is an empiricist able to determine how many households are in the tail of a wealth distribution.

As Figure 3.10 shows, while the aggregate wealth of the population is underestimated by 1.4 trillion, the shares of the wealthiest 1 % and 0.1 % of households are likely to be improved. To be more precise, it is likely that the corrected estimates overshoot the mark: as the re-assessment of the top tail is fed with both wealth data that is biased to the middle-class and rich list data at the very top, a (relative) redistribution from the lower deciles to the top takes place (see sample shares and median of the wealth held by 90th to 99th percentiles), resulting in top wealth shares that are too high and wealth levels held by the upper middle-class that are too low. As the 90th to 99th percentiles are equally important for the aggregate wealth, but their net worth is falsely assessed, the aggregate wealth is still biased downward; the top wealth shares are biased upward.

3.2.4 The impact of biased rich list data

The dubious nature of data taken from rich lists published in magazines largely remains unresolved. Assuming that mistakes in the journalistic black box are merely random would have a negligible effect on the estimated Pareto indices of the top tail. However, if the lists' entries are too high or too low, they have a significant impact on the estimations. Admittedly, since neither the sources of data nor the method of obtaining the information are made public, the details of such lists ultimately cannot be verified. There are results

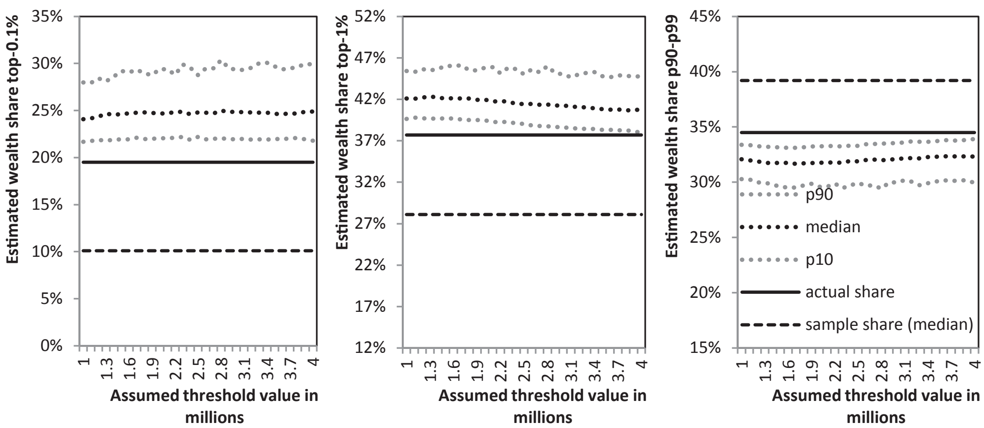


Figure 3.10: Specification 4. Including rich list data in the regression estimation. Top wealth shares plotted as a function of the value assumed for w_m. Uninformed survey weights. 1,000 samples drawn from the test distribution, Eq. 3.8, see also Fig. 3.6.

hinting at an overvaluation of assets in the Forbes magazine. When US federal tax authority researchers compared the tax data of deceased persons and the Forbes list, they discovered that the list overestimated net worth by approximately 50 percent, primarily due to assessment difficulties, fiscal distinctions, and poor assessment of liabilities (Raub et al., 2010).

In the last specification of this simulation study, the possible impact of an overestimation of billionaires' wealth in rich lists on the relevant estimates is assessed. In Specification 5a it is assumed that the survey weights are informed about the non-response mechanisms, in Specification 5b the survey weights are non-informative. The Pareto index is estimated using the weighted regression estimator including the top 50 rich list entries, however, they are multiplied with a random normal variable with a mean of 1.4 and a standard error of 0.15, resulting in an overestimation of billionaire's wealth by 40 % on average.

The impact of biased rich list entries on the estimation of the Pareto index is severe. Table 3.1 shows the results, depending on the assumed value of the threshold parameter,

Table 3.1: Specifications 5a and 5b.

	5a - survey weights **informed**				5b - survey weights **uninformed**			
Assumed threshold w_m **in millions**	Pareto index	Total net worth in billions	Wealth share top-1% in %	Wealth share p90-p99 in %	Pareto index	Total net worth in billions	Wealth share top-1% in %	Wealth share p90-p99 in %
1.0	1.341	10,910	43.1	31.8	1.332	9,192	46.0	30.0
1.2	1.339	10,900	43.3	31.6	1.322	9,198	46.7	29.5
1.4	1.337	10,910	43.2	31.6	1.315	9,046	46.3	29.4
1.6	1.334	10,910	43.5	31.4	1.310	9,029	46.0	29.4
1.8	1.332	10,910	43.5	31.4	1.305	9,009	46.2	29.4
2.0	1.330	10,870	43.2	31.6	1.302	8,984	46.2	29.4
2.2	1.328	10,730	43.2	31.7	1.299	8,979	45.9	29.3
2.4	1.327	10,880	43.4	31.4	1.296	8,948	45.9	29.5
2.6	1.328	10,830	43.1	31.5	1.293	8,865	45.6	29.7
2.8	1.328	10,770	43.0	31.7	1.289	8,821	45.4	29.7
3.0	1.329	10,780	43.1	31.6	1.286	8,955	45.6	29.6
3.2	1.327	10,720	43.0	31.6	1.283	8,849	45.1	29.9
3.4	1.327	10,760	42.8	31.7	1.279	8,808	44.9	30.1
3.6	1.326	10,710	42.7	31.6	1.275	8,774	44.9	30.0
3.8	1.325	10,750	42.8	31.6	1.274	8,833	44.9	30.0
4.0	1.324	10,730	42.9	31.6	1.273	8,753	44.7	30.1
Actual values	1.400	9,875	37.7	34.5	1.400	9,875	37.7	34.5
Before correction (at w_m = 1 million)	1.547	6,800	28.1	39.2	1.547	6,800	28.1	39.2

Median estimates weighted regression method using biased rich list data. Estimated Pareto index, total net worth and shares as a function of the value assumed for w_m. Test distributions given by Eq. 3.8, see also Fig. 3.6.

as well as the actual values and median values calculated from the raw samples (at $w_m = 1$ million). At the threshold value, the index is about 0.07 units too low and appears to decrease slightly thereafter. If the survey weights are non-informative, the Pareto index quickly decreases, the higher the threshold value is assumed to be. The impact on the corrected total net worth is substantial. Against the median the total net worth is either about 1 trillion too high, or at least 700 billion to low, which in turn means that, if both the rich list is biased upward and the weights are uninformed, the two biases somewhat offset each other. However, in both cases this directly translates to top wealth shares that are consistently too high; the aggregate wealth is then missing from the upper middle class (p90-p99). Assuming that an empiricist's survey weights are informed about the correct non-response mechanism, the impact of biased rich list data is less severe than in Specification 4. Empirically, an empiricist might want to factor in both effects in the estimations, as the rich list data are likely to be biased upward and the survey data probably suffer from differential non-response. Overall, it remains questionable whether researchers might want to put a lot of confidence in the results of such an exercise. Moreover, in a study by Brzezinski (2014), Pareto distributions were consistent with the distribution of rich list data covering U.S., China, and Russia in only about one-third of the cases. This leads to the question of whether there might be alternative distributional assumptions or estimation techniques for simulating the super-rich population that is otherwise omitted from survey data. On the other hand, Capehart (2014) finds that the results of goodness-of-fit tests might change for the positive once researchers take measurement error into account.

3.3 Application: German survey data

The primary goal of the German Socio-Economic Panel (SOEP), and other similar surveys, is not to measure the share of the top-1-percent wealth holders. These surveys serve multiple purposes that might not even be related to wealth (for an overview of the SOEP survey, see Wagner et al., 2007). However, the claim to be representative for the whole population and certain features of the surveys make them useful for the task at hand. Additionally, as there is no tax or register data available in Germany–similar to most other countries–survey data remains as the last trustworthy and publicly available source for a scientific analysis of the distribution of wealth and wealth inequality.

As shown in Section 3.2, once survey data suffers from differential non-response, both ML estimation without rich list data and regression estimation including rich list data will yield biased results, as both the Pareto index and the number of households exceeding the threshold value are biased. However, both biases may be corrected if the survey weights are informed about the non-response mechanism. To be more precise, the survey weights need to be explicitly informed about the relationship between non-response and wealth. Thus, this Section provides an illustration: How to inform survey weights about wealth-related non-response using exogenous information and, hence, obtain the correct distribution of household net worth. Notwithstanding this announcement, it turns out that valid exogenous information is of utmost importance, as depending on the source the exercise produces wildly different estimates.

The German Socio-Economic Panel Study (SOEP) is a longitudinal representative survey collecting socio-economic information on private households in Germany. Additionally, a module collecting wealth information was included in 2002, 2007 and 2012. In 2002, the SOEP sampled high-income individuals for the last time, it is reasonable to assume that the precision at the top of the wealth distribution was much higher in 2002, and slowly decreased afterwards due to panel attrition. Table 3.2 summarizes the data with regard to net worth of private households.

The framework we use to estimate the upper margin of wealth distribution is, as in the simulation study, twofold and based on estimation from survey data alone and a combination of survey data with data on the absolute peak of distribution derived from all those with the respective citizenship on the list of billionaires published annually by the US Forbes magazine. However, the Forbes lists does not provide sufficient details every year to be able to determine whether these individuals are also living in the respective

Table 3.2: Summary statistics: Net worth of private households in Germany, according to SOEP survey 2002, 2007 and 2012.

Survey wave		2002		2007		2012
Mean		149838		153998		154380
Median		37247		39220		46680
p90		361239		372899		380740
p95		538470		562386		563100
p99		1272189		1375940		1349640
Share of top-1%		21.1 %		21.6 %		18.2 %
Share of top-0.1%		7.6 %		7.1 %		5.3 %
Max. in million euros		62.7		31.7		45.5
Aggregate private wealth in billion euros		5.800		6.116		6.278
Number of households	n	N	n	N	n	N
> €500000	1089	2342967	986	2522275	862	2516656
> €1000000	334	620910	304	683088	270	708424
> €3000000	47	88204	56	133175	42	108366

Source: SOEPv30, own calculations, means over 5 implicates of multiply imputed data.

country.[18] Likewise, billionaires who are living in one of the countries, but did not hold the respective citizenship, were excluded from the analysis (Table 3.3). In this process, it is assumed that each individual on the Forbes list represents a single household.[19]

In the first step, the ML estimator (see Spec. 3 in Section 3.2.2) and the regression estimator including rich list data (see Spec. 4 in Section 3.2.3) are applied to the 2002, 2007, and 2012 SOEP wealth data. It is shown how the estimation of the Pareto index, as a function of the threshold parameters, yields shapes that are reminiscent of the results in

Table 3.3: Entries in the Forbes list of billionaires at the time of SOEP survey wealth modules.

	Germany 2002	Germany 2007	Germany 2012
Number of entries	34	55	55
Aggregate wealth in billion euros	159.8	185.4	188.7
Max. in billion euros	30.9	15.1	19.1

US Dollar-Euro exchange rates as of March 1 of the respective years. Source: own calculations based on Forbes magazine's yearly-published list of billionaires.

18 Moreover, there may also be individuals living in Germany who are not German nationals but should be classified together with other private households.

19 It is not possible to tell from the Forbes list whether the households of these individuals include other members or not.

Section 3.2.2, leading to the assumption that the data suffers from differential non-response. The empirical results are highly volatile, often driven by very few observations leaving the estimation, as the threshold parameter w_m is set higher (Figure 3.11, left panel). Furthermore, the left hand side, depicting the weighted maximum likelihood results, shows that any regular shape hinting at an empirical w_m is missing. However, the shapes are reminiscent of what is shown in the Monte Carlo experiment simulating survey data with differential non-response: After decreasing estimations for α, there is a local minimum between €1 and €1.6 million. However, to locate w_m and α in this corridor would certainly be bold. The margins of error are large, given the sample size, and the Pareto index α certainly is a good deal too high (see Spec. 3). Including the survey weights in the estimation would only offset the effects of differential non-response, if the weights reflect the true response probabilities along the distribution of wealth (see Spec. 3b). Once the weighted regression method is applied and the respective German rich list members from the Forbes magazine are incorporated, the curves align and do not vary a lot between the survey waves (Figure 3.11, right panel). There are two main results for regressions including rich list data: (1) once survey weights are informed about the non response mechanism, which results in an unbiased estimation of the Pareto index α after the true w_m is reached, then the curves turn into a straight line at the true Pareto index in the simulation (Vermeulen, 2014, see also Spec. 4b in Appendix 3.C). (2) If the survey weights are uninformed, the estimates for Pareto index α decreased steadily, while already being downward biased at w_m (Spec. 4). The empirical results using SOEP wealth data are more reminiscent of the latter case, hinting at differential non-response in the data.

Next, the calculated Pareto indices, as a function of the Pareto threshold value shown in Figure 3.11, are used to compute the aggregate private wealth in Germany. Using weighted maximum likelihood estimators and re-simulating the data barely impacts the estimates for aggregate wealth. Originally, the total net worth varied between €5.8 trillion in 2002 and €6.2 trillion in 2012. As shown in the left panel of Figure 3.12, the estimated values are rather close to the ones observed without correction for the missing rich. This

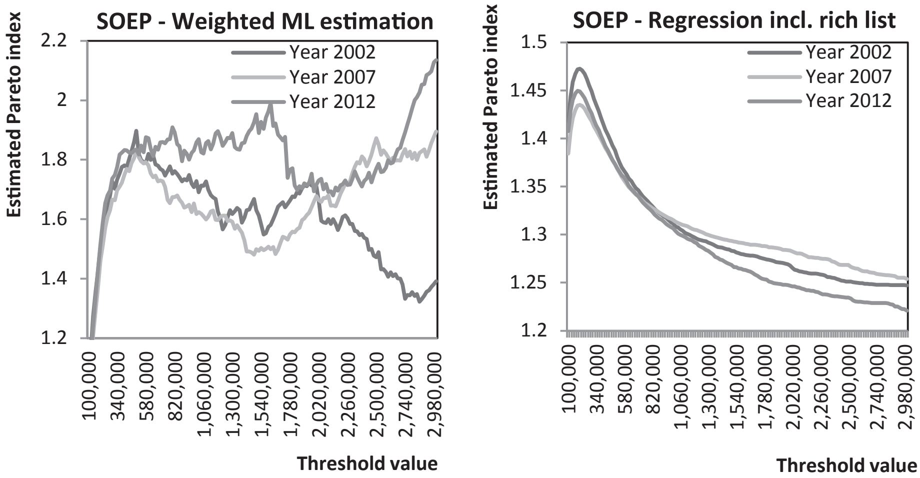


Figure 3.11: SOEP survey 2002, 2007, 2012. Estimation of Pareto index as a function of threshold parameter w_m. Means over 5 implicates of multiply imputed data. Source: SOEPv30, Forbes' list of billionaires, own calculations.

is to be expected if one solely tries to correct for non-observation bias but the data suffers from differential non-response. If differential non-response is assumed and the regression method plus rich list data from the Forbes magazine are used, one can expect that the resulting values are less, but still downward, biased (Spec. 4). If it is assumed that w_m is at €1 million (and not varying over time), the re-assessment adds about €1 trillion in 2002, €1.2 trillion in 2007 and €1.4 trillion in 2012.

Re-calibrating survey wealth data based on external sources

In order for a re-assessment of top wealth with survey data to work, the most obvious first solution is to reweight the data, as it is biased to the middle class using non-informative weights (see Figure 3.6). The steps include: (1) Based on the response as a function of wealth each household is assigned both the probability to respond and the inverse probability. (2) A household's inverse probability is multiplied by its uninformed frequency weight.[20] (3) The aggregate number of households is divided by the sum of (2), yielding a

20 The uninformed frequency weight refers to the household's weight as provided through the survey distributor. For instance, in the SOEP data there is an oversample for East German households and one would want to preserve the ratio between East and West after re-calibration.

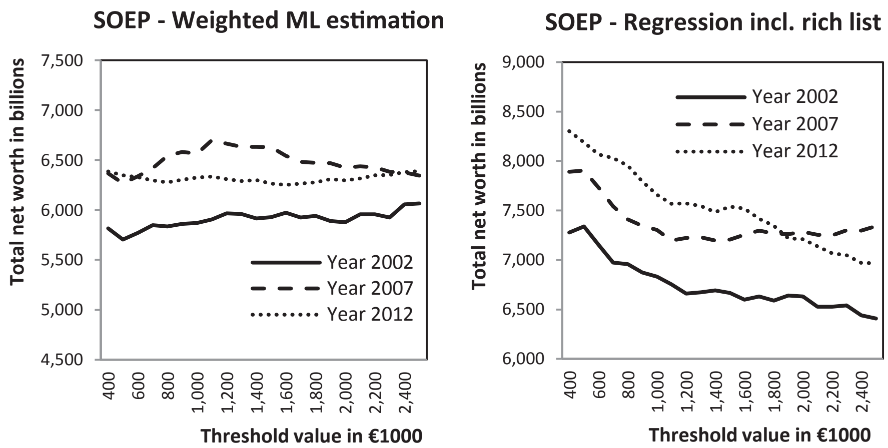


Figure 3.12: SOEP survey 2002, 2007, 2012. Total net worth and top wealth shares based on corrected data using regression method including rich list data, as a function of threshold parameter w_m. Means over 5 implicates of multiply imputed data. Source: SOEPv30, Forbes' list of billionaires, own calculations.

scalar to adjust the weights to the population size. (4) Each household's new frequency weight is then given by product of a household's specific value of (2) and the scalar (3). For this reweighting to be applicable, the functional form of non-response needs to be known, which is not the case for Germany and the SOEP data. As they are the only source of information on the subject matter, the Spanish EFF strata (Bover et al., 2014) or the U.S. SCF strata (Kennickell and McManus, 1993) may be used to describe response probabilities as a function of (log) wealth. Alternatively, some financial institutions release their own reports on the (global) distribution of wealth. For instance, Credit Suisse's report delivers enough information to determine the varying response probabilities depending on the level of wealth in Germany. Hence, using the equation below, the SOEP survey weights can be transformed to mirror the number of households within the strata provided by Credit Suisse for 2012 for the respective reporting year (Credit Suisse, 2012).

- **Spanish EFF:** $\mathrm{P(Response)} = 0.622 + 0.020 * \ln(w_i) - 0.00242 * \ln(w_i)^2$
- **U.S. SCF:** $\mathrm{P(Response)} = 0.425 + 0.021 * \ln(w_i) - 0.00183 * \ln(w_i)^2$
- **Credit Suisse:** $\mathrm{P(Response)} = 0.340 + 0.023 * \ln(w_i) - 0.00256 * \ln(w_i)^2$

In all three cases including a quadratic term substantially increases the fit with the empirical strata. Note that such a re-calibration mathematically only affects the relative probabilities to respond between two households with differing wealth levels as it adjusts the frequency weights according to the functional form, even though the overall non-response rate in the SOEP might differ. The re-calibration is completely relative. It does not affect statistical power or anything else. It is best illustrated as revoking the net sample to the gross sample in Figure 3.6, thus, eliminating the effect of differential non-response on the household level.

The exemplary result for the 2012 SOEP data is shown in Figure 3.13. For once, the Pareto indices are still decreasing after re-calibration using EFF and SCF strata, hinting at either the wrong functional form of the non-response generating mechanism–the middle-class bias might be more severe–or the Pareto distribution does not yield a good fit for the top tail of the German wealth distribution. For the Credit Suisse re-calibration, the Pareto indices are slightly, but steadily, increasing. Assuming the threshold value would be at about €2 million, inequality at the top is lowest for Credit Suisse and highest for the SCF strata. Note that it is the same for aggregate wealth (Figure 3.13, top left panel), meaning that the Credit Suisse strata yields a substantially less unequal Pareto distribution, but also substantially higher wealth levels overall. Trusting in their data would mean that the SOEP survey underestimates upper middle class wealth (say, p95-p99), as compared to the EFF and SCF non-response assumptions, while the shares of the top-1% and top-0.1% are about the same or lower. Likewise, if one does put trust into the non-response mechanisms as reported for Spanish or U.S. surveys, the top wealth shares are considerably higher than for the raw survey data and slightly higher than for the Credit Suisse strata. Note that in every case, the Pareto indices are higher than without re-calibrating the weights and using rich list data (cf. Figure 3.12), meaning less inequality at the top. The aggregate wealth is also higher, depending on the re-calibration between €1.5 and €5 trillion.

As shown, re-calibrating the survey data using exogenous sources has a much more severe impact on the estimates than choosing a threshold value or setting the Pareto

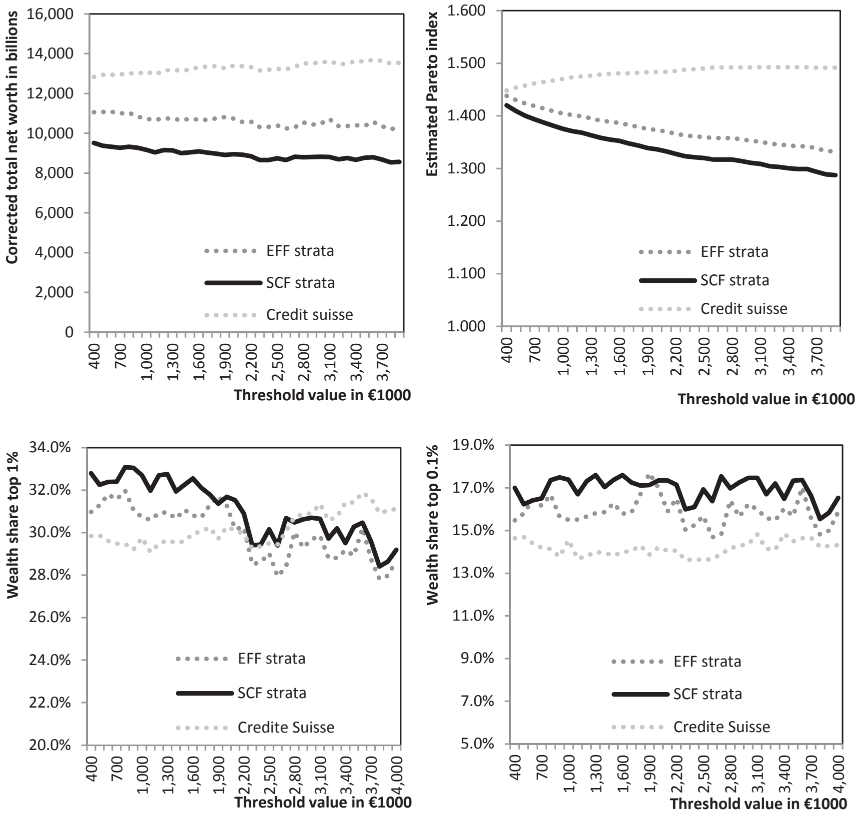


Figure 3.13: SOEP survey 2012. Total net worth, estimated Pareto index and top wealth shares after re-calibration of survey weights. Data sources: SOEPv30, Forbes' list of billionaires as of March 2012, Bover et al., 2014; Kennickell and McManus, 1993; Credit Suisse, 2012, own calculations.

indices (see variation in Figures 3.11 and 3.12). Thus, the source of bias impacting the measurement of top wealth shares with survey data the most are non-informative survey weights. The results in this section show that the resulting corridor, depending on the non-response assumptions, is very broad. Moreover, from an empiricist's point of view, none of the results above could reasonably be ruled out from the outset. The balance sheet data, as provided by the German Federal Statistical Office (Destatis, 2015), uses different definitions and delimitations of household wealth as compared to survey data, which means it is not advisable to compare it directly to survey data (Grabka and Westermeier, 2015). For now, there is no benchmark available. As all three assumptions on the non-response are within the realm of possibilities, the results leave us with a glaringly wide corridor of possible values for aggregate private wealth and top wealth shares. The validity of exogenous data–and the rich list data–remains a matter of trust on the part of empiricist.

3.4 Summary and conclusion

It is safe to say that in countries without reliable tax return data, or otherwise obtained register data, on the distribution of wealth, policy makers remain largely uninformed about the extent of wealth concentration at the top. Empiricists started improving upon lacking survey data by assuming a Pareto distribution at the top and computing new estimates with and without rich list data.

As other simulation studies show, once the survey weights are informed about the relationship between response and wealth, the weighted maximum likelihood estimator is unbiased. Regression estimation including rich list data then improves the precision of the estimates (Vermeulen, 2014). However, the simulations conducted in this study show that if the data suffer from differential non-response and the survey weights are uninformed, then the maximum likelihood estimator yields estimates of the aggregated wealth and top wealth shares that are biased downwards. Adding rich list data and switching to a regression estimation falls short of compensating for the bias, as it underestimates the total net worth, while it overestimates top wealth shares, as too few households in the tail are distributed too unequally. Moreover, there is strong indication that rich list data, such as the yearly published Forbes list, actually overestimate billionaires' wealth, which in turn yields estimates of the top wealth shares and aggregate wealth that are systematically too high. Researchers are not readily able to assess these biases.

The best remedy for lacking survey data is a re-weighting of the survey weights based on either additional assumption or valid data. As such data is not available for Germany, the Spanish EFF and the US SCF response probabilities as a function of wealth were applied in the application of this study to hypothetically show how additional data might be used to compensate for differential non-response. Additionally, households are re-weighted to match their distribution in Credit Suisse's 2012 wealth report. Before re-calibration, both the ML estimator without rich list data and the regression including rich list data yield estimates of the aggregate wealth that might still be biased downwards. After recalibration, the aggregate wealth was higher by more than €1.5 or €5 trillion, depending on the assumed non-response mechanism. The 2012 top wealth shares of the top-1% and the top-0.1% increase by more than 10%. However, all estimations depend on the empirically unknown threshold parameter, the assumed relative response probabilities of the households–which might shape up differently to other countries such as Spain or the US–and the assumption that wealth is actually distributed following a Pareto distribution at the top.

If anything, the findings emphasize the need to use exogenous information in sample design, which allows for creating appropriate weights taking non-response into account. Survey providers must know the exact response probabilities to offset the effects of differential non-response as well as to calculate totals and top wealth shares reliably. Only then can developments in the long run be reasonably analyzed. Until more exhaustive data sources are accessible to researchers–or tax authorities are willing to cooperate more closely with survey providers–it might be a more viable choice to put the efforts in steadily well-run surveys that include dedicated oversamples of high-net-worth-households.

3.A Simulation: Pareto index as a function of threshold parameter without non-response (ML estimation)

As in Clauset et al. (2009), I compute the estimates for α as a function of w_m, using the maximum likelihood estimator a_{ml}, and examine the resulting plot. In this Monte Carlo experiment, simple random samples of 30000 households are drawn from equation (3.8), without non-response. Only the median estimates for varying Pareto indices α and threshold values w_m are of interest. This simulation was carried out 1,000 times for each value of w_m.

In Figure 3.14, w_m is fixed at 1,000,000 and α varies between 1.2 and 1.7, in Figure 3.15, α is fixed at 1.4 and w_m varies between 750,000 and 1,750,000. We observe that the estimates for Pareto index α are increasing with w_m until the true Pareto index value is reached, at which point none of the observations of the log-normal distributed samples are included in the estimations. Plotting ML estimates of the Pareto index as a function of w_m with survey data from the German Socio-economic Panel Survey or Euro-area HFCS data yields some results that are fairly close to the shapes in Figure 3.8 for some countries (see Section 3.2.2 and Appendix 3.B).

The simulation shows that the estimated value of α as a function of w_m exhibits a robustly straight line, if the data truly follow a Pareto distribution and there is no non-response. At least a range of values could be given, which, with a very high probability, also includes the threshold value w_m. One would like to choose the value shortly after the plot becomes a straight line. In this case, setting w_m too low leads to results that overestimate the concentration of wealth in the top area (as α is too high). However, in this example it is by set-up of the simulation data that the parameters of the Pareto distribution are easily identified using a plot. In this case, the mode of data generation makes sure that there is a relatively hard transition between log-normal and Pareto distributed wealth. In empirical data the identification of the parameters is hardly as straightforward as in this Monte Carlo experiment and involves a battery of other problems, Clauset et al. (2009)

offer a very detailed review of the estimation techniques and possible pitfalls. As the shape of the plot depends on the distribution below w_m in specifications 3 to 5, the resulting Pareto indices below the true w_m are not plotted.

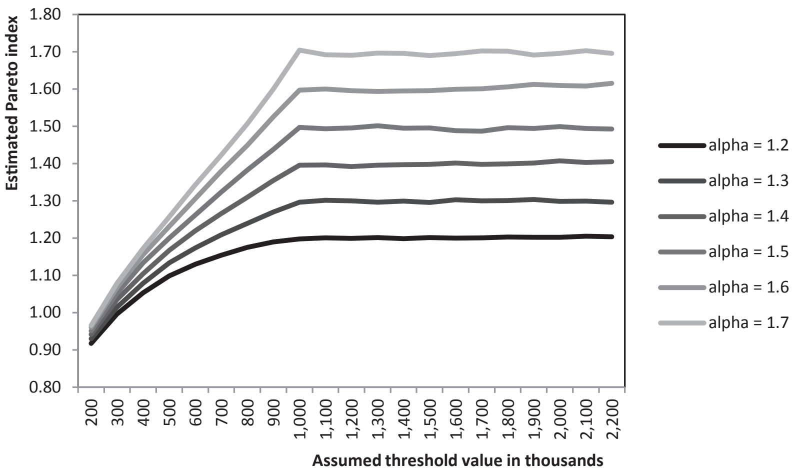


Figure 3.14: ML estimation without non-response. Pareto index α plotted as a function of the value assumed for w_m, for various actual α. 1,000 samples each, drawn from test distributions, Eq. 3.8 , see also Fig. 3.6.

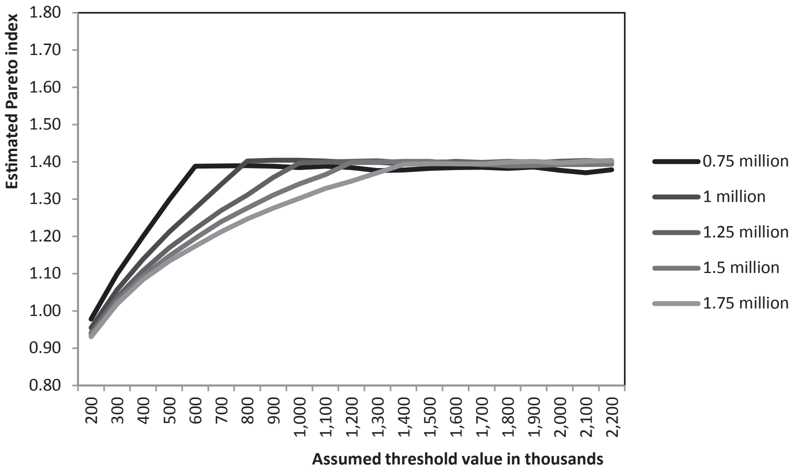


Figure 3.15: ML estimation without non-response. Pareto index α plotted as a function of the value assumed for w_m, for various actual w_m. 1,000 samples each, drawn from test distributions, Eq. 3.8 , see also Fig. 3.6.

3.B Empirical results: Pareto index as a function of threshold parameter using HFCS data (weighted ML estimation)

As seen in Appendix 3.A, estimation of the Pareto index as a function of the threshold parameter may give a hint at the value of the threshold parameter. Here are results of (weighted) maximum likelihood estimates using the Eurosystem Household Finance and Consumption Survey (HFCS, wave 1). Empirically the resulting shapes vary: some are reminiscent of the plots resulting in the simulation, such as net wealth data from Austria, Belgium, France or Italy. Others exhibit a global maximum before seeing a decrease of the Pareto index, such as Finland and Spain. The specific shapes depend on the distribution of wealth below the threshold parameter–assuming that the data are indeed Pareto distributed. Fewer observations are included in the estimation of the Pareto index as the threshold parameter is set higher, hence, the estimates typically become more erratic.

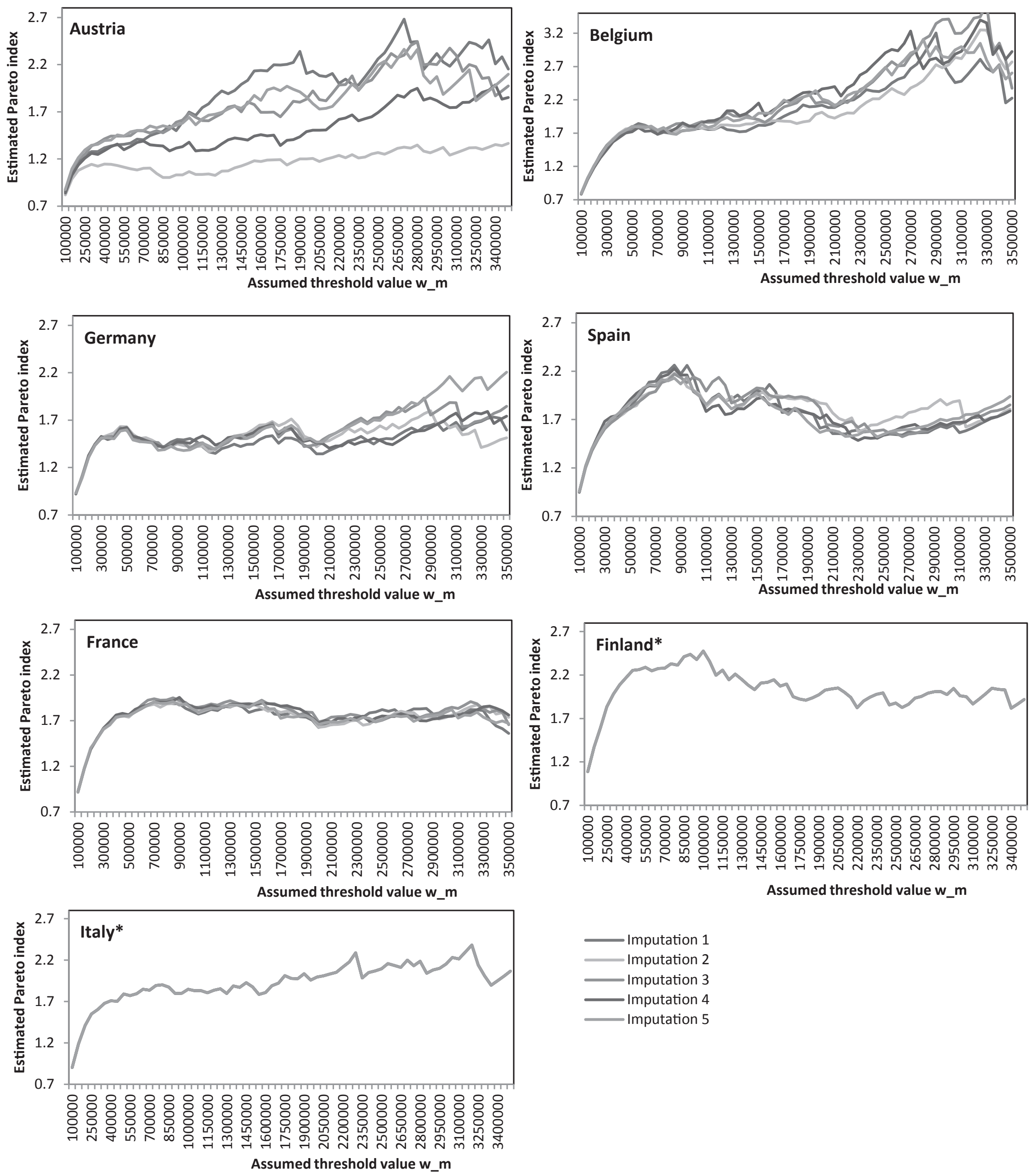


Figure 3.16: ML estimation in HFCS 2013. Household net worth HFCS, Pareto index α as a function of the threshold parameter w_m. Source: HFCS 2013, own calculations. *) No multiply imputed data was provided by Finland and Italy.

3.C Replication of Specifications 3 and 4 with informative weights

Since survey weights are allowed for, when calculating the Pareto parameter with a regression estimator or maximum likelihood estimator, these should, as far as possible, take into account the structure of the differential non-response. Here, for comparison's sake, it is assumed that the researcher knows about the non-response mechanism and survey weights are informed accordingly. This means the alternative weighted maximum likelihood estimator is used instead of the unweighted ML estimator. N_i is the frequency weight of a household i, $N_{w>w_m}$ is the combined frequency weights of all households exceeding the threshold value w_m:

$$a_{wml} = 1 + \frac{N_i}{N_{w>w_m}} \left[\sum_{i=1}^{n} \ln \frac{w_i}{w_m} \right]^{-1}.$$

If the survey weights are (perfectly) informed about the differential non-response the median estimates of the Pareto index are unbiased at the true value of w_m and thereafter (Figure 3.17, left panel). However, the margin of error is rather high as seen by the 10th and 90th percentiles. The weighted maximum likelihood estimator turns out to be unbiased, but not efficient. Overall, even if the researcher knows the exact mechanism of the differential non-response–which is usually not the case–, estimates of the Pareto index α vary strongly due to a lack of precision. This lack of precision directly translates to corrected values of aggregate private wealth (Figure 3.17, right panel) and top wealth shares (Figure 3.18), which are unbiased against the median but not very precise given the net sample size of roughly 15,000 households.

Specification 3b: Replication of Specification 3 with informed weights. Weighted maximum likelihood estimation of Pareto index α as a function of w_m. Assumption: differential non-response.

Next, it is illustrated, how informative survey weights change the results of specification 4. As in Vermeulen (2014) the incorporation of informed survey weights, a weighted regression

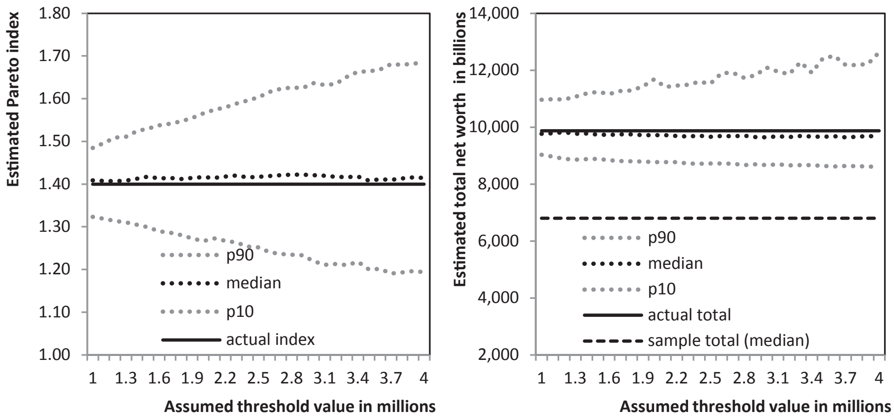


Figure 3.17: Specification 3b. Informative weights, ML estimator: Impact of differential non-response on the maximum likelihood estimates for the Pareto index α and total net worth, plotted as a function of the value assumed for w_m. 1,000 samples each, drawn from test distributions, Eq. 3.8 , see also Fig. 3.6.

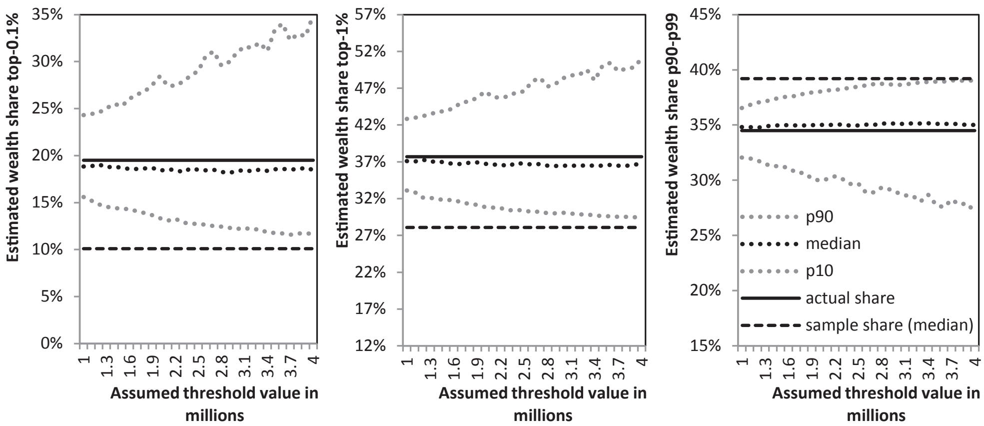


Figure 3.18: Specification 3b. Informative weights, ML estimator: Impact of differential non-response on the top wealth shares before and after Pareto correction, plotted as a function of the value assumed for w_m. 1,000 samples each, drawn from test distributions, Eq. 3.8, see also Fig. 3.6.

estimator and including the top 50 entries from a rich list (assuming they are unbiased) greatly improves the precision as compared to the ML estimator in specification 3b without rich list data. Furthermore, estimation precision of the Pareto index increases (almost) independently of the chosen threshold value (Figure 3.19, left panel). The corrected totals using this rich list estimation are unbiased and efficient. This serves to illustrate that the real problem an empiricist faces is not the estimation of the parameters but to obtain survey data that is informed about the underlying mechanism of non-response. Only then the number of household exceeding the threshold is correct. The slight under- or overestimation of top wealth shares as depicted in Figure 3.20 disappears as the sample size grows larger.

Specification 4b: Replication of Specification 4 with informed weights. Weighted regression estimator (including top 50 rich list entries) of Pareto index α as a function of w_m. Assumption: differential non-response.

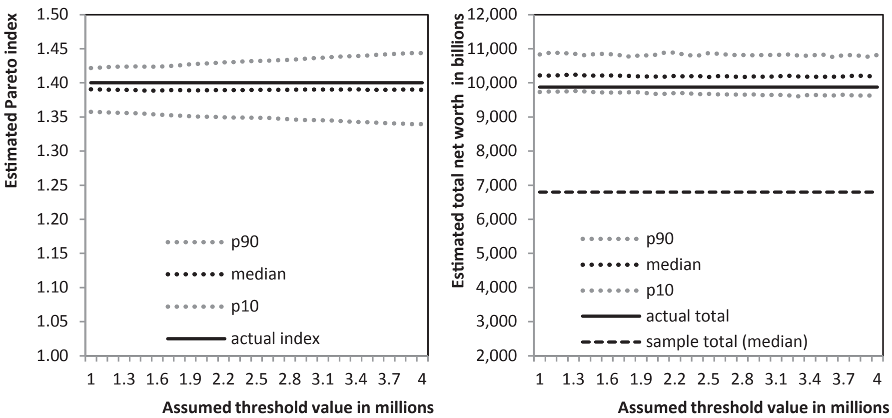


Figure 3.19: Specification 4b. Informative weights, REG estimator plus rich list: Impact of differential non-response on the maximum likelihood estimates for the Pareto index α and total net worth, plotted as a function of the value assumed for w_m. 1,000 samples each, drawn from test distributions, Eq. 3.8 , see also Fig. 3.6.

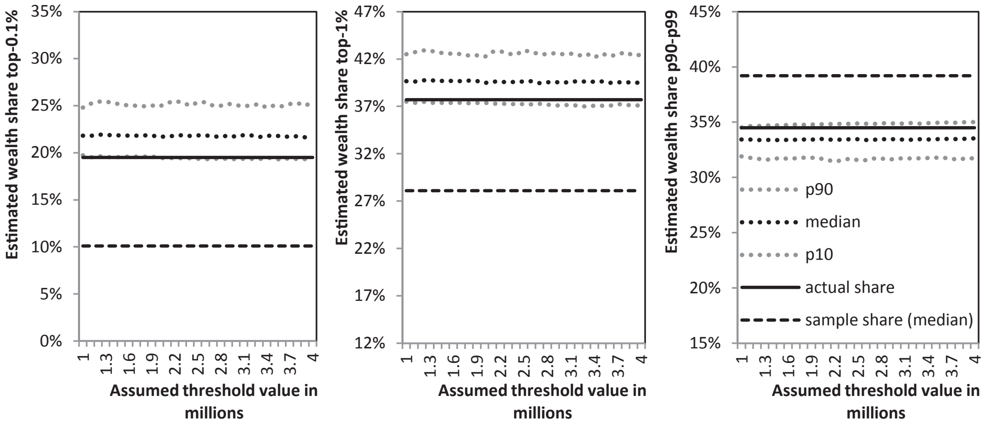


Figure 3.20: Specification 4b. Informative weights, REG estimator plus rich list: Impact of differential non-response on the top wealth shares before and after Pareto correction, plotted as a function of the value assumed for w_m. 1,000 samples each, drawn from test distributions, Eq. 3.8, see also Fig. 3.6.

3.D On the progressivity of non-response rates and the estimation bias

The problems when correcting for the missing rich are also closely related to the question of exactly how many households exceed the value of w_m. This number is unknown if the survey weights are uninformed. This section serves as an illustration for the observation that the overall non-response rates impact top wealth shares and aggregate wealth less than the factor, how quickly the response rates decrease depending on the households' net worth. In order to accomplish this, specification 4 is repeated exactly as in Section 3.2.3, but the assumed non-response mechanism is changed. From mechanisms NR(1) to NR(4) the overall response rates are increased, as indicated by the term independent of wealth in the formulas below. However, in NR(1) the response rates decrease slower than for NR(4), as the quadratic term is smaller.

- **NR(1)** $\text{P(Response)} = 0.4 + 0.02 * \ln(w_i) - 0.0016 * \ln(w_i)^2$
- **NR(2)** $\text{P(Response)} = 0.5 + 0.02 * \ln(w_i) - 0.0020 * \ln(w_i)^2$
- **NR(3)** $\text{P(Response)} = 0.6 + 0.02 * \ln(w_i) - 0.0024 * \ln(w_i)^2$
- **NR(4)** $\text{P(Response)} = 0.7 + 0.02 * \ln(w_i) - 0.0028 * \ln(w_i)^2$

As in specification 4 the resulting Pareto indices and the number of households exceeding the threshold value are used to extrapolate the aggregate private wealth. Figure 3.21 depicts the median relative bias of the estimates as compared to the population's target value of roughly 9.9 trillion. While the overall response rates in NR(1) are much lower than in NR(4), the underestimation of aggregate wealth is more severe in the latter case. As mentioned above, the driving mechanism is not the Pareto index itself (much less the threshold value), but the number of households an empiricist assumes to be exceeding the threshold value. This number is the lower the quicker the response rates increase.

Specifications 4c: Various non-response generating mechanisms. Weighted regression estimation of Pareto index α as a function of w_m, including rich list data. Assumption: differential non-response, underlying response mechanism unknown.

The biased aggregate wealth after correction directly translates to a bias when computing the top wealth shares (Figure 3.22). The top wealth shares are less overestimated under NR(1) and seem to somewhat increase with steeper non-response functions. However, since the top wealth shares are a relative measure having the aggregate wealth in the denominator, there is less variation than for the aggregate wealth for different non-response

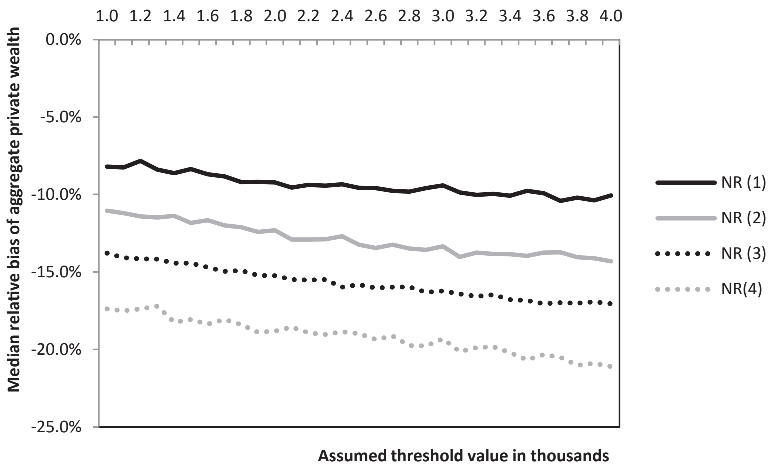


Figure 3.21: Specification 4c. REG estimator plus rich list: Impact of differential non-response on the maximum likelihood estimates for the Pareto index α and total net worth, plotted as a function of the value assumed for w_m. 1,000 samples each, drawn from test distributions, Eq. 3.8 , see also Fig. 3.6.

assumptions. As explained in Section 3.2.3, in this specification are two biases, which somewhat counter each other: (1) the inequality at the top is overestimated, as the Pareto index is too low; but (2) the number of Pareto-distributed households is too low. The effect on the aggregate wealth is a downward bias; the effect on the top wealth shares is an upward bias.

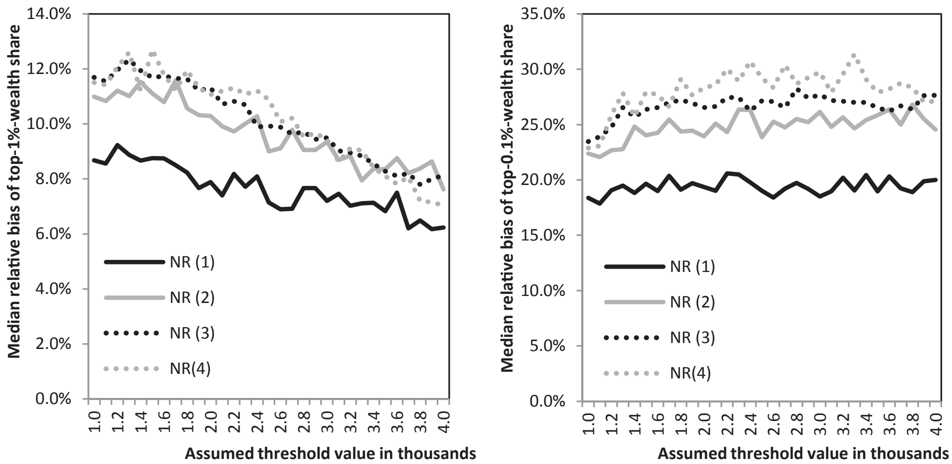


Figure 3.22: Specification 4c. REG estimator plus rich list: Impact of differential non-response on the top wealth shares, plotted as a function of the value assumed for w_m. 1,000 samples each, drawn from test distributions, Eq. 3.8 , see also Fig. 3.6.

3.E Replication of Specification 4: Are the patterns changing for varying Pareto indices or threshold parameters?

In this last section of this appendix, specification 4 is repeated once more, but this time with varying Pareto indices and threshold parameters. For an explanation of the simulation we refer to Section 3.2.2, in Figures 3.23 and 3.24 the parameter $w_m = 1{,}000{,}000$ is fixed and the Pareto index is varying between 1.3 and 1.6. In Figures 3.25 and 3.26 the Parameter $\alpha = 1.4$ is fixed and the threshold value w_m varies between 0.75 and 1.75 million (see also Eq. 3.8, Fig. 3.6).

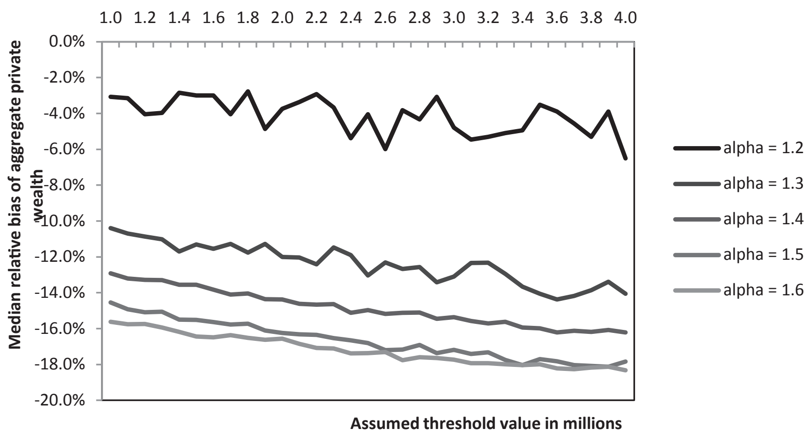


Figure 3.23: Impact of Pareto index α. Impact of differential non-response on Pareto-corrected aggregate private wealth (regression including rich list data) for various values of Pareto index, plotted as a function of the value assumed for w_m. 1 000 samples each, drawn from test distributions, Eq. 3.8 , see also Fig. 3.6.

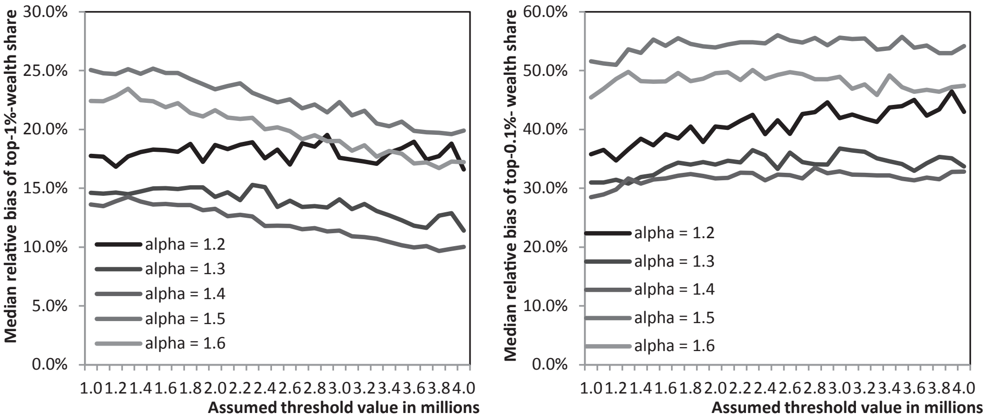


Figure 3.24: Impact of Pareto index α. Impact of differential non-response on Pareto-corrected top wealth shares (regression including rich list data) for various values of Pareto index, plotted as a function of the value assumed for w_m. 1 000 samples each, drawn from test distributions, Eq. 3.8 , see also Fig. 3.6.

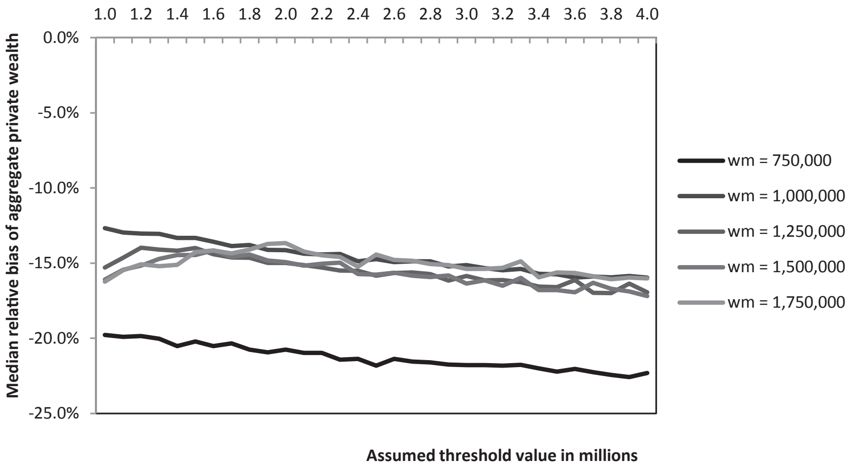


Figure 3.25: Impact of (population) threshold value w_m. Impact of differential non-response on Pareto-corrected aggregate private wealth (regression including rich list data) for various values of w_m, plotted as a function of the value assumed for w_m. 1 000 samples each, drawn from test distributions, Eq. 3.8 , see also Fig. 3.6.

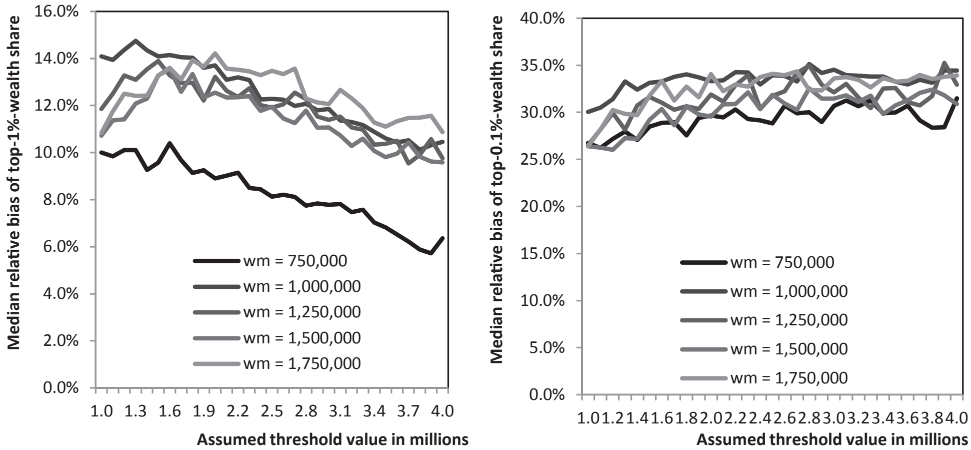


Figure 3.26: Impact of (population) threshold value w_m. Impact of differential non-response on Pareto-corrected top wealth shares (regression including rich list data) for various values of w_m, plotted as a function of the value assumed for w_m. 1 000 samples each, drawn from test distributions, Eq. 3.8 , see also Fig. 3.6.

4 Breaking down Germany's private wealth into inheritance and personal efforts – A distributional analysis

4.1 Introduction

The achievement principle and the equality of opportunities are two core pillars of a modern, liberal social order. Accordingly, one's position in the distribution of wealth is to be determined by the personal efforts and should not reflect the privilege of birth. Whether, and to what extent, the distribution of assets really reflects citizen's economic efforts remains debated. Some argue that the heavy concentration of wealth is the result of the unequal distribution of inheritances and gifts–and, therefore, criticize the dominance of wealth without effort. On the other hand, some praise those with large fortunes as capable 'self-made men' and dismiss criticisms of wealth inequality as being rooted in envy.[21] Accordingly, the ongoing media sensation surrounding Thomas Piketty's 2014 book 'Capital in the 21st Century' is motivated by his prediction that our society is heading for a new 'rentier' society, in which wealth will be concentrated in the hands of a few beneficiaries.

The aim of this study is to shed light on the importance of inheritances and gifts in Germany by means of empirical analysis, and to fuel the debate on the legitimation of

21 A broader analysis of this discourse–to what extent inheritances are consistent with the achievement principle and which normative assumptions shape the German inheritance legislation and taxation–can be found in Beckert (2008, 2013).

wealth inequality with new facts: we hope that empirically robust results de-emotionalize the ongoing discourse. Furthermore, our results might prove useful due to their implications for fiscal policy. A re-introduction of a wealth tax in Germany, for instance, surely depends on the assessment of the roots of existing fortunes.

We consider the following problems: What share of the observed private wealth of the German population is attributable to the receipt of inheritances and gifts, and, on the other hand, what share is attributable to the savings off of the income generated through personal efforts? We divide households into separate groups, which are determined by their current net worth position, and compute the relative importance of inherited wealth along the distribution of wealth. For the sake of simplicity, henceforth we use 'inheritance' synonymously to 'inheritance and gift.'

For Germany, this is the first time that such an assessment of inheritances is conducted.[22] This stems mainly from a lack of appropriate data. However, starting in 2010, the German Panel on Household Finances (PHF) included questions that make this analysis possible. We in underline this in Section 4.2, where we describe the pitfalls of alternative data. The PHF survey is commissioned and prepared for analysis by the German Federal Bank (*Deutsche Bundesbank*). A description of the data set can be found in Section 4.3. Subsequently, we present the descriptive results concerning the distribution of wealth and inheritances, as projected by the survey, in the same section.

Another reason for this study is that recently a new and effective conceptual classification of the term 'inherited wealth as a percent of net worth' was introduced. Traditionally, the methodology introduced by Kotlikoff and Summers (1981) and Kotlikoff (1988) was opposed by Modigliani (1986, 1988). However, neither concept provided researchers with a satisfying formula to compute inherited wealth as a percent of net worth–and the methods yield wildly different results. Thus, our analysis relies on a new approach by Piketty et al. (2014), which we deem superior to the prior approaches. In Section 4.4 we explain the

22 Albeit Piketty and Zucman (2015) offer a macroeconomic assessment of the share of inherited wealth on the aggregate private wealth in Germany. See Section 4.7.

conceptual differences in more detail.

Sections 4.5 and 4.6 constitute the core of our contribution. We find that roughly one-third of the current aggregate private wealth is attributable to inheritances and gifts, whereas two-thirds are the result of self-generated savings. Moreover, these findings hardly vary along the distribution of wealth. In particular, our analysis rejects the hypothesis that the relative importance of inherited wealth systematically increases with household net worth. Admittedly, the data do not cover millionaires in the three-digit area (or higher), and therefore they are not suited to draw conclusions for ultra-high-net-worth-individuals' fortunes.[23] We offer an alternative way to attribute their wealth in Section 4.6: the combined results from the PHF data and other sources may give some useful hints. In Section 4.7 we compare our results to various comparable studies. The concluding Section 4.8 highlights the implications for fiscal policy.

4.2 Data sources: Wealth and wealth transfers in Germany

4.2.1 Wealth

As long as individual wealth taxes were raised in Germany–until 1997–wealth tax statistics provided information on the distribution of wealth. Nonetheless, high and highly diversified tax allowances as well as biased valuation methods never allowed for distributional analyses without constraints.[24]

With the launch of an income and consumption sample (*Einkommens- und Verbrauchstichprobe*; EVS), which is surveyed in five-year intervals, researchers have access to micro data on private wealth from 1978 onwards. A major drawback is that land and real property

23 A study focusing on high-net-worth-individuals by Wealth-X (2014) (more than 30 million Dollars net worth) concludes that the share of German multimillionaires, who gained their fortunes exclusively through inherited wealth, is 28 %, which is particularly high in an international perspective. Furthermore, 31 % build their wealth through both own efforts and inheritance, 41 % are 'self-made men.' As the study draws from a proprietary data base, its sources and methods are not verifiable.

24 Due to high tax allowances after World War II, only two percent of the private households were subject to taxation and recorded in the statistics. Assets were recorded in several asset classes that are not refined enough to correct for their different valuations *ex post* (Bönke et al., 2015).

was valuated with a uniform price (*Einheitswert*) instead of market values, resulting in limited comparability until EVS 1993 (Statistisches Bundesamt, 2014). Figure 4.1 depicts aggregate private wealth in Germany according to the EVS. It also depicts extrapolations and results from alternative sources discussed in this section.

From 1991 onwards, a time series on aggregate private wealth is published jointly by the German Federal Statistical Office and the German Federal Bank (Statistisches Bundesamt, 2013). However, it combines private households with the private non-profit sector. Another drawback is the valuation of property values with replacement values, which frequently deviate from the market values.

In its 2002, 2007 and 2012 survey waves, the Socio-Economic Panel (SOEP) study collected micro data on individual's wealth situation. However, some assets, such as household effects including vehicles, are not recorded. Other assets are difficult to measure in a survey context. For instance, it is challenging for respondents to state the market value for their real property, particularly if the acquisition was a long time ago. The SOEP survey providers impute missing wealth information.[25]

4.2.2 Inheritances

Statistical recording of inheritances and gifts used to be particularly difficult in Germany. As wealth transfers are subject to taxation, the tax statistics records the aggregate of taxable wealth (*Reinnachlass*). However, the taxable wealth only accounts for a mere fraction of the actually transferred assets, as their valuation differs substantially from market values, and only inheritances exceeding a certain amount–corresponding to generous tax allowances–are subject to taxation in the first place (Bartels and Bönke, 2015).

The SOEP included special surveys on intergenerational transfers in 1988 and 2001; the results can be found in Kohli et al. (2006). From 2001 onwards the SOEP only records the receipt of inheritances and gifts if they happened during the preceding year, leading to a

25 Grabka and Westermeier (2014) offer a detailed description of the SOEP wealth module and compare its coverage with the national accounts. For a more thorough review of the distribution of wealth in Germany according to the SOEP study see Frick et al. (2010a).

situation that inhibits the computation of an aggregate inheritance value for any samples drawn after 2001.

The deficiencies of both the inheritance tax statistics and the inheritance data provided by the SOEP survey led researchers to try and estimate the numbers using different sources and methods. Their results, however, diverge significantly: Bach et al. (2014b) estimate the current annual transfer volume to be €64 billion, an estimate by Beckert (2013) amounts to €100 billion, an older estimate is for €150–€200 billion (Beckert, 2008), Schinke (2013) gives an estimate of €220 billion, and Braun et al. (2011) estimate €300 billion.[26]

Figure 4.1 depicts aggregate wealth and inheritance volume according to the data sources discussed above. It is striking that not only do the absolute levels depend on the respective data source, the trends are also different.

4.3 Wealth and inheritance in the PHF study

The data situation for micro data covering wealth and wealth transfers in Germany was dissatisfying all around, however, a pan-European initiative lead to the introduction of a new panel study focusing on household financial situations. Germany was part of the initiative and provided the data set Panel on Household Finances (PHF), supervised by the German Federal Bank, as part of the Eurosystem Household Finance and Consumption Survey (HFCS). The HFCS is a network of euro-area countries and aims to harmonize and improve the data situation on private households' finances in the whole euro-area (European Central Bank, 2013a,b). Within the HFCS questionnaire, the household's current wealth position is a core pillar. Moreover, in a separate module the survey collects information on inheritances and gifts the households received until the time of the survey. Thus, at a household's reporting date both its current net worth and all wealth transfers

26 Bach et al. (2014b) are pulling from both the SOEP and the inheritance tax statistics, both sources yield comparatively low estimates. In contrast, studies drawing from national accounts and mortality tables for their estimates find the transfer volume to be much larger (Schinke, 2013; as well as Braun et al., 2011).

that entered a household's asset portfolio can be compared.

4.3.1 Wealth

In an attempt to determine the data quality of the PHF survey, we compare its aggregate values with the Federal Bank's national accounts. The gross financial assets accumulate to €3,131 billion according to PHF data, the official statistics record a volume of €4,370 billion as of December 2010. This means that the coverage of generally harder to collect information on financial assets is quite high (72 %). The gross property assets exhibit a coverage of 95 %. For the aggregate net worth it is 89 % (Deutsche Bundesbank, 2013).

In the PHF a household's asset portfolio is surveyed in much more detail than in the SOEP. For instance, it includes tangible assets, such as vehicles. The average per capita net worth amounts to €95,000 for the reporting years 2010/2011 in the PHF, while the SOEP study reports €86,000 for 2012 (Grabka and Westermeier, 2014).

Table 4.1 shows the distribution of wealth along the net worth deciles according to the PHF study, separately for East and West Germany, on the household level. The net worth is very unequally distributed: In West Germany the median household's net worth corresponds to roughly 43 % of the mean net worth. In East Germany the inequality is even more pronounced: median net worth corresponds to 22 % of the mean net worth. The regional discrepancy is considerable: the average East German household holds only about 27 % of the assets of the average West German household.

Note that the PHF survey suffers from the same drawbacks as any other surveys collecting wealth micro data, albeit it greatly improved the richness of detail compared to all other data sources for Germany: this implies a bias to the middle class, which generally is overrepresented in surveys in comparison to the distributions upper or lower bounds. The PHF's survey design tries to compensate for this with an oversample for wealthy households. However, it is noteworthy that the non-response rates of households is unusually high. Out of a gross sample of over 20,000 households, only 3,565 were realized (von Kalckreuth et al., 2012, p. 8). Moreover, the refusal by respondents to answer

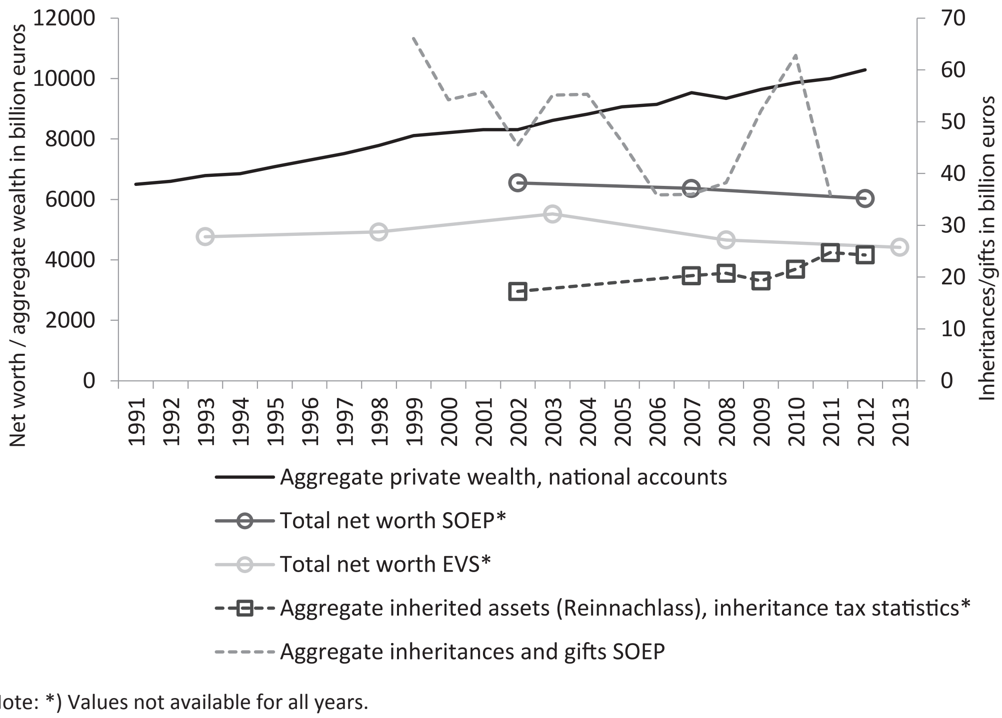


Figure 4.1: Wealth, inheritances and inheritance tax statistics 1991–2013.

questions regarding their financial situation poses a problem, as the possible selectivity of non-response may have critical impact on the distribution of wealth. In the PHF, missing values are treated with multiple imputation (Eisele and Zhu, 2013). The adjusted computations of the standard errors based on multiply imputed data reflect the uncertainty of this procedure. They are calculated using bootstrapping and the provided replicate weights by the PHF.

For our purposes, the fragmentary sampling of households in the top percentile proves problematic. Utilizing a combination of PHF and exogenous sources, Vermeulen (2014) estimates that the wealth share of the top-1% is about 32-33 percent.[27] Using only raw data, the top wealth share in the PHF is a mere 24 percent. There seems to be a substantial underestimation of the top percentile's fortune in the PHF.

27 Alternatively, Bach et al. (2014a), combining SOEP data with *Manager-Magazin*'s list of German multimillionaires, conclude that the wealth share is 36 percent. Westermeier and Grabka (2015) arrive at a share of 31-34 percent based on the 2012 SOEP wave and the *Forbes* list of billionaires.

Table 4.1: The distribution of household net worth in East and West Germany by wealth deciles.

Net worth decile	Aggregate wealth in billion euros	(std. err.)	Mean value in €	(std. err.)	Percentiles in €
East Germany*					
1st - 5th	13.5	(3.4)	3,647	(882)	
6th	21.6	(3.5)	29,262	(775)	23,520
7th	30.3	(4.9)	40,969	(980)	34,640
8th	52.8	(7.8)	71,546	(2,731)	48,828
9th	98.9	(20.2)	133,777	(6,193)	94,250
10th	267.3	(59.4)	364,314	(31,704)	202,344
Total	484.0	(64.6)	65,500	(8,048)	
Top-5%	176.1	(51.2)	493,775	(47,899)	303,980
Top-1%	49.2	(26.0)	897,964	(214,076)	525,400
West Germany*					
1st - 5th	300.4	(22.0)	20,924	(1,346)	
6th	323.5	(33.4)	113,270	(1,399)	88,820
7th	517.5	(42.0)	180,732	(2,880)	140,820
8th	761.1	(64.7)	265,676	(2,468)	220,619
9th	1136.8	(91.3)	397,429	(3,907)	321,370
10th	3949.6	(495.0)	1,381,009	(145,593)	511,419
Total	6992.0	(476.0)	243,859	(16,561)	
Top-5%	3069.8	(471.0)	2,152,300	(279,031)	766,390
Top-1%	1618.6	(452.0)	5,778,196	(1,031,929)	2,567,874

Source: Panel on Household Finances 2013, own calculations.
*) Definition: Head of the household was living in East/West Germany in 1989.
Note that standard errors are calculated using bootstrapping and PHF replicate weights.

The ranking of the wealthiest Germans, as published by the German *Manager-Magazin*, provides further indication of the missing top asset holders from the PHF sample. According to the October 2010 issue, the 500 richest Germans hold assets worth a total of €455.5 billion. The most wealthy German holds a fortune worth €17.1 billion, whereas at the bottom of the list are about 45 families–each having about €200 million. In contrast, the highest recorded net worth in the PHF study is worth roughly €76 million.[28] Yet, although PHF data does not cover ultra-rich households, it is the best existing source on the joint distribution of wealth and wealth transfers.

28 The richest individual in the 2012 SOEP study has assets totaling around €45 million.

4.3.2 Inheritances

The PHF only records inheritances and gifts if they were bequeathed from an individual not living within the same household. Thus, the calculations primarily–but not exclusively–cover intergenerational transfers of wealth. From a macroeconomic or fiscal point of view, the significant amount of transfers within private households, for instance between spouses, are not subject of this analysis.

According to the PHF, about one-third of the households living in Germany received at least one wealth transfer from outside the household of any sort before the day of the interview. As Figure 4.2 depicts, net worth is correlated with the claim of an inheritance or gift. With regard to West Germany, in the bottom half of German's wealth distribution one-fifth of the households record a wealth transfer, while in the richest decile two-thirds of the households inherited some or all of their assets.

The east-west gap observed for household net worth also spreads to inheritances: not only is the frequency much higher in West Germany than in East Germany, the level is higher as well. The aggregate volume of transfers amounts to €1,622 billion (nominal value) in households where the head of the household was living in West Germany in 1989. In contrast, if the head was living in East Germany, the transfers amount to €139 billion. The average transfer received in East Germany is half of the West German average; against the median, transfers received in East Germany are worth one-fourth of the West German transfers.

The PHF also surveys the year in which an inheritance was received. Hence, the real volume of inheritances and gifts in 2010 euros can be determined. Figure 4.3 depicts the volume of inheritances and gifts since 1951 as well as number of cases as a five-year average value. Over the course of the last 60 years, the volume of transfers received continuously increased and reaches a maximum in the 2006-2010 time period, with a volume of €120 billion. In earlier times, the transferred volume was significantly lower.

The underrepresentation of high-net-worth-households mentioned above has the potential to bias both the aggregate inheritance volume and share of households that received a

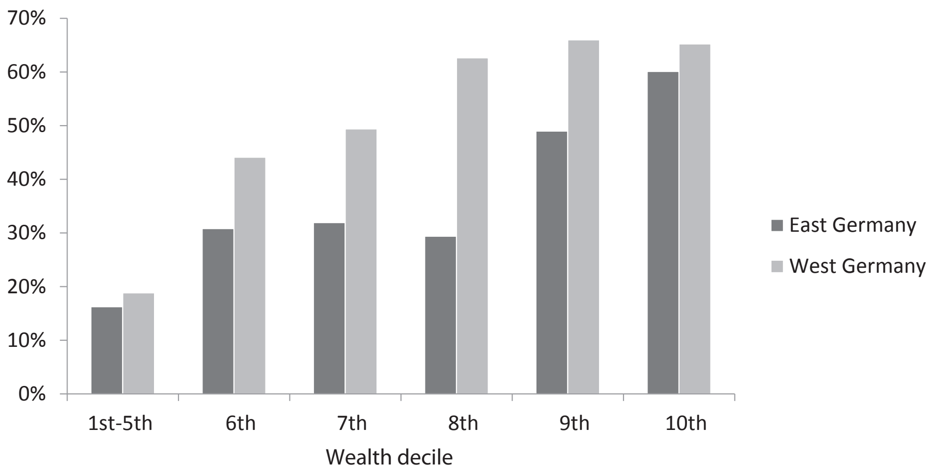


Figure 4.2: Households and inheritances received by wealth deciles, PHF data.

transfer in the upper class: it might be even higher in reality. In comparison to the record of wealth assets, non-response and implausible values prove to be less problematic here (Eisele and Zhu, 2013; Le Blanc, 2014). More information on the questionnaire and data quality are provided in Appendix 4.D.

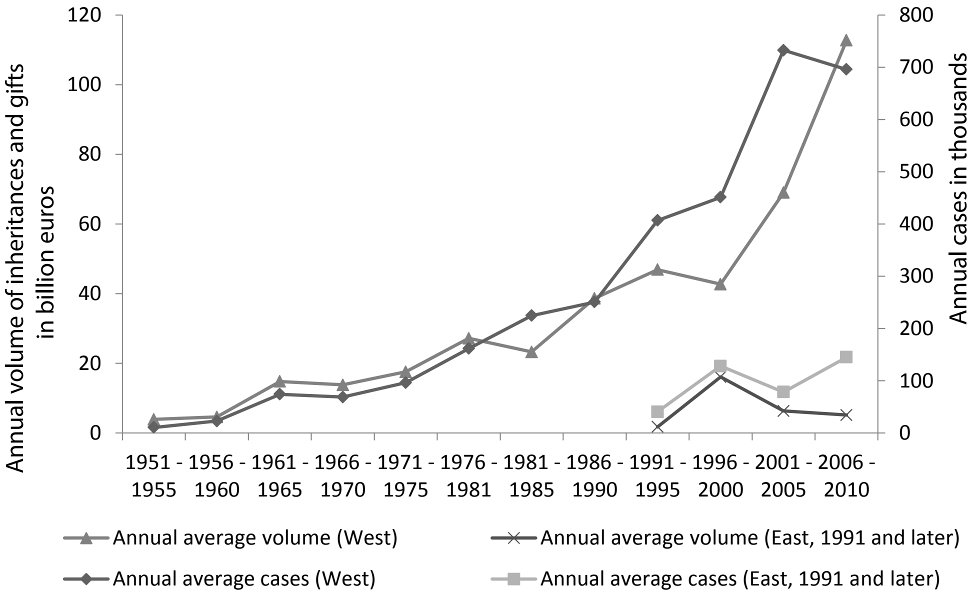


Figure 4.3: Volumes and case numbers of inheritances received in the 2010/11 PHF sample, in 2010 prices.

4.4 Definitions

It is our aim to calculate the value of inheritances received as a percent of net worth, broken down into households' position in the distribution of wealth. To properly define this ratio, we apply a micro-economic concept, which is immediately resulting from the intertemporal budget constraint of households, and is explained in more detail in Piketty et al. (2014). The theoretical framework is the text book model of a multi-periodic market economy with perfectly protected property rights, which can be traded under perfect competition without transaction costs at any time.[29] We also assume that all households hold a non-negative asset portfolio. If we observe any households net worth and inheritance history jointly at any point in time, we may divide those households in two groups. One group holds assets totaling at a higher value than the value of their capitalized inheritances. The intertemporal budget constraint implies that the difference between net worth and

29 In reality, search costs, differing debit and credit interests, uncertain capital income, and asset losses due to theft, accidents and force majeure, for instance, play a role. We neglect these in our definition, as they are not recorded in the data.

capitalized inheritances is accumulated through a household's own savings. The household has net savings, i.e. until the day of the interview the persons within this households consumed less than they would have been able to from their labor and transfer income.[30] Piketty et al. (2014) defines these households as *savers*. The inherited wealth as a percent of net worth is therefore given by b^*/w, with b^* defined as the capitalized inheritances and with w as the household's net worth. In contrast to the *savers*, Piketty et al. (2014) defines *rentiers* as households with net assets lower than their capitalized value of all inheritances. Here, the intertemporal budget constraint implies that this group of people consumed more than they would have been able from their labor and transfer income alone. As *rentiers* did not accumulate any wealth through their own efforts, their net worth is capped at 100 % by the not yet consumed part of their inheritance.

The inheritances received as a percent of net worth is now well-defined for each single household. By aggregating the values, we derive the ratio for various groups or the whole population. For instance, the inheritance volume as a percent of aggregate private wealth for the whole economy is defined as

$$\beta = \frac{\sum \min(b^*, w)}{W}, \tag{4.1}$$

by summation over all households; W is the aggregate net worth of the economy.

4.5 Personal efforts versus inheritance

What is the role of inherited wealth for private households' wealth in Germany? In order to be able to divide the households into *savers* and *rentiers*, according to our definitions, we must capitalize all observed inheritances and gifts. We capitalize past wealth transfers utilizing the yearly-averaged nominal interest rate on German government bonds (Deutsche

30 We define consumption as the difference between income and savings henceforth. Note that the difference might not necessarily result from consumption; households can invest in human capital, donate for charitable purposes, or pass wealth to the next generation.

Bundesbank, 2014) as it is a proper approximation to a risk-free capital market interest rate.[31]

For the time before re-unification, serious valuation issues prevent us from accounting for inheritances and gifts in East Germany; hence, we concentrate on West Germany. In addition, as the time series of government bond yields only date back to 1949, two cases are not readily covered by our method of capitalizing inheritances and gifts.[32]

Table 4.2 shows the result of dividing households into *savers* and *rentiers*. We find that more than 80 % are *savers*, i.e. their net worth is higher than their capitalized inheritances.

Rentier households are spread rather equally along the distribution of wealth. This might seem like an remarkable observation at first, as the frequency of wealth transfers increases with the level of wealth (Figure 4.1). However, households in the bottom half of the distribution barely hold assets, resulting in the consumption of their inheritance in order to maintain their standard of living. Consequently, they qualify as *rentiers*. Anyhow, by our definition Table 4.1 overstates the share of *rentiers*, as households with a negative net worth are classified as such, even though they never received an inheritance.[33]

We now concentrate on the upper half of the distribution of wealth. Next to the disclosure of our results by net worth deciles, and in accordance with Piketty et al. (2014), we divide the richer 50 % of households into *middle class*, *upper middle class* and *upper class*. The *middle class* is made up of households positioned in deciles 6 through 9; the *upper middle class* is formed by households from wealth percentiles 90 through 99; for the *upper class* only the remaining richest one percent of households qualify. Thus, to qualify

31 For the sake of simplicity, the taxation of capital income is neglected. In the literature we find the common pattern to capitalize with a real interest rate of 3 percent. Our main results remain the same, if we adopt this approach (see Table 4.9 in Appendix 4.B). Note that we capitalize all transfers independently of the specifics of the asset portfolios, i.e. real property, shares or tangible assets. In a well-run market economy heirs have the option to trade their asset portfolio for risk-free investments and benefit from the risk-free interest rate.

32 Both cases report the receipt of owner-occupied property, which coincidentally gives us the opportunity to valuate the cases with their current market value as surveyed by the questionnaire. Thus, we put the value of the inheritance received on level with the home's market value.

33 Excluding the group of households with negative net worth from the group of *rentiers* results in a decrease of the respective share of 8.85 %–albeit without increasing the share of *savers* as households in debt are not covered by the definitions in Piketty et al. (2014). This weakness regarding net borrowers extends to Tables 4.4 and 4.5 as well as Table 4.9 in the appendix.

as *upper class* a household needs at least €2.5 million in net assets resulting in a group of about 290,000 households. It is striking that the frequency of *rentier* households decreases dramatically in the upper class as compared to the rest of the population, we return to this observation later.

First, for the whole population, the inherited wealth as a percent of net worth is roughly 34 %.[34] Consequently, more than two-thirds of German's aggregate private wealth is accumulated through the households' own resources. This ratio barely varies for different positions across the distribution of wealth. In the bottom half, net borrowers drag down the observed aggregate net worth. The first two wealth deciles consist almost exclusively of net borrowers or zero-wealth households, while in wealth deciles 3, 4 and 5, the role of inheritance is negligible. In the middle class, 37 % of private wealth is inherited. We do not find a significant difference to the upper middle class. The ratio is lower for the upper class: 27 % of the aggregate private wealth is attributable to past inheritances and gifts.

Note that the average values we provide hide a lot of the heterogeneity within groups. The richest decile combines households that exclusively build their wealth from transfers (roughly one-fifth of the households) with households that never received any kind of wealth transfer (one-third of the households).

Additionally, Table 4.3 shows the capitalized inheritances as a percent of net worth for the richest households sorted by their net worth levels. Media attention frequently seems to revolve around the group of millionaires; they also are a reference group for fiscal policy. We find that 31 % of all assets in the group of households between €1 and €3 million is the result of a wealth transfer. Households holding more than €3 million in net assets exhibit about the same ratio (29 %).

Ultimately, our results show that the pronounced inequality of the distribution of wealth in Germany might not be the result of a dominance of 'effortless wealth.' However, this result is restricted by the fact that our data do not cover the group of ultra-rich households.

As the distribution of wealth is strongly correlated with age, we repeat the analysis

34 If we exclude households without inheritances from the analysis, the ratio is 55 %.

Table 4.2: Inherited wealth as a percent of net worth, share of *rentiers* on population, in %.

	Share of *rentiers*		Inheritance-wealth ratio	
	Share	(Std. err.)	Ratio	(std. err.)
1st - 5th	19.40	(1.94)	17.21	(2.77)
6th	22.40	(5.43)	31.93	(5.13)
7th	19.07	(4.09)	34.75	(4.42)
8th	21.33	(3.50)	40.05	(3.06)
9th	19.50	(3.99)	38.30	(3.47)
10th	17.55	(3.16)	32.57	(4.89)
Total	19.68	(1.31)	33.76	(3.05)
Middle class	20.57	(1.86)	37.32	(2.19)
Upper middle class	19.23	(3.43)	36.18	(3.27)
Upper class	2.09	(1.78)	27.03	(10.82)

Source: Panel on Household Finances 2013, own calculations.
*) Definition: Head of the household was living in East/West Germany in 1989. Middle class: deciles 6–9, upper middle class: percentile 90–99, upper class: top-1% percentile.
Note that standard errors are calculated using bootstrapping and PHF replicate weights. Capitalized inheritances using yearly averages of long-term yields on government bonds (Deutsche Bundesbank, 2014).

Table 4.3: Inherited wealth as a percent of net worth by net worth in %.

Household net worth in €	Inheritance-wealth ratio	
	Ratio	(std. err.)
500,000 or under	35.42	(1.99)
500,000–999,999	38.74	(4.50)
1,000,000–2,999,999	30.79	(5.15)
3,000,000 or over	28.56	(11.57)

Source: Panel on Household Finances 2013, own calculations.
*) Definition: Head of the household was living in East/West Germany in 1989.
Note that standard errors are calculated using bootstrapping and PHF replicate weights. Capitalized inheritances using yearly averages of long-term yields on government bonds (Deutsche Bundesbank, 2014).

separately for age groups. Figure 4.4 depicts the results graphically. We need to take the ‘age of the household’s head’ in order to classify households by age groups. The inheritances received as a percent of net worth are U-shaped: Inheritances are relatively more important for young and old households, whereas middle-aged households are the least dependent upon intergenerational wealth transfers. For younger households, this observation stems from their inability to already have accumulated wealth on their own, instead they depend on consumer credits and student loans, resulting in low levels of net worth, whereas many young households already received transfers, particularly gifts, from older generations. Older households also exhibit low levels of net worth, they are already actively transferring wealth to the next generation.[35]

Figure 4.4 also proves that the volume of capitalized inheritances depends much less on household age than the volume of net worth. There are two contradictory effects at play: on the one hand, the ratio of households with wealth transfers over all households increases with age, on the other hand the value of inheritances received per case is lower

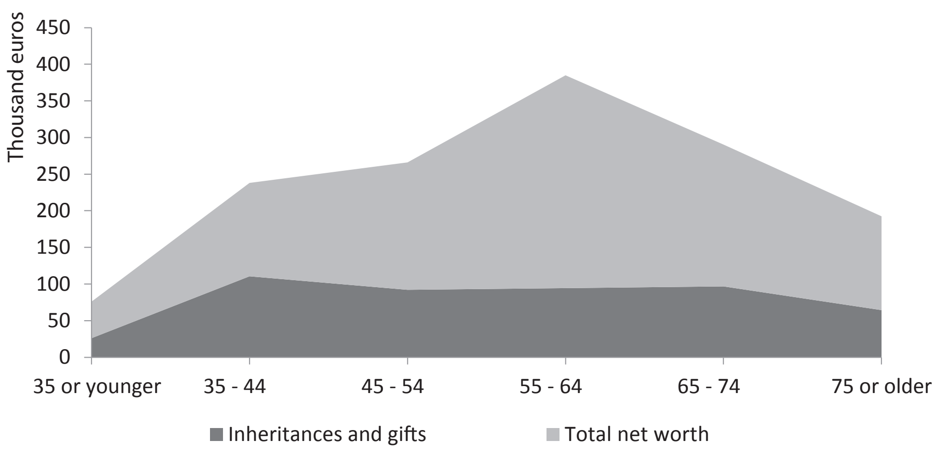


Figure 4.4: Mean capitalized inheritances and net worth by age class.
Source: Panel on Household Finances 2013, own calculations. Capitalized inheritances using annual average yields on long-term government bonds (Deutsche Bundesbank, 2014).

35 Leopold and Schneider (2010) investigate whether households are planning wealth transfers in advance and when heirs might want to expect them.

for older households, as older households are made up of the descendants of a much poorer society.[36]

The number of observations in the PHF is not sufficiently high enough to warrant analyses of the role of inheritances for different birth cohorts and wealth deciles simultaneously. However, larger age groups surely can be analyzed with greater scrutiny. Particularly interesting are households for which the division into *savers* and *rentiers* proves to be the most telling. Table 4.4 sums up the results for all households with the household's head older than 65. The pattern barely deviates from the results considering the whole population in Table 4.2. The inheritances received as a percent of net worth remains at one-third. At the top of the distribution, this value decreases similarly to the decrease in the highest percentile of the overall population.

36 In West Germany, between 1960 and 1990, the pre-unification era, the ratio of aggregate private wealth over national income continually increased. At the time of the re-unification it was twice as high as in 1961. The increase continued in re-unified Germany. Simultaneously, life-time incomes of the cohorts increased rapidly. Within one generation–from the 1935 cohort to the 1965 cohort–real life-time incomes increased by two-thirds. For the ratio of wealth over national income, we refer to Piketty and Zucman (2015); the evolution of life-time incomes is documented in Bönke et al. (2015) and Corneo (2015).

Table 4.4: Inherited wealth as a percent of net worth, share of *rentiers* on population for the cohort 65+, West Germany, in %.

	Share of *rentiers*		Inheritance-wealth ratio	
	Share	(Std. err.)	Ratio	(std. err.)
1st - 5th	15.40	(3.04)	30.30	(5.83)
6th	28.00	(10.12)	35.71	(8.81)
7th	21.93	(7.83)	48.79	(7.59)
8th	26.94	(9.58)	43.28	(8.26)
9th	10.67	(4.25)	25.98	(7.05)
10th	22.89	(6.72)	30.09	(5.67)
Total	18.77	(2.27)	33.31	(3.33)
Middle class	21.98	(3.78)	36.78	(4.57)
Upper middle class	24.99	(7.44)	32.19	(6.73)
Upper class	3.66	(3.81)	25.95	(9.94)

Source: Panel on Household Finances 2013, own calculations.
*) Definition: Head of the household was living in East/West Germany in 1989. Middle class: deciles 6–9, upper middle class: percentile 90–99, upper class: top-1% percentile.
Note that standard errors are calculated using bootstrapping and PHF replicate weights. Capitalized inheritances using yearly averages of long-term yields on government bonds (Deutsche Bundesbank, 2014).

4.6 Robustness checks

In this section, we evaluate how changes of the assumptions affect the joint distribution of wealth and inheritances. First, we define wealth differently, adding pension wealth, from the public pension scheme (*Gesetzliche Rentenversicherung* GRV) in particular, to the equation in order to see how the outcomes change for the cohort 65+. Second, we define inheritances and gifts differently, applying the definitions by Modigliani (1986, 1988) and Kotlikoff and Summers (1981)/Kotlikoff (1988). Third, we assess the robustness of the results in the richest percentile with two different approaches: we abandon the assumption of a uniform capitalization for all households and assume that richer households achieve better interest rates. In addition, we trim the PHF data at its 99th percentile and combine those presumably reliable findings with results by Piketty and Zucman (2015) and Vermeulen (2014), leading to an alternative estimate for the *upper class*.

4.6.1 Considering pension wealth

Research on the distribution of wealth regularly discards pension wealth, excluding it from the definition of household net worth, as households must not trade their pension entitlements–from the statutory pension scheme, at the very least–freely on the market. Accordingly, we did not include those entitlements in our calculations. We now apply a broader definition of wealth, including pension rights, and estimate the inheritances received as a percent of net worth anew.[37]

We limit the analysis to households that are led by persons older than 65, as only this group allows us to transparently forecast their pension entitlements. Only for the cohort 65+ is it reasonable to assume that, technically, all persons with pension entitlements are receiving pensions. The PHF surveyed annual pension payments from statutory, company, and private pension schemes. We compute present values as the annual payment combined with the individual life-expectancy (Wolff, 2015).[38] We then add the present value of pension wealth to the household net worth.

Table 4.5 shows the results if we take pension wealth from statutory, company and private pension schemes into consideration. For the sake of comparability, we sort the households by wealth deciles without pension wealth as in Table 4.4. Apparently, and in relative terms, households in the bottom half of the distribution of wealth benefit the most. For this population the value of inheritances received as a percent of net worth drops by 18 percentage points. Moreover, the relative value of pension wealth decreases with the level of household net worth. For the *middle class*, the inheritance-wealth ratio drops by 10 percentage points, for the *upper middle class* it decreases by 7 percentage points, in the *upper class* the impact of pension wealth becomes almost irrelevant (3 percentage points).

37 Frick and Grabka (2010) estimate the average present value–with a real interest rate of 3 percent–for pension entitlements to be €307,000 for civil servants and €125,000 for regular retirees.

38 The present value of pension wealth PW is given by $PV = R((1+i)^N - 1)/i(1+i)^N$; R is the annual pension payment, i the discount rate and N the statistical life-expectancy (Statistisches Bundesamt, 2011).

Table 4.5: Inherited wealth as a percent of net worth, share of *rentiers* on population for the cohort 65+, including pension wealth, West Germany, in %.

	Share of *rentiers*		Inheritance-wealth ratio	
	Share	(Std. err.)	Ratio	(std. err.)
1st - 5th	5.21	(2.27)	12.39	(3.01)
6th	21.30	(9.39)	21.35	(6.97)
7th	15.39	(7.00)	33.24	(5.76)
8th	21.46	(9.00)	30.41	(7.46)
9th	5.64	(2.78)	20.79	(4.96)
10th	15.71	(5.38)	24.73	(4.68)
Total	10.57	(1.88)	22.53	(2.34)
Middle class	16.03	(3.39)	26.40	(3.35)
Upper middle class	17.15	(5.91)	25.41	(5.53)
Upper class	2.53	(3.02)	22.97	(8.76)

Source: Panel on Household Finances 2013, own calculations.
Definitions: Head of the household was living in East/West Germany in 1989. Middle class: deciles 6–9, upper middle class: percentile 90–99, upper class: top-1% percentile.
Note that standard errors are calculated using bootstrapping and PHF replicate weights. Capitalized inheritances using yearly averages of long-term yields on government bonds (Deutsche Bundesbank, 2014).
For the computation of the pension wealth the annual payments from statutory, company and private pension schemes are included, with statutory and civil servants' pension schemes predominating the annual payment flows. We approximate the net present value assuming a real interest rate = 3 %. Results are robust for interest rates between 0 % and 4 %.

Overall, for the cohort already retired, the addition of pension wealth significantly decreases the impact of inheritances on household net worth in the lower wealth deciles. We conclude that, in relative terms, in deciles 7 and 8 inheritances and gifts remain a substantial source of wealth.

4.6.2 Modigliani and Kotlikoff-Summers

So far we defined inheritances received as a percent of household net worth according to Piketty et al. (2014), with the characteristic feature of this definition being that households are divided in two separate groups: *rentiers* and *savers*. The inheritance-wealth ratio depends on a household's balance, whether the household net worth w is higher than the capitalized inheritances b^*. If it is lower, the household consumed more than it would have been able from labor and transfer income alone, i.e. the capitalized inheritance is best approximated by the current household net worth w. Consequently, the ratio of inheritances and gifts over net assets becomes 100 %.

The previous approaches by both Modigliani (1986, 1988) or Kotlikoff and Summers (1981) and Kotlikoff (1988) neglect this distinction and invariably put the accumulated inheritances in the numerator. The two approaches differ by their mode of capitalizing past inheritances though. Modigliani (1986, 1988) does not capitalize the inheritances, he replaces b^* with the real value of the past inheritance b^0, i.e. he merely adjusts for inflation. For the total population, the inheritances received as a percent of net worth are then given by

$$\beta^M = \frac{B^0}{W}. \tag{4.2}$$

This approach mostly is target of criticism for its implication on asset holders, who do not work, but have high capital income due to an inherited fortune. They count as individuals, whose wealth partially results from personal efforts. This happens as their real returns are positive, but they do not entirely consume it, hence, $b^0 > w$ holds.

Therefore, most empirical studies adopt the approach by Kotlikoff and Summers (1981)/Kotlikoff (1988) and capitalize past inheritances. They define inheritances received as a percent of net worth as the ratio of capitalized inheritances and gifts over aggregate private wealth:

$$\beta^{KS} = \frac{B^*}{W}. \tag{4.3}$$

Table 4.6: Inherited wealth as a percent of net worth, according to alternative definitions, in %.

	Kotlikoff and Summers (1981) (Capitalzing with government bonds)		Modigliani (1986) (Adjusting for inflation only)	
	Ratio	(Std. err.)	Ratio	(std. err.)
1st - 5th	71.07	(22.77)	34.40	(8.31)
6th	193.07	(81.44)	59.64	(17.08)
7th	98.89	(34.98)	39.19	(9.17)
8th	92.89	(12.58)	39.26	(4.43)
9th	65.36	(13.61)	30.71	(4.81)
10th	49.97	(8.67)	26.64	(4.65)
Total	68.61	(8.65)	31.54	(3.30)
Middle class	94.95	(13.21)	38.23	(3.68)
Upper middle class	64.41	(11.09)	29.75	(3.67)
Upper class	27.81	(10.93)	21.87	(10.26)

Source: Panel on Household Finances 2013, own calculations.
Definition: Head of the household was living in East/West Germany in 1989. Middle class: deciles 6–9, upper middle class: percentile 90–99, upper class: top-1% percentile.
Note that standard errors are calculated using bootstrapping and PHF replicate weights. Capitalized inheritances using yearly averages of long-term yields on government bonds (Deutsche Bundesbank, 2014).

As long as the real rate of return is positive, $\beta^{KS} > \beta^{M}$ holds. A major implication of this approach by Kotlikoff and Summers is that households, which qualify as *rentiers* according to Piketty et al. (2014), exhibit an inheritance-wealth ratio exceeding 100 %. In theory, this might also apply to the definition by Modigliani. The ratio exceeding 100 % informs researchers on the extent that the households consumed received inheritances and gifts. It is, however, not suited to split the aggregate wealth into inherited wealth and wealth accumulated through households' efforts.

Despite these reservations, we compute the inheritance-wealth ratio following the definitions by Modigliani (1986, 1988) and Kotlikoff and Summers (1981)/Kotlikoff (1988). Household net worth is defined according to our initial approach without pension wealth. As Table 4.6 shows, the approach by Kotlikoff and Summers results in a substantial increase of inheritances as a percent of net worth: for the overall population the ratio is up from 34 percent to roughly 69 percent. In contrast, Modigliani's method results in a slight decrease to 32 percent.

The empirical relationship between inheritances received and household net worth turns out to be more negative in comparison to Piketty et al. (2014). Following the definition by Kotlikoff and Summers (1981), only 5 percent of the observed aggregate private wealth is the result of the households' own efforts, 95 % are attributable to past inheritances.

4.6.3 The role of inheritances for the *upper class*

All the results we present have in common that, in relative terms, past inheritances become less important for the *upper class*; in comparison to the *middle class* the ratio is strikingly small. In this section we critically scrutinize this finding.

In the finance literature, higher net worth typically correlates with higher risk taking. Wealthy individuals are usually financially educated and exhibit less risk-aversion. Higher risk-taking, in return, yields higher returns on investments. Taking this into account, the assumption of a uniform interest rate for the whole population seems unlikely to hold. It is more plausible that wealthy households positioned in the *upper class* achieve higher real interest rates. Whereas a risk-free investment into federal bonds might be representative of the majority, risk-taking in the *upper class* is presumably higher.

How does the finding that inherited wealth is less important for the *upper class* change once we consider riskier investments? To maximize the impact of the effect we now assume that households belonging to the upper class invested 100 % of their past inheritances in broad stock market indices. This means the upper class is capable of realizing an *equity risk premium* (ERP), which puts them above the risk-free interest rate on government bonds.

As we merely try to evaluate the impact, and are not pursuing an exact measurement, we simplify the methods and assume a constant ERP. Mehra (2008, p. 7) offers some viable magnitudes: for Germany he proposes an ERP of 9.1 %. In the time period 1946–2005 the ERP was 7.48 %, with regard to the United States though (Mehra, 2008, p. 8).

Thus, for this robustness check we offer results with an ERP varying between 7 and 9 %. Accordingly, we calculate inheritances received as a percent of net worth anew, for the upper class only, under the assumption that the households realized these ERPs.

Table 4.7 summarizes the results. The inheritance-wealth ratio in the *upper class* increases to 36 percent: in other words, it aligns with the ratio observed in the *middle class*. The comparably small effect of marginal increases of the ERP is explained by capping the ratio at 100 % for individual households. Overall, the observation that the inherited wealth only explains a small portion of *upper class* wealth is not an artifact of a low interest rate for the upper class.

However, this robustness check is still based on the assumption that the upper class is accurately covered by the PHF data. In Section 4.3 we offered plenty of reasons for doubt. That is why we now consider the possible effects of a trimming of the data after the 99th percentile, for which the PHF might actually be representative. In order to carry this analysis out, we need to draw from alternative sources, and determine the result for the *upper class* as a residual.

Once we assume the PHF data covers the highest percentile's population insufficiently, the results provide not only false estimates for their top wealth share. Our findings for

Table 4.7: Inherited wealth as a percent of net worth, including ERP for *upper class*, West Germany, in %.

Upper middle class	Upper class		
Est. in Table 2	Est. in Table 2	ERP	Ratio with ERP
		7 %	35.12
36.18 %	27.03 %	8 %	35.68
		9 %	36.20

Source: Panel on Household Finances 2013, own calculations.
Definitions: Head of the household was living in East/West Germany in 1989. Middle class: deciles 6–9, upper middle class: percentile 90–99, upper class: top-1% percentile.

the overall inherited wealth as a percent of net worth does not hold either, as the highest percentile contributes disproportionally to our calculation. We need exogenous sources for both values. Vermeulen (2014, p. 29) offers an estimate for the top wealth share of the richest one percent of the population. His estimate is based on external sources and arrives at $\alpha = 33\%$, which is 9 percentage points higher than the estimate from the raw PHF data. For the overall ratio of inherited wealth over the aggregate private wealth Piketty and Zucman (2015, online appendix) offer an estimate of $\beta = 51\%$. This is about 17 percentage points higher than our estimate in Table 4.2.

The ratio of inherited wealth over aggregate private wealth in the *upper class*, γ, is then given by

$$\gamma = \frac{1}{\alpha}\left[\beta - \frac{1-\alpha}{1-\hat{\alpha}}(\hat{\beta} - \hat{\gamma}\hat{\alpha})\right]. \tag{4.4}$$

$\hat{\alpha} = 24\%$, $\hat{\beta} = 33.76\%$, and $\hat{\gamma} = 27.03\%$ are estimates based on the PHF data.[39]

Plugging in the numerical values into formula (4.4) yields $\gamma = 81.69\%$. It means, in effect, that the inheritance-wealth ratio in the upper class is higher than four-fifth and, moreover, substantially higher than for the rest of the population. This result suggests that inherited wealth is the driving force behind the asset accumulation of the ultra-rich population.

4.7 Comparing the results with previous studies

In the original paper by Piketty et al. (2014), from which we heavily draw, the significance of inherited wealth for Paris in 1872–1937 was estimated; the authors focused on the upper half of the distribution of wealth. With regard to the *middle class*, our results are strikingly similar to population of Paris one century ago. This refers to both the share of *rentiers* and the significance of inherited wealth. For the *upper middle class*, in 1912 Paris,

39 For the derivation of equation (4.4) see Appendix 4.C.

the inheritance-wealth ratio was much higher than 2011 Germany: 65 %. The respective ratio was 80 % for the *upper class*–very similar to our residual estimate in the previous section.[40]

The only study similar to ours that focuses on the present day and uses similar data and methods is by Wolff and Gittleman (2014) and concentrates on the US. Based on data of the Survey of Consumer Finances (SCF), which is collected in a similar fashion to the PHF, and a uniform real rate of return of 3 %, Wolff and Gittleman (2014) find that the share of inherited wealth tends to decrease for higher levels of household net worth. In 2007, they estimate the US inheritance-wealth ratio to be roughly 15 %. Our results using a similar capitalization of past inheritances are in Appendix 4.B (Table 4.9). They barely differ from the results in Table 4.2.

Our study, and the studies by Piketty et al. (2014) and Wolff and Gittleman (2014), rely on the joint distribution of capitalized past inheritances and net worth, which allows us to estimate the significance of intergenerational transfers for various quantiles of the distribution of wealth. Some studies draw from similar data, however, more often than not similar data are not available, in which case researchers base their studies on theoretical considerations to provide macroeconomic estimates. As mentioned above, Piketty and Zucman (2015) estimate that 51 % of the present German aggregate private wealth is attributable to inherited assets.[41] For France they estimate a somewhat higher share of 55 %.[42] For Sweden, Ohlsson et al. (2014) estimate this ratio to be almost 50 %.[43] With regard to the US, Gale and Scholz (1994) estimate that the lower bound is 51 %. This presents to be a sharp contrast to the estimates by Wolff and Gittleman (2014) mentioned above, as their SCF micro data based estimate is in the 20–25 % area.

With regard to Germany, Kohli et al. (2006) investigate the relationship between wealth

40 The overall share of inheritances received as a percent of aggregate private wealth was more than 70 %; in the upper class the share amounted to more than 60 %.
41 Another estimate by Reil-Held (2004) arrives at 34 %. Reil-Held (2004) also gives a more exhaustive overview of studies covering Germany.
42 In contrast, Kessler and Masson (1989) estimate that inheritance-wealth ratio in France is 35 %.
43 Klevmarken (2004), meanwhile, arrives at a ratio of 19 % for Sweden.

and inheritances using SOEP data. Their evaluation indicates that inheritances and gifts have an equalizing effect on the distribution of wealth, even though a positive correlation between households' net worth position and their inherited assets is observed. However, the authors forgo capitalizing past inheritances and instead combine nominal values from the 2001 SOEP survey with information on households' net worth from the 2002 wave. Thus, we argue the results may hardly be interpreted as an accurate estimate for inherited wealth as a percent of net worth.

4.8 Summary and conclusion

The PHF data suggest that approximately one-third of the German private wealth stems from past intergenerational inheritances and gifts. This finding is stable throughout the distribution of wealth. In particular, the significance of wealth transfer does not increase with higher wealth levels. Our basic scenario with uniform capitalization suggests the opposite: For the *middle class* about 36 % of the net worth is attributable to inheritances and gifts. In the *upper class*, meanwhile, this value merely is 27 %.

For the cohort of retirees, the findings turn out to be similar to the whole population. In particular, the addition of pension wealth reduces the significance of past wealth transfers for the poorer wealth deciles, whereas the pension wealth of the *upper middle class* or the *upper class* reduces the significance only by a few percentage points. We conclude that the consideration of pension wealth decreases the significance of inheritances only for small and medium fortunes. Nonetheless, the importance of inherited wealth is highest for the *middle class*.

We also prove that the findings for the *upper class* are not the result of our assumed uniform interest rate, as we modified this assumption to an investment into risky stock markets with higher returns. Not until an equity risk premium of 9 % would their inherited wealth share align with the value observed for the middle class.

It is noteworthy that our results are based on survey data. Although the PHF markedly improves the data situation with regard to wealth and inheritances in Germany, it still applies that participation is voluntary and false information is not penalized. Overall, the PHF exhibits a rather high non-response rate coming in at 80 %, and the selectivity might affect the representativeness of our results. In particular, we assume that the richest one percent of the population has fundamentally different assets portfolios (primarily valuable business assets), which are not accounted for by our study. A bias to the middle class might also result in an under-representation of particularly poor households; however, as presumably both their inheritance volume and their net worth are low, their absence should have miniscule impact.

Notwithstanding these deficiencies, we believe that our results for the remaining 99 % of the households remain intact and are firmly rooted in the empirical framework. Our analysis finds that inheritances and gifts are not the predominant source of wealth for the vast majority of the German population; moreover, the significance does not increase for higher wealth deciles.

The results for the top percentile of the wealth distribution are substantially less backed by the data, as this population is not sufficiently covered by the PHF data. Assuming that the results for the remaining 99 % are representative, and combining our results with exogenous sources, we arrive at the conclusion that more than 80 % of their wealth is inherited. This is in line with the observation of Piketty et al. (2014) for Paris right before World War I–and a blatantly higher estimate than the PHF data alone suggest.

Despite this discrepancy concerning the richest percent of the population our results contribute to a higher level of objectivity in the debate on wealth inequality in Germany. The highest percentile starts at €2.5 million–and presumably the boundary of representativeness of our data is much higher. We now may assess a hypothetical re-introduction of an individual wealth tax, which more often than not is justified by the argument that it would target wealth that is not earned from the households' own efforts.

Our analysis suggests that this argument is not sustainable by the data: raising a wealth tax today, targeting a population with more than €500,000, for example, would affect some *rentiers*, whose wealth is 100 % inherited–about one-fifth of taxpayers–, it would, however, also affect *savers*, who never received an inheritance or gift–about one-third of the hypothetical taxpayers.

To limit the tax burden to inherited wealth, a re-introduction of an individual wealth tax is the inferior instrument, policy makers are instead advised to revert to an inheritance tax with a broad assessment basis. In this regard, large scale exemptions for business assets are detrimental to the viability, as the richest families are enabled to bequeath their assets free of tax.

Furthermore, the aggregate wealth transfer volume increased rapidly during the last 50 years. According to macroeconomic estimates by Piketty and Zucman (2015), the volume increased from 2 % of the GDP in 1960 continually to 11 % in 2010, and it is bound to increase further. Concurrently, the taxation of inheritances decreased from 3.5 % in 1960 to 1.7 % in 2010. The German inheritance and gift tax not only presents an opportunity to increase equality of opportunity, it also is an opportunity for policy-makers to reduce labor-related taxes. The political support of a target-oriented restructuring of the inheritance and gift tax can be improved, if the general public is made aware of the actual facts.

4.A Sample sizes of the PHF

Table 4.8: Sample sizes in West Germany*

Net worth decile	Number of households
1st - 5th	963
6th	203
7th	288
8th	345
9th	423
10th	607
Top-5%	342
Top-1%	56

Source: Panel on Household Finances 2013.
Definition: Head of the household was living in West Germany in 1989.
Due to weighting and, in particular, oversampling of wealthy households, they are spread unequally across wealth deciles.

4.B Results for real rate of return $r = 3\%$

Table 4.9: Inherited wealth as a percent of net worth, share of *rentiers* on population, in %.

	Share of *rentiers*		Inheritance-wealth ratio	
	Share	(Std. err.)	Ratio	(std. err.)
1st - 5th	19.71	(1.98)	18.25	(2.76)
6th	22.76	(5.50)	33.24	(5.16)
7th	18.54	(4.06)	34.64	(4.40)
8th	21.20	(3.43)	37.78	(3.05)
9th	18.84	(3.90)	38.14	(3.43)
10th	16.17	(2.97)	32.10	(4.80)
Total	19.61	(1.27)	33.34	(2.99)
Middle class	20.33	(1.98)	36.80	(2.76)
Upper middle class	17.72	(3.23)	35.69	(3.18)
Upper class	1.87	(1.75)	26.96	(10.61)

Source: Panel on Household Finances 2013, own calculations.
Definitions: Head of the household was living in East/West Germany in 1989. Middle class: deciles 6–9, upper middle class: percentile 90–99, upper class: top-1% percentile.
Note that standard errors are calculated using bootstrapping and PHF replicate weights. Capitalized inheritances using real interest rate $r = 3\%$.

4.C Proof of ratio (3.4)

Let the actual ratio of inherited wealth over aggregate private wealth be given by $\beta = \frac{B}{W}$, let the actual top wealth share of the richest percentile be given by α. Further on, $\hat{\alpha}$ and $\hat{\beta} = \frac{\hat{B}}{\hat{W}}$ are the respective values estimated from the PHF data. Assuming that the distribution of wealth and inheritances is accurately surveyed by the PHF up the 99th percentile, it holds for the aggregate private wealth that

$$W = \hat{W}(1 - \hat{\alpha}) + \alpha W \text{ or } \hat{W} = \frac{1 - \alpha}{1 - \hat{\alpha}} W. \tag{4.5}$$

Let γ be the inheritance-wealth ratio of the *upper class*, and $\hat{\gamma}$ the respective PHF estimate. The aggregate capitalized inheritances are then given by

$$B = \hat{\beta}\hat{W} - \hat{\gamma}\hat{\alpha}\hat{W} + \gamma\alpha W. \tag{4.6}$$

Rewriting (4.6) and utilizing (4.5) yields the overall inherited wealth as a percent of net worth:

$$\beta = \frac{B}{W} = \frac{\hat{\beta}\hat{W} - \hat{\gamma}\hat{\alpha}\hat{W} + \gamma\alpha W}{W} = \hat{\beta}\frac{1-\alpha}{1-\hat{\alpha}} - \hat{\gamma}\hat{\alpha}\frac{1-\alpha}{1-\hat{\alpha}} + \gamma\alpha. \tag{4.7}$$

From (4.7) it immediately follows that the share of inherited wealth in the *upper class* is given by

$$\gamma = \frac{1}{\alpha}\left(\beta - \hat{\beta}\frac{1-\alpha}{1-\hat{\alpha}} + \hat{\gamma}\hat{\alpha}\frac{1-\alpha}{1-\hat{\alpha}}\right) = \frac{1}{\alpha}\left(\beta - \frac{1-\alpha}{1-\hat{\alpha}}(\hat{\beta} - \hat{\gamma}\hat{\alpha})\right) \tag{4.8}$$

or equation (4.4) as in Section 4.4. QED

4.D Inheritances, data quality and non-response in the PHF

Within the PHF questionnaire, Section 6 contains the module entitled 'Intergenerational transfers / gifts'. First, a filter question determines whether the household received any inheritance or gift before the time of the survey:

- ***6.01*** *(Have you / Have you or another member of your household / Has any member of the household) ever received a substantial gift or inheritance, e.g. money or any other assets, from someone who is not a part of the household?*

It is the respondent's task to assess whether an inheritance or gift might be 'substantial.' Next, the respondent provides the number of inheritances and gifts the household received:

- ***6.01A*** *How many substantial gifts or inheritances were received?*

In a maximum of three loops the respondent is then requested to answer the following questions regarding the year in which the transfers were received, its portfolio, and its values at time they were received:

- ***6.02*** *Was that a gift or an inheritance?*
- ***6.03*** *In what year did (you / your household / the household) receive the [inheritance/gift] that was most important for (your current wealth / the current wealth of your household / the current wealth of the household)?*
- ***6.04*** *What type of asset was the [inheritance/gift]?*
- ***6.05*** *At the time (you / your household / the household) received the [inheritance/gift], how much was it worth?*

Question 6.02 is modified depending on the number of inheritances and gifts. However, if a household received more than three inheritances or gifts, only the three most important (with respect to a household's current financial situation) are surveyed. Thus, it is to be expected that the majority of respondents choose the most valuable transfers first. Furthermore, the household's main residence is surveyed separately in Section 3 'Real assets and their financing,'

With regard to this questionnaire module on inheritances, neither item-nonresponse nor implausible values pose a significant problem. Eisele and Zhu (2013) report that respondent easily recall the values of past inheritances received. They also provide researches with comprehensive information on the imputation of missing values. The overall highest share of imputed values is observed for the value of the first transfer and amounts to 8 %, which is far below the values in other parts of the questionnaire (Eisele and Zhu, 2013, p. 33ff). Le Blanc (2014) notes, with regard to editing of the data, that few values needed editing, but some values needed to be converted from DM into €.

The refusal of whole households to participate (unit non-response) poses a much greater threat to representativeness of the survey. The selectivity of non-response might bias the estimates. Overall, at 18.6 %, the response rate of PHF households was rather low (von Kalckreuth et al., 2012, p. 8); however, the survey providers used the available paradata and sampling probabilities to re-weight the data, additionally, in an attempt to ensure representativeness, the household weights were calibrated to match the German micro census by the German Federal Statistical Office (von Kalckreuth et al., 2012, p. 17).

5 Comparing the joint distribution of intergenerational transfers, income and wealth across the Euro area

5.1 Introduction

Private wealth is a crucial factor of economic well-being for individuals and households. Research suggests that saving rates from income and intergenerational wealth transfers (inheritances and gifts) are two key determinants of wealth held by private households (for an overview see Davies and Shorrocks, 2000; for more recent research see Semyonov and Lewin-Epstein, 2013; Arrondel et al., 2014; Mathä et al., 2014; Fessler and Schürz, 2015; among others). Since the 1980s there is an ongoing debate over which of the two determinants contributes more to the current net worth of private households (Modigliani, 1986, 1988; Kotlikoff and Summers, 1981; Kotlikoff, 1988). Research stresses that intergenerational transfers are a dominant factor (Piketty, 2011, 2014; Piketty and Zucman, 2015), thus fueling the discussion about the legitimacy of wealth without effort. Some economists argue that this development may even pose a threat to democracy (Piketty, 2014).

We investigate the current role of wealth transfers in the Euro-area (Austria, Belgium, France, (West) Germany, Cyprus, Greece, Portugal, and Spain). As the availability of data was limited, this is the first time that cross-country comparisons focusing on Europe are possible. We analyze the percentages of households with a transfer as well as the conditional present values of transfers received. Additionally, we tackle the crucial question of how important are wealth transfers for the current distribution of household net worth[44] in Europe, computed as inheritance-wealth ratios.

The paper is structured as follows: In Section 5.2 we give an overview of the literature about wealth transfers in absolute and relative terms in developed countries. In Section 5.3 we describe the data we are using, the Household Finance and Consumption Survey (HFCS), as well as our reasoning concerning the country selection. We also give an overview of the inheritance and gift taxation in each country (see also Appendix 5.A). In Section 5.4 we present the distribution of intergenerational transfers in the Euro-area in absolute terms and analyze the sociodemographic characteristic of heirs applying logit and OLS regression analyses. Additionally, we analyze the role of past intergenerational transfers for current net worth using recently established methods by Wolff and Gittleman (2014) and Piketty et al. (2014) as well as a fractional logit model explaining the relative importance of transfers received. Section 5.5 summarizes and concludes.

5.2 Literature

5.2.1 The role of inheritance and inter-vivos transfers in absolute terms

Künemund and Vogel (2011) provide an overview of the studies for Germany (for example, works by Kohli et al., 2006, 2005), finding that transfers are positively correlated with education, income and wealth of both the donors as well as the recipients. For Germany, it is well established that parents of children with higher education usually also hold a higher degree, which, in turn, results in higher income and more possibilities to accumulate wealth to bequest (Baumert et al., 2001). In addition, the offspring also typically cash in on their higher education, profiting from higher earnings and savings. Szydlik and Schupp (2004) find that there are no differences between genders. Albuquerque (2014) describes a downward flow of monetary gifts from parents to their children for several countries in Europe, which may either be motivated by altruism, an accident, or in a strategic

44 Definition: Assets minus liabilities.

manner (Brunner, 2014). In the first case parents gain utility from knowing that their children will enjoy their bequest. In the second it is assumed that lifetime is uncertain and, thus, parents accidentally leave bequests if they die younger than expected. In the last case parents expect something from their children, such as visits, in exchange for a bequest. For Austria, Fessler et al. (2008) find that workers receive wealth transfers less often than the average household; while entrepreneurs receive, on average, the highest transfers. Karagiannaki (2015) and Wolff and Gittleman (2014) report similar findings for the UK and the US, respectively.

Cross-country comparisons are rare: Semyonov and Lewin-Epstein (2013) report the percentage of households older than 50 that received inheritances for many European countries, Israel, and the US. The data (for most countries SHARE) was collected between 2004 and 2007. The prevalence ranges between 46.2% in Switzerland, followed by Belgium with 42%, to 17% in Austria, and 4.4% in the UK. Fessler et al. (2008) compare means and medians for heir and non-heir households and conclude that beneficiaries are better educated, have higher incomes, and more wealth. They use LWS data, which was surveyed around the year 2000.

5.2.2 The role of inheritance and inter-vivos transfers in relative terms

Analyzing inheritances and inter-vivos transfers in relative terms, the inheritance-wealth ratio, requires decisions that imply methodological differences. Namely, Modigliani (1986, 1988) solely adjusts past wealth transfers for inflation to compute the present value of wealth transfers. Conversely, in Kotlikoff and Summers (1981) and Kotlikoff (1988), past wealth transfers are additionally capitalized, based on the assumption that transfers are usually invested in some kind of portfolio and are not held in cash. The first case results in quite low inheritance-wealth ratios (at most 25%). The second approach yields ratios that are considerably higher (45 to 80%). However, both approaches have in common that the share of wealth transfers due to past wealth transfers can exceed 100%, as the summarized past transfers are not capped at a household's net worth. Piketty et al. (2014) explicitly combine the two rival approaches (for details see Section 5.4). However, as Piketty et al. use data from the late 19th and early 20th century, their results are only of historical interest and not immediately relevant to the 21st century. Wolff and Gittleman (2014), using a similar method, find for the US in 2007 that the present value of transfers as a percent of net worth varies between 20 and 25%. Corneo et al. (2016) analyze, in a study similar to this one, the role of inheritances and gifts for the total net worth of (West) Germany in 2010. They conclude that one-third of wealth is attributable to capitalized wealth transfers.

Our analyzes in Section 5.4, as well as the studies from Piketty et al. (2014) and Wolff and Gittleman (2014), are based on the joint distribution of wealth and capitalized wealth transfers. Only a few studies use comparable data; some studies need additional assumptions in order to apply macroeconomic estimation techniques. Reil-Held (2004) estimates that inheritances and gifts account for approximately 34% of Germany's total net worth;[45] another macroeconomic estimate, from Piketty and Zucman (2015), is considerably higher: 51%. For France, Kessler and Masson (1989) estimate that the share of wealth transfers is 35%. The value computed by Klevmarken (2004) for Sweden is 19%. To the best of our knowledge, cross-country comparisons analyzing the impact of intergenerational wealth transfers on the distribution of wealth in absolute and relative terms are not available yet.

5.3 Data, country selection and institutional environment

The Household Finance and Consumption Survey (HFCS) contains information about households'[46] net worth, income and indicators of consumption, and credit constraints from

45 Note that the HFCS only surveys inheritances and gifts that are received from a person not living within the same household. Any macroeconomic estimate includes tax-relevant transfers within households (e.g. widowhood) and should be, logically, higher than results based on the HCFS for intergenerational transfers.

46 Our unit of analysis is, therefore, the household and not the individual. However, we provide a robustness check applying a per (adult) capita definition for the total present value of transfers in Appendix 5.B (Table 5.11). In the multivariate analyzes we control for household structure.

almost all Euro-countries[47] around the year 2010 (European Central Bank, 2013a,b). In addition, it contains information about intergenerational wealth transfers from outside the household. Each household's reference person[48] retrospectively answered a question about how many inheritances or substantial gifts the household received from any person who was not a member of the same household.[49] Consequently, the total number and amount of wealth transfers is underestimated because, among others, transfers due to the death of a partner who was part of the same household are not included. In addition, it affects the comparisons of countries with different household structures e.g. adult children still living with their parents. In the HFCS survey, the value of up to three intergenerational transfers was collected. In a separate module the mode of acquisition of the household main residence was collected; the choices include 'inherited' and 'gifted'.[50] The respondents sorted all transfers according to their subjective importance for their current financial situation.[51] It is also collected in which year the household received the transfer, what kind of assets the portfolio contained, if it was a gift or inheritance, and from whom it was received.

5.3.1 Country selection

The HFCS 'is a milestone for cross-country comparisons' and its data quality with regard to institutional environment, relevance, coherence, timeliness, accessibility, comparability and accuracy is quite high (Tiefensee and Grabka, 2016, p. 137). Nevertheless, they also show that net worth positions are not unlimitedly comparable between all countries due to methodological differences. Based on their analysis and the fact that not all countries

47 Austria, Belgium, Cyprus, Finland, France, Greece, Germany, Italy, Luxembourg, Malta, Netherlands, Portugal, Slovakia, Slovenia and Spain. Estonia, Ireland and Latvia will take part in the next wave.

48 For selection criteria see European Central Bank (2013a, pp. 16–17).

49 As past wealth transfers are collected retrospectively, it is highly likely that the data is plagued by under-reporting problems and the estimates are biased downwards. This is even more probable the more members live in a household. We do not know, and it is hard to quantify, whether under-reporting varies systematically for different age classes or demographic characteristics of the respondents.

50 In France, household main residence is part of the same intergenerational transfers module and not collected separately.

51 This implies that the sorting does not generally reflect the absolute value of the transfer, but it should be closely related.

surveyed wealth transfers, we include the following countries in our analysis: Austria, Belgium, France, (West) Germany, Cyprus, Greece, Portugal and Spain.[52]

Household are, on average, larger in Mediterranean countries; larger households also tend to accumulate more wealth than smaller ones (European Central Bank, 2013b). Furthermore, owner-occupied real estate, which is especially common in Mediterranean countries and usually represents the largest share of net worth, is likely to be transferred as inheritance, while financial wealth might be passed on to the next generation as inter-vivos transfers. Fessler and Schürz (2015) show that welfare state spending is negatively correlated with household wealth. Though the effect on transfers is uncertain.

To account for the most obvious differences, we divide our country selection into two groups. The core European countries (Austria, Belgium, France, (West) Germany) possess a generous welfare state regime with high social expenditures[53] at least since the 1980s and, on average, smaller households with similar structures (based on Figure 5.1, European Central Bank, 2013b, and Fessler et al., 2014). The Mediterranean countries (Cyprus, Greece, Portugal and Spain) comprise the second group with, on average, larger households and less generous welfare state expenditures. In addition, for several years following World War II these countries were without stable financial markets–and consequently, without comparable investment opportunities–due to e.g. civil wars and military dictatorships.

5.3.2 Inheritance and gift taxation

Effective average inheritance and gift tax rates depend on tax rates, allowances, exemptions etc. and are complex to calculate and not available for all countries over time. Based on

52 For Germany, we base our analysis on the western part due to problems of capitalization for past intergenerational transfers that date from before the fall of the wall. For the rest of this analysis, we use ‘Germany’ and ‘(West)’ Germany synonymously. We restrict the analysis to households with a head of at least 21 years of age. Additionally, not all countries in the HFCS oversample wealthy households. Therefore, our analysis for most countries is likely not representative for the very top (Vermeulen, 2014). To account for missing values, the data is multiply imputed (five implicates) by the data providers (European Central Bank, 2013b). Our calculations are based on standard applications for multiply imputed data; we use the provided replicate weights and all standard errors are bootstrapped.

53 These includes: public, mandatory and voluntary private social expenditure in the following fields: old age, survivors, incapacity-related benefits, health, family, active labor market programs, unemployment, housing, and other social policy areas.

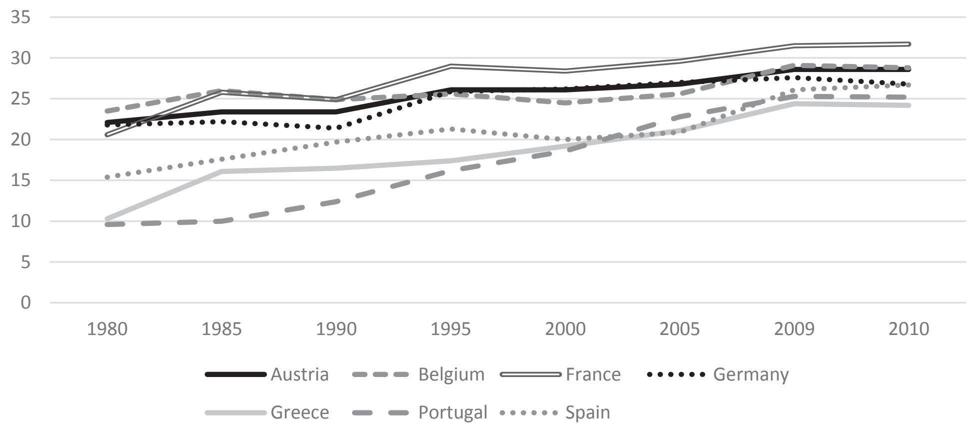


Figure 5.1: Social expenditure as percentage of GDP.
Source: OECD.Stat (2015). Note: Data for Cyprus is not available from OECD.Stat.

the tax rates, the thresholds for the maximum tax rate, the maximum tax allowances (see Appendix 5.A) and the tax revenues (Figure 5.2) we define three types of gift and inheritance taxation regimes: (1) no or low inheritance and gift tax; (2) moderate inheritance and gift tax with moderate or high allowances; and (3) high inheritance and gift tax with low or moderate allowances. As demonstrated by Piketty (2014) the wealth transfer flow collapsed following the shocks of 1914-1945, but again gained momentum starting in the 1990s in several European countries (namely France, Britain and Germany). In addition, Figure 5.2 demonstrates that tax revenues diverged, particularly in the 2000s. Therefore, our analysis of the institutional settings starts in 2000 and ranges through the time of the survey (year 2010). For a more thorough summary, we refer to Appendix 5.A, where all key information is provided in table form.

The first group (no or low inheritance and gift tax) consists of Cyprus, Austria and Portugal. Cyprus and Austria abandoned the taxation of inheritances and gifts completely after 2000/2008, respectively, with only a land transfer tax levied, which is in the one-digit area. In Austria, before 2008 the taxation depended on the level of relationship between testator and heir, with tax rates moderate or high, but tax allowances low. In Portugal, since 2004 only a stamp duty is levied on the respective documents. Transfers between spouses, or other immediate relatives, are largely exempt. Before the changes occurred, tax rates were moderate and tax allowances low.

The second group (moderate taxation of inheritances and gifts with moderate or high tax allowances) consists of Greece and Germany. In both countries the tax rate varies depending on the relationship and the value of the transfers received. The tax rates are lower in Greece, the tax allowances higher in Germany.

The third group (high inheritance and gift tax with low or moderate tax allowances) consists of Spain, France and Belgium. In Spain the applicable tax rate varies not only depending on the relationship and the value of the transfers received, but it also takes into account the net worth of the heir. However, since 2004 some regional governments factually abandoned the taxation of wealth transfers. The tax system in France is similar to that in Germany, but with higher tax rates and lower allowances. In Belgium we observe varying gift taxes depending on the region, the relationship, and the value since 2001; and for inheritance tax since 2002. Another peculiarity in Belgium is a considerable difference between the taxes on inheritances and gifts.

Almost all countries we consider have more or less extensive exemption clauses applying to the transfer of businesses and owner-occupied property.

In summary, the inheritance and gift tax regimes might not strongly influence the incidence of wealth transfers, because for transfers within the closer family tax rates never exceed 50%[54] and are accompanied by allowances and additional exemptions. Ceteris paribus, among the Mediterranean countries the levels of inheritances and gifts are probably the highest in Cyprus (no inheritance and gift tax for several years) and the lowest in Spain. In the Core European countries we expect them to be lower in Belgium and France than in Germany or Austria.

54 For Austria, Belgium, France and (West) Germany this is already the case since the 1950s (Scheve and Stasavage, 2012).

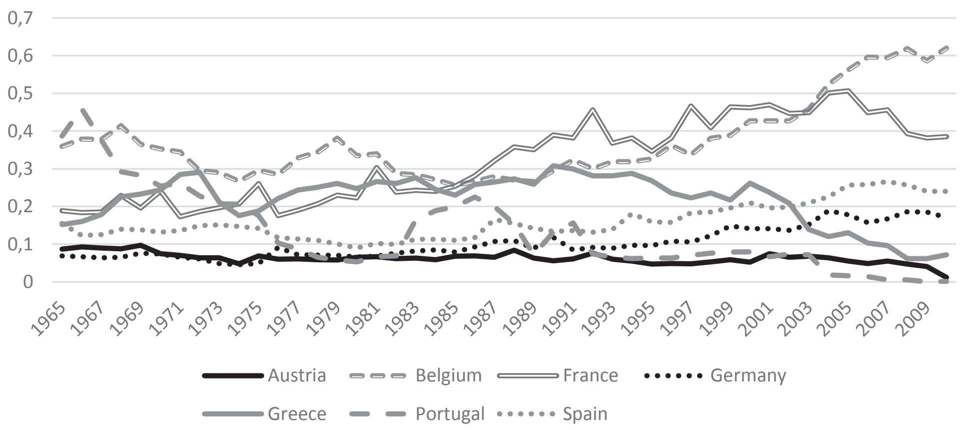


Figure 5.2: Inheritance and gift tax revenue as percentage of GDP.
Source: OECD.Stat (2015). Note: Data for Cyprus is not available from OECD.Stat.

5.4 Who receives wealth transfers and what is the value of the transfers received?

In the first step of our empirical analysis (Sections 5.4.1 and 5.4.2) we give an overview of the distribution of intergenerational wealth transfers[55] from outside the household (as defined by the HFCS questionnaire) for eight European countries. We first tabulate the incidence as well as the conditional mean values of inherited wealth. We calculate the present value of all past wealth transfers that a household received, in 2010 prices, and capitalize the past wealth transfers using a real annual rate of return of three percent. The analysis relies on the intertemporal budget constraint of private households, it is described by Piketty et al. (2014) in more detail. In short, the idea is as follows: for all households we observe the joint distribution of all past wealth transfers and net worth at time y. Note that $y = 2010$ on average for the surveyed households in our analysis. We capitalize the past wealth transfers using a real annual rate of return r, which yields the present value

55 Gifts and inheritances are analyzed together. If only looking at gifts, the sample sizes are quite small in some countries. This is probably due to missing tax incentives in these countries and different asset portfolios (e.g. if households mainly possess a household main residence it will be most likely be passed on after death).

of wealth transfers $PVWT_{yi}$ for all households i in any sample or subsample at time y. This deserves an explanation: we assume that at the point in time any household receives a wealth transfer it always has the option to make a secure investment yielding a real rate of return r. Hence, similar to Wolff and Gittleman (2014), we calculate the present value of wealth transfers

$$PVWT_{yi} = \sum_{t=t_0}^{y} WT_{ti} \cdot e^{r(y-t)}. \tag{5.1}$$

For each single household, i, in our sample we determine the actual sum of inheritances and gifts based on our assumptions: If a household's net worth is larger than the present value of transfers, it follows that the household has real savings as high as the residual $(W_{yi} - PVWT_{yi})$. If a household has a net worth less than the present value of wealth transfers, we conclude that the household consumed part (or all) of the wealth transfer instead of choosing a secure financial market investment (or lost over time). The residual resulting from the secure investment is, therefore, interpreted as the household's savings, as it was the investment decision of the household to either invest differently (and potentially more risky) or consume the wealth transfer. The total present value of wealth transfers for any country or subsample in year y is then given by

$$TPVWT_y = \sum_i \min(PVWT_{yi}, w_{yi}). \tag{5.2}$$

Additionally, we are interested in calculating the total value of wealth transfers as a percent of positive net worth, which, according to the literature, are computed at the aggregate level as the total inherited wealth divided by the total current wealth

$$\beta_y = \frac{TPVWT_y}{\sum_i w_{yi}} = \frac{TPVWT_y}{W_y}. \tag{5.3}$$

However, in our application we calculate the ratio on the household levels and average within countries or subpopulations, as this definition seems more useful with regard to our typology of households: In line with Piketty (2014), any population can be divided into three groups of households. For those households that (1) never received a transfer or has negative net worth, β_{yi} is always zero. For those households that (2) received a transfer and the present value falls below the net worth, β_{yi} is the ratio of the present value to net worth w_{yi}. For the third group of households that (3) received a transfer but the present value exceeds the net worth in year y, it follows that β_{yi} is 100%, hence all of the net worth can be attributed to the transfers, as the household consumed more than it could have afforded from its own labor or investment decisions. Based on this reasoning we conclude that the residual that cannot be attributed to the inherited portion of the net worth must be the result of a household's saving decision and attributed to the portion resulting from its own efforts.

The most arbitrary assumption in our analysis is the choice of the real rate of return r. The base rate we choose is $r = 3\%$ in accordance to Wolff and Gittleman (2014). We add a few robustness checks (see Appendix 5.B) in order to identify systematic changes, if we vary the real rate of return between $r = 1\%$ and $r = 5\%$. Additionally, we check the variation of β_{yi} depending on a wealth related rate of return, as it seems reasonable that richer households are financially better educated, have the possibility to invest more diverse and, therefore, might realize higher rates of return (these results are presented in Appendix 5.B). For some countries (Belgium, France and Germany) it would also be possible to use the yields of investments in long-term government bonds, as these investments are in line with our definition of a secure investment. As the time series are not available for all countries from the 1950s onward, we add the results to Appendix 5.B and note that the differences to a real interest rate of 3% are negligible.

5.4.1 Incidence and levels of past intergenerational transfers

As shown in Table 5.1, the incidence of transfers received varies slightly across the European countries we analyze. In Portugal, the share of households that received at least one wealth transfer is the lowest (27%), with the highest prevalence observed in (West) Germany (38%) and France (roughly 40%).

In the core European countries, we find that with increasing household income[56] the probability that a household already received a wealth transfer increases. Households finding themselves in the highest income quintile record double the incidence of transfers (more than 50% of all households) as compared to the first quintile. The Mediterranean countries on the other hand do not exhibit similar variation along the distribution of income. For instance, in Portugal the incidence varies independently of income (around 25%). This is, amongst other things, explained by the expansion of secondary and tertiary education since the 1960s, which has greatly improved the educational mobility for the current generation of heirs.

The higher the observed net worth of a household is, the higher is the likelihood that it reports a wealth transfer. The picture is very similar for all countries in our analysis. For the population reporting a net worth below €20,000 it is well below 20% and then it quickly rises to 70%-75% in countries where this correlation seems to be the most pronounced (Austria, France and Germany).

In general, the likelihood that a household[57] reports a wealth transfer increases with age.[58] However, in addition to lifecycle effects, cohort effects can be identified. Due to lifecycle effects, the age classes between 45 and 64 have significantly higher percentages of

56 The current gross household income refers to the last 12 months / the last calendar year before the time of the survey and is composed of the following components: all earned income, pensions (public, occupational and private), unemployment benefits and other regular social transfers, regular private transfers, rental income, income from financial assets, income from private companies / partnerships plus additional other income.

57 Most sociodemographic characteristics of the households are referring to its head. We use ‘household’ and ‘household’s head’ synonymously.

58 Age class according to the age of the head of the household as reported in the HFCS survey data. In the multivariate part we investigate the last two age classes together due to the low numbers of cases.

households with a wealth transfers than the younger age classes, as their older relatives (especially parents and grandparents) more likely already deceased. The age classes over 65, on the other hand, have decreasing percentages of households that report a wealth transfer. Their older relatives, of which the majority is likely to be deceased already, presumably lived in much poorer conditions (e.g. due to the two World Wars) and did not bequest (large) fortunes. For instance, in Cyprus the effect is extremely pronounced, as the oldest cohort reports only half as many inheritances and gifts compared to the second oldest cohort. The patterns are very similar across Europe with a few exceptions. Some countries do not experience a drop for the oldest cohorts.

In the next step, we look at the capitalized conditional mean present value of wealth transfers across Europe (see Table 5.2). Therefore, we limit the sample to all households reporting at least one transfer, adjust the values for inflation, capitalize them and sum them up by households (see equation 5.1). Belgium and Greece are fairly close to each other (€155,000 and €152,000, respectively). The conditional mean present values in Austria and Germany are considerably higher (€230,000 and €193,000). Spain records €174,000 and France €137,000. There are two outliers: Portugal at only €85,000 and Cyprus at €274,000. Linking this to the inheritance and gift tax regimes, we find indeed that the present values are highest for Cyprus among the Mediterranean countries. However, they are significantly lower in Portugal than in Spain in spite of the much more steep taxation in Spain, this is probably because the overall wealth levels are much lower in Portugal for historical reasons. In addition, as expected based on the tax regimes, Belgium and France do have the lowest wealth transfer values among the core European countries.

With regard to the joint distribution of income, the capitalized present value is highest in the highest income quintile. This confirms the strong relationship between a household's income position and the expected wealth transfers from previous generations indicating low intergenerational mobility. While the incidence does vary less for Mediterranean countries, the absolute value does increase with income as in the core European countries.

Not surprisingly, most countries experience a sharp rise in the conditional mean present value of transfers received from the second highest wealth level to the highest with household net worth above €1 million. For all countries the value at least doubles. Generally, the conditional present value of the transfers seems to increase monotonically with the wealth level. The wealth levels above €250,000 show values in the six to seven-digit euros region, whereas for the lowest wealth level below €20,000 the conditional value never exceeds €10,000.

The conditional mean present value peaks only in Belgium and France for the oldest cohort aged 75 or older. In Portugal and Spain the variation across the age classes is rather low. In Austria we observe a spike for the age class 45 to 54 (€285,000), in Germany it only increases slightly for cohorts older than 44. The rather liberal legislation concerning the taxation of gifts clearly left its mark in the distribution for younger households: Austria, Greece, (West) Germany and Cyprus all exhibit a reversely U-shaped pattern. This is in line with the observations of the percent of households with transfers, i.e. not only did the middle aged households report having received a wealth transfer considerably more often, those transfers were considerably higher as well. This is the result of the cohort effect offsetting the life-cycle effect in wealth transfers in those countries.

As the observed patterns, and especially mean and median present values, might depend on the household size, we add a robustness check in Appendix 5.B (Table 5.5) and check for the variation in per (adult) capita wealth transfers instead. Using per capita transfers expectedly reduces the values, but does not change the patterns reported in this section.

5.4.2 Correlates of the prevalence and value of transfers received

We estimate a logit model characterized by the specification

$$p_j = F(\alpha + \beta X_j + \varepsilon_j), \tag{5.4}$$

with p_j denoting the probability of households in country j of having received a transfer, α is an intercept, ε_j are unobservable variables. X_j is the matrix of all explanatory

Table 5.1: Percent of households with a transfer.

	I. Core European countries								**II. Mediterranean countries**							
	Austria		Belgium		France		(W) Germany		Cyprus		Greece		Portugal		Spain	
All households	35.7	(1.3)	31.7	(1.2)	39.9	(0.7)	38.1	(1.7)	31.5	(1.7)	30.7	(1.5)	26.7	(1.3)	30.1	(1.1)
A. Income quintiles																
1st quintile	26.2	(2.3)	25.3	(2.9)	31.0	(1.5)	24.6	(3.1)	22.8	(3.7)	28.3	(3.0)	26.5	(2.2)	32.9	(2.0)
2nd quintile	29.7	(2.7)	32.5	(3.2)	33.8	(1.6)	32.2	(3.8)	30.8	(4.1)	33.7	(2.6)	30.4	(2.6)	29.9	(2.1)
3rd quintile	34.3	(2.9)	27.6	(3.0)	38.2	(1.6)	37.6	(3.4)	30.3	(3.8)	31.4	(2.7)	26.6	(2.6)	25.2	(2.6)
4th quintile	38.0	(2.7)	35.0	(2.9)	43.1	(1.5)	44.6	(3.0)	40.3	(4.0)	29.2	(2.9)	26.2	(2.3)	29.8	(2.4)
5th quintile	50.3	(3.1)	37.9	(2.8)	53.2	(1.3)	51.8	(3.0)	33.1	(3.8)	31.0	(2.8)	24.1	(1.9)	32.9	(2.3)
B. Wealth levels																
Under €20,000	11.6	(1.6)	12.9	(2.3)	17.9	(1.1)	13.1	(2.2)	7.9	(2.8)	4.1	(1.1)	11.8	(1.3)	7.9	(1.7)
€20,000-€99,999	31.3	(2.7)	27.6	(4.2)	35.5	(1.9)	28.2	(3.2)	18.7	(4.4)	34.8	(2.9)	28.1	(2.1)	24.8	(2.7)
€100,000-€249,999	45.8	(2.7)	27.6	(2.9)	44.5	(1.4)	49.3	(3.2)	30.3	(3.8)	39.4	(2.2)	34.9	(2.2)	27.4	(1.8)
€250,000-€499,999	54.4	(3.2)	39.1	(2.7)	56.5	(1.5)	65.3	(2.9)	36.5	(4.3)	37.5	(3.9)	34.1	(3.5)	39.2	(2.5)
€500,000-€999,999	71.6	(4.3)	48.8	(3.7)	69.0	(2.1)	63	(5.8)	38.1	(4.9)	42.7	(5.6)	33.4	(4.3)	46.4	(3.9)
€1,000,000 or over	68.4	(6.8)	51.3	(5.0)	75.1	(2.3)	69.7	(5.8)	51.7	(4.8)	51.1	(15.6)	44.5	(6.3)	62.1	(5.3)
C. Age classes																
21-35	22.9	(2.4)	16.1	(2.8)	24.8	(1.6)	22.3	(3.8)	28.7	(4.0)	22.5	(1.9)	12.9	(2.3)	16.0	(2.3)
35-44	34.8	(3.1)	25.3	(2.9)	32.0	(1.5)	36.1	(3.0)	31.0	(3.8)	34.3	(2.6)	20.8	(2.4)	20.4	(2.1)
45-54	38.6	(2.5)	29.2	(2.8)	38.3	(1.6)	46.8	(3.1)	38.3	(3.6)	33.8	(2.8)	28.0	(2.3)	33.0	(2.2)
55-64	44.4	(2.4)	43.0	(3.1)	51.7	(1.7)	46.2	(3.4)	33.3	(4.2)	33.4	(3.3)	30.5	(2.3)	40.6	(2.6)
65-74	37.1	(3.1)	40.0	(3.2)	51.9	(1.7)	39.9	(3.6)	31.5	(4.7)	30.4	(3.0)	29.9	(2.3)	40.7	(2.3)
75 and older	35.1	(4.5)	42.2	(3.4)	46.1	(1.9)	33.5	(4.2)	17.2	(4.9)	30.6	(3.6)	34.2	(2.5)	32.7	(2.2)
Sample size (n)		2,337		2,307		14,929		2,826		1,234		2,915		4,393		6,188
Weighted in Mio. (N)		3.71		4.66		27.51		28.64		0.30		4.06		3.92		16.97

Note: Standard errors are shown in parentheses. Means over 5 implicates, standard errors bootstrapped. The figures record the proportion of households who indicate receiving a wealth transfer at any time before the time of the survey.
Source: own computations from the HFCS survey wave 1 (2013).

Table 5.2: Mean present value of transfers received (in €1,000), in 2010 prices and capitalized with $r = 3\%$, recipients only.

	I. Core European countries								**II. Mediterranean countries**							
	Austria		Belgium		France		(W) Germany		Cyprus		Greece		Portugal		Spain	
Mean present value	230	(19)	155	(10)	137	(4)	193	(13)	274	(23)	152	(8)	85	(7)	174	(11)
Median present value	110		77		46		107		165		113		38		77	
A. Income quintiles																
1st quintile	119	(28)	116	(26)	73	(6)	97	(21)	157	(57)	98	(9)	50	(5)	98	(8)
2nd quintile	140	(21)	114	(14)	95	(8)	130	(20)	154	(26)	119	(10)	60	(6)	126	(14)
3rd quintile	205	(27)	142	(18)	95	(8)	158	(20)	266	(78)	151	(19)	63	(7)	148	(43)
4th quintile	226	(34)	173	(22)	113	(7)	194	(21)	344	(49)	167	(29)	65	(8)	180	(19)
5th quintile	361	(47)	208	(28)	252	(11)	304	(33)	389	(61)	226	(22)	201	(37)	310	(36)
B. Wealth levels																
Under €20,000	6	(1)	6	(1)	5	(0)	6	(1)	6	(2)	10	(2)	6	(1)	6	(1)
€20,000-€99,999	42	(3)	34	(5)	31	(2)	33	(3)	47	(7)	59	(2)	38	(2)	40	(3)
€100,000-€249,999	118	(6)	98	(9)	73	(3)	116	(5)	133	(12)	141	(3)	82	(5)	85	(6)
€250,000-€499,999	231	(13)	135	(12)	143	(5)	204	(12)	199	(22)	246	(13)	116	(14)	141	(9)
€500,000-€999,999	435	(33)	220	(23)	256	(14)	414	(29)	277	(40)	436	(54)	252	(44)	300	(36)
€1,000,000 or over	904	(145)	478	(74)	739	(44)	818	(105)	584	(79)	931	(278)	696	(198)	734	(108)
C. Age classes																
21-35	176	(48)	60	(15)	45	(5)	116	(38)	244	(37)	139	(10)	42	(8)	149	(31)
35-44	197	(31)	131	(30)	97	(7)	188	(28)	287	(42)	152	(9)	81	(13)	164	(24)
45-54	285	(28)	136	(19)	133	(9)	196	(18)	296	(40)	193	(21)	65	(6)	171	(24)
55-64	239	(34)	154	(19)	141	(9)	201	(30)	310	(79)	191	(28)	83	(19)	190	(25)
65-74	245	(51)	170	(18)	176	(13)	233	(23)	242	(73)	93	(9)	104	(21)	173	(16)
75 and older	181	(49)	226	(33)	200	(14)	182	(22)	154	(36)	109	(18)	104	(18)	185	(48)

Note: Standard errors are shown in parentheses. Means over 5 implicates, standard errors bootstrapped. The figures show the present value of all transfers as of the survey year which were received up to the time of the survey in prices of 2010 using country specific inflation rates.
Source: own computations from the HFCS survey wave 1 (2013).

variables: age, education, work and marital status as well as gender of the reference person, income[59] of the household and its size.[60] Additionally, we estimate the following OLS specification:

$$y_j = \alpha + \beta X_j + \varepsilon_j \tag{5.5}$$

with y_j denoting the capitalized present value of all wealth transfer for households in country j. We sum up all past wealth transfers in prices of 2010. α is the intercept and ε_j denotes unobservables. X_j is the matrix of all explanatory variables, which are the same as for the logit estimation.

The results regarding the probability of receiving a transfer in the individual countries are shown in Table 5.3. Table 5.4 shows the results for the OLS regressions regarding the mean intergenerational wealth transfer value (as log) in each country for the heir population only.

For the household income the following pattern emerges: The higher the income, the higher the probability that the household reports a wealth transfer. This is especially pronounced in the core European countries. Both findings also hold for the average amount of transfers a household receives: Households of higher income quintiles tend to report higher transfers. The pattern is most salient at the edges of the income distribution. The findings are connected to those regarding education and intergenerational mobility: In the core European countries, we find for all countries that households with primary education had a smaller propensity to receive a transfer compared to those with secondary

59 In the HFCS gross income was collected, usually referring to the calendar year prior to the survey year or the 12 months preceding the interview.

60 Except for income, all explanatory variables relate to the time of the interview (around 2010). Due to endogeneity, net wealth is not used as an explanatory variable. Further information about the transfers cannot be used in the analysis due to the pooled estimation of the transfers. Information about the tax regimes are only available on the country level and can therefore not be considered due to the low number of cases. To use the household head as reported in the survey is standard in the literature. However, in an alternative specification we used the oldest person in the household as its head. The results suggest that the estimates are fairly robust.

education. Households with tertiary education, on the other hand, are characterized by higher propensities. Interestingly, in Cyprus households with lower education had a higher chance to receive a transfer as compared to secondary education. This might be a hint that intergenerational mobility is still comparatively high. Considering the present values, the relationship between education levels and the value of transfers received is very pronounced in France, Portugal and Greece, i.e. those households that received a transfer expect a higher value if their head has tertiary education. Research suggests that children of parents with higher education usually also hold a higher degree, which in turn results in higher income and more possibilities to accumulate wealth to bequest (see for example Baumert et al., 2001).

Again life-cycle effects are visible: With increasing age, the likelihood of losing family and friends and, thus, receiving a wealth transfer is monotonically increasing for most countries. For the age classes between 45 and 64 we identify significantly higher probabilities of having already received a transfer than the younger ones. The cohort effect, decreasing transfers for old cohorts due to poorer living conditions, which are reported in Table 5.1, is not visible or significant once we control for other sociodemographic variables. In Belgium, France and Spain the lifecycle patterns of transfer recipients are the most pronounced and significant. For the mean present value (Table 5.4) of those who received a transfer, in many countries the 45 to 54 age cohort has received higher transfers than younger cohorts.

We find that self-employed households (compared with employed ones) have, in the majority of the countries, a higher chance to receive a transfer and, also, report larger transfer values. One explanation for this might be that the self-employed often inherit the business that they are working for. Compared with the status married, households led by widowed or divorced persons have smaller chances of having received an inheritance or gift. Keep in mind that the inheritance from the deceased spouse is not reported in the survey, if the spouse used to be part of the same household (see Section 5.3). In the case of a divorce, it is logical that the incidence is reduced because high transfers mostly come from (grand-)parents(-in-law), and after a divorce the chances naturally halved for a household. Differences between genders are only significant in Austria, Germany, Cyprus and Spain. In these countries men have a smaller probability to receive a wealth transfer or the wealth transfers are lower than that for women.

Overall, the patterns we find for the probability to receive a transfer and the patterns for the average transfer value are quite similar over the countries. Specifically, the correlations between education and income with the present values of transfers received are high for all countries. The question arises, what exactly is the role of wealth transfers for the overall wealth situation of households in Europe? In the next section we explore household's net worth and transfers simultaneously by computing the transfers received as a percent of observed net worth, thereby obtaining an indicator for the impact of wealth transfers on the distribution of wealth.

5.4.3 Intergenerational wealth transfers and the distribution of wealth

In the previous section, we find that the prevalence of transfers received differs greatly between socio-economic groups. In addition, some households have not yet received a gift or inheritance, while others may never receive one. In this section, we investigate past intergenerational transfers as a percent of net worth (see Table 5.5).

Overall there are basically two tiers of countries. The first consists of the core European countries Austria and (West) Germany, and the Mediterranean country Greece. For these countries, the share is around 31%, meaning the share of inheritances and gifts is just under one-third in those countries. Rather low shares are computed for the second tier: Belgium, Portugal, Spain and Cyprus. In Portugal both the percent of households with a transfer and the conditional present values of those transfers tend to be lower than in the core European countries, resulting in an overall lower inheritance-wealth ratio (15%). In Spain the mean present values tend to be on par with the rest of Europe (Table 5.2), however, households receive the wealth transfers later in their lifecycle. In combination with an overall higher net worth level for Spanish households, the result is a rather low

Table 5.3: Average marginal effects of the logit estimations for probability of wealth transfer received.

Logit	**AT**	**BE**	**FR**	**(W)DE**	**CY**	**GR**	**PT**	**ES**
1st income quintile	-0.090*	-0.061	-0.088***	-0.074*	-0.059	-0.034	-0.033	0.006
	(0.039)	(0.037)	(0.014)	(0.039)	(0.058)	(0.029)	(0.025)	(0.023)
2nd income quintile	-0.042	0.017	-0.048***	-0.035	-0.058	0.030	-0.001	0.008
	(0.035)	(0.035)	(0.013)	(0.034)	(0.050)	(0.026)	(0.023)	(0.022)
4th income quintile	0.045***	0.057***	0.031***	0.037	0.001	-0.011	-0.014	0.050***
	(0.034)	(0.034)	(0.013)	(0.033)	(0.046)	(0.030)	(0.027)	(0.022)
5th income quintile	0.106***	0.093***	0.111***	0.048	0.033	-0.004	0.015	0.088***
	(0.034)	(0.035)	(0.012)	(0.030)	(0.046)	(0.029)	(0.025)	(0.020)
Age 21-34	-0.153***	-0.131***	-0.155***	-0.199***	-0.223***	-0.085***	-0.122***	-0.191***
	(0.034)	(0.041)	(0.016)	(0.039)	(0.047)	(0.028)	(0.034)	(0.033)
Age 35-44	-0.056**	-0.058*	-0.044***	-0.061*	-0.127***	0.019	-0.058**	-0.121***
	(0.030)	(0.033)	(0.013)	(0.030)	(0.038)	(0.025)	(0.024)	(0.022)
Age 55-64	0.097***	0.171***	0.078***	0.056*	-0.077	0.034	0.010	0.040**
	(0.034)	(0.034)	(0.014)	(0.029)	(0.050)	(0.031)	(0.023)	(0.018)
Age 65plus	0.059	0.232***	0.101***	0.045	0.052	0.006	0.041	0.062***
	(0.042)	(0.051)	(0.018)	(0.045)	(0.098)	(0.040)	(0.029)	(0.022)
Education primary	-0.042	-0.074***	-0.079***	-0.069**	0.009	0.036*	0.006	-0.027
	(0.028)	(0.027)	(0.010)	(0.036)	(0.042)	(0.021)	(0.022)	(0.017)
Education tertiary	0.083***	0.063***	0.099***	0.096***	-0.032	-0.005	0.021	0.023
	(0.030)	(0.023	(0.010)	(0.020)	(0.034)	(0.024)	(0.029)	(0.017)
Work status self-employed	0.155***	0.066	0.100***	0.083**	0.031	0.132***	0.108***	0.159***
	(0.034)	(0.044	(0.012)	(0.032)	(0.043)	(0.024)	(0.023)	(0.019)
Work status unemployed/other	-0.020	-0.009	0.007	0.023	0.024	0.031	0.012	0.046***
	(0.036)	(0.035	(0.016)	(0.032)	(0.051)	(0.025)	(0.025)	(0.018)
Work status retired	0.025	-0.045	0.086***	0.005	-0.182**	0.038	0.045*	0.059***
	(0.035)	(0.044	(0.016)	(0.040)	(0.087)	(0.035)	(0.025)	(0.021)
Marital status single	0.017	-0.010	-0.069***	0.031	-0.055	-0.003	-0.028	0.124***
	(0.033)	(0.033	(0.013)	(0.036)	(0.077)	(0.029)	(0.028)	(0.020)
Marital status widowed	0.007	0.028	-0.087***	-0.111***	-0.233***	0.006	-0.069**	-0.019
	(0.043)	(0.038	(0.017)	(0.043)	(0.079)	(0.036)	(0.028)	(0.022)
Marital status divorced	-0.081**	-0.048	-0.083***	-0.047	-0.056	-0.008	-0.140***	-0.036
	(0.036)	(0.036	(0.015)	(0.036)	(0.064)	(0.036)	(0.031)	(0.027)
Gender man	-0.044**	-0.010	0.011	-0.056**	-0.022	-0.009	0.022	-0.006
	(0.020)	(0.020	(0.009)	(0.020)	(0.032)	(0.019)	(0.019)	(0.013)
HH size 1	0.051	-0.030	0.032**	-0.074**	0.036	-0.010	0.050**	-0.018
	(0.033)	(0.031	(0.013)	(0.033)	(0.072)	(0.031)	(0.024)	(0.021)
HH size 3	0.116***	-0.040	-0.062***	0.029	0.039	0.004	-0.015	0.003
	(0.032)	(0.032	(0.013)	(0.029)	(0.049)	(0.025)	(0.020)	(0.016)
HH size 4	0.051	-0.047	-0.038***	0.058*	0.053	0.043	-0.014	0.019
	(0.037)	(0.036	(0.014)	(0.033)	(0.049)	(0.026)	(0.023)	(0.019)
HH size 5plus	0.189***	-0.049	-0.085***	0.066	0.026	0.068*	-0.022	-0.049*
	(0.046)	(0.045	(0.017)	(0.045)	(0.051)	(0.036)	(0.032)	(0.027)
Sample size (n)	2,380	2,296	15,004	2,828	1,220	2,971	4,399	6,197
Weighted in Mio. (N)	3.77	4.61	27.86	28.66	0.30	4.11	3.93	17.02

Reference groups: 3rd income quintile, age 45-54, education secondary, work status employed, marital status married, gender women, HH size.

Standard errors in parentheses. All 5 implicates are used, standard errors bootstrapped. *** p<0.01, ** p<0.05, * p<0.1

Source: Own computations from the HFCS survey wave 1 (2013).

Table 5.4: OLS regression for present value of wealth transfer received (heir population).

OLS	**AT**	**BE**	**FR**	**(W)DE**	**CY**	**GR**	**PT**	**ES**
1st income quintile	-0.677**	-0.612*	-0.480***	-0.422	-0.0180	-0.251**	-0.253	-0.577***
	(0.263)	(0.317)	(0.136)	(0.314)	(0.431)	(0.119)	(0.242)	(0.213)
2nd income quintile	-0.356*	-0.413*	-0.171	-0.120	-0.184	-0.204*	0.0641	-0.103
	(0.201)	(0.236)	(0.119)	(0.226)	(0.403)	(0.108)	(0.196)	(0.212)
4th income quintile	-0.0106	0.328	0.190*	0.337*	0.605*	0.00263	-0.0905	0.241
	(0.214)	(0.225)	(0.100)	(0.203)	(0.309)	(0.163)	(0.277)	(0.205)
5th income quintile	0.313	0.509*	0.712***	0.634***	0.455	0.228*	0.475**	0.427**
	(0.227)	(0.265)	(0.0905)	(0.192)	(0.346)	(0.128)	(0.219)	(0.210)
Age 21-34	-0.652**	-0.567*	-0.968***	-0.534	0.0316	-0.253**	-0.492**	-0.460
	(0.272)	(0.342)	(0.160)	(0.361)	(0.328)	(0.125)	(0.246)	(0.332)
Age 35-44	-0.480**	-0.450	-0.324**	0.124	0.0659	-0.130	-0.00676	-0.222
	(0.202)	(0.281)	(0.127)	(0.214)	(0.260)	(0.0799)	(0.212)	(0.214)
Age 55-64	0.0404	-0.0545	0.148	0.135	0.353	-0.0532	0.0332	0.107
	(0.231)	(0.297)	(0.126)	(0.194)	(0.341)	(0.165)	(0.194)	(0.176)
Age 65plus	0.116	0.105	0.781***	0.903***	0.397	-0.319	0.380	0.174
	(0.279)	(0.355)	(0.159)	(0.269)	(0.648)	(0.210)	(0.257)	(0.252)
Education primary	0.113	-0.206	-0.350***	-0.242	-0.363	-0.192**	-0.619***	-0.229
	(0.212)	(0.224)	(0.0914)	(0.213)	(0.278)	(0.0951)	(0.199)	(0.172)
Education tertiary	0.00948	0.208	0.334***	0.213	0.176	0.299***	0.530*	0.0562
	(0.166)	(0.154)	(0.0765)	(0.148)	(0.215)	(0.108)	(0.281)	(0.195)
Work status self-employed	0.632***	0.195	0.727***	0.608***	0.368	0.0939	0.638***	0.712***
	(0.198)	(0.414)	(0.112)	(0.193)	(0.291)	(0.113)	(0.190)	(0.207)
Work status unemployed/other	-0.363	-0.394	-0.246	-0.0712	-0.326	0.0603	0.236	0.165
	(0.268)	(0.343)	(0.184)	(0.248)	(0.353)	(0.0931)	(0.250)	(0.199)
Work status retired	0.0265	0.594**	0.0697	-0.0881	-0.429	-0.00858	0.268	0.304
	(0.256)	(0.265)	(0.129)	(0.227)	(0.623)	(0.188)	(0.210)	(0.278)
Marital status single	-0.0362	-0.411	-0.0772	-0.259	-0.702	-0.0802	0.0581	0.427**
	(0.226)	(0.268)	(0.118)	(0.312)	(0.530)	(0.112)	(0.212)	(0.210)
Marital status widowed	-0.512	-0.0266	-0.261	-0.217	-0.507	-0.0101	-0.122	0.145
	(0.313)	(0.313)	(0.162)	(0.326)	(0.585)	(0.143)	(0.237)	(0.216)
Marital status divorced	-0.441	-0.145	-0.379***	-0.361	-0.307	-0.118	0.0700	-0.177
	(0.306)	(0.274)	(0.130)	(0.280)	(0.492)	(0.155)	(0.334)	(0.264)
Gender man	0.162	-0.0346	-0.113	-0.0907	-0.452**	0.0522	-0.0958	-0.269**
	(0.129)	(0.160)	(0.0827)	(0.147)	(0.209)	(0.0814)	(0.182)	(0.132)
HH size 1	-0.00521	-0.0127	0.123	-0.135	0.524	-0.102	0.0299	0.0682
	(0.261)	(0.245)	(0.118)	(0.269)	(0.532)	(0.125)	(0.245)	(0.198)
HH size 3	0.416**	0.128	-0.0553	0.0987	0.427	0.0945	0.286*	0.0523
	(0.204)	(0.235)	(0.117)	(0.173)	(0.278)	(0.0874)	(0.166)	(0.166)
HH size 4	0.481**	0.0113	0.0489	-0.194	0.280	0.0680	0.188	0.0203
	(0.239)	(0.311)	(0.135)	(0.262)	(0.333)	(0.111)	(0.216)	(0.214)
HH size 5plus	0.989***	-0.415	-0.255	0.383	0.108	-0.0421	0.334	-0.107
	(0.247)	(0.363)	(0.203)	(0.367)	(0.353)	(0.140)	(0.315)	(0.356)
Constant	11.28***	10.86***	10.38***	11.01***	11.65***	11.74***	10.32***	10.94***
	(0.269)	(0.285)	(0.147)	(0.218)	(0.500)	(0.143)	(0.361)	(0.281)
Sample size (n)	813	777	6,663	1,251	410	844	1,042	2,404
Weighted in Mio. (N)	1.30	1.42	10.34	10.71	0.91	12.4	1.01	5.09

Reference groups: 3rd income quintile, age 45-54, education secondary, work status employed, marital status married, gender women, HH size.
Standard errors in parentheses. All 5 implicates are used, standard errors bootstrapped. *** p<0.01, ** p<0.05, * p<0.1
Source: Own computations from the HFCS survey wave 1 (2013).

inheritance-wealth ratio.[61] In Cyprus, the low inheritance-wealth ratio is the result of a very high mean net worth (European Central Bank, 2013b,c); the capitalized values of the transfers are rather low in comparison. The result for Belgium is surprisingly similar to most of the Mediterranean countries and differs greatly from France and Germany, which deserves an explanation. First, the percentage of households with a transfer is significantly lower than in Germany or France, in particular for households with a fortune larger than €1 million. Second, since the overall wealth level in Belgium is rather high–the median net worth of all households is almost four times as high as in Germany (see European Central Bank, 2013b,c)–, this results in an overall comparatively small fraction of the Belgian private wealth that can be attributed to inheritances and gifts.

While the correlation between absolute inheritance values and household income positions is quite high, the inheritance-wealth ratios barely differ across income quintiles. With increasing income the opportunities to save parts of the income increase as well. This results in stable inheritance-wealth ratios across income quintiles (see Austria, Belgium and France). In Germany, for the highest income quintile the inheritance-wealth ratio drops by about 17 percentage points as compared to the second highest income quintile. Overall, the high-income households receive significantly higher wealth transfers, but are equally capable of saving significant amounts through their personal efforts, resulting in a decreasing relevance of inheritances and gifts for their wealth position.

The analysis of the relative importance of transfers along the distribution of wealth reveals two sets of countries. In Austria and (West) Germany the share of capitalized wealth transfers is highest for the wealth level €500,000 to €1 million and quickly decreases for the net worth above €1 million. Cyprus exhibits a similar picture, albeit on an overall lower level. On the other hand in Belgium, France, Portugal and Spain the shares do not vary a lot between the wealth levels and stay approximately at their overall level. In Greece we observe a pronounced U-shaped pattern. We conclude that, as in most applications,

61 Keep in mind that at the time the survey was conducted in Spain, the aftermath of the financial crisis was not yet fully in effect; housing prices were still high. A repetition of the survey with more up-to-date data presumably would reveal another pattern.

the relative importance of wealth transfers does not significantly increase with the level of wealth. For the core European countries plus Cyprus it even decreases with a net worth higher than €1 million. This observation presents a stark contrast to some observations made before (Tables 5.1 and 5.2)–whereas the percentages of households with a transfer as well as the conditional present value of those transfers are increasing with the wealth level–the value of transfers as a percent of net worth drops for the wealthiest households. On the one hand, this result shows that those households accumulated a lot more of their large fortunes through their own efforts, independent of transfers. On the other hand, financially educated persons tend to have better options for investment, are less risk averse, and realize higher rates of return on their investment. The assumed real rate of return (3%) might be too low for those households. However, as can be seen in Appendix 5.B, the patterns are largely robust to both overall higher interest rates and wealth-related interest rates.

Transfers as a percent of net worth are steadily increasing over the lifecycle in Belgium and France, as well as in Portugal and Spain. This is in line with the result that the cohort effect does not offset the lifecycle effect in those countries (see Tables 5.3 and 5.4). As expected from those same results, the connection is less clear in Austria and Germany, for the older cohorts inheritance-wealth ratio is at around one-third. The younger cohorts exhibit rather high ratios as well, but inter-vivos transfers drive them: more than 50% of the transfers received are gifts. The high shares for younger generations hardly come as a surprise in Germany with rather generous tax exempt amounts (since 2009, €400,000 per child for an inter vivos transfer from each parent every ten years are free of tax, up from €205,000 before).

In several countries young households have a comparably high share, because of low initial savings levels. The shares are also high again for old households because of high absolute values of the capitalized wealth transfers. For middle-aged households the value of their own savings tends to be higher than their relative low absolute transfer value. The differences between the age classes are minimal though. In Belgium, France and Portugal (as well as in Spain) the inheritance-wealth ratio increases with age and peaks for the oldest cohort. In Austria, (West) Germany, Cyprus and Greece the ratio is surprisingly high for some or all young cohorts. One of the reasons why the ratio is not substantially higher for older cohorts might be the Second World War and its aftermath, resulting in a situation where there simply was not much to inherit by heirs of the war generation.

5.4.4 Correlates of the relative value of intergenerational transfers

Using a fractional logit model we further investigate the share of current wealth due to past wealth transfers for those who received a transfer. The advantage of this model is that it explicitly accounts for proportions in the $(0,1)$ interval. We estimate the following equation:

$$q_j = F(\alpha + \beta X_j + \varepsilon_j), \tag{5.6}$$

where q_j denotes the sum of past wealth transfers as a percent of current net worth for households, which received a transfer in country j. In addition to the inflation adjustment we capitalize transfers as a percentage of net worth–with a cap at 100%, i.e. the sum of capitalized wealth transfers within a household cannot be possibly higher than the net worth of a household. α is an intercept, ε_j denotes unobservables. X_j is the matrix of all explanatory variables: age, education, work and marital status as well as gender of the reference person, income of the household, and household size.

Table 5.6 shows the results for the fractional logit regressions analyzing capitalized inherited wealth in prices of 2010 as a percent of current household wealth. The sample is limited to households that received at least one gift or inheritance. The income of the household matters: Compared with the third income quintile, the first and second quintiles show a positive relationship and the fourth and fifth a negative one. This means that with increasing income, wealth transfers exhibit a decreasing impact on inherited wealth as a percent of net worth. Naturally, with higher incomes it is easier to save income and

Table 5.5: Present value of wealth transfers received as a percent of net worth, capitalized with $r = 3\%$.

	I. Core European countries								**II. Mediterranean countries**							
	Austria		Belgium		France		(W) Germany		Cyprus		Greece		Portugal		Spain	
All households	30.9	(4.2)	14.4	(1.0)	23.2	(0.8)	31.4	(2.6)	12.8	(1.5)	31.4	(1.8)	14.8	(1.4)	18.0	(1.1)
A. Income quintiles																
1st quintile	41.4	(7.6)	18.7	(4.1)	26.8	(2.2)	39.3	(9.4)	13.3	(5.7)	36.8	(4.1)	19.5	(2.3)	22.6	(1.8)
2nd quintile	30.6	(9.1)	14.7	(2.4)	26.8	(2.1)	36.1	(5.6)	14.5	(4.0)	40.1	(3.9)	21.5	(2.5)	20.8	(2.3)
3rd quintile	34.0	(7.6)	12.4	(2.0)	23.0	(1.9)	39.6	(4.4)	19.2	(6.5)	33.9	(4.1)	15.2	(2.1)	15.8	(3.7)
4th quintile	30.3	(6.6)	15.8	(2.3)	21.5	(1.5)	37.7	(5.0)	19.5	(3.7)	30.1	(4.6)	12.5	(1.7)	18.5	(2.1)
5th quintile	29.2	(5.0)	13.3	(1.8)	22.7	(1.3)	25.4	(3.5)	8.0	(1.9)	26.1	(3.0)	13.2	(2.6)	16.7	(1.9)
B. Wealth levels																
Under €20,000	-		-		-		-		-		-		-		-	
€20,000 - €99,999	25.8	(2.8)	16.6	(3.5)	21.0	(1.6)	17.5	(2.5)	14.7	(4.3)	34.5	(3.1)	18.4	(1.5)	16.3	(2.0)
€100,000 - €249,999	31.6	(2.6)	15.5	(2.3)	18.8	(0.9)	34.8	(2.9)	23.4	(3.4)	35.1	(2.0)	18.2	(1.4)	13.7	(1.5)
€250,000 - €499,999	36.1	(2.7)	14.8	(1.6)	23.1	(1.1)	38.5	(2.7)	20.4	(3.5)	27.2	(2.9)	11.8	(1.6)	15.9	(1.5)
€500,000 - €999,999	45.9	(4.6)	16.0	(2.3)	25.6	(1.6)	39.2	(4.5)	14.9	(2.8)	27.3	(5.1)	12.1	(2.8)	20.1	(3.2)
€1,000,000 or over	23.9	(6.7)	12.2	(2.0)	24.5	(2.1)	22.6	(4.6)	10.0	(2.0)	34.9	(13.6)	12.8	(4.6)	21.4	(3.5)
C. Age classes																
21-35	35.7	(8.4)	8.9	(2.4)	16.3	(2.2)	34.5	(8.1)	23.4	(3.8)	32.1	(3.0)	8.5	(2.1)	16.5	(3.9)
35-44	24.0	(7.1)	12.6	(3.0)	15.9	(1.3)	36.7	(3.9)	13.3	(2.4)	33.1	(2.8)	13.2	(2.7)	15.3	(2.3)
45-54	28.0	(5.4)	10.7	(1.8)	18.6	(1.3)	34.5	(3.0)	11.6	(2.9)	35.3	(3.6)	12.2	(1.5)	16.4	(2.0)
55-64	34.9	(6.8)	15.1	(2.0)	21.0	(1.8)	24.2	(5.3)	11.2	(3.0)	31.3	(4.8)	11.0	(2.6)	17.5	(2.6)
65-74	37.3	(6.3)	13.6	(1.9)	27.7	(1.7)	32.1	(4.0)	12.3	(5.0)	21.2	(2.8)	18.9	(3.8)	21.4	(1.8)
75 and older	34.8	(9.6)	21.9	(3.2)	38.5	(2.3)	31.7	(5.1)	11.5	(4.0)	30.2	(4.7)	25.7	(3.9)	22.3	(4.6)

Note: Standard errors are shown in parentheses. Means over 5 implicates, standard errors bootstrapped.
The figures show the present value of all wealth transfers as of the survey year which were received up to the time of the survey and accumulated at a real interest rate of 3.0% as a ratio to the respective net worth in the overall population or subpopulations.
Source: own computations from the HFCS survey wave 1 (2013).

accumulate wealth, thus, even though the absolute present value of transfers is higher for high income households, their relative importance is decreasing.

With regard to the age classes, the results do not reveal a unified pattern. It seems like the households over 65 have higher shares of current wealth due to transfers in comparison with the middle aged ones (45 to 54). However, this finding is only significant in France and Portugal. For the households under 45 the coefficients point into both directions, no matter if they are located in core or Mediterranean countries. Positive correlations give a hint that in Belgium and Spain younger households have already received large fortunes. So far they have had less time to accumulate wealth off their own income. Hence, transfers have a much higher impact on their financial situation than for older cohorts.

Self-employed households have a lower inheritance-wealth ratio than employees (except for Spain). However, in the analysis it is assumed that all accumulated wealth exceeding the capitalization is due to own efforts, if business owners inherited their business and consistently generate a higher rate of return, the resulting wealth is defined as savings. For the self-employed population, an initial transfer might be the reason for the latter wealth though. In the majority of the countries studied, singles have a higher share of current wealth due to past intergenerational transfers compared to households led by a person in marriage. For households led by a widowed or divorced person the share of past intergenerational transfers also tend to be higher, being divorced or widowed diminishes the possibilities to increase savings and accumulate wealth thorugh personal efforts. The gender of the household head does matter significantly, especially in the southern European countries and France. Men have a smaller inheritance-wealth ratio than women. As is shown in the first part, there are no significant differences for absolute present value of transfers between men and women, resulting in the overall conclusion that, all things equal, the lower ratio is the result of a gender pay gap, which enables men to save larger sums.

In summary, as compared to the analysis in absolute terms, many results are reversed. Especially the finding that the share of current wealth due to past intergenerational transfers is actually decreasing with income needs to be emphasized. Remember from the first part of this empirical analysis that those households with higher income have higher chances of receiving inheritances and gifts while also receiving larger transfers in absolute terms. This suggests that these households are able to build up wealth out of both their annual income as well as substantial inheritances and inter-vivos transfers.[62]

62 Keep in mind that the income variable is only a proxy for life-time earnings, as it does refer to the calendar year prior to the survey year (or the 12 months preceding the survey).

Table 5.6: Fractional logit regressions for share of current wealth due to past intergenerational transfers (heir population).

Fractional logit	AT	BE	FR	(W)DE	CY	GR	PT	ES
1st income quintile	0.457*	0.271	0.621***	0.006	1.059***	0.208	0.545***	0.382***
	(0.279)	(0.237)	(0.079)	(0.211)	(0.365)	(0.349)	(0.209)	(0.128)
2nd income quintile	0.105	-0.004	0.342***	0.120	0.419	0.222	0.266	0.123
	(0.237)	(0.200)	(0.072)	(0.192)	(0.314)	(0.289)	(0.187)	(0.133)
4th income quintile	-0.038	0.083	-0.001	-0.265**	0.211	-0.184	-0.280	-0.019
	(0.226)	(0.187)	(0.068)	(0.152)	(0.260)	(0.302)	(0.208)	(0.126)
5th income quintile	-0.295	-0.174	-0.158***	-0.670***	-0.393*	-0.147	-0.277	-0.349***
	(0.214)	(0.196)	(0.062)	(0.136)	(0.275)	(0.306)	(0.195)	(0.114)
Age 21-34	-0.092	0.468*	0.059	-0.608**	0.216	-0.684**	0.069	0.035
	(0.265)	(0.267)	(0.105)	(0.229)	(0.281)	(0.338)	(0.310)	(0.216)
Age 35-44	-0.244	-0.059	-0.066	0.018	-0.133	-0.526**	0.286	0.199*
	(0.209)	(0.212)	(0.072)	(0.149)	(0.205)	(0.274)	(0.209)	(0.133)
Age 55-64	-0.051	0.068	-0.045	-0.069	-0.042	0.071	0.140	-0.079
	(0.224)	(0.201)	(0.070)	(0.140)	(0.302)	(0.351)	(0.189)	(0.102)
Age 65plus	0.256	0.463	0.502***	0.232	0.440	0.140	0.661***	0.173
	(0.268)	(0.286)	(0.090)	(0.211)	(0.595)	(0.463)	(0.243)	(0.125)
Education primary	0.162	0.031	0.138***	0.026	-0.084	0.326	-0.275	-0.116
	(0.225)	(0.159)	(0.051)	(0.202)	(0.239)	(0.241)	(0.195)	(0.104)
Education tertiary	0.159	-0.103	0.156***	-0.177	0.110	0.055	-0.310	-0.137
	(0.214)	(0.134)	(0.049)	(0.096)	(0.191)	(0.280)	(0.255)	(0.102)
Work status self-employed	-0.176	-0.401	-0.477***	-0.079	-0.275	-0.727**	-0.274	0.186*
	(0.209)	(0.250)	(0.060)	(0.157)	(0.254)	(0.280)	(0.189)	(0.109)
Work status unemployed/other	-0.190	0.023	0.196*	-0.025	0.123	-0.510	-0.367*	0.099
	(0.242)	(0.222)	(0.097)	(0.159)	(0.313)	(0.296)	(0.216)	(0.111)
Work status retired	0.029	-0.221	-0.073	-0.247	-0.074	-0.040	0.225	0.075
	(0.228)	(0.239)	(0.076)	(0.189)	(0.584)	(0.407)	(0.211)	(0.121)
Marital status single	0.005	0.448**	0.188**	0.291	-0.075	0.728*	0.466*	0.435***
	(0.210)	(0.206)	(0.073)	(0.193)	(0.629)	(0.344)	(0.257)	(0.124)
Marital status widowed	0.080	0.729***	0.118	0.288	-0.449	0.726*	-0.075	0.216*
	(0.370)	(0.235)	(0.089)	(0.242)	(0.583)	(0.408)	(0.250)	(0.133)
Marital status divorced	0.054	0.609***	0.116	0.420**	0.726	0.616	0.210	0.114
	(0.260)	(0.221)	(0.083)	(0.181)	(0.440)	(0.437)	(0.292)	(0.152)
Gender man	-0.178	-0.095	-0.108**	-0.044	-0.331	-0.478**	-0.421**	-0.170**
	(0.143)	(0.119)	(0.046)	(0.100)	(0.181)	(0.217)	(0.179)	(0.081)
HH size 1	-0.027	0.035	0.126*	-0.028	-0.206	-0.696**	0.253	-0.030
	(0.242)	(0.206)	(0.073)	(0.186)	(0.565)	(0.372)	(0.221)	(0.127)
HH size 3	0.145	0.211	-0.058	0.012	-0.216	0.054	0.259*	0.069
	(0.228)	(0.192)	(0.067)	(0.136)	(0.291)	(0.299)	(0.165)	(0.096)
HH size 4	0.000	0.188	-0.006	-0.212	0.116	0.041	0.355*	0.031
	(0.333)	(0.227)	(0.074)	(0.154)	(0.287)	(0.292)	(0.201)	(0.110)
HH size 5plus	0.103	-0.125	0.035	0.034	0.009	0.164	0.882***	0.229
	(0.339)	(0.290)	(0.095)	(0.220)	(0.336)	(0.391)	(0.285)	(0.162)
Constant	0.688**	-0.703***	-0.394***	0.747***	-0.003	2.363***	0.260	-0.204
	(0.279)	(0.233)	(0.088)	(0.181)	(0.391)	(0.391)	(0.332)	(0.161)

Reference groups: 3rd income quintile, age 45-54, education secondary, work status employed, marital status married, gender women, HH size.

Standard errors in parentheses. All 5 implicates are used, standard errors bootstrapped. *** p<0.01, ** p<0.05, * p<0.1

Source: Own computations from the HFCS survey wave 1 (2013).

5.5 Conclusion

We conduct a detailed investigation of the distribution of wealth transfers in eight countries in the Euro-area (Austria, Belgium, France, (West) Germany, Cyprus, Spain, Greece, and Portugal). Our main finding is that high-income households in European countries have, in the past, inherited significantly higher amounts than low-income households. Using a series of country-specific regressions, we confirm these findings and additionally discover that high education levels strongly correlate with both the probability of receiving a transfer and the value of those transfers. At the same time, capital transfers seem to be less relevant to the current asset position for high-income households. Through their strong income positions and persistently low intergenerational mobility, these households are presumably able to build wealth both from their regular incomes *and* from inheritances and gifts.

Overall we observe quite similar patterns in all European countries included in our sample. The share of households that received at least one wealth transfer varies between 27% (Portugal) and roughly 40% (France). The capitalized conditional mean present value of wealth transfers falls between €274,000 (Cyprus) and €85,000 (Portugal). Expressing the mean present value of transfers in relative terms, as a percent of current net worth, it never exceeds 32% and shares are lower in the Mediterranean countries (Greece deviates from the other Mediterranean countries, as does Belgium compared to the rest of core Europe). In most countries the percentages of households with a transfer as well as the mean present value of those transfers is expectedly increasing along the distribution of net worth. However, the importance of those transfers for the current net worth level does not increase with the level of wealth. In addition, we find that self-employed households tend to have a higher prevalence, compared with employees, to have received a transfer, and those transfers tend to be higher than those of employees.

Overall inheritances and gifts may be considered as a channel through which the existing inequality of opportunity and the resulting economic inequality are amplified. In Germany, for instance, taxes on inheritances and gifts are virtually regressive due to its comprehensive exemptions on large assets (Bach and Thiemann, 2016). In Austria, Cyprus, Portugal, and some regions of Spain, inheritance and gift tax has been abolished or abandoned de facto. We observe the pattern that households from higher income quintiles are able to accumulate more wealth through an increased capacity to save on their own. In addition, once high income households report an inheritance or inter vivos transfer, the values are substantially higher than for low income households, thereby increasing the gap between rich and poor households. If policy aims to reduce wealth inequality and, more generally, economic inequality, it must therefore revisit the strong link between high incomes and high expected values of wealth transfers.

5.A Taxation of inheritances and gifts: a European comparison

Country	Reference period: 2000-2010	Tax rate depending on level of relation(1)		Max. tax rate threshold	Max. tax allowance (renewed)	Exemptions / special regulations
I. No or low inheritance and gift tax						
Cyprus	since 2000	**No inheritance or gift tax, but land transfer tax for gifts**				
		Spouses, children	3-8%	€170,860 (since 2008, €100,000 before)	—	business transfers within families
		Other persons	3-8%	€170,860	—	—
Austria	since 2008	**No inheritance or gift tax, but land transfer tax**				
		Spouses, children	2%	—	€1,100	business transfers
		Other persons	2-3.5%	—	€1,100	business transfers
	before 2008	**Moderate inheritance and gift tax with low allowances**				
		Spouses, children	2-15%	€4,380,000	€2,200 (10 yrs.)	business transfers
		Other persons	4-60%	€4,380,000	€110/440/ 2,200 (10 yrs.)	business transfers
Portugal	since 2004	**Stamp duty**				
		Spouses, children	0% inheritance; 0.8% property gift	—	—	business transfers (tax rate 25%)
		Other persons	0/10% inheritance; 0.8/10.8% property gift	—	—	business transfers (tax rate 25%)
	before 2004	**Moderate inheritance and gift tax with low allowances**				
		Spouses, children	3-24%	€355,343	€3,641 tax free, children under age tax free (never)	—
		Other persons	7-50%	€355,343	€374, plus €1,820 if inheritance in ascending line (never)	—
II. Moderate inheritance and gift tax rate with moderate or high allowances						
Greece	since 2010	**inheritance and gift tax**				
		Spouses, children	1-10%	€600,000	€400,000 if inheritance - married at least 5 years, only children under age	primary residence, shares and business transfers

Country	Reference period: 2000-2010	Tax rate depending on level of relation(1)		Max. tax rate threshold	Max. tax allowance (renewed)	Exemptions / special regulations
		Other persons	1-40%	€267,000	€6,000-€30,000 from this amount on taxes are due, depending on level of relation	shares and business transfers
	before 2010	Numerous changes, e.g. tax allowances (2004: €19,076 spouses and children) tax rates (2004: 5-25% and up to 60% for other persons, 2008: depending on asset: for spouses and children property max. 1%, shares max. 0.6%)				
Germany	since 2010	**inheritance and gift tax**				
		Spouses, children	7-30%	€26,000,000	€500,000, €400,000 for children, (10 yrs.)	owner-occupied property, business transfers
		Other persons	7-50%	€26,000,000	€20,000/ 100,000/ 200,000 (10 yrs.)	business transfers
	before 2010	less exemptions, lower tax allowances, thresholds in tax brackets lower lower tax rate for some 'other persons'				

III. High or moderate inheritance and gift tax rate with low or moderate allowances

Spain	since 2010	**inheritance and gift tax (on national level, regional differences)**				
		Spouses, children	7.65-34% + multiplier: 1-1.2%*	€797,555, multiplier depending on heir's wealth (max. threshold €4,020,770)	€15,956, €47.858 for children under age (3 yrs.)	business transfers, property
		Other persons	7.65-34% + multiplier: 1.59-2.4%*	same as spouses	€0/ €7.993/ €15,956 (3 yrs.)	business transfers, property
		*) The corresponding tax rate (amount of transfer relevant) is applied to the taxable amount. The resulting balance is then multiplied with the corresponding multiplier (results from the existing assets of the heir and the degree of relationship).)				
	before 2010	hardly changes (e.g. lower allowances), but regional governments may deviate from national legislation since 2004, this resulted in tax exemptions of up to 99% of estate value				
France	since 2000	**inheritance and gift tax**				
		Spouses, children	5-45% (except for spouses since 2008)	€1,805,677	€156,956 (10 yrs.)	business transfers, tax reduced if three children under age
		Other persons	5-60%	€0-1,805,677	€1,520-€156,359 (10 yrs.)	
Belgium	since 2010	**inheritance tax (regional differences)**				

Country	Reference period: 2000-2010	Tax rate depending on level of relation(1)		Max. tax rate threshold	Max. tax allowance (renewed)	Exemptions / special regulations
		Spouses, children	3-30%	€250,000-€500,000	€15,000-€25,000, €65,000-€75.000 for children under age (3 yrs.)	owner-occupied property, business assets, and others depending on region
		Other persons	3-80%	€75,000-€500,000	€620-1,250, €15,000-25,000 (3 yrs.)	—
	since 2010	**gift tax (regional differences)**				
		Spouses, children	1-30% (max. 7.7% for movable assets)	€500,000	—	owner-occupied property, business assets, and others depending on region
		Other persons	1-80% (max. 7.7% for movable assets)	€75,000-€500,000	—	—
	before 2010	Regional legislation of gift tax possible since 2001, inheritance tax since 2002				

Table 5.7: Taxation of inheritances and gifts: a European comparison.
(1) In some countries spouses and partners have the same legal rights. This is not documented here.
Sources: Legal texts from individual countries, Mennel and Förster (2014), Schupp and Szydlik (2004) and EY (2014)).

5.B Robustness checks

In Table 5.5 we assume that the time invariant interest rate on the investment of all wealth transfers is 3% for all households. In order to check the impact of this assumption on the relative importance of wealth transfers for the net worth along the distribution, we conduct a series of robustness checks. Additionally, as the wealth transfers received are not independent of the number of adults in a household, we check by how much the present value of transfers would change, if we apply a per-capita definition of transfers.

Long-term interest rates on government bonds

In Section 5.4 we argue that a secure investment would be in line with a rate of return of three percent ($r = 3\%$), as this is a capitalization rate quite common in the literature (Wolff and Gittleman, 2014). Alternatively, one might assume that the most secure investment a citizen may choose is a long-term investment in government bonds (cf. Corneo et al., 2016). The nominal rate of return then is the (yearly average) nominal yield of such an investment. The data does not allow us to compute the resulting capitalized values of inheritances and gifts, as the time series are not entirely available for any of the Mediterranean countries. However, they are available for Belgium, France and Germany from the 1950s onward. In Table 5.8 the results are shown for a capitalization of past inheritances and gifts using the nominal yields of long-term government bonds.

This change of method would have almost no effect on the overall inheritance-wealth ratio, the maximum deviation would be in France with +0.9%. For the individual wealth classes all changes are bellow one percent, no patterns are visible. For household income this change would affect lower quintiles slightly more, but again the changes are negligible. The shares are somewhat higher for the older cohorts, probably due to higher interest rates on government bonds in the 1970s and 1980s as compared to a real interest rate of 3%. The variation for both the conditional mean and median value of transfers received is below €5,000. In summary, applying government bonds instead of a flat real interest rate hardly affects the outcomes for the countries where time series are available.

Real interest rates $r = 1\%$ versus $r = 5\%$

The second robustness check assesses the impact of a flat low versus a flat high interest rate. We compare the different outcomes of $r = 1\%$ and $r = 5\%$ on the wealth transfers received as a percent of net worth and conditional mean and median present values (Table 5.9). Most importantly, the general patterns we observe along the distribution of wealth are largely independent of the chosen real interest rate, even though the higher wealth classes are affected more by a higher rate of return. For the income quintiles there are no changes of the patterns visible either. The overall increase of share is the lowest in Portugal (3.9 percentage points) and highest in Austria (9.5 percentage points). However, in most countries the increase is spread almost equally among the income quintiles. Only in Belgium the lowest quintile seems to be affected slightly more, in (West) Germany and Cyprus the middle income classes are experiencing a slightly sharper surge. The conditional mean values are varying considerably between low and high interest rates: in Cyprus the mean is up by about €121,000, in Greece and Portugal it is affected the least (around +€25,000). For the remaining country the difference varies between €47,000 and €71,000.

Wealth related interest rates

However, assuming that the interest rate is the same no matter the position along the distribution of wealth may not seem reasonable. It is more likely that households with a higher level of wealth are better informed about financial markets and investment opportunities. In addition they hold enough money to be able to divide it into different investments; consequently they might take higher risks and realize higher rates of returns than the middle class or households from the bottom half of households. Hence, in this last step we assume that the real interest rate correlates with the net worth position: The wealth class below €20,000 includes a significant number of net borrower and zero wealth observations and is excluded from the analysis. The next class realizes an interest rate of 3%, which then is increasing with every wealth class by 1%, thus leading to an interest rate of 7% for households with net worth higher than €1 million. We then compare the results for the assumption that all realize the same real interest rate (3%) to the wealth related interest rate in Table 5.10.

As expected, the changes in percentage points are highest for the highest wealth class. In comparison to a flat real interest rate the changes vary between 2.8 percentage points in Portugal and 10.8 in Greece. In the core countries Germany, Austria and Belgium and in Cyprus the second wealthiest class stays ahead of the top class after adjusting to a wealth related interest rate, only in France we observe a change, albeit the difference is not statistically significant. As for the Mediterranean countries, compared to a real annual interest rate of 3% we do not observe any considerable structural differences. The conclusion that there is relatively small variation in the importance of inheritances and gifts for net worth between the wealth classes is still viable. We conclude that this observation is not the result of an arbitrarily chosen interest rate.

Per capita present value

Computing the per capita wealth transfers (Table 5.11) instead of the total present values per household (Table 5.2), yields a similar picture, albeit on an expectedly lower level. Mean and median values are depending on the household sizes, which are somewhat higher in some Mediterranean countries. The patterns we observe do not change: the present values are increasing with the level of wealth, exhibit various patterns depending on the age class and are again highest for the highest income quintile. We conclude that switching to per capita transfers would have reduced the numbers, but not changed the outcomes of our study.

Table 5.8: Present value of wealth transfers received as a percent of net worth, capitalized using country-specific yields of long-term government bonds.

	Belgium		France		Germany	
All households in %	14.7	(1.0)	24.1	(0.8)	31.9	(2.7)
Cond. mean present value in €	158,412	(10,140)	142,615	(4,205)	196,039	(13,096)
Cond. median present value in €	79,177		46,665		106,981	
A. Wealth levels in %						
Under €20,000	-		-		-	
€20,000 - €99,999	16.3	(3.4)	21.3	(1.7)	17.5	(2.5)
€100,000 - €249,999	15.5	(2.3)	19.2	(0.9)	35.0	(2.8)
€250,000 - €499,999	15.3	(1.7)	23.9	(1.1)	39.0	(2.7)
€500,000 - €999,999	16.4	(2.3)	26.4	(1.6)	39.3	(4.5)
€1,000,000 or over	12.6	(2.0)	25.8	(2.2)	23.5	(4.9)
B. Income quintiles in %						
1st quintile	19.3	(4.0)	27.9	(2.3)	39.9	(9.5)
2nd quintile	15.7	(2.5)	28.0	(2.2)	36.6	(5.6)
3rd quintile	12.8	(2.1)	23.8	(1.9)	40.1	(4.5)
4th quintile	15.9	(2.3)	22.1	(1.5)	38.0	(5.0)
5th quintile	13.4	(1.8)	23.6	(1.3)	26.0	(3.7)
C. Age classes in %						
21-35	19.3	(4.0)	27.9	(2.3)	39.9	(9.5)
35-44	15.7	(2.5)	28.0	(2.2)	36.6	(5.6)
45-54	12.8	(2.1)	23.8	(1.9)	40.1	(4.5)
55-64	15.9	(2.3)	22.1	(1.5)	38.0	(5.0)
65-74	13.4	(1.8)	23.6	(1.3)	26.0	(3.7)
75 and older	19.3	(4.0)	27.9	(2.3)	39.9	(9.5)

Source: own computations from the HFCS survey wave 1 (2013).
Standard errors are shown in parentheses. All 5 implicates are used, standard errors bootstrapped.
The figures show the present value of all wealth transfers as of the survey year which were received up to the time of the survey and accumulated at a nominal interest rate of long-term government bonds as a ratio to the respective net worth in the overall population or subpopulations.

Table 5.9: Present value of wealth transfers received as a percent of net worth, real interest rate = 1 versus real interest rate = 5.

	I. Core European countries								**II. Mediterranean countries**							
	Austria		Belgium		France		(W) Germany		Cyprus		Greece		Portugal		Spain	
	r=1%	r=5%	r=1%	r=5%	r=1%	r=5%	r=1%	r=5%	r=1%	r=5%	r=1%	r=5%	r=1%	r=5%	r=1%	r=5%
All households	25.3	34.8	11.7	17.2	19.2	27.2	26.3	35.8	10.2	15.9	28.5	33.7	12.7	16.6	14.5	21.4
Cond. mean present value in €1000	188	259	125	185	114	161	162	220	218	339	138	164	73	95	141	207
Cond. median present value in €1000	94	130	58	94	37	56	89	125	114	199	104	119	31	42	60	97
A. Wealth levels																
Under €20,000	-		-		-		-		-		-		-		-	
€20,000 - €99,999	23.6	27.6	15.5	17.5	18.6	23.0	15.2	19.5	12.0	15.3	33.4	35.0	16.8	19.6	14.3	17.6
€100,000 - €249,999	28.2	34.1	13.3	17.0	15.8	21.5	31.0	38.0	20.2	25.8	33.0	36.4	15.8	20.0	11.4	15.9
€250,000 - €499,999	31.0	39.6	12.3	17.4	19.0	26.8	32.8	43.5	16.8	23.7	23.7	30.1	9.8	14.0	13.1	18.5
€500,000 - €999,999	38.1	50.7	12.6	19.2	21.4	29.9	33.4	42.8	11.1	18.7	23.6	30.4	9.6	14.0	16.1	23.8
€1,000,000 or over	17.9	28.4	9.5	15.5	19.9	29.6	17.5	27.9	8.0	13.0	28.7	41.8	10.6	14.3	16.8	26.5
B. Income quintiles																
1st Quintile	35.9	46.6	14.6	22.8	22.5	30.8	36.0	41.6	9.9	15.4	34.1	38.1	16.6	22.1	18.1	26.1
2nd Quintile	25.9	34.2	11.1	18.2	22.4	31.1	30.9	40.9	11.1	18.4	37.5	42.3	18.7	24.3	17.7	23.8
3rd Quintile	28.8	37.4	9.7	15.1	19.5	26.4	33.6	43.8	15.5	23.8	31.2	35.7	13.1	17.3	12.6	19.4
4th Quintile	25.6	34.0	13.3	18.3	17.5	25.1	32.0	42.1	15.8	23.7	28.0	32.0	10.9	13.9	15.1	21.0
5th Quintile	22.9	33.4	11.2	15.7	18.7	26.9	20.7	30.0	6.4	10.2	22.5	29.3	11.1	14.5	13.2	20.5

Note: Standard errors are shown in parentheses. Means over 5 implicates, standard errors bootstrapped.
The figures show the present value of all wealth transfers as of the survey year which were received up to the time of the survey and accumulated at a real interest rate of rate either $r = 1\%$ or $r = 5\%$ as a ratio to the respective net worth in the overall population or subpopulations.
Source: own computations from the HFCS survey wave 1 (2013).

Table 5.10: Present value of wealth transfers received as a percent of net worth, real interest rate = 3 versus wealth related interest rates.

	Austria	Belgium	France	(W)Germany	Cyprus	Greece	Portugal	Spain
Wealth levels								
A. real interest *rate* = 3%								
Under €20,000	-	-	-	-	-	-	-	-
€20,000 - €99,999	25.8	16.6	21.0	17.5	14.7	34.5	18.4	16.3
€100,000 - €249,999	31.6	15.5	18.8	34.8	23.4	35.1	18.2	13.7
€250,000 - €499,999	36.1	14.8	23.1	38.5	20.4	27.2	11.8	15.9
€500,000 - €999,999	45.9	16.0	25.6	39.2	14.9	27.3	12.1	20.1
€1,000,000 or over	23.9	12.2	24.5	22.6	10.0	34.9	12.8	21.4
B. wealth related interest rate								
Under €20,000	-	-	-	-	-	-	-	-
€20,000 - €99,999	25.8	16.6	21.0	17.5	14.7	34.5	18.4	16.3
€100,000 - €249,999	32.9	16.3	20.2	36.6	24.7	35.8	19.2	14.8
€250,000 - €499,999	39.6	17.4	26.8	43.5	23.7	30.1	14.0	18.5
€500,000 - €999,999	52.6	20.8	31.9	44.3	20.1	31.8	14.7	25.4
€1,000,000 or over	32.7	18.9	34.3	31.7	16.7	45.6	15.5	30.6

Note: Standard errors are shown in parentheses. Means over 5 implicates, standard errors bootstrapped.
The figures show the present value of all wealth transfers as of the survey year which were received up to the time of the survey and accumulated at a real interest rate either $r = 3\%$ or wealth related, i.e. from €100,000 onwards the interest rate is increasing in steps of one, yielding an interest rate of $r = 7\%$ for the highest wealth level.
Source: own computations from the HFCS survey wave 1 (2013).

Table 5.11: Per capita mean present value of transfers received (in €1,000), in 2010 prices and capitalized with $r = 3\%$, recipients only.

	I. Core European countries								**II. Mediterranean countries**							
	Austria		Belgium		France		(W) Germany		Cyprus		Greece		Portugal		Spain	
A. Income quintiles																
1st quintile	90	(19)	90	(16)	62	(5)	84	(20)	97	(29)	72	(7)	38	(4)	78	(8)
2nd quintile	100	(13)	85	(11)	71	(6)	100	(17)	80	(15)	68	(6)	32	(3)	70	(9)
3rd quintile	132	(19)	98	(16)	68	(8)	90	(12)	157	(57)	70	(11)	30	(4)	76	(21)
4th quintile	110	(18)	100	(16)	68	(4)	98	(11)	152	(24)	71	(14)	26	(3)	81	(9)
5th quintile	156	(23)	114	(23)	136	(7)	137	(14)	153	(28)	85	(7)	87	(15)	131	(15)
B. Wealth levels																
Under €20,000	5	(1)	5	(1)	5	(0)	6	(1)	6	(2)	11	(1)	7	(1)	7	(1)
€20,000 - €99,999	42	(2)	32	(3)	27	(1)	38	(2)	45	(4)	48	(1)	29	(1)	30	(2)
€100,000 - €249,999	105	(6)	67	(6)	66	(2)	97	(5)	91	(10)	117	(4)	67	(5)	68	(5)
€250,000 - €499,999	203	(15)	129	(12)	141	(7)	205	(14)	153	(22)	192	(26)	146	(26)	157	(12)
€500,000 - €999,999	448	(52)	232	(37)	281	(18)	357	(49)	170	(43)	429	(117)	174	(49)	245	(26)
€1,000,000 or over	552	(161)	483	(122)	787	(77)	501	(92)	436	(109)	786	(485)	1097	(403)	909	(255)
C. Age classes																
21-35	100	(28)	33	(8)	28	(3)	65	(20)	142	(22)	74	(7)	29	(6)	75	(18)
35-44	99	(16)	65	(13)	57	(5)	92	(13)	157	(25)	76	(5)	36	(5)	84	(12)
45-54	132	(15)	66	(9)	66	(5)	96	(10)	97	(13)	73	(10)	27	(3)	82	(13)
55-64	128	(15)	98	(15)	83	(5)	106	(15)	140	(40)	83	(10)	35	(6)	86	(9)
65-74	150	(34)	100	(10)	114	(8)	137	(13)	150	(68)	53	(5)	56	(11)	89	(8)
75 and older	115	(26)	186	(32)	147	(11)	129	(18)	100	(32)	73	(12)	58	(9)	114	(25)

Note: Standard errors are shown in parentheses. Means over 5 implicates, standard errors bootstrapped.
The figures show the present value of all transfers as of the survey year which were received up to the time of the survey in prices of 2010 using country specific inflation rates.
Source: own computations from the HFCS survey wave 1 (2013).

Summary

All research contributions within this thesis are guided by the assumption that evidence about the distribution of wealth ought to be indispensable for concrete decision making in social and fiscal policy. This holds equally true for official statistics, scholars, and the private sector. In Germany, a flagrant lack of official register or tax data for scholarly use leads to a situation wherein survey data is the last remaining source of evidence about the distribution of wealth. Two of the four research chapters in this thesis aim to evaluate methods for the improvement of available survey data both in Germany and internationally. The other two contributions discuss the possibilities and limitations of survey data for the analysis of the joint distribution of wealth and wealth transfers using both German and ex-ante harmonized Euro-area data.

The first research contribution 'Longitudinal wealth data and multiple imputation – An evaluation study' is a simulation exercise that intends to improve the imputation of missing values in wealth surveys. One distinctiveness of the statistical distribution of wealth assets is their high skewness. Since in panel studies such as the Socio-economic Panel study (SOEP) values from past or future waves might be available, the question arises how the imputation method may be modified–explicitly considering this longitudinal information–in order to improve the estimates of trend, inequality, and mobility analyses. Through blanking out of actually observed data points under several assumptions concerning the non-response generating mechanism–missing at random (MAR) or differential non-response for rather low or high values–we generate simulation data sets with missing values. We test the state-of-the-art imputation method *multiple imputation by chained equations* (MICE), a specification based on regressions with Heckman correction for sample selection, and

a third procedure specifically designed for panel data called *row-and-column imputation*. MICE and the regression approach serve as fallback methods, if only cross-sectional data are available. We choose six evaluation criteria that are particularly tailored to wealth data; additionally, we add three criteria for the evaluation of effects on wealth mobility and overall statistical inference. Considering trend and inequality estimates, the univariate row-and-column methods performs surprisingly well. Its combination with MICE as fallback procedure unanimously improves the imputation quality for all asset types considered. However, researchers interested in wealth mobility might prefer the imputation with MICE as basic *and* fallback, as it best replicates the mobility structures observed in the original data. Overall, we conclude that, for highly skewed panel data, data providers and users are advised to not dismiss the row-and-column method beforehand.

The chapter 'Estimating top wealth shares using survey data – An empiricist's guide' picks up where the previous one left off: the treatment of missing values (item non-response) is not sufficient if individuals or complete households refuse to participate in a survey (unit non-response). Systematic unit non-response in surveys leads to a middle class bias, particularly in income and wealth data, as the participation is correlated with a household's financial situation. The absence of wealthy households yields estimates of the aggregate net worth and top wealth shares that are heavily biased downward. In a series of Monte Carlo experiments this contribution shows that using maximum likelihood techniques to simulate a Pareto distributed top tail does not improve the estimates, as aggregate wealth and shares are still biased downward. The addition of rich list data does improve–and potentially overestimate–top wealth shares, while still yielding aggregate numbers that are too low. The reason is that rich list data merely replicates the billionaires of the distribution, whereas the percentiles 90 to 99.9 are underrepresented. Moreover, a systematic overvaluation of billionaires' fortunes by rich list data editors yields, ceteris paribus, aggregate numbers and shares that are heavily biased upward. In an application to German wealth data it is shown that a re-calibration of the provided frequency weights based on exogenous information has a much stronger impact on the results than choosing the right parameters of a Pareto distributed tail.

The fourth chapter ‘Breaking down Germany’s private wealth into inheritance and personal efforts – A distributional analysis’ takes advantage of the records of both household wealth and wealth transfers received from inheritances and gifts in the questionnaire of the survey *Panel on Household Finances*, which enables us to compute the significance of inheritance for different quantiles of the distribution of wealth in Germany. We define inherited wealth as a percent of net worth at the household level and split them into two categories: ‘rentiers’, whose assets are completely inherited, and ‘savers’, who saved off their regular income to accumulate wealth. This definition is useful to compute an improved inheritance-wealth ratio, albeit still many assumptions are attached. We show that wealth inequality, at least for 99 % of the German population, is hardly the result of an unequal distribution of inherited wealth: the ratio is one-third and barely changes along the distribution of wealth. This observation is nigh on identical for the sub population of retirees. The addition of pension wealth reduces the significance of inherited wealth for a household’s financial situation particularly in the bottom half of the wealth distribution, whereas the numbers in the upper class (the top percentile) decrease by a few percentage points only. In a series of robustness checks we show that the low importance of inherited wealth in the top percentile is not an artifact of a conservatively chosen rate of return. However, once we combine PHF data with exogenous sources on aggregate and inherited wealth we arrive at an inheritance-wealth ratio of over 80 % for the wealthiest households.

The final research contribution ‘Comparing the joint distribution of intergenerational transfers, income and wealth across the Euro area’ expands upon the previous chapter with additional analyses in a European context. With the release of the Eurosystem Household Finance and Consumption Survey we are able to determine the inheritance-wealth ratio comparatively for Austria, Belgium, France, West Germany, Cyprus, Greece, Portugal, and Spain for the first time. This contribution presents prevalence rates of households to receive wealth transfers, their absolute figures as well as their relative importance for household wealth. Using logit, OLS and fractional logit regressions we control for a multitude of socio-economic characteristics simultaneously. The relationship between the propensity to receive an inheritance and income is stronger in core Europe than in the Mediterranean countries, indicating less intergenerational mobility, whereas the relationship between an inheritance's value and a household's income is high in all countries. As expected, the present value of inheritances received increases with household net worth, however, considering the inheritance-wealth ratios we see some of the results reversed: the ratio does not correlate with wealth or income. In fact, the fractional logit regression shows that the significance of wealth transfers for household wealth is negatively correlated with income.

German summary

Allen Forschungsbeiträgen dieser Dissertation liegt die Annahme zugrunde, dass die korrekte Erfassung der Verteilung von Vermögen unerlässliche Grundlage konkreter sozial- und steuerpolitischer Entscheidungsfindung sein sollte. Ob die Erfassung durch die offizielle Statistik, die Wissenschaft oder durch privatwirtschaftliche Institutionen erfolgt, spielt hierbei keine Rolle. Ein eklatanter Mangel an Register- oder Steuerdaten zur Nutzung von Forschern in Deutschland führt zur Situation, dass Umfragedaten die einzig nutzbaren Anhaltspunkte zur Vermögensverteilung geben. Zwei der vier Forschungsbeiträge dieser Dissertation beschäftigen sich mit der Frage, wie zum einen die Qualität der Umfragedaten – sowohl in Deutschland als auch international – verbessert werden kann. Zwei weitere Beiträge erörtern die Möglichkeiten und Limitationen von Forschungsdaten zur gemeinsamen Verteilung von Vermögen und Erbschaften sowohl am Beispiel von Deutschland als auch mittels eines ex-ante harmonisierten europäischen Datensatzes.

Der erste Forschungsbeitrag „Longitudinal wealth data and multiple imputation – An evaluation study“ ist ein Simulationsprojekt, das die bessere Imputation von fehlenden Werten in Vermögenssurveys zum Ziel hat. Eine Besonderheit der statistischen Verteilung von Vermögenskomponenten ist ihre extreme Schiefe. Da in Panelstudien wie dem Sozio-ökonomischen Panel (SOEP) aber Werte aus vergangenen oder zukünftigen Wellen für die Imputation von fehlenden Werten zur Verfügung stehen können, stellt sich die Frage, wie durch eine Imputationsmethode, welche explizit die Panelstruktur berücksichtigt, bessere und präzisere Schätzungen für Trend-, Ungleichheits- und Mobilitätsanalysen erreicht werden können. Durch das Ausblenden eigentlich beobachteter Datenpunkte nach verschiedenen Annahmen zum bestimmenden Mechanismus – Missing at Random (MAR) oder differenziellen Non-Response für besonders hohe oder niedrige Vermögenswerte – generieren wir Simulationsdatensätze mit fehlenden Werten. Wir testen das im Moment meistgenutzte Imputationsverfahren *Multiple Imputation by Chained Equations* (MICE), eine Spezifikation basierend auf einer Regression mit Heckman-Korrektur für Selektivität, sowie die speziell für Paneldaten entwickelte *Row-and-Column*-Imputation. Auf MICE und Regression greifen wir zurück, wenn nur Querschnittsdaten zur Verfügung stehen. Die sechs Evaluationskriterien sind speziell auf Vermögensdaten zugeschnitten, drei zusätzliche Kriteren dienen der Abschätzung der Effekte auf Vermögensmobilität und Inferenz. Die univariate *Row-and-Column*-Imputation schneidet entsprechend der Trend- und Ungleichheitsanalysen erstaunlich gut ab. Ihre Kombination mit MICE als Fallback-Methode hat für alle betrachteten Vermögenskomponenten die Imputationsqualität erhöht. Für Vermögensmobilität zeigt sich auf der anderen Seite, dass ein reiner MICE-Ansatz die bevorzugte Wahl darstellt, da es die in den Originaldatensätzen beobachteten Mobilitätsstrukturen am besten repliziert. Insgesamt sollten Datenbereitsteller oder Anwender die Imputation durch die Row-and-Column-Methode nicht ohne Überprüfung verwerfen, falls die Impuationsvariablen stark schief sind und Paneldaten zur Verfügung stehen.

Das nächste Kapitel „Estimating top wealth shares using survey data – An empiricist's guide“ knüpft daran an, da die Behandlung fehlender Werte (Item-Non-Response) noch nicht ausreicht, um der Tatsache Rechnung zu tragen, dass Individuen oder Haushalte komplett die Teilnahme verweigern (Unit-Non-Response). Systematischer Unit-Non-Response führt in Umfragen zu einer Verzerrung zur Mittelschicht, besonders mit Blick auf Einkommens- und Vermögensdaten, da die Teilnahme mit der finanziellen Situation des Haushalts korreliert ist. Da Hochvermögende (fast) gar nicht teilnehmen, werden sowohl das Gesamtvermögen als auch die Anteile der oberen Perzentile am Gesamtvermögen in Umfragen stark unterschätzt. In einer Reihe von Monte-Carlo-Simulationen zeigt dieser Beitrag, dass die Zuschätzung Top-Vermögender durch einen Pareto-simulierten Rand mittels Maximum-Likelihood-Methode sowohl Gesamtvermögen als auch Anteilswerte deutlich unterschätzt. Die Hinzunahme von Daten aus sogenannten Reichenlisten führt zu einer Verbesserung bis hin zur Überschätzung der Vermögensanteile der Top-Vermögenden, obwohl das Gesamtvermögen immer noch unterschätzt wird, da zwar durch diese Methode die absolute Spitze repliziert wird, die 90. bis 99.9 Perzentile aber trotzdem unzureichend repräsentiert bleiben. Zusätzlich wird gezeigt, dass eine Überschätzung der Vermögen in Reichenlisten, ceteris paribus, zu einer eklatanten Überschätzung der Gesamtvermögen führt. Die Anwendung auf deutsche Paneldaten zeigt, dass die Korrektur der Gewichte der Umfragedaten auf der Basis externer Angaben viel stärkere Effekte auf die Schätzer hat als die Wahl der Parameter der Pareto-Verteilung.

Das vierte Kapitel „Breaking down Germany's private wealth into inheritance and personal efforts – A distributional analysis" betrachtet die gemeinsame Verteilung von Privatvermögen und erhaltener Vermögenstransfers aus Erbschaften und Schenkungen im neuen Haushaltspanel *Private Haushalte und ihre Finanzen* zur Berechnung der Bedeutung von Erbschaften entlang der Vermögensverteilung. Durch Definition des Anteils der Erbschaften und Schenkungen am Privatvermögen auf der Haushaltsebene können selbige in zwei Kategorien getrennt werden: Erben, welche ihr gesamten Vermögen aus Übertragungen bezogen haben, sowie Sparer, welche zusätzlich aus eigenen Mitteln Vermögen angehäuft haben. Diese Definition erlaubt eine realitätsnähere, wenngleich mit Annahmen behaftete, Berechnung der Rolle von Erbschaften als klassischerweise in der Literatur üblich. Es zeigt sich, dass die Vermögensungleichheit für 99 % der Bevölkerung kaum durch erhaltene Erbschaften zu erklären ist: ihr Anteil am Vermögen liegt bei etwa einem Drittel und ändert sich minimal entlang der Vermögensverteilung. Für Rentner und Pensionäre ist dasselbe Muster zu beobachten. Die Hinzunahme des erwarteten Rentenvermögens reduziert die Bedeutung von Erbschaften vor allem für ärmere Dezile, während sich die Werte in der Oberschicht (dem reichsten Perzentil) nur um wenige Prozentpunkte nach unten verschieben. Verschiedene Robustheitsanalysen zeigen, dass die geringe Bedeutungen von Erbschaften für die Vermögendsten kein Artefakt einer zu konservativen Kapitalisierung ist. Sobald die Daten des PHF allerdings mit anderen Schätzwerten zum Gesamtvermögen und Erbschaftsaufkommen verknüpft werden, ergibt sich ein Erbschaftsanteil für das

Top-Vermögensperzentil von über 80 %.

Der letzte Forschungsbeitrag dieser Dissertation „Comparing the joint distribution of intergenerational transfers, income and wealth across the Euro area“ erweitert die Anwendung aus dem vorangehenden Kapitel um zusätzliche Analysen und stellt die Ergebnisse im europäischen Kontext dar. Für die Länder Österreich, Belgien, Frankreich, Westdeutschland, Zypern, Griechenland, Portugal und Spanien kann durch die Verfügbarkeit des *Eurosystem Household Finance and Consumption Surveys* erstmals die Bedeutung von Erbschaften und Schenkungen für das Haushaltsvermögen im europäischen Vergleich berechnet werden. Der Beitrag zeigt sowohl die Prävalenzraten des Empfangs von Vermögenstransfers, die absoluten Höhen, als auch die relative Bedeutung gemessen am Haushaltsvermögen. In Logit-, OLS und *Fractional Logit* Regressionen kontrollieren wir für die wichtigsten sozio-ökonomischen Merkmale gleichzeitig. Es zeigt sich, dass der Zusammenhang zwischen Erbschaftsempfang und Einkommen größer in Kerneuropa als in den Mittelmeerländern ist, was auf eine niedrigere intergenerationale Mobilität hinweist, während die Korrelation zwischen Einkommen und der absoluten Erbschaftshöhe groß in ganz Europa ist. Während die Erbschaftshöhe erwartungsgemäß mit dem Haushaltsvermögen steigt, kehren sich diese Ergebnisse in der relativen Betrachtung der Erbschaft als Teil des Haushaltsvermögens um. Die relative Bedeutung der Erbschaft korreliert nicht mit der Vermögenshöhe oder dem Einkommen. Die Analyse mittels Fractional Logit Regression zeigt tendenziell das Gegenteil, die relative Bedeutung nimmt für höhere Einkommen ab.

Bibliography

Albuquerque, P. C. (2014). Intergenerational private transfers: Portugal in the european context. *European Journal of Ageing*, 11(4):301–312.

Allison, P. D. (1987). Estimation of linear models with incomplete data. *Sociological Methodology*, 17:71–103.

Arrondel, L., Roger, M., and Savignac, F. (2014). Wealth and income in the euro area: Heterogeneity in households' behaviours? Working Paper Series 1709, European Central Bank.

Atkinson, A. B. (2013). Wealth and inheritance in Britain from 1896 to the present. LSE Research Online Documents on Economics 58087, London School of Economics and Political Science.

Bach, S., Beznoska, M., and Steiner, V. (2014a). A wealth tax on the rich to bring down public debt? revenue and distributional effects of a capital levy in germany. *Fiscal Studies*, 35:67–89.

Bach, S., Houben, H., Maiterth, R., and Ochmann, R. (2014b). *Aufkommens- und Verteilungswirkungen von Reformalternativen für die Erbschaft- und Schenkungsteuer: Endbericht; Forschungsprojekt im Auftrag der Bundestagsfraktion Bündnis 90/Die Grünen*, volume 83 of *DIW Berlin: Politikberatung kompakt*. DIW Berlin, German Institute for Economic Research.

Bach, S. and Thiemann, A. (2016). Inheritance Tax Revenue Low Despite Surge in Inheritances. *DIW Economic Bulletin*, 6(4/5):41–48.

Bach, S., Thiemann, A., and Zucco, A. (2016). The top tail of the wealth distribution in Germany, France, Spain, and Greece. Discussion Papers 1502, Deutsches Institut für Wirtschaftsforschung, Berlin.

Barcelo, C. (2006). Imputation of the 2002 wave of the Spanish survey of household finances (EFF). Occasional Papers 0603, Banco de Espana.

Bartels, C. and Bönke, T. (2015). Die statistische Erfassung hoher Einkommen, Vermögen und Erbschaften in Deutschland. In Horn, G., Schmidt, K., van Treek, T., and Bofinger, P., editors, *Thomas Piketty und die Verteilungsfrage: Analysen, Bewertungen und wirtschaftspolitische Implikationen für Deutland*, pages 159–92. SE Publishing.

Bartels, C. and Schröder, C. (2016). Development of top incomes in germany since 2001. *DIW Economic Bulletin*, 6(1+2):3–8.

Baumert, J., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W., Stanat, P., Tillmann, K., and Weiß, M. (2001). *PISA 2000 – Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich*. Leske + Budrich, Opladen.

Beckert, J. (2008). *Inherited wealth*. Princeton University Press: Princeton.

Beckert, J. (2013). *Erben in der Leistungsgesellschaft*. Campus: Frankfurt.

Bönke, T., Corneo, G., and Lüthen, H. (2015). Lifetime Earnings Inequality in Germany. *Journal of Labor Economics*, 33(1):171 – 208.

Bönke, T. and Schröder, C. (2014). European-wide inequality in times of the financial crisis. *Journal of Income Distribution*, 23(3):3–24.

Bover, O. (2004). The spanish Survey of Household Finances (EFF): description and methods of the 2002 wave. Occasional Papers 0409, Banco de Espana.

Bover, O. (2008). The spanish Survey of Household Finances (EFF): description and methods of the 2005 wave. Occasional Papers 0803, Banco de Espana.

Bover, O. (2011). The spanish Survey of Household Finances (EFF): description and methods of the 2008 wave. Occasional Papers 1103, Banco de Espana.

Bover, O., Coronado, E., and Velilla, P. (2014). The spanish Survey of Household Finances (EFF): description and methods of the 2011 wave. Occasional Papers 1407, Banco de Espana.

Braun, R., Pfeiffer, U., and Thomschke, L. (2011). *Erben in Deutschland – Volumen, Verteilung, Verwendung*. Deutsches Institut für Altersvorsorge, Cologne.

Bricker, J., Dettling, L. J., Henriques, A., Hsu, J. W., Moore, K. B., Sabelhaus, J., Thompson, J., and Windle, R. A. (2014). Changes in U.S. family finances from 2010 to 2013: Evidence from the Survey of Consumer Finances. Federal Reserve Bulletin 4, Federal Reserve Board, Washington, DC.

Brunner, J. K. (2014). Die Erbschaftsteuer – Bestandteil eines optimalen Steuersystems? *Perspektiven der Wirtschaftspolitik*, 15(3):199–218.

Brzezinski, M. (2014). Do wealth distributions follow power laws? Evidence from 'rich lists'. *Physica A: Statistical Mechanics and its Applications*, 406:155–162.

Capehart, K. (2014). Is the wealth of the world's billionaires not Paretian? *Physica A: Statistical Mechanics and its Applications*, 395:255–260.

Chambers, R. L. (2001). Evaluation criteria for statistical editing and imputation. National Statistics Methodological Series 28, University of Southampton.

Chatterjee, A., Yarlagadda, S., and Chakrabarti, B. (2005). *Econophysics of Wealth Distributions*. Springer, Milan.

Christelis, D. (2011). Imputation of missing data in waves 1 and 2 of share. Retrieved from http://dx.doi.org/10.2139/ssrn.1788248, SHARE study.

Clauset, A., Shalizi, C. R., and Newman, M. E. J. (2009). Power-law distributions in empirical data. *SIAM Review*, 51(4):661–703.

Corneo, G. (2015). Income inequality from a lifetime perspective. *Empirica*, 42:225–39.

Corneo, G., Bönke, T., and Westermeier, C. (2016). Erbschaft und Eigenleistung im Vermögen der Deutschen: Eine Verteilungsanalyse. *Perspektiven der Wirtschaftspolitik*, 17(1):35–53.

Credit Suisse (2012). *Credit Suisse Global Wealth Data Book 2012*. Credit Suisse Group AG, Zurich. Retrieved from https://www.credit-suisse.com/ch/en/about-us/research/research-institute/publications.html.

Davies, J. and Shorrocks, A. F. (2000). The distribution of wealth. In Atkinson, A. and Bourguignon, F., editors, *Handbook of Income Distribution*, volume 1, chapter 11, pages 605–675. Elsevier, 1 edition.

Destatis (2015). *Sektorale und Gesamtwirtschaftliche Vermögensbilanzen 1999 - 2014*. Statistisches Bundesamt Wiesbaden. Retrieved from https://www.bundesbank.de/Navigation/DE/Statistiken/Gesamtwirtschaftliche_Rechenwerke/Vermoegensbilanzen/vermoegensbilanzen.html.

Deutsche Bundesbank (2013). Vermögen und Finanzen privater Haushalte in Deutschland: Ergebnisse der Bundesbankstudie. Monatsbericht Juni 2013, Deutsche Bundesbank.

Deutsche Bundesbank (2014). *Zeitreihe WU0004: Umlaufrenditen inländischer Inhaberschuldverschreibungen / Anleihen der öffentlichen Hand / Monatsdurchschnitte*. Deutsche Bundesbank. Retrieved from https://www.bundesbank.de/Navigation/DE/Statistiken/Zeitreihen_Datenbanken/Makrooekonomische_Zeitreihen/.

Deutsche Bundesbank (2016). Vermögen und Finanzen privater Haushalte in Deutschland: Ergebnisse der Vermögensbefragung 2014. Monatsbericht März 2016, Deutsche Bundesbank.

Eckerstorfer, P., Halak, J., Kapeller, J., Schütz, B., Springholz, F., and Wildauer, R. (2015). Correcting for the missing rich: An application to wealth survey data. *Review of Income and Wealth*. Online first.

Eisele, M. and Zhu, J. (2013). Multiple imputation in a complex household survey - the German Panel on Household Finances (PHF): challenges and solutions. Econstor preprints, ZBW - German National Library of Economics.

European Central Bank (2013a). The Eurosystem Household Finance and Consumption Survey. Methodological report for the first wave. Statistics Paper Series 1, European Central Bank, Franfurt am Main.

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

European Central Bank (2013c). Statistical tables. Statistics Paper Series 1, European Central Bank, Franfurt am Main.

EY (2014). Cross-country review of taxes on wealth and transfers of wealth. Revised Final report for the European Commission, Brussels. Retrieved from https://ec.europa.eu/taxation_customs/sites/taxation/files/docs/body/2014_eu_wealth_tax_project_finale_report.pdf.

Fessler, P., Lindner, P., and Segalla, E. (2014). Net wealth across the euro area - why household structure matters and how to control for it. Working Paper Series 1663, European Central Bank.

Fessler, P., Mooslechner, P., and Schürz, M. (2008). How inheritances relate to wealth distribution? – Theoretical reasoning and empirical evidence on the basis of LWS data. LWS Working Paper Series 6, LWS.

Fessler, P. and Schürz, M. (2015). Private wealth across European countries: the role of income, inheritance and the welfare state. Working Paper Series 1847, European Central Bank.

Frick, J. and Grabka, M. (2010). Alterssicherungsvermögen dämpft Ungleichheit – aber große Vermögenskonzentration bleibt bestehen. *DIW Wochenbericht*, 77(3):2–12.

Frick, J., Grabka, M. M., and Hauser, R. (2010a). *Die Verteilung der Vermögen in Deutschland. Empirische Analysen für Personen und Haushalte*. Edition Sigma.

Frick, J. R. and Grabka, M. M. (2005). Item nonresponse on income questions in panel surveys: incidence, imputation and the impact on inequality and mobility. *Allgemeines Statistisches Archiv*, 89:49–61.

Frick, J. R., Grabka, M. M., and Marcus, J. (2007). Editing and multiple imputation of item-non-response in the 2002 wealth module of the German Socio-Economic Panel (SOEP). Retrieved from http://hdl.handle.net/10419/86162, DIW Berlin Data Documentation 18.

Frick, J. R., Grabka, M. M., and Marcus, J. (2010b). Editing und Multiple Imputation der Vermögensinformation 2002 und 2007 im SOEP. Retrieved from http://hdl.handle.net/10419/86163, SOEP Survey Papers 146.

Gale, W. G. and Scholz, J. K. (1994). Intergenerational Transfers and the Accumulation of Wealth. *Journal of Economic Perspectives*, 8(4):145–160.

Goebel, J., Grabka, M. M., and Schröder, C. (2015). Income inequality remains high in germany: Young singles and career entrants increasingly at risk of poverty. *DIW Economic Bulletin*, 5(25):325–339.

Grabka, M. M. and Westermeier, C. (2014). Anhaltend hohe Vermögensungleichheit in Deutschland. *DIW Wochenbericht*, 81(9):151–164.

Grabka, M. M. and Westermeier, C. (2015). Real net worth of households in Germany fell between 2003 and 2013. *DIW Economic Bulletin*, 5(34):441–450.

Hauser, R. (2007). Sampling for household financial characteristics using frame information on past income. In *Integrierte Analyse von Einkommen und Vermögen – Forschungsstand und Ausblick*, pages 12–29, Institut für Sozialforschung und Gesellschaftspolitik, Köln. Bundesministeriums für Arbeit und Soziales.

Hayes, C. and Watson, N. (2009). HILDA imputation methods. Hilda project technical paper series 2/09, Melbourne Institute of Applied Economic and Social Research.

Henrekson, M. and Waldenström, D. (2016). Inheritance taxation in Sweden, 1885-2004: the role of ideology, family firms, and tax avoidance. *Economic history review*, 69(4):1228–1254.

Hoogland, J. J. and Boomsma, A. (1998). Robustness studies in covariance structure modeling: an overview and meta-analysis. *Sociological Methods and Research*, 26:329–367.

Kalton, G. (1998). Handling wave nonresponse in panel surveys. *Journal of Social Statistics*, 3:303–314.

Karagiannaki, E. (2015). Recent trends in the size and the distribution of inherited wealth in the UK. *Fiscal Studies*, 36(2):181–213.

Kennickell, A. B. (1991). Imputation of the 1989 Survey of Consumer Finances: stochastic relaxation and multiple imputation. Federal Reserve Working Paper Series, Federal Reserve Board, Washington, DC.

Kennickell, A. B. (1998). Multiple imputation in the Survey of Consumer Finances. Federal Reserve Working Paper Series, Federal Reserve Board, Washington, DC.

Kennickell, A. B. (2007). The role of oversampling of the wealthy in the Survey of Consumer Finances. Survey of Consumer Finances Working Paper, Federal Reserve Board.

Kennickell, A. B. (2009). Getting to the top: Reaching wealthy respondents in the SCF. Paper prepared for the 2009 Joint Statistical Meetings, Washington, DC.

Kennickell, A. B. and McManus, D. A. (1993). Sampling for household financial characteristics using frame information on past income. In *JSM Proceedings, Survey Research Methods Section*, volume 47, pages 88–97, Alexandria, VA. American Statistical Organization.

Kennickell, A. B. and Woodburn, R. L. (1997). Consistent weight design for the 1989, 1992, and 1995 SCFs, and the distribution of wealth. Survey of Consumer Finances Working Paper, Federal Reserve Board.

Kessler, D. and Masson, A. (1989). Bequest and Wealth Accumulation: Are Some Pieces of the Puzzle Missing? *Journal of Economic Perspectives*, 3(3):141–52.

Klass, S. O., Biham, O., Levy, M., Malcai, O., and Solomon, S. (2007). The Forbes 400, the Pareto power-law and efficient markets. *The European Physical Journal B*, 55(2):143–147.

Klevmarken, N. A. (2004). On The Wealth Dynamics Of Swedish Families, 1984-98. *Review of Income and Wealth*, 50(4):469–491.

Kohli, M., Künemund, H., Schäfer, A., Schupp, J., and Vogel, C. (2006). Erbschaften und ihr Einfluss auf die Vermögensverteilung. *Vierteljahrshefte zur Wirtschaftsforschung / Quarterly Journal of Economic Research*, 75(1):58–76.

Kohli, M., Künemund, H., Vogel, C., Gilles, M., Heisig, J., Schupp, J., Schäfer, A., and Hilbrich, R. (2005). Zusammenhänge und Wechselwirkungen zwischen Erbschaften und Vermögensverteilung. Bundesministerium für Gesundheit und Soziale Sicherung. Policy report, Bundesministerium für Gesundheit und Soziale Sicherung.

Kotlikoff, L. J. (1988). Intergenerational Transfers and Savings. *Journal of Economic Perspectives*, 2(2):41–58.

Kotlikoff, L. J. and Summers, L. H. (1981). The Role of Intergenerational Transfers in Aggregate Capital Accumulation. *Journal of Political Economy*, 89(4):706–32.

Künemund, H. and Vogel, C. (2011). Erbschaften und Vermögensungleichheit. Vortrag zur Frühjahrstagung 2011 der Sektion Wirtschaftssoziologie.

Le Blanc, J. (2014). Editing the Panel on Household Finances (PHF). Technical report, Deutsche Bundesbank.

Leopold, T. and Schneider, T. (2010). Schenkungen und Erbschaften im Lebenslauf - Vergleichende Längschnittanalyse zu intergenerationalen Transfers. *Zeitschrift für Soziologie*, 39(4):258–80.

Little, R. J. A. (1988). Missing-data adjustments in large surveys. *Journal of Business and Economic Statistics*, 6:287–296.

Little, R. J. A. and Rubin, D. B. (1987). *Statistical analysis with missing data*. New York: Wiley.

Little, R. J. A. and Su, H. L. (1989). Item nonresponse in panel surveys. In Kasprzyk, D., Duncan, G., Kalton, G., and Singh, M. P., editors, *Panel surveys*, pages 400–425. New York: Wiley.

Mahalanobis, P. C. (1936). On the generalised distance in statistics. proceedings of the national institute of sciences of india. *Proceedings of the National Institute of Sciences of India*, 2:49–55.

Mathä, T. Y., Porpiglia, A., and Ziegelmeyer, M. (2014). Household wealth in the euro area: the importance of intergenerational transfers, homeownership and house price dynamics. Working Paper Series 1690, European Central Bank.

Mehra, R. (2008). *Handbook of the equity risk premium*. Elsevier: Amsterdam.

Mennel, A. and Förster, J. (2014). *Steuern in Europa, Amerika und Asien*. NWB, Hamm.

Modigliani, F. (1986). Life Cycle, Individual Thrift, and the Wealth of Nations. *American Economic Review*, 76(3):297–313.

Modigliani, F. (1988). The Role of Intergenerational Transfers and Life Cycle Saving in the Accumulation of Wealth. *Journal of Economic Perspectives*, 2(2):15–40.

OECD (2015). *In It Together: Why Less Inequality Benefits All*. Organisation for Economic Co-operation and Development, Paris.

Ohlsson, H., Roine, J., and Waldenström, D. (2014). Inherited wealth over the path of development: Sweden, 1810–2010. Working Paper Series, Center for Fiscal Studies 2014:7, Uppsala University, Department of Economics.

Panel Study of Income Dynamics (2011). Multiple imputation in the Survey of Consumer Finances. Documentation for the 2007 PSID supplemental wealth file. Release 2: march 2011. Retrieved from http://psidonline.isr.umich.edu/Data/Documentation/wlth2007.pdf, Panel Study of Income Dynamics.

Pfeffermann, D., Krieger, A. M., and Rinott, Y. (1998). Parametric distributions of complex survey data under informative probability sampling. *Statistica Sinica*, 8:1087–1114.

Piketty, T. (2011). On the Long-Run Evolution of Inheritance: France 1820–2050. *The Quarterly Journal of Economics*, 126(3):1071–1131.

Piketty, T. (2014). *Capital in the 21st Century*. Harvard University Press, Cambridge, MA.

Piketty, T., Postel-Vinay, G., and Rosenthal, J.-L. (2014). Inherited vs self-made wealth: Theory & evidence from a rentier society (Paris 1872–1927). *Explorations in Economic History*, 51(C):21–40.

Piketty, T. and Zucman, G. (2015). Wealth and inheritance in the long run. In Atkinson, A. and Bourguignon, F., editors, *Handbook of Income Distribution*, volume 2B, pages 1303–1368. Elsevier.

Raub, B., Johnson, B., and Newcomb, J. (2010). A comparison of wealth estimates for America's wealthiest descendants using tax data and data from the Forbes 400. In *National Tax Association Proceedings, 103rd Annual Conference on Taxation*, pages 128–135, Chicago, IL.

Reil-Held, A. (2004). Die Rolle intergenerationaler Transfers in Einkommen und Vermögen älterer Menschen in Deutschland. mea Studies 02, mea.

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

Royston, P. (2004). Multiple imputation of missing values. *Stata Journal*, 4:227–241.

Rubin, D. B. (1976). Inference and missing data. *Biometrika*, 63:581–592.

Rubin, D. B. (1986). Statistical matching using file concatenation with adjusted weights and multiple imputations. *Journal of Business and Economic Statistics*, 4:87–94.

Rubin, D. B. (1987). *Multiple Imputation for Nonresponse in Surveys*. John Wiley & Sons.

Saez, E. and Zucman, G. (2016). Wealth inequality in the United States since 1913: evidence from capitalized income tax data. *Quarterly Journal of Economics*, 131(2):519–578.

Scheve, K. and Stasavage, D. (2012). Democracy, war, and wealth: Lessons from two centuries of inheritance taxation. *American Political Science Review*, 106(1):81–102.

Schinke, C. (2013). Inheritance in Germany 1911 to 2009: A Mortality Multiplier Approach. Working paper, PSE.

Schupp, J. and Szydlik, M. (2004). Erbschaften und Schenkungen in Deutschland: wachsende fiskalische Bedeutung der Erbschaftsteuer für die Länder. *DIW Wochenbericht*, 71(5):59–65.

Semyonov, M. and Lewin-Epstein, N. (2013). Ways to richness: Determination of household wealth in 16 countries. *European Sociological Review*, 32(5).

Statistisches Bundesamt (2011). Generationensterbetafeln für Deutschland: Modellrechnungen für die Geburtsjahrgänge 1896–2009. Technical report, Statistisches Bundesamt. Available at https://www.destatis.de/DE/Publikationen/Thematisch/Bevoelkerung/Bevoelkerungsbewegung/Generationssterbetafeln.html.

Statistisches Bundesamt (2013). Sektorale und gesamtwirtschaftliche Vermögensbilanzen 1991–2012. Technical report, Statistisches Bundesamt. Available at https://www.destatis.de/DE/Publikationen/Thematisch/VolkswirtschaftlicheGesamtrechnungen/Vermoegensrechnung/Vermoegensbilanzen.html.

Statistisches Bundesamt (2014). Einkommens- und Verbrauchsstichprobe. Geld- und Immobilienvermögen sowie Schulden privater Haushalte 2013. Technical report, Statistisches Bundesamt, Fachserie 15, Heft 2. Available at https://www.destatis.de/DE/Publikationen/Thematisch/EinkommenKonsumLebensbedingungen/Ueberschuldung/EVS_GeldImmobilienvermoegenSchulden2152602139004.pdf?__blob=publicationFile.

Szydlik, M. and Schupp, J. (2004). Who inherits more? Inheritances, social structure, and old age provision. *KZfSS Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 56(4):609–629.

Tiefensee, A. and Grabka, M. M. (2016). Comparing Wealth – Data quality of the HFCS. *Survey Research Methods*, 10(2):119–142.

Uhrig, N., Bryan, M., and Budd, S. (2012). UKHLS Innovation Panel household wealth questions: preliminary analysis. Working Paper Series 2012, Retrieved from https://www.iser.essex.ac.uk/publications/working-papers/understanding-society/2012-01, Understanding Society.

van Buuren, S., Boshuizen, H. C., and Knook, D. L. (1999). Multiple imputation of missing blood pressure covariates in survival analysis. *Statistics in Medicine*, 18:681–694.

van Buuren, S., Brand, J. P. L., Groothuis-Oudshoorn, C. G. M., and Rubin, D. B. (2006). Fully conditional specification in multivariate imputation. *Journal of Statistical Computation and Simulation*, 76:1049–1064.

Vermeulen, P. (2014). How fat is the top tail of the wealth distribution? Working Paper Series 1692, European Central Bank.

Vermeulen, P. (2016). Estimating the top tail of the wealth distribution. *American Economic Review*, 106(5):646–650.

von Kalckreuth, U., Eisele, M., Le Blanc, J., Schmidt, T., and Zhu, J. (2012). The PHF: A comprehensive panel survey on household finances and wealth in Germany. Discussion Papers 13/2012, Deutsche Bundesbank, Research Centre.

Wagner, G. G., Frick, J. R., and Schupp, J. (2007). The German Socio-Economic Panel Study (SOEP) - Scope, evolution and enhancements. *Schmollers Jahrbuch : Journal of Applied Social Science Studies / Zeitschrift für Wirtschafts- und Sozialwissenschaften*, 127(1):139–169.

Waldenström, D. (2016). The national wealth of Sweden, 1810-2014. *Scandinavian Economic History Review*, 64(1):36–54.

Watson, N. and Mark, W. (2002). The Household, Income and Labour Dynamics in Australia (HILDA) Survey: Wave 1 Survey Methodology. HILDA Project technical papers series 1/02, May 2002 (Revised October 2002). Retrieved from https://www.melbourneinstitute.com/hilda-research/Survey%5C_Methods%5C_and%5C_Data.html, Melbourne Institute.

Watson, N. and Starick, C. (2011). Evaluation of alternative income imputation methods for a longitudinal survey. *Journal of Official Statistics*, 27:693–715.

Wealth-X (2014). Wealth-X and UBS World Ultra Wealth Report 2014. Wealth-x, New York.

Westermeier, C. and Grabka, M. M. (2015). Significant statistical uncertainty over share of high net worth households. *DIW Economic Bulletin*, 5(14/15):210–219.

Wolff, E. and Gittleman, M. (2014). Inheritances and the distribution of wealth or whatever happened to the great inheritance boom? *The Journal of Economic Inequality*, 12(4):439–468.

Wolff, E. N. (2015). U.S. pensions in the 2000s: The lost decade? *Review of Income and Wealth*, 61(4):599–629.

List of Figures

List of Tables

Vorveröffentlichungen

Kapitel 2: Longitudinal wealth data and multiple imputation – An evaluation study

- SOEPpapers Nr. 790, Deutsches Institut für Wirtschaftsforschung
- *Survey Research Methods* 10(2016), S. 327-352

Kapitel 3: Estimating top wealth shares using survey data – An empiricist's guide

- Diskussionspapiere des Fachbereichs Wirtschaftswissenschaften der Freien Universität Berlin, Nr. 21/2016

Kapitel 4: Breaking down Germany's private wealth into inheritance and personal efforts – A distributional analysis

- Diskussionspapiere des Fachbereichs Wirtschaftswissenschaften der Freien Universität Berlin, Nr. 10/2015
- *Perspektiven der Wirtschaftspolitik* 17(2016), S. 35-53

Kapitel 5: Comparing the joint distribution of intergenerational transfers, income and wealth across the Euro area

- Diskussionspapiere des Fachbereichs Wirtschaftswissenschaften der Freien Universität Berlin, Nr. 4/2016
- Diskussionspapiere des Deutschen Instituts für Wirtschaftsforschung, Nr. 1556

Erklärung gem. §10 Abs. 3 der Promotionsordnung

Hiermit erkläre ich, dass ich neben der in der Bibliographie gelisteten Literatur folgende Hilfsmittel und Hilfen verwendet habe:

- Stata, StataCorp.
- Excel, Microsoft Corp.
- Illustrator, Adobe Systems Inc.
- Wolfram Alpha, Wolfram Research Inc.

Auf dieser Grundlage habe ich die Arbeit selbstständig verfasst.

Berlin, 19. Juni 2017

Christian Westermeier

Erklärung gem. §4 Abs. 2 der Promotionsordnung

Hiermit erkläre ich, dass ich mich noch keinem Promotionsverfahren unterzogen oder um Zulassung zu einem solchen beworben habe, und die Dissertation in der gleichen oder einer anderen Fassung bzw. Überarbeitung einer anderen Fakultät, einem Prüfungsausschuss oder einem Fachvertreter an einer anderen Hochschule nicht bereits zur Überprüfung vorgelegen hat.

Berlin, 19. Juni 2017

Christian Westermeier