I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 10573

An Integrated Approach for Top-Corrected Ginis

Charlotte Bartels
Maria Metzing

FEBRUARY 2017

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 10573

An Integrated Approach for Top-Corrected Ginis

Charlotte Bartels
DIW Berlin and IZA

Maria Metzing
DIW Berlin

FEBRUARY 2017

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

An Integrated Approach for Top-Corrected Ginis*

Household survey data provide a rich information set on income, household context and demographic variables, but tend to under report incomes at the very top of the distribution. Administrative data like tax records offer more precise information on top incomes, but at the expense of household context details and incomes of non-filers at the bottom of the distribution. We combine the benefits of the two data sources and develop an integrated approach for top-corrected Gini coefficients where we impute top incomes in survey data using information on top income distribution from tax data. We apply our approach to European EU-SILC survey data which in some countries include administrative data. We find higher inequality in those European countries that exclusively rely (Germany, UK) or have relied (Spain) on interviews for the provision of EU-SILC survey data as compared to countries that use administrative data.

JEL Classification: C46, C81, D31, H2

Keywords: Gini coefficient, top income shares, survey data, tax record data, Pareto distribution

Corresponding author:
Charlotte Bartels
SOEP research infrastructure
DIW Berlin
Mohrenstraße 58
10117 Berlin
Germany
E-mail: cbartels@diw.de

* This paper has greatly benefited from the comments of Carsten Schröder. We thank Robin Jessen for providing simulated net household incomes from STSM (Steuer-Transfer-Simulations-Modell) as robustness check for our approximation. Furthermore, we thank participants of the 2016 conferences of the Society for Social Choice and Welfare (SSCW), the European Economic Association (EEA), the International Institute of Public Finance (IIPF) and the German Economic Association (Verein für Socialpolitik) for valuable comments.

1 Introduction

Has inequality of living standards in European countries increased in recent years? The answer is far from conclusive, varying as we look at different inequality measures and different data sources. A well-known and intensively discussed reason for diverging trends is the inequality measure's sensitivity to changes in the top, middle or bottom of the income distribution. Another reason for diverging trends is much less investigated: the different nature of the data employed to estimate inequality measures. Whereas the top income share literature based on tax data produces wide evidence of rising inequality in recent decades, survey data based inequality studies find less clear trends.[1]

Tax and survey data are substantially different in the definition of income and unit of observation. Household surveys usually apply a comprehensive income concept, while tax data contain income subject to taxation.[2] While incomes in survey data are aggregated at the household level, the income-receiving unit in tax data is the tax unit. If household members pool their income, the narrower sharing unit of a tax unit usually produces higher inequality. Furthermore, survey and tax data are affected differently by time-variant factors such as survey response and reporting behavior, tax filing behavior as well as economic, demographic and legislative changes. Undercoverage and underreporting of top incomes may produce a downward-bias for survey-based inequality measures. Tax filing behavior is sensitive to changes in the income tax law creating downward- or upward-bias before or during reform years. Top income earners tend to benefit disproportionately from economic growth (Roine et al., 2009), which in turn produces higher inequality estimates in tax data

[1]The top incomes literature produces internationally comparable measures for income concentration at the top of the distribution based on taxable incomes received by tax units, which are assembled in the World Wealth and Income Database (WID) available online at `http://www.wid.world/`. Top income shares and survey data based Gini coefficients, e.g., collected in the OECD database, indicate deviating inequality trends for some countries. In Germany and the United Kingdom, the income share of the top 1% has increased since the mid-2000s, whereas the Gini remains rather stable. In Spain, while the top 1% income share falls after peaking in 2006, the Gini has steadily increased since 2006 (see Appendix Figures A.1 and A.6).

[2]Not only do household surveys document a variety of market income sources, they also incorporate private transfers. In contrast, tax incomes ever more frequently exclude capital income due to the introduction of dual income taxation where capital income is taxed separately. This is the case for Germany since 2009.

than in survey data where top income earners are underrepresented. Changes in the number of unmarried couples affects tax-based inequality measures in countries with joint taxation where the direction of the effect depends on the degree of assortative mating.

For the United States and the United Kingdom, a growing number of studies investigates these differences by reconciling estimates from administrative and survey data (Burkhauser et al., 2012; Armour et al., 2013; Bricker et al., 2015; Burkhauser et al., 2016; Jenkins, 2016) or adjusting survey-based Gini coefficients with tax data-based top income shares (Atkinson et al., 2011; Alvaredo, 2011). However, these contributions draw on access to tax record microdata which require substantial knowledge of the country's tax rules to harmonize income concepts and are usually difficult to access. This makes cross-country comparisons rather difficult. Furthermore, most of these studies document inequality trends of tax income over tax units that do not necessarily reflect how inequality of living standards evolved for the entire population.

We develop a new method to obtain top-corrected Gini coefficients by combining easily available information from tax and survey data. We replace the top 1% of the survey income distribution with Pareto-imputed incomes using information on the top incomes' distribution from the World Wealth and Income Database (WID).[3] Our approach is easily applicable by relying on information publicly available from the WID for the upper tail of the distribution and easily accessible survey data, such as the German SOEP or EU-SILC, for the middle and bottom of the distribution. Neither access to tax record microdata, which is limited and difficult to obtain in many countries, nor record linkage, which is often not allowed, is needed.[4] In contrast to the decomposition approach for top-corrected Gini coefficients (Atkinson, 2007; Alvaredo, 2011) which exclusively relies on tax incomes of tax units, our integrated approach allows for producing inequality measures for a variety of income definitions and for the entire population of a country, e.g., analyzing inequality in

[3]Another example of a top income imputation approach is in Lakner and Milanovic (2013). They distribute the gap between national accounts and survey means over the top decile according to a fitted Pareto distribution in order to obtain a global Gini coefficient.

[4]Bach et al. (2009) is an example where the authors integrate both survey and tax record micro data to obtain Gini coefficients over the whole spectrum of the population in Germany.

households' needs.

First, we reconcile German survey and tax data, examining the extent to which differences in top income share estimates from household surveys and tax returns arise from differences in income concepts, observation units or from the coverage of top incomes. Second, we compare our integrated approach for top-corrected Ginis on German survey data with the decomposition approach (Atkinson, 2007; Alvaredo, 2011). Third, we apply our integrated approach to EU-SILC data and estimate top-corrected Gini coefficients for those European countries where information on the top incomes' distribution is available in the WID.

Our results are the following. First, reconciled German survey data show that the top 10-5% and top 5-1% income shares are of similar magnitude in both tax return and survey data. In contrast, survey data report a substantially lower top 1% income share which suggests that this group is not sufficiently captured. We find that different definitions of income and observation unit yield substantially different inequality levels in Germany: the Gini of tax income by tax unit is about 10%-points higher than the Gini of equivalent gross household income by household unit. The selected income concept is responsible for the largest part of this gap, whereas the observation unit changes inequality only slightly as most German households form a single tax unit anyway. Second, our top-corrected Gini for 2001-2012 Germany is 4% higher for gross household income and about 2% higher for net equivalent household income than the unadjusted Gini. Our top-correction method indicates similar trends and slightly lower inequality levels than the decomposition approach (Atkinson, 2007; Alvaredo, 2011). Third, the application of our top-correction approach to EU-SILC survey data shows remarkably higher inequality levels in those countries that exclusively rely (Germany, UK) or have relied (Spain) on interviews for the provision of EU-SILC data. I.e., replacing the top of the survey incomes with Pareto-imputed incomes has a bigger effect on inequality which implies that top incomes are not sufficiently covered by the survey in these countries. For most countries using register data, the gap between top-corrected and unadjusted Ginis is negligible.

The paper is structured as follows. In Section 2, we reconcile German house-

hold survey data with income tax data definitions, then compute top income shares and Gini coefficients contrasting original and reconciled data. Our new integrated approach for top-corrected Gini coefficients is explained in Section 3. In Section 4, inequality trends according to top-corrected Gini coefficients in European countries are presented. Section 5 concludes.

2 Reconciling household survey and income tax return data

Two major differences between household survey data and income tax return data call for reconciling the data before comparing inequality measures across data sources. First, survey data and administrative data differ in what is counted as income. Second, data discord in the definition of the income receiving unit. Household survey based inequality measures include incomes collected on the questionnaires before and after taxes as well as transfers. Incomes aggregated at the household level are then usually adjusted to differences in households' needs using an equivalence scale. Income tax return data document taxable incomes before taxes paid and transfers received by the tax unit which may consist of an individual or a married couple (plus their children) depending on the country's income tax legislation.

We reconcile survey data from the German Socio-Economic Panel (SOEP)[5] and German income tax records.[6] Using microsimulation we construct tax units and income in the SOEP data according to the governing income tax law for each year from 2001 through 2012.[7] The opposite direction is not possible since tax records offer very limited information on household context such that tax units cannot be summed up to households. In the reconciled SOEP data, a household with a married couple is treated as one unit and a household with an unmarried couple as two units.

[5]For further details on German SOEP data see Wagner et al. (2007) or Gerstorf and Schupp (2016).

[6]Since the data requirement for reconciling data is large and a microsimulation model incorporating the frequent changes of the tax law and transfer regulations must be at hand, we restrict this step of our analysis to Germany.

[7]We choose this period because German income tax data became annually available in 2001; 2012 is currently the last available year.

The income concept used in the income tax statistics is the total amount of income (*Gesamtbetrag der Einkünfte*) defined by the German Income Tax Act, which is the sum of the seven income categories (agriculture and forestry, business, self-employment, employment, capital income,[8] renting and leasing, as well as other), plus tax-relevant capital gains less income type-specific income-related expenses, savings allowances, and losses. Old-age lump-sum allowance and exemptions for single parents are deducted.[9] Since a number of large tax-deductible items, such as special expenses for social security contributions, are not deducted at that stage, the total amount of income is considerably higher for most tax units than the eventual taxable income to which the tax rate is applied. For reasons of simplicity, we refer to tax income instead of the total amount of income in the following.

We then compare the estimated share of total income accruing to the top of the income distribution based on reconciled SOEP data and income tax records. The observation unit is the tax unit and the income concept is tax income in both data sources. Figure 1 shows how income accruing to the top decile in Germany is split among the bottom half (10-5%), the upper 4% (5-1%) and the top 1% and contrasts results from the two data sources.

Three findings stand out: First, the estimates of the income share of the top 10-5% and top 5-1% are of similar magnitude in both data sources. The income share of the bottom half (10-5%) is around 12 % in the SOEP data and between 11.2 to 11.8 % in the income tax data.[10] The upper 4 % do not differ significantly until 2008 in both datasets and are between 13.4 and 15 %.

Second, there are large quantitative differences for the top 1% between SOEP and tax data. Tax data measure 3 to 6 %-points higher income shares for the top

[8]Since the introduction of dual income taxation in Germany in 2009, capital income is taxed separately at a flat rate and, hence, is no longer readily visible in tax data. However, it is still beneficial to declare capital income in their income tax declaration for some tax units, e.g., if the flat rate exceeds their personal income tax rate. But the size of reported capital income is negligible.

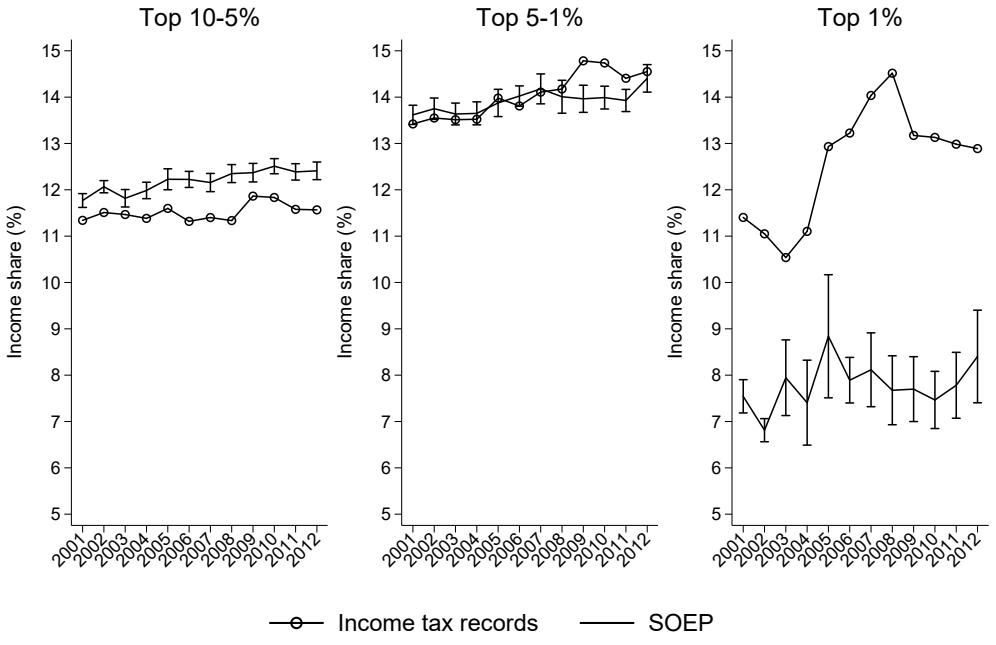
[9]The total amount of income is modeled in the SOEP data by deducting the allowances from the gross income of the tax unit and adding the taxable share of the pension income. It should be noted, however, that the total amount of income can only be approximately simulated in the SOEP data because incomes, such as self-employment income, are differently recorded across data sets.

[10]The result that income share of the bottom half of the top decile is significantly higher in the SOEP data than in the tax records indicates a potential middle class bias in the SOEP data.

1%. The income share in the tax data is between 10.6 % to 15 % whereas the income share in the SOEP data is between 7 % and 8.8 %. The mismatch between the data sources for the top 1% does not come as a surprise as average incomes of the top 1% in the two data sources differ by more than 100,000 Euros. This result also applies two other countries' survey data: sizable larger gaps for the top 1% income share are found by Burkhauser et al. (2012) for the US using March Current Population Survey (CPS) and Internal Revenue Service (IRS) data and by Jenkins (2016) for the UK using Family Resources Survey (FRS) and income tax return data. Based on this finding, we decide to replace the top 1% of the survey income distribution with Pareto-imputed incomes.

Third, both data sources document a trend of rising income concentration over the period. But whereas the tax data show a steep increase until 2008, particularly for the top 1%, and then a strikingly stable path following the Great Recession in 2009, SOEP data indicate a decline since 2005 and an increase since 2010.

Figure 1: Top income shares in income tax and survey data, Germany

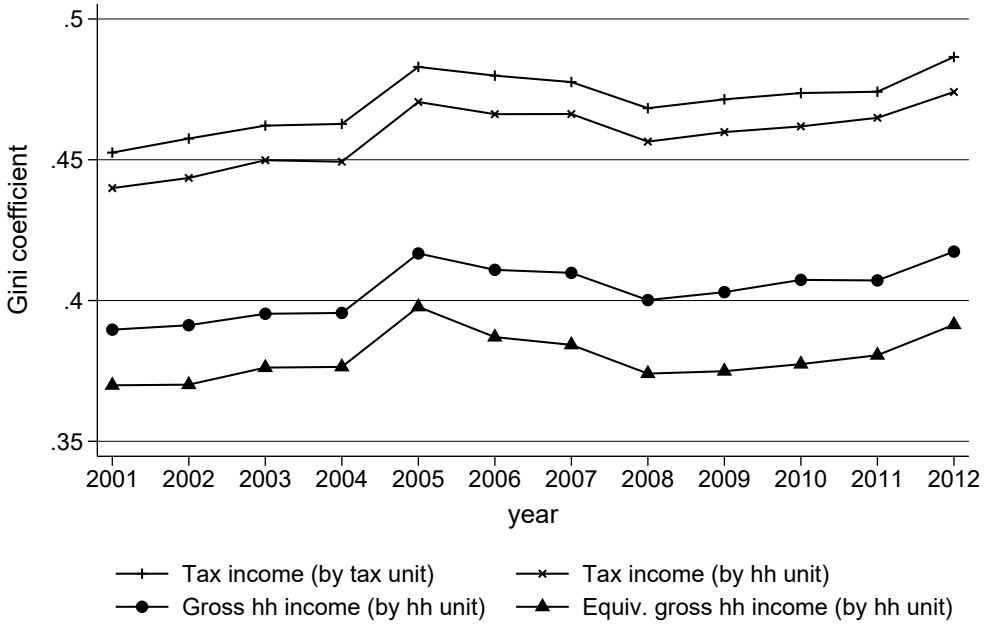


Source: SOEP (own calculations) and income tax records (Bartels and Jenderny, 2015) also available in WID.
Note: The observation unit is the tax unit and the income concept is tax income in both data sources. Vertical lines show bootstrap confidence intervals at the 95%-level based on 200 drawings.

Cross-walking from income tax data definitions to survey data definitions using German SOEP data reveals a gradual decline in inequality measured by the Gini coefficient as shown by Figure 2. The Gini based on tax income per tax unit (*Tax income by tax unit*) exhibits the highest level of inequality. If we then aggregate tax income at the household level (*Tax income by hh unit*), we the Gini coefficient is reduced by about 3%. Considering gross household income (*Gross hh income by hh unit*)[11] instead of households' tax income yields a Gini reduction of about 12%. Finally, when we equivalize gross household income to account for differences in households' needs (*Equiv. gross hh income by hh unit*), the Gini declines by 5 to 8%. Applying different definitions of income and observation unit yields substantial differences in inequality levels: the Gini of tax income by tax unit is about 10%-points higher than the Gini of equivalent gross household income by household unit. All in all, the income concept is of major importance for the inequality level measured. The unit of observation accounts only for a small change because most households in Germany consist of a single tax unit. In contrast, tax income as defined by German tax law is substantially more unequally distributed than gross household income. As explained above, tax income is obtained after income type-specific income-related expenses, savings allowances, old-age lump-sum allowance, and exemptions for single parents are deducted. If these reductions are relatively more important for middle and low-income households, this contributes to a more unequal distribution of tax income. Furthermore, gross household income includes social security pensions and private transfers that contribute to equalizing the income distribution.

[11] Gross household income includes household social security pensions in order to increase comparability with tax income. In Germany, an increasing share of social security pensions is subject to income taxation and, thus, included in tax income.

Figure 2: Cross-walking from tax to survey data definitions, Germany



Source: SOEP (own calculations).
Note: Gross household income includes social security pensions as they are partly included in taxable income under German tax law.

3 An integrated approach for top-corrected Gini coefficients

Building on the assumption that top incomes are Pareto distributed, we replace the incomes of the top 1% of the survey income distribution with Pareto-imputed incomes.[12] We opt to replace the top 1% since the comparison of the top income shares in Section 2 reveals that this group is under represented in the survey data whereas the lower 4% of the top twentieth seem to match the tax data distribution quite well.[13] A nice feature of the Pareto distribution is its small number of parameters that need to be estimated. The top income shares documented in the World Wealth and Income Database (WID) suffice to obtain an estimate of the central parameter α. The Pareto distribution function can be written as follows

[12] A large literature shows that incomes follow a Pareto distribution, e.g., Clementi and Gallegati (2005a) for Germany, Piketty (2003) for France, Clementi and Gallegati (2005b) for Italy, Atkinson (2007) for United Kingdom and Piketty and Saez (2003) for United States.

[13] Jenkins (2016) finds that under-coverage of top incomes in UK survey data varies over the years starting above P95 in the 2000s and above P99 in the 1990s. This check, however, requires access to microdata and Jenkins (2016) recommends making a judicious choice of the cut-off. Burkhauser et al. (2012) supports under-coverage of the P99 percentile.

$$1 - F(y) = \left(\frac{k}{y}\right)^{\alpha}, \tag{1}$$

where α is the Pareto coefficient and k is the income threshold above which incomes are Pareto distributed. We estimate the Pareto coefficient α following Atkinson (2007) as

$$\alpha = \frac{1}{\left(1 - \frac{log(S_j/S_i)}{log(P_j/P_i)}\right)} \tag{2}$$

where P_j is the population share of group j and S_j is the income share of group j documented in the World Wealth and Income Database (WID). Top income shares for Germany in the WID are produced by Bartels and Jenderny (2015).

Empirically, α increases when moving the Pareto threshold from the middle of the distribution to the top (see, e.g., Jenkins (2016); Atkinson (2007)). We use α estimated for $P_i = 0.1\%$ and $P_j = 1\%$. It seems reasonable to calculate α for the top percentile of the distribution, which is less well represented in survey data as shown in Figure 1.[14] Threshold k is then obtained from rearranging Eq. 1 to

$$k = (1 - F(y))^{1/\alpha} \cdot y, \tag{3}$$

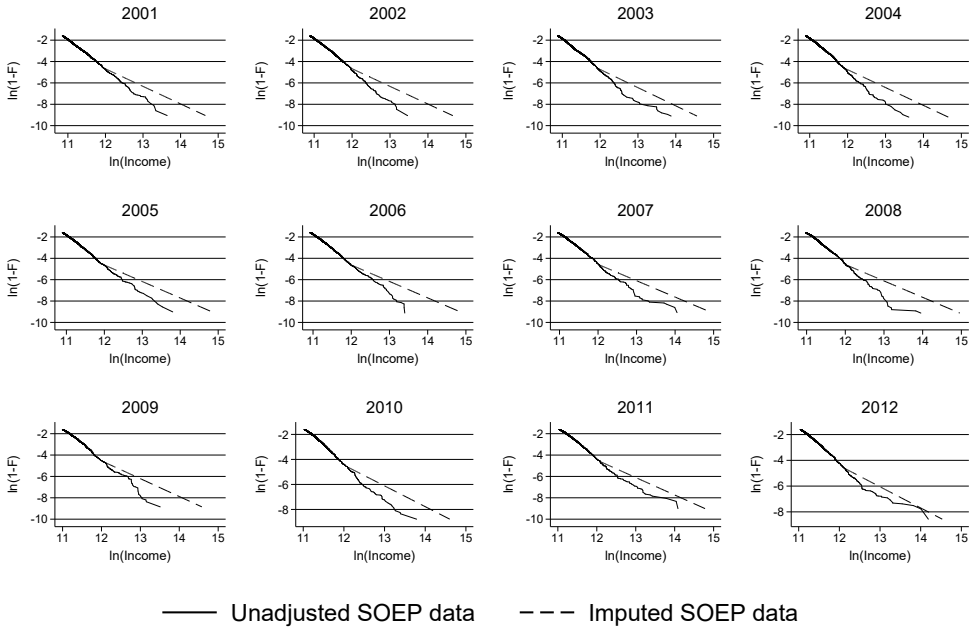
where $F(y)$ and y are taken from the survey data distribution. Since we replace the top 1% of the distribution, y is the P99 percentile.[15] Our results for α and k for Germany are presented in Appendix Table A.2. We then replace the top 1% of incomes observed in the survey data with incomes following the Pareto distribution characterized by our estimated parameters.

[14] Appendix Figure A.3 shows that α estimated for $P_i = 0.1\%$ and $P_j = 1\%$ produces the best fit of the top 1% income share in Germany. Using α estimated for $P_i = 1\%$ and $P_j = 5\%$ or $P_i = 1\%$ and $P_j = 10\%$, which creates a less heavy tail, we obtain a substantially lower top 1% income share in comparison to income tax data. Moreover, α estimated for $P_i = 0.1\%$ and $P_j = 1\%$ yields the best fit for the income share of the lower half of the upper decile (see Appendix Figure A.4). Our α estimates for $P_i = 0.1\%$ and $P_j = 1\%$ in Germany are around 1.6, whereas estimates for $P_i = 5\%$ and $P_j = 10\%$ are mostly greater than 2 (see Appendix Table A.2).

[15] Thresholds between P95 and P99.5 are commonly used. Jenkins (2016) provides an extensive discussion of the choice of the Pareto threshold and shows that choosing different Pareto thresholds has noticeable impacts on estimates of inequality among the rich, but overall inequality trends in the UK are broadly robust to the choice of the threshold.

If one plots $log(1 - F(y))$ against $log(y)$, Pareto distributed incomes produce a straight line with the slope $-\alpha$ (so-called *Zipf plot*). The smaller α (the flatter the line), the more unequal is the income distribution. Figure 3 shows this plot for both unadjusted SOEP data and SOEP data with imputed top incomes. Replacing top incomes with Pareto-imputed incomes generates a more unequal income distribution reflected by the flatter curve than original SOEP incomes. Assuming that tax data provide a more accurate picture of the very top, we would underestimate the tail of the income distribution using Pareto parameters fitted to survey data.[16] In most of the years, original SOEP top incomes do not seem to follow a Pareto distribution. However, in 2002 and 2006 we obtain rather straight lines from original SOEP incomes.

Figure 3: Fit of the Pareto distribution



Source: SOEP (own calculations).
Note: Lines cross at the income level of P99 above which we impute top incomes.

[16] Jenkins (2016) also states that replacing the top of survey distribution with Pareto-imputed values fitted from the same source may not produce reliable results and tax return data should be used instead.

The approach derived by Atkinson (2007) and extended by Alvaredo (2011) is based on the Gini decomposition for two non-overlapping subgroups by Dagum (1997)

$$G = \sum_{j=1}^{k} G_{jj} P_j S_j + \sum_{j=1}^{k} \sum_{h=1}^{j-1} G_{jh}(P_j S_h + P_h S_j), \tag{4}$$

where G_{jj} is the Gini coefficient of the j-th group, G_{jh} is the Gini coefficient between the j-th and h-th group, P_j is the population share of group j and S_j is the income share of group j. Assuming that the population can be divided into two groups – the top covered by tax records (e.g., the top 1%) and the rest of the population covered by survey data – we can rearrange Eq. 4 using the notation from Alvaredo (2011) to

$$G = G^{**} PS + G^{*}(1 - P)(1 - S) + S - P, \tag{5}$$

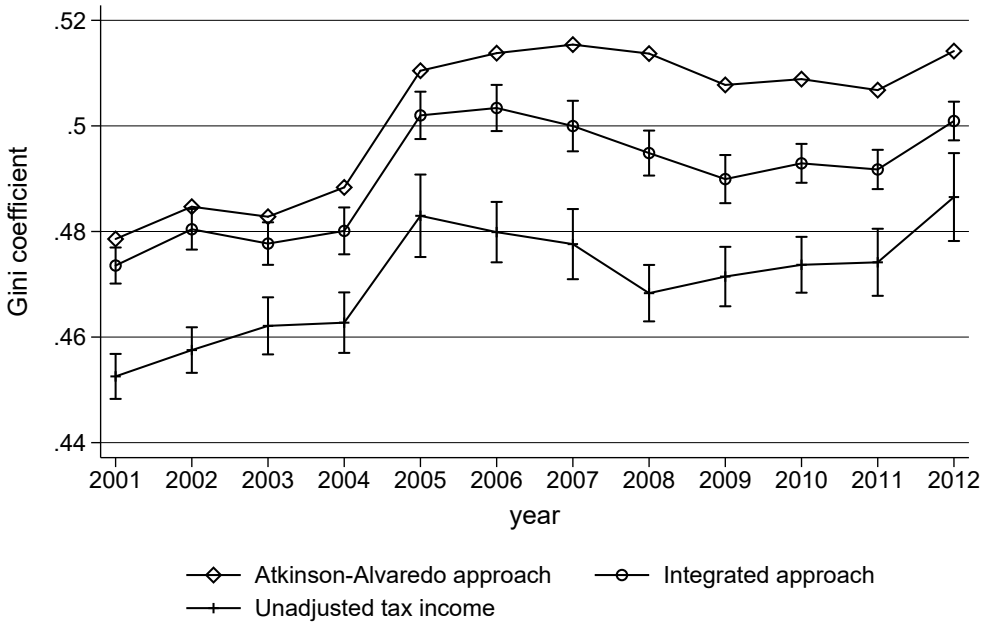
where P and S are population and income share of the top, respectively, and $1 - P$ and $1 - S$ are population and income share of the rest of the population. G^{*} is the Gini for the population without the top group. Assuming that top incomes are Pareto distributed, the Gini of the top is computed as $G^{**} = \frac{1}{2\alpha - 1}$, where α is the Pareto coefficient obtained from the tax income distribution documented by tax data applying Eq. 2.

We now turn to the comparison of the two approaches for top-corrected Gini coefficients. As can be taken from Figure 4, Gini coefficients of both top-correction methods are substantially higher than Ginis based on unadjusted survey data income. But where the Gini based on unadjusted SOEP data shows a peak of inequality in 2005 and a low point in 2008, the top-corrected approaches rather hint at a plateau between 2005 and 2007 and a low point in 2009.[17] Between 2005 and 2008, incomes of the top 1% grew especially rapidly, which is not sufficiently captured by survey data where this group is underrepresented. The Great Recession hitting Germany in 2009 primarily affected top income earners whose business incomes

[17] As shown in Biewen and Juhasz (2012), the rise in inequality in Germany until 2005 is largely driven by high unemployment.

collapsed (Bartels and Jenderny, 2015). Therefore, top-corrected Ginis exhibit a decline in inequality whereas unadjusted Ginis show a stable path. Interestingly, both top-corrected approaches show a rise in inequality after 2011, even though the income share of the top 1% remained rather stable since 2009. All in all, we find that both correction approaches produce rather similar levels and trends of income inequality as measured by the Gini coefficient. The decomposition approach puts more weight on the development of the top incomes than does our integrated approach and, thereby, produces higher inequality levels.[18]

Figure 4: Top-corrected Gini coefficients, Germany



Source: SOEP (own calculations).
Note: Gini coefficients are based on tax income. The integrated approach Gini is based on $P_i = 0.1\%$ and $P_j = 1\%$. The Atkinson-Alvaredo approach Gini is based on P=1%. Vertical lines show bootstrap confidence intervals at the 95%-level based on 100 drawings.

For calculating top-corrected Ginis reflecting the inequality of living standards of the German population, we undertake two steps: First, we have to impute gross household incomes for the top. We argue that the α parameter estimated from tax records' top of the distribution can be used to impute both the top of the tax and

[18]Theil coefficient and Half Squared Coefficient of Variation (HSCV) of both imputed and unadjusted income exhibit similar trends like the Gini and are displayed in Appendix Figure A.9.

the household income distribution regardless of the unit of observation.[19] Second, we have to compute (equivalent) net household incomes from the imputed gross household incomes. We use an approximation of the tax-benefit-system introduced by Feldstein (1969):

$$y^{net} = \lambda(y^{gross})^{1-\tau}, \tag{6}$$

where y^{net} presents the net household income and y^{gross} the gross household income. Parameter τ is the degree of progressiveness[20] and λ is an indicator for the average level of the household taxation.

We estimate the following equation by household type h[21] and year t for observed gross and net household incomes:

$$ln(y_{h,t}^{net}) = ln(\lambda) + (1-\tau)ln(y_{h,t}^{gross}) + \epsilon_{h,t}. \tag{7}$$

Our estimates for τ are between 0.14 and 0.29, depending on household type and year (see Appendix Table A.4). The model fits the observed relationship between gross and net household income quite well (R^2 is between 0.78 and 0.96). We collapse our five samples into 50 quantiles and plot the imputed gross household income against our predicted net household income to demonstrate the good representation of the German tax scheme. Our estimates for τ are between 0.14 and 0.26 for all household types (see Appendix Table A.4). The model fits the observed relationship between gross and net household income quite well (R^2 is between 0.78 and 0.96).[22]

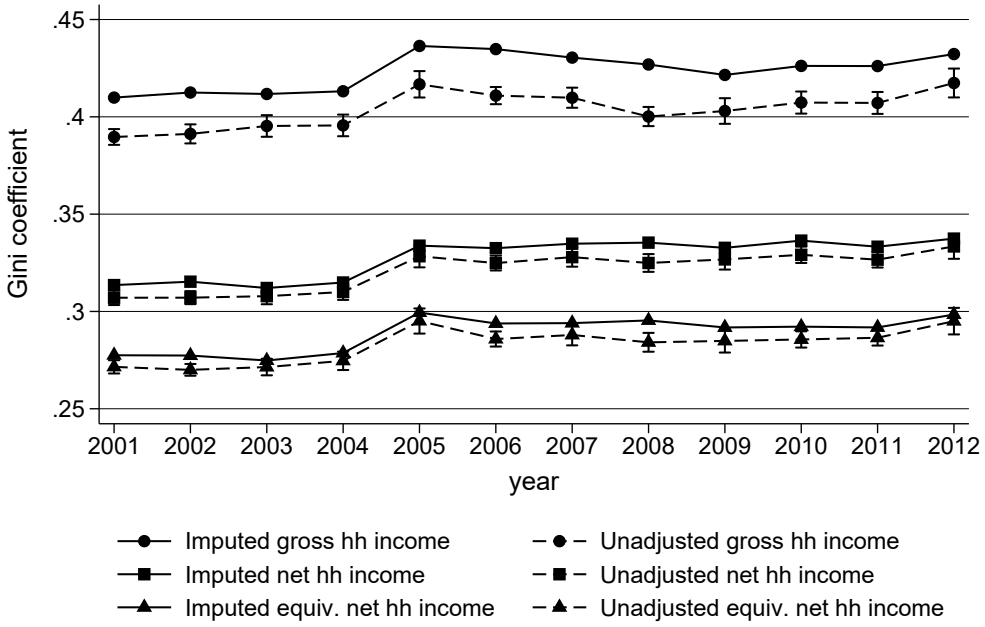
[19] Appendix Figure A.2 shows that the steepness of the log-log-curve for unadjusted tax incomes by tax unit is quite similar to household incomes by household unit in the German SOEP data.

[20] A positive τ indicates a progressive tax schedule, whereas a negative τ indicates a regressive tax schedule.

[21] We regress the equation separately for five different household types – singles without children, singles with children, couples without children, couples with children, and other household types – in order to account for different tax allowances and exemptions. Only tax-paying households with a minimum household income of 20,000 Euro are included.

[22] See Appendix Figure A.8 for a graphical presentation of the fit for the five different household types. We collapse each of our five samples into 50 income quantiles and plot the unadjusted gross household income against our unadjusted net household income. This check is also used in Heathcote et al. (ming) and show the good fit of our simple regression model. As a robustness check, we also compute net household income with the *STSM* microsimulation model for the

Figure 5: Top-corrected Gini coefficients (gross, net, equivalent net income), Germany



Source: SOEP (own calculations).

Figure 5 presents the top-corrected Ginis for gross, net, and equivalent net household income. The top-corrected Gini for gross household income is about 4% higher than the unadjusted. The top-corrected Ginis for net and equivalent net household income are about 2% higher. Apart from that, the observed trends do not reverse. The smaller gap between the top-corrected and unadjusted net income Ginis is due to the progressive tax system in Germany.

4 An application to European survey data

We apply our integrated approach for top-corrected Ginis to other European countries where both EU-SILC survey data[23] and top income shares are available from

German tax-benefit system. Net household incomes predicted by the Feldstein-formula are very close to *STSM*-simulated net household incomes (see Appendix Figure A.7).

[23]EU Statistics on Income and Living Conditions (EU-SILC) is coordinated by Eurostat and was launched in 2003 in seven countries (Austria, Belgium, Denmark, Greece, Ireland, Luxembourg, and Norway). In 2004, EU-SILC was introduced in fifteen further countries and in 2005, it was

WID. The WID offers long-run series of top income shares for nine European countries: Denmark, France, Germany, Ireland, Italy, the Netherlands, Norway, Spain, Sweden, Switzerland and United Kingdom.[24] Computing the Pareto parameter α from the country-specific top income shares documented in the WID, we then replace the top 1% of the country's gross household income distribution in EU-SILC survey data with Pareto imputed incomes.[25]

Figure 6 shows trends of Gini coefficients for gross household income in the nine European countries, for which both EU-SILC and WID-data are available, contrasting Ginis based on unadjusted data and imputed top income data.[26] The gap between top-corrected and unadjusted Ginis differs greatly between the countries and is mostly explained by the use or non-use of register data for EU-SILC provision.[27] The gap is negligible for countries, that use register information for income, like Denmark, Norway, the Netherlands, and Ireland. In these countries, top incomes seem to be well represented in EU-SILC data. The rapid increase in Norway's top-corrected Gini in 2005 is explained by an increase in dividends for top income earners in this year before the implementation of a permanent dividend tax in 2006 (Aaberge and Atkinson, 2010).

Surprisingly, Sweden and Switzerland reveal a sizable gap between top-corrected and unadjusted Ginis even though both rely on incomes from register data. Törmälehto (2014) also finds that Swiss EU-SILC data do not capture top incomes very well in a cross-country comparison with other register countries. He adjusts EU-SILC incomes to tax income definitions and finds a substantial difference for Swiss top income shares on adjusted tax incomes from EU-SILC and WID data. As for Sweden,

---

expanded to all EU-25 Member States. Until 2007, Bulgaria, Romania, Switzerland and Turkey joined EU-SILC.

[24]The WID-series for Portugal is only available until 2005, when EU-SILC was first conducted in Portugal.

[25]See Appendix Figure A.6 for income shares of the top 1% in European countries as provided by the WID. The provided top income shares in Italy, Germany and Sweden are including capital gains whereas capital gains are excluded in Switzerland, Denmark, France and Greece. For the remaining countries, Spain, Ireland, Norway, and the United Kingdom, information on the inclusion of capital gains is not provided in WID.

[26]WID years and EU-SILC years do not always coincide. Hence, top-corrected Ginis can only be computed for a subset of EU-SILC data years.

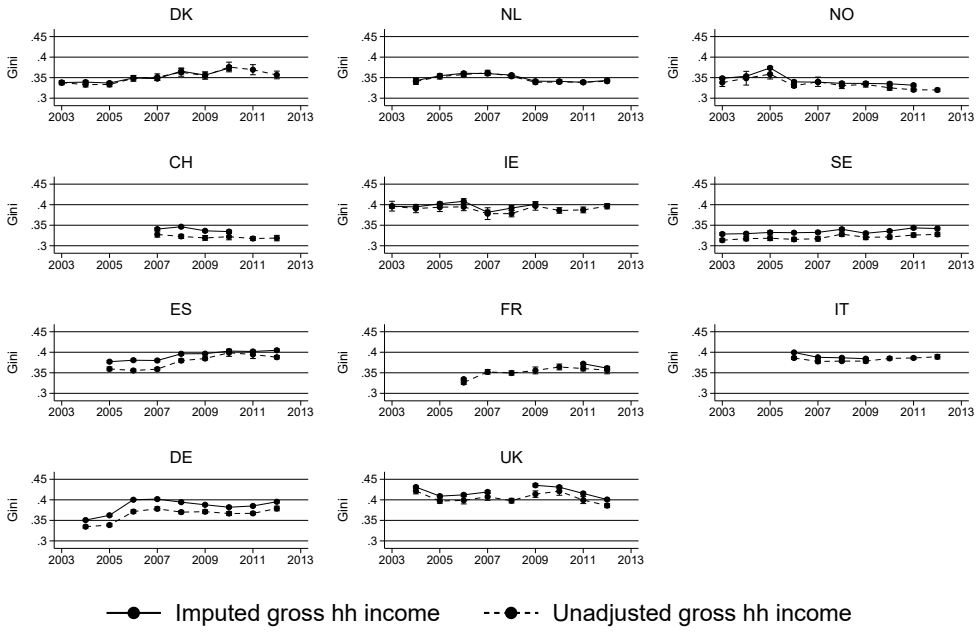
[27]See Jäntti et al. (2013) for an overview on the use of register and interview data in EU-SILC.

Frick et al. (2015) find large annual fluctuations of poverty rates in Sweden and a poverty rate in cross-sectional EU-SILC in 2006 that is twice as high as the poverty rate measured with longitudinal EU-SILC. They speculate that the complete elimination of households where income from a household member is missing (partial unit non-response (PUNR)) might lead to a misrepresentation of low and top income earners (which are more likely to refuse to reply) if no appropriate weighting takes place.

Countries that use both interview and register information also show quite a good picture of the top of the income distribution. The importance of at least partly using register data is stressed by the case of Spain: The top-correction gap in Spain decreases in 2008 when register information was first used.

Not surprisingly, the gap between top-corrected and unadjusted Ginis is largest in Germany and the UK, where EU-SILC is based on survey data only. Top corrected Ginis are 4 to 7% higher in Germany and 2 to 5% in the United Kingdom.

Figure 6: Top-corrected Gini of gross household income, European countries

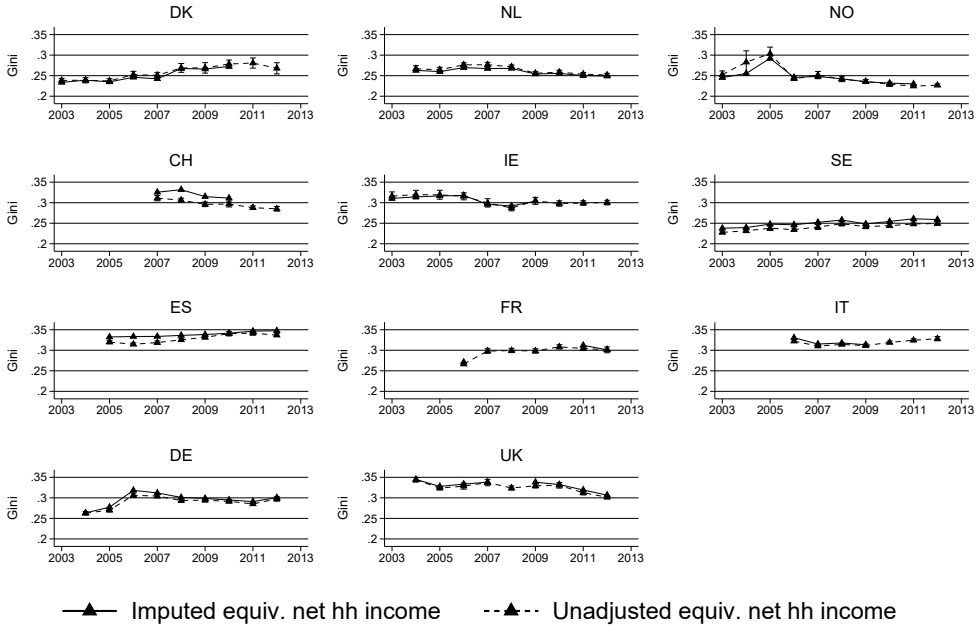


Source: EU-SILC (own calculations).
Note: For Ireland and the Netherlands the Pareto α is calculated with the income share ratios of top 1 % and top 0.5 %, since the income share of the top 0.1 % is currently not available in WID.

Figure 7 shows trends of Gini coefficients for living standards (equivalent net household income) in the same set of countries. The inequality difference induces by our integrated approach is smaller than for gross household incomes. Progressive tax systems in the European countries studied here play a major role in reducing the increased inequality in the gross income distribution. As for gross household income, the gap between top-corrected and unadjusted Ginis is almost negligible in most of the register countries and is largest in German EU-SILC data, which exclusively uses interviews to assess incomes. Top-corrected Ginis are 1 to 4% higher in Germany.

All in all, our top-correction approach merging information on the top 1% of the distribution from tax data with the bottom 99% of the distribution from survey data produces remarkably higher inequality levels in those countries that exclusively rely (Germany, UK) or have relied (Spain) on interviews for the provision of EU-SILC data.

Figure 7: Top-corrected Gini of living standards, European countries



Source: EU-SILC (own calculations).
Note: For Ireland and the Netherlands the Pareto α is calculated with the income share ratios of top 1 % and top 0.5 %, since the income share of the top 0.1 % is currently not available in WID.

5 Conclusion

This paper provides a new picture of recent inequality trends in EU countries using a novel top income imputation approach for survey data. We merge information on the top 1% of the distribution from tax data with the bottom 99% of the distribution from survey data. Inequality levels based on our top-corrected Gini coefficients are higher in those countries that exclusively rely (Germany, UK) or have relied (Spain) on interviews for the provision of EU-SILC data.

We first reconciled German survey and tax data and examined the extent to which differences in top income share estimates from household surveys and tax returns arise from differences in income concepts, observation units or from the ability to capture top incomes. We found that the top 1% is underrepresented in German SOEP data compared to tax data, but the lower percentiles of the top decile match very well. We find that different definitions of income and observation unit yield substantially different inequality levels in Germany: the Gini of tax income by tax unit is about 10%-points higher than the Gini of equivalent gross household income by household unit. The selected income concept is responsible for the largest part of this gap, whereas the observation unit changes inequality only slightly as most German households form a single tax unit anyway.

For our integrated approach for top-corrected Ginis, we estimated parameters of the Pareto distribution from top income shares and then replaced the top 1% of the survey income distribution by Pareto-imputed incomes. Our approach is easily applicable by relying on information publicly available at WID and easily accessible EU-SILC survey data. Neither access to tax record microdata, which is limited and difficult to obtain in many countries, nor record linkage, which is often not allowed, is needed. Of course, the applicability of the approach is restricted by the number of countries and years for which top income shares are available at the WID. However, we expect the WID to grow in the years to come such that our approach becomes usable for many additional countries and years. Furthermore, our integrated approach allows for producing a variety of measures for the inequality of living standards in the entire population of a country also considering differences in households' needs. Our top-correction method indicates similar trends and slightly

lower inequality levels than the decomposition approach (Atkinson, 2007; Alvaredo, 2011).

Finally, we applied our integrated approach to German SOEP data and European EU-SILC data. Our top-corrected Gini based on German SOEP data 2001-2012 is 4% higher for gross household income and about 2% higher for net equivalent household income than the unadjusted Gini. We estimated top-corrected Gini coefficients for European countries where information on the shape of the top of the income distribution is available in the World Wealth and Income Database (WID). The gap between unadjusted and top-corrected Ginis is highest in countries that rely (Germany, UK) or have relied (Spain) on interviews for the provision of EU-SILC data. Top corrected Ginis of gross household income are 4 to 7% higher in Germany and 2 to 5% in the United Kingdom. Top-corrected Ginis of equivalent net household income are 1 to 4% higher in Germany. This means, that German SOEP data provide a comparably better picture of top incomes than German EU-SILC data since inequality levels change less using our integrated approach. For most countries using register data, the gap between top-corrected and unadjusted Ginis is negligible since top incomes are already well-represented.

References

Aaberge, R. and A. B. Atkinson (2010). Top incomes in norway. In A. B. Atkinson and T. Piketty (Eds.), *Top incomes: a global perspective*, pp. 448–482. Oxford University Press.

Alvaredo, F. (2011). A Note on the Relationship between Top Income Shares and the Gini Coefficient. *Economics Letters 110*, 274–277.

Armour, P., R. V. Burkhauser, and J. Larrimore (2013). Deconstructing Income and Income Inequality Measures: A Crosswalk from Market Income to Comprehensive Income. *American Economic Review: Papers & Proceedings 2013 103*(3), 173–177.

Atkinson, A. B. (2007). Measuring Top Incomes: Methodological Issues. In A. B. Atkinson and T. Piketty (Eds.), *Top Incomes over the Twentieth Century: A Contrast Between Continental European and English-Speaking Countries*, Chapter 2, pp. 18–42. Oxford: Oxford University Press.

Atkinson, A. B., T. Piketty, and E. Saez (2011). Top Incomes in the Long Run of History. *Journal of Economic Literature 49*(1), 3–71.

Bach, S., G. Corneo, and V. Steiner (2009). From Bottom to Top: The Entire Income Distribution in Germany, 1992-2003. *Review of Income and Wealth 2*, 303–330.

Bartels, C. and K. Jenderny (2015). The Role of Capital Income for Top Income Shares in Germany. *World Top Incomes Database Working Paper Nr.1/2015*.

Biewen, M. and A. Juhasz (2012). Understanding Rising Inequality in Germany, 1999/2000 - 2005/06. *Review of Income and Wealth 58*(4), 622–647.

Bricker, J., A. Henriques, J. Krimmel, and J. Sabelhaus. (2015). Measuring Income and Wealth at the Top Using Administrative and Survey Data. *Finance and Economics Discussion Series 2015-030, Federal Reserve Board, Washington*.

Burkhauser, R., S. Feng, S. P. Jenkins, and J. Larrimore (2012). Recent Trends in Top Income Shares in the United States: Reconciling Estimates from March CPS and IRS Tax Return Data. *Review of Economics and Statistics 94*(2), 371–388.

Burkhauser, R., N. Hrault, S. P. Jenkins, and R. Wilkins. (2016). What Has Been Happening to UK Income Inequality since the Mid-1990s? Answers from Reconciled and Combined Household Survey and Tax Return Data. *IZA Discussion Paper No. 9718*.

Clementi, F. and M. Gallegati (2005a). Pareto's Law of Income Distribrution: Evidence for Germany, the United Kingdom, and the United States. In S. Y. A. Chakrabarti and B. Chatterjee (Eds.), *Econophysics of Wealth Distribution*, pp. 3–14. Springer Verlag.

Clementi, F. and M. Gallegati (2005b). Power Law Tails in the Italian Personal Income Distribution. *Physica A: Statistical Mechanics and its Applications 350*, 427–438.

Dagum, C. (1997). A New Approach to the Decomposition of the Gini Income Inequality Ratio. *Empirical Economics 22*, 515–531.

Feldstein, M. (1969). The Effects of Taxation on Risk Taking. *Journal of Political Economy 77(5)*, 755–764.

Frick, J., M. Grabka, and K. Krell (2015). Measuring the consistency of cross-sectional and longitudinal income information in eu-silc. *Review of Income and Wealth*. doi: 10.1111/roiw.12202.

Gerstorf, S. and J. Schupp (Eds.) (2016). *SOEP Wave Report 2015*.

Heathcote, J., K. Storesletten, and G. L. Violante (forthcoming). Optimal Tax Progressivity: An Analytical Framework. *The Quarterly Journal of Economics*.

Jäntti, M., V.-M. Törmälehto, and E. Marlier (2013). *The Use of Registers in the Context of EU-SILC: Challenges and Opportunities*. Publications Office of European Union.

Jenkins, S. P. (2016). Pareto Models, Top Incomes, and Recent Trends in UK Income Inequality. *IZA Discussion Paper No. 10124*.

Lakner, C. and B. Milanovic (2013). Global Income Distribution: from the Fall of the Berlin Wall to the Great Recession. *World Bank Policy Research Working Paper No. 6719*.

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

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

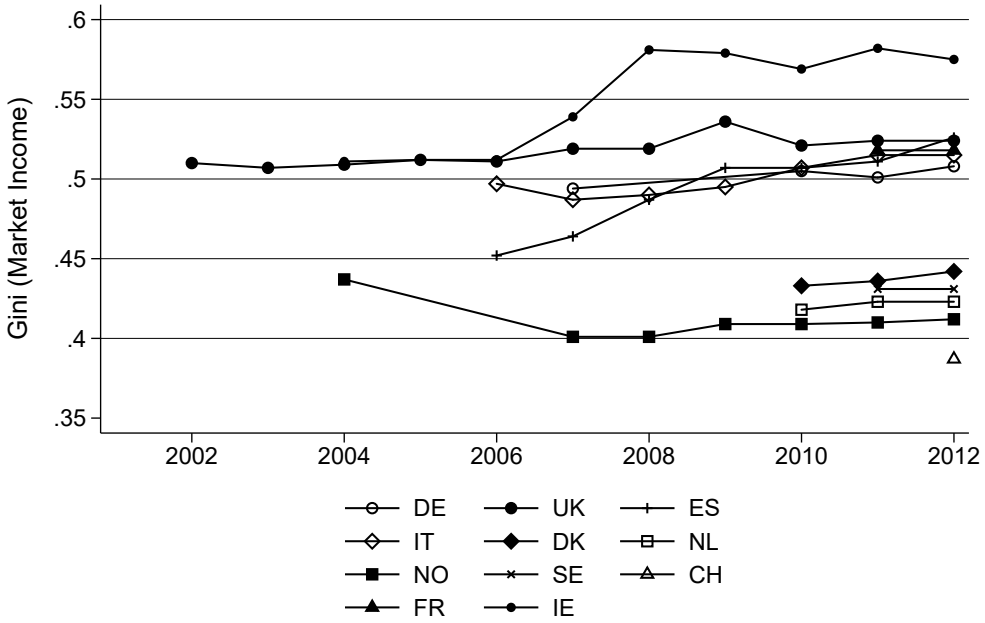
Roine, J., J. Vlachos, and D. Waldenström (2009). The Long-Run Determinants of Inequality: What Can We Learn from Top Income Data? *Journal of Public Economics 93*, 974–988.

Törmälehto, V.-M. (2014). High Incomes and Affluence: Evidence from EU-SILC. Unpublished.

Wagner, G. G., J. R. Frick, and J. Schupp (2007). The German Socio-Economic Panel Study (SOEP): Scope, Evolution and Enhancements. *Schmollers Jahrbuch 127*(1), 139–169.

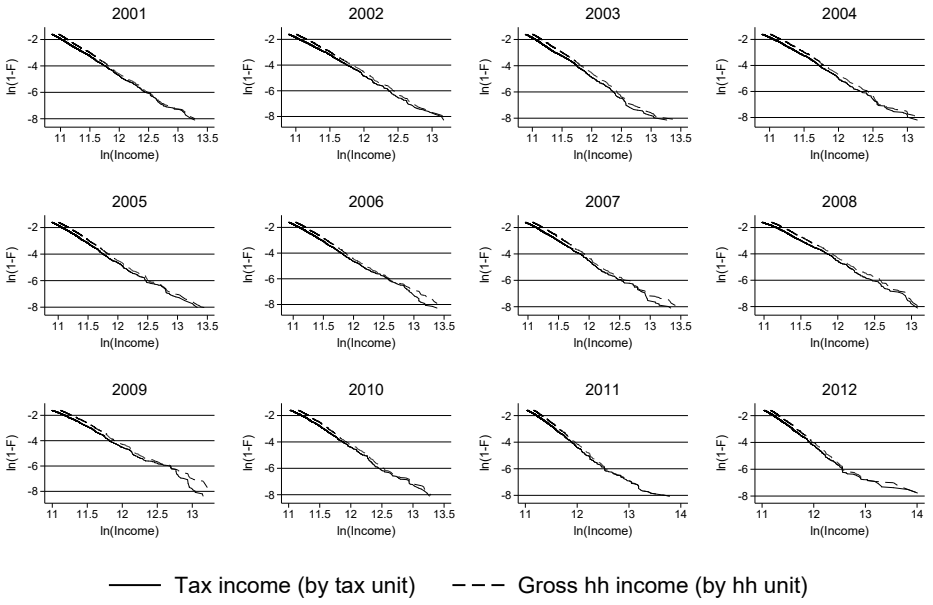
A Appendix

Figure A.1: Gini (Market Income), European countries



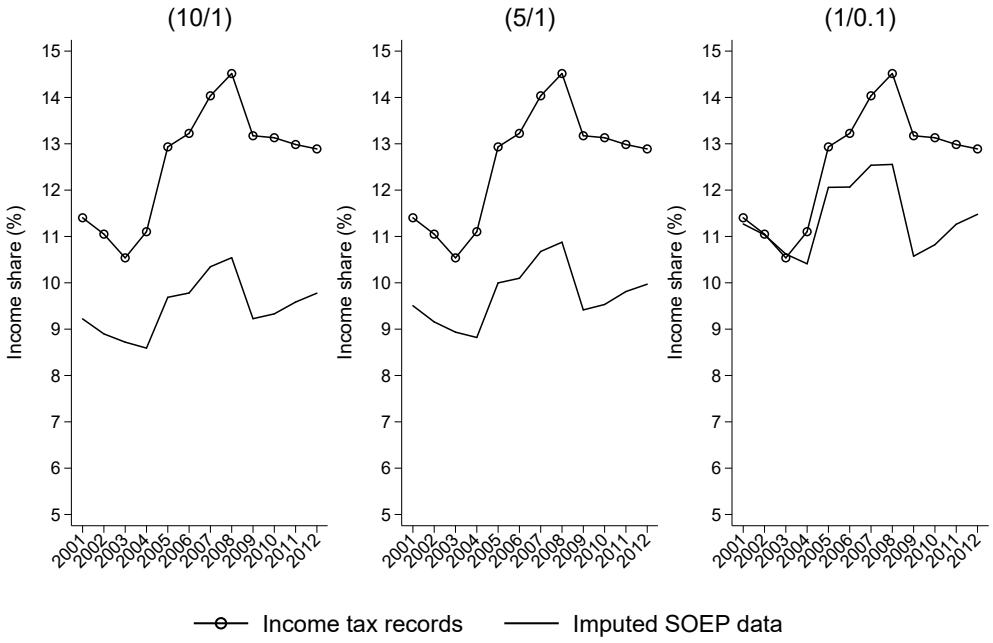
Source: OECD data.

Figure A.2: Tax income vs. gross household income distribution



Source: SOEP (own calculations).

Figure A.3: Income share of top 1 % with varying α specifications

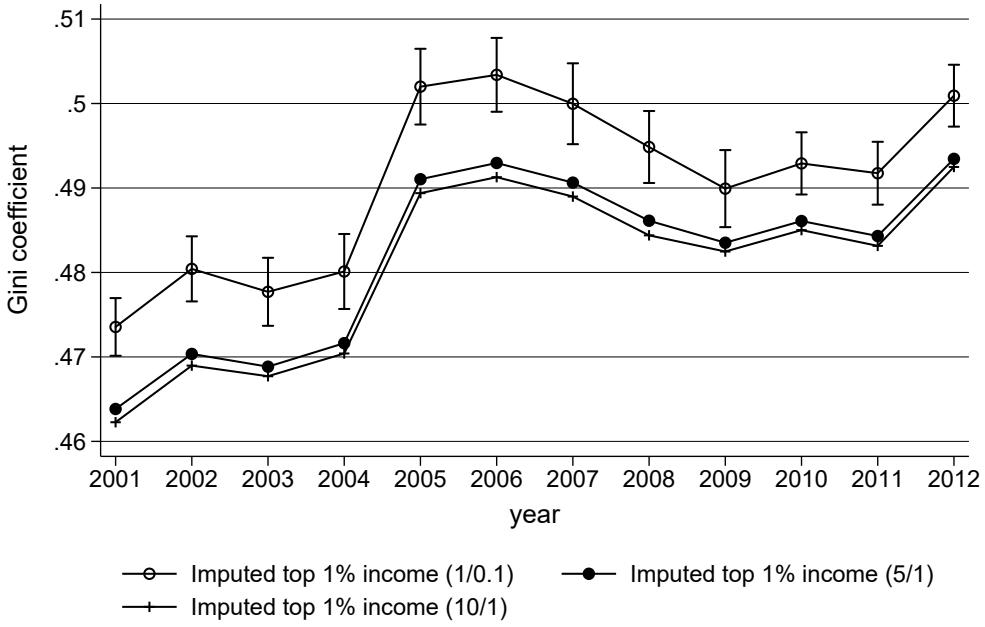


Source: SOEP (own calculations) and income tax records (Bartels and Jenderny, 2015) also available in WID.

Figure A.4: Top income shares (α 1/0.1)

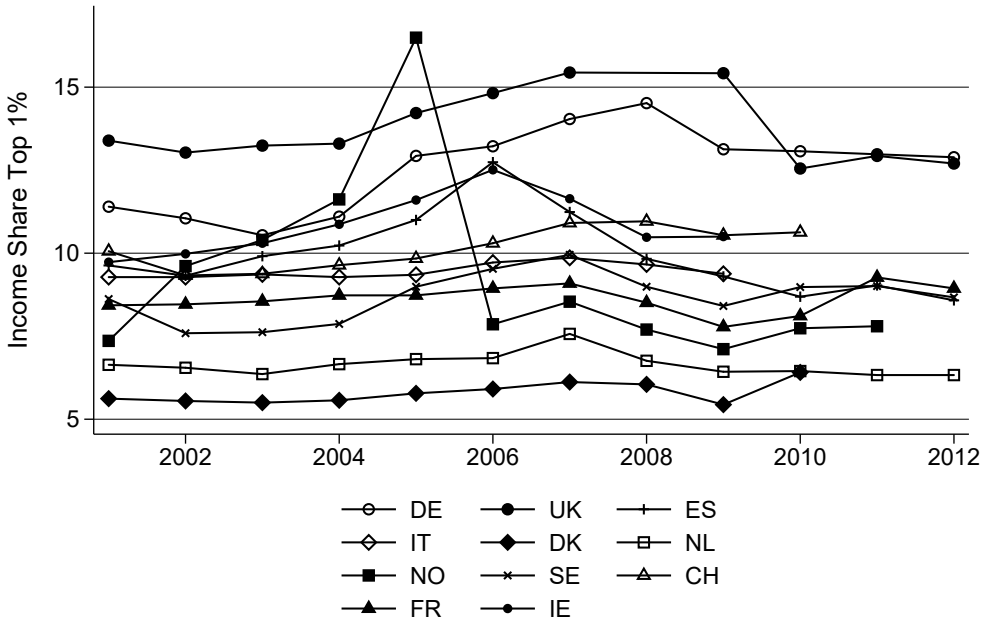
Source: SOEP (own calculations) and income tax records (Bartels and Jenderny, 2015) also available in WID.

Figure A.5: Top-corrected Gini coefficients, Germany



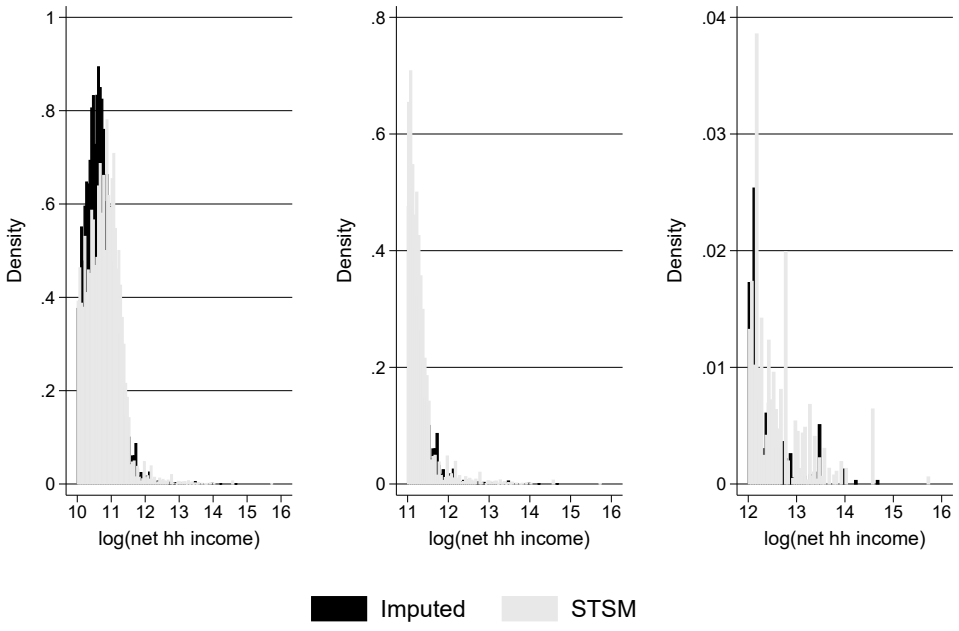
Source: SOEP (own calculations).

Figure A.6: Income share of top 1%, European countries



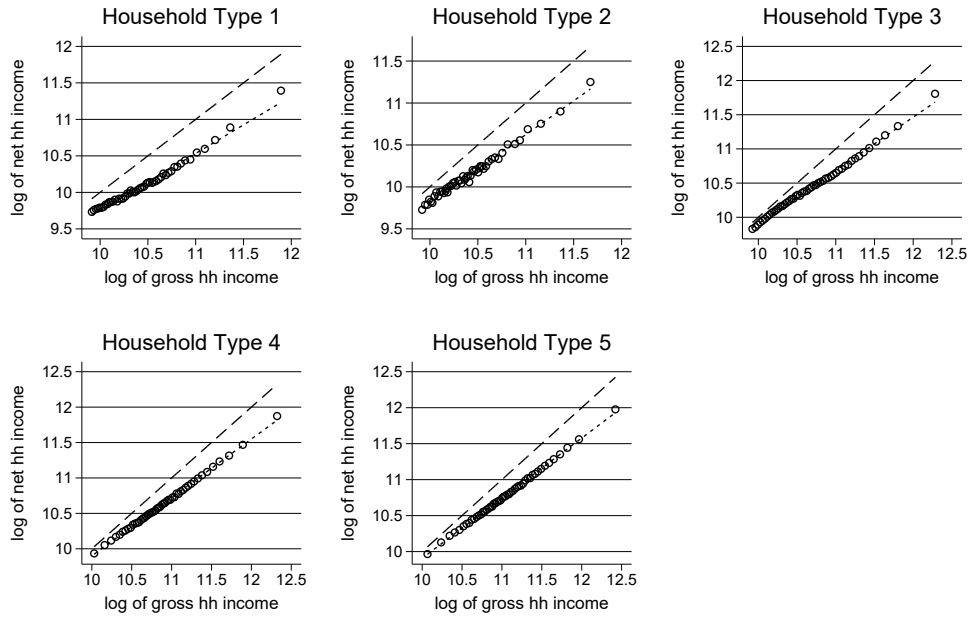
Source: World Wealth and Income Database (WID).

Figure A.7: Prediction of net household income: Feldstein vs. STSM (2011)



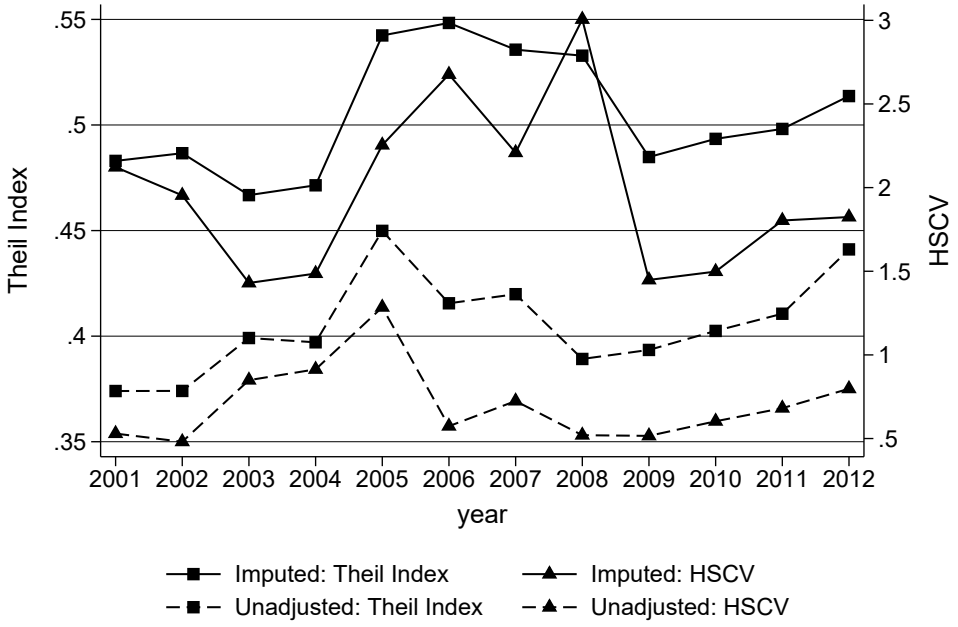
Source: SOEP (own calculations and STSM-calculations by Robin Jessen).

Figure A.8: Prediction of net income by household type



Source: SOEP (own calculations).
Note: Household types are defined as type 1=singles, type 2=singles with children, type 3=couples without children, type 4=couples with children, and type 5=other household compositions.

Figure A.9: Top-corrected Entropy Measures (imputed tax income), Germany



Source: SOEP (own calculations).

Table A.1: Relative change between income concepts (see Figure 2)

	tax^1 vs. tax^2	tax^1 vs. gross	tax^1 vs. net	tax^1 vs. equiv. net
2001	0.03	0.16	0.47	0.67
2002	0.03	0.17	0.49	0.69
2003	0.03	0.17	0.50	0.70
2004	0.03	0.17	0.49	0.68
2005	0.03	0.16	0.47	0.64
2006	0.03	0.17	0.48	0.68
2007	0.02	0.17	0.46	0.66
2008	0.03	0.17	0.44	0.65
2009	0.03	0.17	0.44	0.66
2010	0.03	0.16	0.44	0.66
2011	0.02	0.16	0.45	0.66
2012	0.03	0.17	0.46	0.65

Source: SOEP (own calculations).
Note: The observation unit for tax income tax^1 is the tax unit and for tax^2 the household unit.

Table A.2: Pareto distribution parameter, Germany (DE)

	$\alpha(10/1)$	$k_{tax}^{0.90}$ (SOEP)	$k_{tax}^{0.95}$ (SOEP)	$k_{tax}^{0.99}$ (SOEP)	$k_{tax}^{0.999}$ (SOEP)
2001	2.01	21548.44	19993.32	15644.77	10990.65
2002	2.06	23192.42	21515.61	16701.20	10740.12
2003	2.12	23331.32	21185.86	17100.62	11082.56
2004	2.05	23118.10	20812.00	15609.37	9739.74
2005	1.90	21402.39	18918.44	14237.16	9810.47
2006	1.86	21369.54	19012.32	13544.45	9975.71
2007	1.82	21355.31	18868.05	13235.94	9053.86
2008	1.79	21707.73	18451.06	12555.10	7896.92
2009	1.93	24322.80	21828.94	15592.65	10507.25
2010	1.93	25065.23	22217.63	16483.68	11697.83
2011	1.91	24927.31	22126.26	15966.65	12042.26
2012	1.93	25326.98	22704.71	16802.88	14714.34
	$\alpha(5/1)$	$k_{tax}^{0.90}$ (SOEP)	$k_{tax}^{0.95}$ (SOEP)	$k_{tax}^{0.99}$ (SOEP)	$k_{tax}^{0.999}$ (SOEP)
2001	1.94	20684.43	18956.71	14415.34	9720.91
2002	1.99	22302.07	20447.26	15443.51	9550.07
2003	2.05	22536.26	20251.44	15955.00	9987.75
2004	1.98	22287.86	19844.89	14508.35	8727.63
2005	1.84	20506.58	17894.79	13070.30	8629.45
2006	1.80	20495.11	18006.45	12458.67	8800.53
2007	1.76	20498.84	17889.53	12195.56	8007.63
2008	1.73	20861.14	17520.41	11594.91	7008.55
2009	1.89	23626.13	21019.03	14712.21	9630.00
2010	1.89	24331.68	21375.44	15533.00	10700.57
2011	1.86	24164.05	21248.92	15003.84	10969.60
2012	1.89	24669.41	21940.78	15941.69	13597.74
	$\alpha(1/0.1)$	$k_{tax}^{0.90}$ (SOEP)	$k_{tax}^{0.95}$ (SOEP)	$k_{tax}^{0.99}$ (SOEP)	$k_{tax}^{0.999}$ (SOEP)
2001	1.64	16722.30	14375.26	9421.70	5136.46
2002	1.65	17623.14	15051.82	9643.21	4712.16
2003	1.70	17928.88	15039.25	10098.10	5029.00
2004	1.67	17977.21	15003.85	9438.99	4579.92
2005	1.54	16187.28	13154.84	8144.17	4244.48
2006	1.54	16505.95	13586.85	8080.77	4597.06
2007	1.53	16814.73	13824.81	8205.84	4419.63
2008	1.53	17562.29	14005.00	8217.76	4181.75
2009	1.67	20152.79	17090.97	10704.42	5976.60
2010	1.66	20590.87	17202.59	11123.99	6485.07
2011	1.63	20309.21	16948.82	10598.61	6512.70
2012	1.64	20613.81	17368.85	11130.97	7933.50

Source: SOEP (own calculations) and income tax records (Bartels and Jenderny, 2015) also available in WID.
Note: α is obtained from top income shares based on income tax returns assuming that top incomes follow the Pareto distribution. Thresholds k are in current Euros.

Table A.3: Pareto distribution parameter, Germany

	$\alpha(1/0.1)$	$k_{tax}^{0.99}$ (WID)	$k_{tax}^{0.99}$ (SOEP)	$k_{hh}^{0.99}$ (SOEP)	$k_{hh}^{0.99}$ (EU-Silc)	$y_{tax}^{0.99}$ (WID)	$y_{tax}^{0.99}$ (SOEP)	$y_{hh}^{0.99}$ (SOEP)	$y_{hh}^{0.99}$ (EU-Silc)
2001	1.64	7747.29	9421.70	9917.31		127792.00	155411.00	163586.00	
2002	1.65	7745.92	9643.21	10174.36		125623.00	156393.00	165007.00	
2003	1.70	8429.08	10098.10	11072.01		125686.00	150573.00	165095.00	
2004	1.67	8517.02	9438.99	10110.41	8804.68	133828.00	148315.00	158865.00	138348.00
2005	1.54	6975.94	8144.17	8742.05	6741.92	137477.00	160499.41	172282.00	132865.00
2006	1.54	7320.62	8080.77	8675.91	9451.80	145810.00	160951.00	172805.00	188259.00
2007	1.53	7637.00	8205.84	8478.48	9337.39	155237.00	166800.00	172342.00	189801.00
2008	1.53	8126.81	8217.76	8800.03	8843.46	163361.00	165189.45	176894.00	177767.00
2009	1.67		10704.42	11555.13	11105.99		169190.00	182636.00	175537.00
2010	1.66		11123.99	11716.97	10496.53		178120.00	187615.00	168073.00
2011	1.63		10598.61	11373.46	10289.35		177274.70	190235.00	172102.00
2012	1.64		11130.97	11963.68	12244.18		183452.00	197176.00	201799.00

Source: SOEP data, EU-Silc data and income tax records (Bartels and Jenderny, 2015) also available in WID.

Note: α is obtained from top income shares based on income tax returns assuming that top incomes follow the Pareto distribution. The index *tax* indicates tax units and tax income whereas *hh* indicates household unit and household gross income. Thresholds k and y are in current Euros.

Table A.4: Coefficients and R^2 of the net household income estimation.

		$\hat{\tau}$		$\hat{\lambda}$		
Year	Household Type	β	SE	β	SE	R^2
2001	Single without children	0.24	0.01	8.64	0.12	0.79
	Single with children	0.21	0.03	6.36	0.27	0.89
	Couple without children	0.25	0.00	10.59	0.04	0.91
	Couple with children	0.20	0.00	6.48	0.05	0.93
	Other household type	0.19	0.00	5.90	0.05	0.94
2002	Single without children	0.24	0.01	8.30	0.12	0.78
	Single with children	0.20	0.04	5.72	0.39	0.82
	Couple without children	0.26	0.00	12.15	0.05	0.90
	Couple with children	0.20	0.00	6.85	0.05	0.93
	Other household type	0.19	0.01	6.19	0.06	0.93
2003	Single without children	0.26	0.01	10.45	0.12	0.76
	Single with children	0.26	0.04	10.83	0.40	0.79
	Couple without children	0.26	0.00	12.27	0.05	0.90
	Couple with children	0.20	0.01	6.95	0.06	0.92
	Other household type	0.20	0.01	6.72	0.07	0.92
2004	Single without children	0.25	0.01	9.49	0.12	0.77
	Single with children	0.24	0.04	9.42	0.40	0.81
	Couple without children	0.24	0.00	10.14	0.05	0.91
	Couple with children	0.20	0.01	6.58	0.06	0.93
	Other household type	0.17	0.01	4.98	0.07	0.93
2005	Single without children	0.19	0.01	5.33	0.10	0.82
	Single with children	0.19	0.03	5.31	0.27	0.91
	Couple without children	0.21	0.00	7.74	0.04	0.92
	Couple with children	0.16	0.00	4.55	0.05	0.94
	Other household type	0.16	0.01	4.30	0.06	0.95
2006	Single without children	0.21	0.01	6.21	0.11	0.81
	Single with children	0.20	0.03	6.28	0.27	0.90
	Couple without children	0.22	0.00	7.86	0.04	0.92
	Couple with children	0.17	0.00	4.69	0.05	0.95
	Other household type	0.15	0.01	3.96	0.06	0.94
2007	Single without children	0.21	0.01	6.36	0.11	0.83
	Single with children	0.22	0.03	7.26	0.28	0.91
	Couple without children	0.21	0.00	7.38	0.04	0.93
	Couple with children	0.17	0.00	4.99	0.05	0.95
	Other household type	0.14	0.01	3.80	0.06	0.94
2008	Single without children	0.19	0.01	5.45	0.10	0.85
	Single with children	0.21	0.03	7.09	0.30	0.90
	Couple without children	0.21	0.00	7.76	0.04	0.94
	Couple with children	0.17	0.01	4.94	0.05	0.95
	Other household type	0.15	0.01	4.07	0.06	0.95
2009	Single without children	0.20	0.01	5.82	0.10	0.85
	Single with children	0.18	0.03	5.07	0.36	0.88
	Couple without children	0.20	0.00	6.62	0.04	0.94
	Couple with children	0.16	0.01	4.60	0.06	0.95
	Other household type	0.16	0.01	4.59	0.06	0.95
2010	Single without children	0.18	0.01	4.93	0.09	0.86
	Single with children	0.24	0.03	9.37	0.28	0.87
	Couple without children	0.20	0.00	6.91	0.03	0.94
	Couple with children	0.17	0.00	4.92	0.05	0.95
	Other household type	0.16	0.00	4.60	0.05	0.96
2011	Single without children	0.19	0.01	5.33	0.09	0.86
	Single with children	0.20	0.02	6.25	0.22	0.91
	Couple without children	0.21	0.00	7.15	0.03	0.94
	Couple with children	0.18	0.00	5.59	0.05	0.95
	Other household type	0.16	0.01	4.70	0.06	0.96
2012	Single without children	0.20	0.01	5.98	0.08	0.86
	Single with children	0.16	0.02	4.37	0.18	0.94
	Couple without children	0.20	0.00	6.70	0.03	0.95
	Couple with children	0.17	0.00	4.78	0.04	0.96
	Other household type	0.17	0.00	4.91	0.05	0.96

Source: SOEP (own calculations).

Table A.5: Relative change of imputed and unadjusted Ginis (see Figure 5 and 6)

	Tax	Gross	Net income	Equiv. net
2001	0.05	0.05	0.02	0.02
2002	0.05	0.05	0.03	0.03
2003	0.03	0.04	0.01	0.01
2004	0.04	0.04	0.02	0.01
2005	0.04	0.05	0.02	0.01
2006	0.05	0.06	0.02	0.03
2007	0.05	0.05	0.02	0.02
2008	0.06	0.07	0.03	0.04
2009	0.04	0.05	0.02	0.02
2010	0.04	0.05	0.02	0.02
2011	0.04	0.05	0.02	0.02
2012	0.03	0.04	0.01	0.01

Source: SOEP (own calculations).

Table A.6: Pareto distribution parameter, Switzerland (CH)

	α (1/0.1)	$y_{tax}^{0.99}$ (WID)	$y_{hh}^{0.99}$ (EU-Silc)	$k_{hh}^{0.99}$ (EU-Silc)
2001	1.70			
2002	1.81			
2003	1.75			
2004	1.73			
2005	1.75	206638.25		
2006	1.73	216216.15		
2007	1.69	231100.34	292216.06	19154.56
2008	1.70	239240.83	333073.28	22185.50
2009	1.71	232507.30	316386.84	21410.55
2010	1.73	238015.94	362083.47	25277.80
2011			412014.25	
2012			403391.69	

Source: EU-Silc and WID.

Note: α is obtained from top income shares based on income tax returns assuming that top incomes follow the Pareto distribution. The index *tax* indicates tax units and tax income whereas *hh* indicates household unit and household gross income. Thresholds k and y are in current Euros.

Table A.7: Pareto distribution parameter, Denmark (DK)

	α (1/0.1)	$y_{tax}^{0.99}$ (WID)	$y_{hh}^{0.99}$ (EU-Silc)	$k_{hh}^{0.99}$ (EU-Silc)
2001	2.50	79219.37		
2002	2.51	80670.63		
2003	2.52	82217.75	196535.98	31675.65
2004	2.44	83809.36	200979.19	30492.32
2005	2.29	86519.79	205302.34	27474.90
2006	2.22	89192.24	235751.30	29525.40
2007	2.13	92963.48	222428.73	25712.87
2008	2.22	95814.45	231410.34	28966.80
2009	2.51	91813.53	218371.66	34939.42
2010	2.16	103980.64	254436.64	30218.34
2011			265319.09	
2012			265808.81	

Source: EU-Silc and WID.

Note: α is obtained from top income shares based on income tax returns assuming that top incomes follow the Pareto distribution. The index *tax* indicates tax units and tax income whereas *hh* indicates household unit and household gross income. Thresholds k and y are in current Euros.

Table A.8: Pareto distribution parameter, Spain (ES)

	α (1/0.1)	$y_{tax}^{0.99}$ (WID)	$y_{hh}^{0.99}$ (EU-Silc)	$k_{hh}^{0.99}$ (EU-Silc)
2001	1.92	63419.66		
2002	1.99	65418.98		
2003	1.87	68431.64		
2004	1.83	71908.82		
2005	1.73	77617.65	101715.89	7146.27
2006	1.61	87352.59	113679.00	6470.52
2007	1.70	88469.53	115596.00	7682.50
2008	1.83	87391.44	150111.91	12046.50
2009	1.87	85021.10	146844.70	12559.35
2010	1.99	82310.22	153389.00	15206.00
2011	1.89	82781.68	147660.70	12858.67
2012	1.96	79038.49	147515.80	14153.24

Source: EU-Silc and WID.

Note: α is obtained from top income shares based on income tax returns assuming that top incomes follow the Pareto distribution. The index tax indicates tax units and tax income whereas hh indicates household unit and household gross income. Thresholds k and y are in current Euros.

Table A.9: Pareto distribution parameter, France (FR)

	α (1/0.1)	$y_{tax}^{0.99}$ (WID)	$y_{hh}^{0.99}$ (EU-Silc)	$k_{hh}^{0.99}$ (EU-Silc)
2001	2.26	101887.44		
2002	2.25	104500.82		
2003	2.26	107123.91		
2004	2.23	110747.82		
2005	2.21	115596.03		
2006	2.12	121725.43	140185.00	15954.57
2007			204764.00	
2008			211679.00	
2009			212357.00	
2010			211602.00	
2011	1.84		231725.00	19086.56
2012	1.96		222801.00	21173.68

Source: EU-Silc and WID.

Note: α is obtained from top income shares based on income tax returns assuming that top incomes follow the Pareto distribution. The index tax indicates tax units and tax income whereas hh indicates household unit and household gross income. Thresholds k and y are in current Euros.

Table A.10: Pareto distribution parameter, Ireland (IE)

	α (1/0.5)	$y_{tax}^{0.99}$ (WID)	$y_{hh}^{0.99}$ (EU-Silc)	$k_{hh}^{0.99}$ (EU-Silc)
2001	1.96			
2002	1.95			
2003	1.94		224313.44	20889.47
2004	1.87		211204.94	17996.39
2005	1.80		240309.39	18606.28
2006	1.75		273320.63	19670.49
2007	1.85		242059.86	20083.61
2008	1.96		252914.66	24130.47
2009	1.98		227311.73	22208.58
2010			227739.13	
2011			208220.94	
2012			238584.83	

Source: EU-Silc and WID.

Note: α is obtained from top income shares based on income tax returns assuming that top incomes follow the Pareto distribution. The index tax indicates tax units and tax income whereas hh indicates household unit and household gross income. Thresholds k and y are in current Euros.

Table A.11: Pareto distribution parameter, Italy (IT)

	α (1/0.1)	$y_{tax}^{0.99}$ (WID)	$y_{hh}^{0.99}$ (EU-Silc)	$k_{hh}^{0.99}$ (EU-Silc)
2001	2.19	74677.68		
2002	2.17	76434.18		
2003	2.14	79285.64		
2004	2.16	80593.19		
2005	2.12	83359.74		
2006	2.03	88476.47	177373.00	18374.02
2007	2.04	91191.19	171725.00	17946.90
2008	2.11	92330.44	171084.00	19247.74
2009	2.18	91747.60	171358.00	20681.61
2010			173543.00	
2011			181935.00	
2012			183420.00	

Source: EU-Silc and WID.

Note: α is obtained from top income shares based on income tax returns assuming that top incomes follow the Pareto distribution. The index tax indicates tax units and tax income whereas hh indicates household unit and household gross income. Thresholds k and y are in current Euros.

Table A.12: Pareto distribution parameter, the Netherlands (NL)

	α (1/0.5)	$y_{tax}^{0.99}$ (WID)	$y_{hh}^{0.99}$ (EU-Silc)	$k_{hh}^{0.99}$ (EU-Silc)
2001	2.63			
2002	2.71			
2003	2.81			
2004	2.84		186423.00	36836.14
2005	2.70		206208.00	37459.74
2006	2.68		228506.00	40985.38
2007	2.89		251792.00	51168.23
2008	2.80		251638.00	48583.69
2009	3.07		224724.00	50139.92
2010	3.11		216192.00	49175.95
2011	3.13		213250.00	48967.89
2012	3.20		218220.00	51748.12

Source: EU-Silc and WID.

Note: α is obtained from top income shares based on income tax returns assuming that top incomes follow the Pareto distribution. The index tax indicates tax units and tax income whereas hh indicates household unit and household gross income. Thresholds k and y are in current Euros.

Table A.13: Pareto distribution parameter, Norway (NO)

	α (1/0.1)	$y_{tax}^{0.99}$ (WID)	$y_{hh}^{0.99}$ (EU-Silc)	$k_{hh}^{0.99}$ (EU-Silc)
2001	1.90	103499.68		
2002	1.57	113174.17		
2003	1.55	121842.87	232907.30	12009.33
2004	1.48	130071.67	225248.69	10090.90
2005	1.43	159333.81	287405.13	11482.28
2006	1.91	129805.26	224271.67	20184.44
2007	1.87	146481.36	268802.84	22992.37
2008	1.96	149459.87	268784.84	25703.21
2009	2.14	151308.68	268884.13	31118.02
2010	1.96	159462.65	307696.75	29493.41
2011	2.02	168755.35	342300.78	34979.82
2012			345246.09	

Source: EU-Silc and WID.

Note: α is obtained from top income shares based on income tax returns assuming that top incomes follow the Pareto distribution. The index tax indicates tax units and tax income whereas hh indicates household unit and household gross income. Thresholds k and y are in current Euros.

Table A.14: Pareto distribution parameter, Sweden (SE)

	α (1/0.1)	$y_{tax}^{0.99}$ (WID)	$y_{hh}^{0.99}$ (EU-Silc)	$k_{hh}^{0.99}$ (EU-Silc)
2001	1.69	82333.76		
2002	1.86	81535.86		
2003	1.81	84349.00	133180.77	10529.55
2004	1.81	91045.17	149718.02	11827.92
2005	1.76	103895.33	143207.31	10437.43
2006	1.67	115404.58	158546.31	10131.24
2007	1.68	127660.06	163094.56	10507.92
2008	1.73	121731.33	168932.00	11741.44
2009	1.86	121737.56	154852.86	13023.09
2010	1.80	131300.45	178906.92	13914.56
2011	1.77	132234.50	210301.00	15676.94
2012	1.78	134064.73	218356.52	16541.93

Source: EU-Silc and WID.

Note: α is obtained from top income shares based on income tax returns assuming that top incomes follow the Pareto distribution. The index tax indicates tax units and tax income whereas hh indicates household unit and household gross income. Thresholds k and y are in current Euros.

Table A.15: Pareto distribution parameter, United Kingdom (UK)

	α (1/0.1)	$y_{tax}^{0.99}$ (WID)	$y_{hh}^{0.99}$ (EU-Silc)	$k_{hh}^{0.99}$ (EU-Silc)
2001	1.82			
2002	1.86			
2003	1.86			
2004	1.82		246227.86	19633.48
2005	1.78		211534.66	15941.31
2006	1.74		253152.59	18050.73
2007	1.69		231953.23	15107.13
2008			203201.11	
2009	1.61		231381.73	13183.46
2010	1.76		248769.66	18043.18
2011	1.76		246300.00	17872.34
2012	1.79	120599.00	201250.00	15340.18

Source: EU-Silc and WID.

Note: α is obtained from top income shares based on income tax returns assuming that top incomes follow the Pareto distribution. The indix tax indicates tax units and tax income whereas hh indicates household unit and household gross income. Thresholds k and y are in current Euros.