the review of income and wealth

Review of Income and Wealth 2015
DOI: 10.1111/roiw.12188

CORRECTING FOR THE MISSING RICH: AN APPLICATION TO WEALTH SURVEY DATA

BY PAUL ECKERSTORFER

Parliamentary Budget Office, National Council of Austria

JOHANNES HALAK

Department of Philosophy and Theory of Science, University of Linz

JAKOB KAPELLER*

Department of Philosophy and Theory of Science, University of Linz

BERNHARD SCHÜTZ

Department of Economics, University of Linz

FLORIAN SPRINGHOLZ

Department of Philosophy and Theory of Science, University of Linz

AND

RAFAEL WILDAUER

Department of Economics, Kingston University London

It is a well-known criticism that if the distribution of wealth is highly concentrated, survey data are hardly reliable when it comes to analyzing the richest parts of society. This paper addresses this criticism by providing a general rationale of the underlying methodological problem as well as by proposing a specific methodological approach tailored to correcting the arising bias. We illustrate the latter approach by using Austrian data from the *Household Finance and Consumption Survey*. Specifically, we identify suitable parameter combinations by using a series of maximum-likelihood estimates and appropriate goodness-of-fit tests to avoid arbitrariness with respect to the fitting of the Pareto distribution. Our results suggest that the alleged non-observation bias is considerable, accounting for about one quarter of total net wealth in the case of Austria. The method developed in this paper can easily be applied to other countries where survey data on wealth are available.

JEL Codes: C46, C81, D31

Keywords: non-observation bias, Pareto distribution, wealth distribution

Note: We thank René Böheim, Bettina Csoka, Yannis Dafermos, Pirmin Fessler, Paul Hofmarcher, Peter Lindner, Maria Nikolaidi, Martin Riese, Matthias Schnetzer, Stefan Steinerberger, and Engelbert Stockhammer for valuable comments. This paper was supported by the Austrian Chamber of Labor and uses data from the Eurosystem Household Finance and Consumption Survey. The opinions expressed in this article are the authors' own and do not reflect the view of the Parliamentary Budget Office of the National Council of Austria.

*Correspondence to: Jakob Kapeller, Department of Philosophy and Theory of Science, University of Linz, Altenbergerstraße 69, 4040 Linz, Austria (jakob.kapeller@jku.at).

1. Introduction

The recent interest in distributional issues came along with the publication of unprecedented data sources such as "The World Top Income Data Base," the wealth series published in Piketty and Zucman (2014), or the Eurozone's *Household Finance and Consumption Survey* (HFCS). In this context, the investigation of the distribution of wealth is either based on surveys, like the HFCS, or on administrative data, which is mostly based on tax return statistics. Both methods suffer from their own limitations, even though administrative data are often considered to be more reliable than survey data (e.g., Piketty, 2014). However, the differences between both data sources are subtle: while, for instance, inheritance tax data allow for estimating representative results for the top age cohort, such data do not necessarily deliver solid estimates related to the entire population. Additionally, the concept of "wealth" used in most administrative sources might be considerably narrower than expected at first hand, simply due to tax exemptions. Tax avoidance might be an additional problem, especially if rich households are more prone to conceal their assets or have more resources at their disposal to minimize taxes within or beyond legal boundaries.

Unfortunately, survey data are far from perfect. Since voluntary participants are not required by law to provide correct answers, the reliability of the collected data is often judged worse as compared to tax reports, where such a legal enforcement is indeed the case. Moreover, sample sizes of typical surveys are much smaller than sample sizes in administrative data, which leads to a downward bias in final estimates, since typical surveys do not capture the subtle distributional properties at the very top of the wealth distribution. This lack of reliability due to small sample sizes—labeled "non-observation bias" in what follows—is especially relevant for the top of the distribution. In a similar vein, evidence from the Fed's Survey of Consumer Finance (SCF) shows that rich households are less likely to participate in surveys about household wealth. Such a differential non-response across different levels of wealth gives rise to sample selection problems (Kennickell and McManus, 1993; Singer, 2006), which lead to a "non-response bias" (ECB, 2013, p. 10). Nevertheless, there are also advantages of survey data. Probably the most important one is that surveys usually collect not only the variable of interest, but also a rich set of supplementary information, and thus allow for asking different and more nuanced research questions. Moreover, high quality surveys such as the SCF and the HFCS, employ administrative data to improve the sample design. Specifically, these data allow for taking non-observation and non-response problems into account by means of oversampling of rich households.[1]

In this context our study focuses on the phenomenon of the "missing rich," that is, the underrepresentation of very rich households in surveys caused by non-observation bias. While in principle the probability of non-observation is the same for all households in the population and independent of the level of net wealth, only omissions at the top cause a significant bias, since in the segment of the very rich (those within the upper 0.5–1 percent), very few households decisively influence estimations for total wealth and wealth inequality due to the power-law

[1]Within the HFCS this strategy was implemented in the case of Belgium, Germany, Greece, Spain, France, Cyprus, Luxembourg, Portugal, and Finland (cf. ECB, 2013, p. 10).

characteristics of the underlying distribution (Avery *et al.*, 1986; Hoeller *et al.*, 2012). In short: wealth at the very top is so skewed, that the few households drawn from this segment are not considered as representative for the underlying population. Correcting for non-observation bias would allow for addressing one of the major drawbacks of survey data compared to administrative data. This latter point is of special importance for those surveys which refrain entirely from oversampling or only use geographical information instead of administrative tax data to design the oversampling strategy. Under the hypothesis that wealth at the very top follows a Pareto distribution,[2] we show that wealth estimates based on the fitted parameters of that distribution are able to correct for non-observation bias arising from small sample sizes.

Since neither the issue of the "missing rich" in wealth survey data, nor the correction via fitted Pareto distributions are entirely novel, these issues have already motivated a series of studies. Cowell (2011), for instance, illustrates how the estimates for wealth inequality depend on the scale parameter (i.e., the minimum value m) of the Pareto distribution using data for the Netherlands, Sweden, the U.K. and the U.S. Bach *et al.* (2010, 2014) and Bach and Beznoska (2011) estimate a Pareto distribution (i.e., the shape parameter α) for Germany based on a journalists' list of the richest German households, where they choose the scale of the distribution on an ad-hoc basis. Vermeulen (2014) uses HFCS data and presents estimates of the shape parameter based on arbitrarily chosen minimum values. Additionally, he compares these results to estimates where he adds observations from the Forbes list of billionaires to the sample. Finally, Eckerstorfer *et al.* (2014) estimate the top of the wealth distribution for Austria in a way that is quite similar to the one in the present study. However, they also choose the scale parameter on a simple ad-hoc basis.

This study proposes a novel method to correct for the absence of the "missing rich" in survey data, which has several advantages compared to previous approaches. First, and most importantly, it employs a non-arbitrary strategy to determine the parameters of the Pareto distribution based solely on statistical testing. Specifically, we do not have to resort to ad-hoc assumptions when choosing the scale parameter, which is a key difference from other recent and similar contributions like Vermeulen (2014) or Eckerstorfer *et al.* (2014). Second, our approach does not rely on additional sample information. External data such as rich lists provided by popular magazines are often not available and entail unresolved concerns about data quality. Third, we motivate our approach solely with reference to a non-observation bias and show that the latter is already a sufficient condition for receiving biased survey estimates. Hence, we abstain from assuming a differential non-response bias—that is, that wealthier households have a higher propensity to refuse participation in surveys—which would reinforce the downward bias in survey data but is hardly observable in practice.[3] However, our proposed approach is nonetheless applicable to situations where non-observation

[2]According to the survey of Davies and Shorrocks (2000), empirical work on personal wealth-holdings shows that the top tail of wealth distributions is indeed well approximated by a Pareto distribution.

[3]Although there exists some empirical evidence pointing to this possibility (see, e.g., Kennickell and McManus, 1993).

and differential non-response biases occur simultaneously: since both biases lead to a lack of information at the top of the distribution, the very same treatment should also suffice to remedy such a collusion of biases.

The remainder of the paper is structured as follows. Section 2 motivates our paper and illustrates the non-observation bias and its emergence via a Monte Carlo experiment. This approach allows us not only to illustrate the bias, which arises when small samples are drawn from skewed distributions, but also to show how estimates can be improved by means of the suggested procedure in a relatively simple setup, given the assumption that a Pareto distribution is indeed valid. The third section focuses on this latter point and explains the more specific method used to test the validity of the underlying Pareto assumption and to correct the original data for the case of the Austrian HFCS sample. Section 4 presents the corrected wealth measures and compares them with the original HFCS statistics as well as non-estimated data from the Austrian counterpart of the Forbes list of billionaires (the so called *Trend list*). Section 5 presents upper and lower bounds with respect to irregularities and outliers within the dataset. Section 6 concludes.

2. Non-Observation Bias in Surveys: An Illustration

Heavily skewed distributions behave very differently compared to commonly used normal distributions. In the case of the Pareto distribution a small number of observations at the very top have a strong impact on aggregate parameter estimates. Thus, a sample from a population following a Pareto distribution needs to be rather large in order to capture some of these few but highly important observations at the very top, which are necessary to obtain representative results. Since actual wealth surveys only cover tiny fractions of the underlying population (SCF 2010: 0.06‰; HFCS net samples: e.g., Austria 0.63‰, Germany 0.09‰, France 0.54‰) standard sampling strategies under these circumstances lead to a significant underestimation of population quantiles.[4] This section is dedicated to illustrating this problem by means of a Monte Carlo simulation. In turn we also demonstrate the effectiveness of estimating a Pareto distribution in order to deal with the underlying bias.

Assume the following setup: the top of a country's wealth distribution consists of $N = 500{,}000$ households. Within this group net wealth is greater than or equal to €100,000 (our scale parameter m) and distributed according to a Pareto distribution with a shape parameter $\alpha = 1.3$. How well can one estimate the population's total wealth based on random sampling?

In order to answer this question, we first compute each household's wealth holdings based on the distribution parameters. This can be done by exploiting the facts that the complementary cumulative distribution function (ccdf) in the Pareto case is given by $P(X > x) = (m/x)^{\alpha}$ and that $P(X > x)$ can be approximated by r/N, where r is the rank of the individual with wealth x and N denotes the population size. As a result each individual's wealth can be computed as a function of its rank

[4]Given that the sample design does not include or only includes low quality oversampling procedures (using geographical instead of information on household income and wealth from tax statistics).

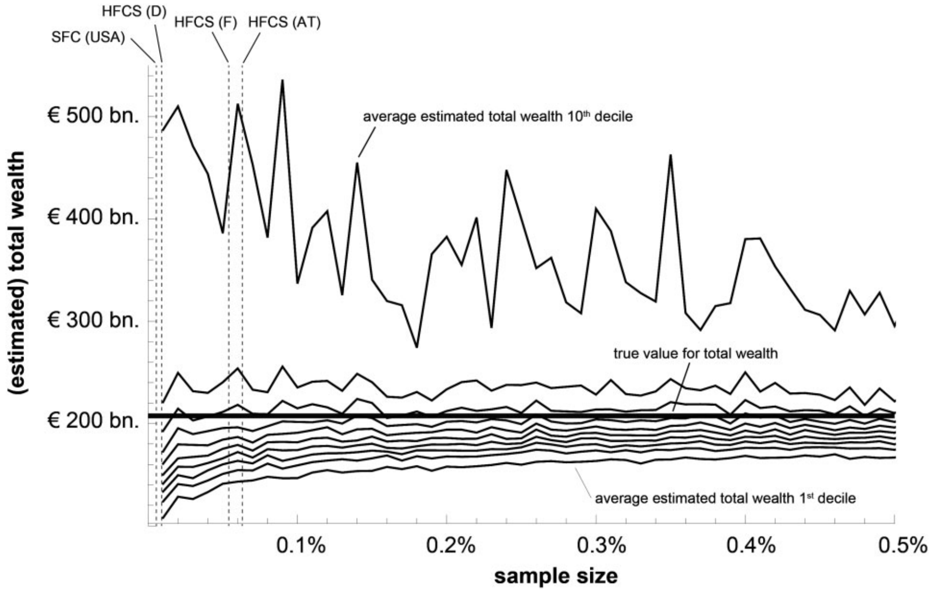

Figure 1. Decile Averages of the Estimated Population Wealth for Sample Sizes from 0.1‰ to 0.5% (estimations were grouped according to estimated population wealth); estimation based on sample mean

r, population size N, and the two parameters of the distribution (m and α). Total wealth of this population is equal to €208 billion.

In a next step we draw random samples with increasingly large sample sizes (n), beginning from 0.1‰ in steps of 0.1‰ up to 0.5 percent of the population. Since we draw 200 independent samples in each step, we arrive at a total of 10,000 synthetic datasets and estimate the population wealth based on the means of these samples. The telling results are summarized in Figure 1: we group the 200 wealth estimates obtained in every step into deciles and plot the decile averages. Accordingly, the lower line in Figure 1 shows the average estimated total wealth in the first (i.e., lowest) decile for increasingly large sub-samples, while the top line shows the average estimated total wealth for the 10th decile. Furthermore, the gray horizontal line represents the true population wealth and the two vertical lines show the actual sample sizes used by the SCF (0.06‰) and the HFCS in Austria (0.63‰). There are three striking results. First, independent of the sample size, estimates in the first seven deciles underestimate total wealth. Second, estimates in the 10th decile yield extremely volatile and exaggerated estimates. Third, underestimation in the lower deciles as well as overestimation in the 10th decile decreases with increasing sample size.

Figure 1 directly visualizes the non-observation problem: in most cases the samples do not contain enough very rich individuals and therefore underestimate total wealth at the top of the distribution. In contrast, when the sample does contain observations of very rich individuals, implicitly the weight assigned to these observations is too high and total wealth is overestimated accordingly. When interpreting these results it is important to keep in mind that surveys such as the SCF and HFCS use samples of 0.06‰ to 0.6‰ of the underlying population. For such small segments biases are most intense and underestimation of top wealth individuals is highly probable.

The next step is to demonstrate how estimates of the population wealth can be improved by using estimated distribution parameters instead of the sample mean. Thus for each sample the shape parameter is estimated based on a maximum likelihood approach.[5] Again we draw on the properties of the ccdf to obtain estimates of the population wealth. Similar to Figure 1, the upper panel of Figure 2 shows the decile averages of the estimated total wealth based on fitted distribution parameters. In the lower panel, the median as well as the 25th and the 75th percentiles of the mean-based (gray lines) and Pareto-based estimates (black lines) are compared. Both graphs depict samples from $n = 50$ (i.e., 0.1‰) up to $n = 2500$ (i.e., 0.5 percent). The decile plot shows that estimates based on distribution parameters do not systematically underestimate population wealth. Rather, over- and under-valuations occur with roughly equal probabilities. Moreover the plot on the right reveals that the median of Pareto-based estimates is very close to the underlying true value. In comparison, in the mean-based approach the median constantly fluctuates well below that. In addition the 25th percentile of the distribution-based estimates is much closer to the true value than its mean-based counterpart. The Pareto-based 75th percentile is larger than the mean-based, especially for small sample sizes up to 1‰. Overall the right panel of Figure 2 demonstrates that the maximum likelihood approach is able to remove the systematic underestimation and to decrease the estimator's variance.

The results presented above are robust to the chosen parameter values of N, m and α. Replications of this exercise with α's in the range of 1.1 to 2 can be found in the online Appendix (part I); further results based on a variation of a broader set of parameter values can also be found in the Appendix. After having demonstrated that the non-observation problem of the "missing rich" can be tackled by fitting a Pareto distribution (under the hypothesis that such a distribution describes the data well), we move on to bring that procedure to a real world application. In doing so we are leaving the neatly defined Monte Carlo environment behind and have to deal with the twin problem of empirically determining statistically adequate distribution parameters and critically asking whether our estimated distribution really represents the underlying data. The remaining paper will be dedicated to these issues.

3. Estimating the Missing Rich: A Methodological Suggestion

In short our methodical approach can be described as follows: in a first step, we fit a Pareto distribution[6] to the upper tail of the HFCS sample. At this stage

[5]Starting from the pdf $p(x) = \alpha \cdot m^{\alpha}/x^{\alpha+1}$ one obtains the log-likelihood function $L(\alpha, x) = \sum_{j}^{n} \ln\alpha + \alpha \ln m - (\alpha+1)\ln x_j$ and maximization yields $\hat{\alpha} = n\big/\left(\sum \ln x_j/m\right)$. For derivations with slightly different notation, see Hill (1975) or appendix B in Clauset *et al.* (2009).

[6]The assumption of a Pareto distribution is widely used in studying the distribution of income (see, e.g., Feenberg and Poterba, 1993; Piketty, 2003; Piketty and Saez, 2003) and wealth (e.g., Kopczuk and Saez, 2004; Alvaredo and Saez, 2009; Bach *et al.*, 2010, 2014; Durán-Cabré and Esteller-Moré, 2010; Bach and Beznoska, 2011; Eckerstorfer *et al.*, 2014; Vermeulen, 2014).

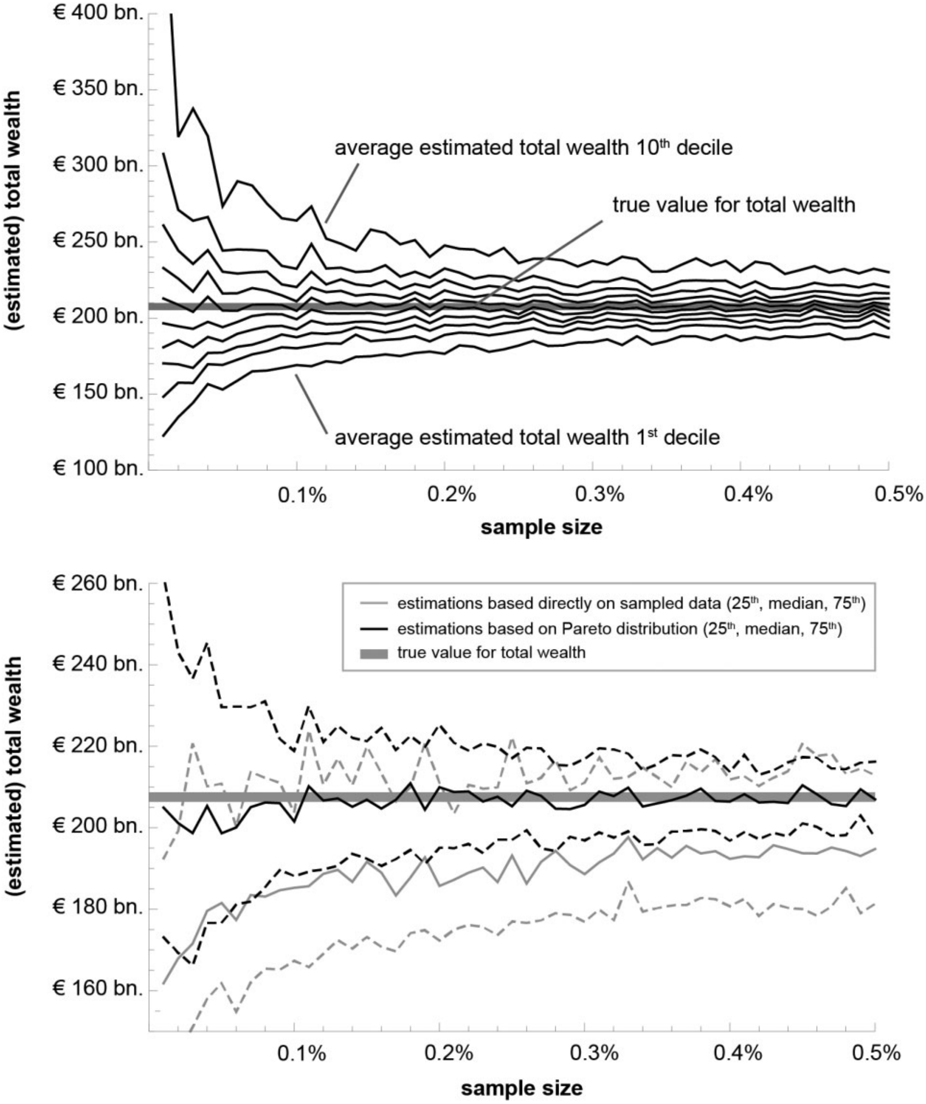

Figure 2. Top Panel: Decile Averages of the Estimated Population Wealth for Sample Sizes from 0.1‰ to 0.5% (based on a Pareto distribution). Bottom Panel: 25th, 50th, and 75th Percentiles for Estimations Based on the Sample Mean (gray lines) and a Pareto Distribution (black lines)

the selection of the adequate scale parameter m is of crucial importance, since choosing a lower bound that is too high ignores useful information (which reduces the precision of the estimate), while choosing a lower bound that is too low leads to a bias in the results (since estimation would involve non-Pareto distributed data). Therefore, we apply in a second step a bootstrap procedure to test the validity of the distribution estimated in the first step. These two steps closely follow the method suggested in Clauset *et al.* (2009). In a third step we eliminate all observations with reported net wealth beyond a €4 million cut-off point from the sample and replace them with data points based on the previously estimated Pareto distribution. We illustrate our method using data from the HFCS for Austria, but in principle this method can be applied to survey data for any other country. Figure 3 provides a graphical illustration of these three steps. The elimination mainly affects observations from the 100th percentile and accounts for the fact that the sample does not contain observations on net wealth exceeding €14 million but claims to be representative for the whole population. In this context the assumption that the data follow a Pareto distribution in the first step is not a restrictive one, since we are only interested in the upper tail. Commonly used alternative distributions for modeling household wealth, such as Dagum or Singh–Maddala, are designed to describe the entire interval of positive household wealth and the upper tail of those distributions converges toward a Pareto distribution. Dagum (2006) even considers this latter property as an essential characteristic for any distribution used to describe the behavior of household wealth.

In what follows we find that accounting for non-observation bias of the very rich households increases aggregate net wealth (compared to the original HFCS data) by roughly 28 percent to €1278 billion. The share of the top 1 percent increases from 23 to 38 percent and the share of the top 5 percent increases from 48 to 59 percent. Although we limit the maximum wealth obtained from the estimated Pareto distribution to €1 billion per household for reasons of conservatism, that is, we do not use the upper part of the distribution's tail (see below), we find that the latter describes the total wealth of the richest families as reported by the Austrian list of billionaires reasonably well (deviating +5 percent for the richest 30 and –0.7 percent for the richest 60; See section 4). Since this list, however, fails to distinguish between households and family clans, and the quality of the data is also difficult to assess, it must be used with caution.

3.1. *Data*

The Household Finance and Consumption Survey (HFCS) is the first comprehensive survey on tangible assets, financial wealth, liabilities, and expenditures of private households carried out in parallel in 15 countries of the Euro Zone (Ireland and Estonia chose to opt out of the HFCS). In Austria the Austrian National Bank conducted the survey in cooperation with the Institute for Empirical Social Studies (IFES). In what follows we give a brief overview of the survey design (for more details, see Albacete *et al*., 2012).

The basic reporting unit in the HFCS is the household, which is represented by the single person within the household who felt most competent with regard to the household finances. The survey is based on personal interviews conducted between September 2010 and May 2011. The initial sample consisted of 4436 households; 2380 households have been successfully interviewed, indicating a response rate of roughly 56 percent. The selection of households is based on a two-stage *stratified sampling design*, to ensure that the randomly drawn participants adequately reflect the composition of the Austrian population. Stratification was based on Austrian NUTS-3 regions and municipality size in order to ensure

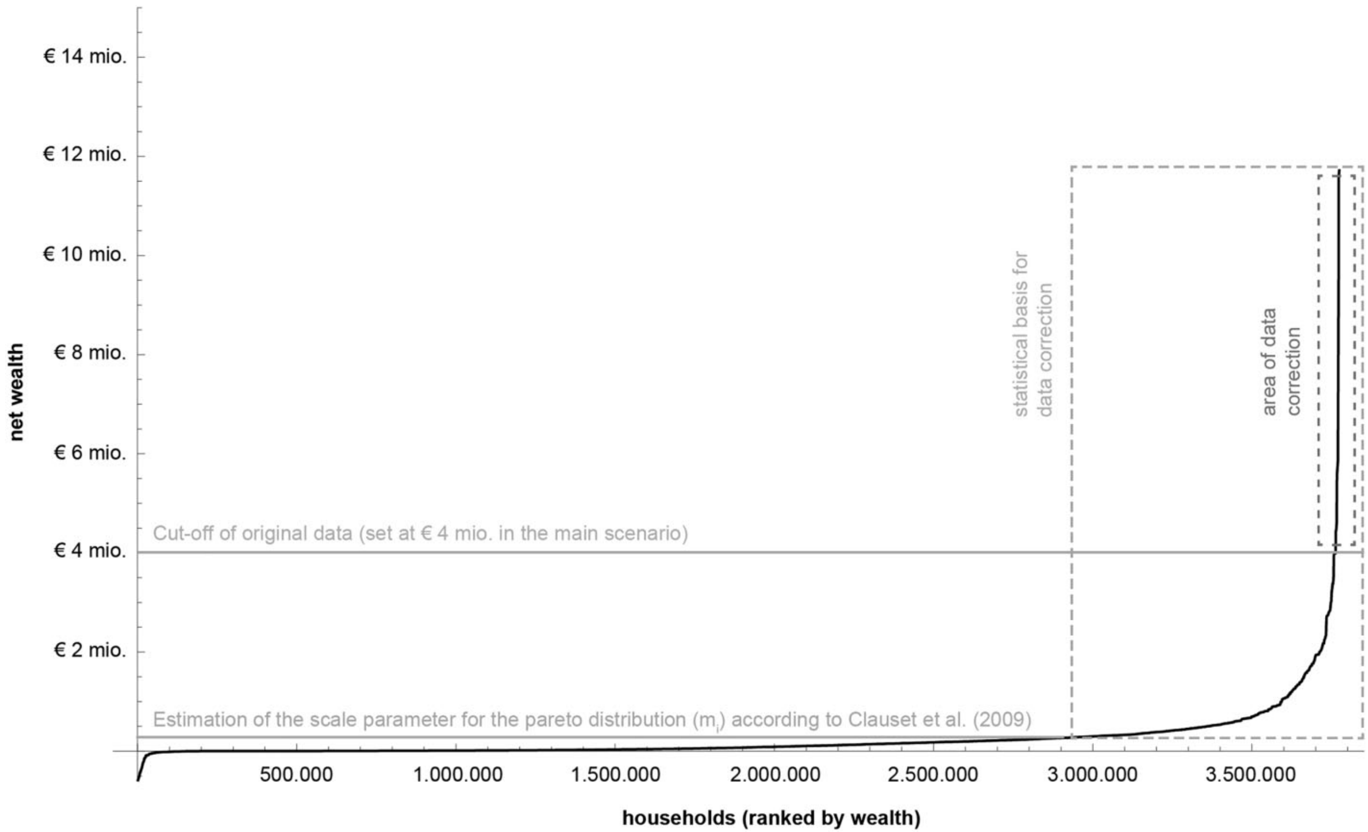

Figure 3. A Sketch of the Methodological Strategy

that households from different regions enter the sample proportionally. Data collection was based on computer-assisted interviews.

It is typical for surveys, especially those that try to evaluate sensitive information such as wealth, income, or debt, that participants refuse or are unable to answer certain questions (item non-response), which can bias the results. In order to reduce such a bias, missing values were inserted ex post using *multiple imputation*. In this process, missing values are replaced through estimated values. This preserves the correlation structure of the dataset, since one does not have to drop all incomplete observations.[7] Imputation was repeated five times, producing five different samples—so called *implicates*. The problem of non-observation of very rich households (problem of *coverage*), however, cannot be compensated for with this method (nor could any potential bias arising from differential non-response be compensated by these means).

Each observation in the dataset received a probability weight in order to adjust the sample to the statistical population and to reduce the sample variance. The final survey weights emerge from the *design weights*, which account for unequal probability to be part of the sample due to the stratified sample design (*unequal probability sampling bias*); the *post-stratification weights*, which try to correct for erroneous exclusion, for instance in the case of a wrong postal address (*frame bias*); and *non-response weights*, which try to correct for unequal probabilities of households not participating in the sample (*non-response bias*). Weight corrections in the latter case only account for factors like the experience of the interviewer but do not correct for the potential linkages between household wealth or income and response probabilities.

These preliminary measures, however, cannot resolve the already discussed problem of downward-biased estimates at the top of the distribution. This claim is easily confirmed by a quick look at the gap between the richest household in the survey (net wealth of €14 million[8]) and the richest individual reported by the Forbes list (€3.6 billion, see Table A1 in part III of the online Appendix) or the 60th rank (i.e., the "poorest" household) of the list of wealthy Austrian individuals and families published by *Trend* magazine (€405 million, see also Table A1 in part III of the online Appendix). In what follows we provide a statistical approach to compensate for this gap and thereby to improve the reliability of estimates by correcting for the "missing rich" at the top end of the distribution.

3.2. *Estimating the Distribution Parameters*

Generally, a Pareto distribution has to be established for each implicate. The associated cumulative distribution function is denoted in the following way:

[7]If the probability of non-response for certain questions correlates with household characteristics, this gives rise to item-non-response bias in sample estimates. By means of multiple imputation this bias is corrected and thus we do not address it further. Throughout the paper we use non-response as a synonym for unit-non-response in contrast to item-non-response.

[8]Average net wealth of the richest household across all five implicates.

$$(1) \qquad P_i(x_i) = \Pr(X_i \le x_i) = 1 - \left(\frac{m_i}{x_i}\right)^{\alpha_i} \ \forall \ implicates \ i = 1 \ldots 5 \wedge x_i \ge m_i.$$

In this context x_i represents observed net wealth of a given household, m_i is the true but unknown scale parameter above which the sample data can be described using a Pareto distribution, and α_i is an as well unknown shape parameter describing the specific form of the underlying power law (*Pareto alpha*). Our approach closely follows Clauset *et al*. (2009) and can be summarized in the following way: the estimators $\hat{m}_i$ and $\hat{\alpha}_i$ are determined by estimating Pareto distributions systematically for increasing subsets of the data and choosing that subset and its corresponding parameters which exhibits the best fit to the data. We employ Cramer–Von-Mises (CvM) test statistics to compare the relative fit of the estimated distributions.[9] The exact procedure is documented in the online Appendix (part II), which contains the Mathematica code we used to carry out the steps described below.

First, we fit Pareto distributions by maximum likelihood to increasingly large subsamples starting from the 100th percentile. Through expanding the subsample by one additional percentile until we reach the 71st percentile, we get 30 different estimates for each α_i and m_i, where m_i is equal to the smallest observation within each subsample. The smallest subset includes only the data points within the 100th percentile and the largest contains all observations between the 100th and 71st percentile. Maximum likelihood (ML) is our preferred estimation method since it is a well-established result that ML estimators are superior to other approaches if the data of interest follow a power law (Clauset *et al*., 2009; Greene, 2012). The ML estimator in our case is equivalent to the so-called Hill estimator (Hill 1975; see also footnote 5 in the previous section).

Second, we perform a goodness of fit test for each of the 30 subsamples per implicate by computing the CvM statistic (ts_i), which increases with the difference between the observed sample and the estimated Pareto distribution. Thus, low test statistics point to relatively good fits. Figure 4 plots these statistics for the upper 30 percentile subsamples[10] and the corresponding $\hat{\alpha}$ across all five implicates. As is evident from Figure 4, the test statistics vary considerably between implicates and even more do the $\hat{\alpha}$'s, especially for the first few subsamples. The huge variation in the top of the distribution is very likely the effect of small sample sizes, since each percentile only contains roughly 24 observations.

Both the estimations of $\hat{\alpha}_i$ as well as the computations of ts_i, were performed using the HFCS data without the corresponding weights. This operation is the only one where we ignore weights and we believe there are good reasons for doing so. First, estimation results hardly change when weights are taken into account. The difference in the average $\hat{\alpha}$ across implicates is equal to –0.00317, not only demonstrating a minor relative effect, but also leading to slightly more conserva-

[9]A rationale for using the test statistics instead of standard p-values available in software packages is provided in Section 3.3.

[10]The results do not change if one considers the upper 50 percentiles instead of the upper 30. Fitting Pareto distributions to subsamples including data below the 50th percentile is not supported theoretically nor does it yield satisfactory fits as the CvM statistics rapidly increase.

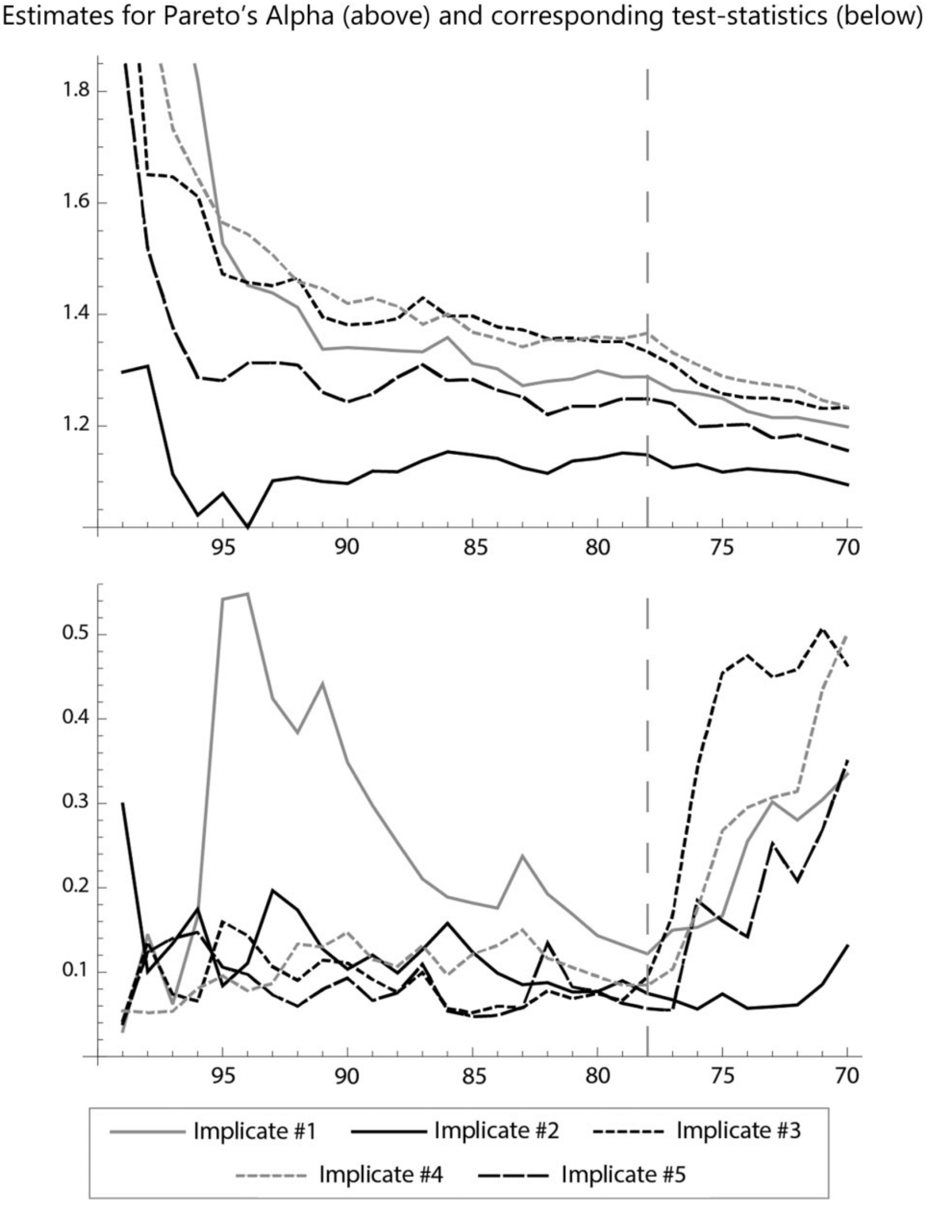

Figure 4. Estimates for Pareto's α and Corresponding Cramer von Mises Test Statistics Across All Implicates

tive estimates (i.e., wealth is distributed more evenly). Second, the construction of sampling weights by the Austrian National Bank involves a battery of unknown regression models and assumptions about the determinants of response probabilities. By only using weights for linking the sample to the underlying population, which is, for instance, required for the definition of wealth percentiles, we strongly limit the influence of those unknown implicit assumptions. Third, if using weighted data to carry out the CvM tests, one needs to handle the variation involved in the construction of weights by using resampling weights. While the question how to combine the CvM test with those resampling weights is far from trivial, such a procedure would also greatly increase complexity and computation time involved, since total operations (estimation of α_i and computation of ts_i for each subsample) would increase from 300 to more than 150,000 if the full set of resampling weights provided by the Austrian National Bank is used.

Our focus was on determining $\hat{m}_i$ such that on the one hand it is not sensitive to sample size problems (as seems to be the case in the highest percentiles), and on the other hand it does not rely on ad-hoc assumptions alone (e.g., Bach and Beznoska, 2011; Bach *et al.*, 2014) or a merely visual inspection of well-known log–log graphs (e.g., Cowell, 2011). In this context Clauset *et al.* (2009) illustrate the effect of using unreliable scale parameters on the estimates of the shape parameter. Choosing a scale parameter below the true value m_i leads to the inclusion of non-Pareto distributed data and thus to downward biased estimates of α. Conversely, choosing a scale parameter above m_i ignores potentially useful information, and thus lowers the statistical precision of the estimates and also biases the results upwards. Although the method provided by Clauset *et al.* (2009) is in principal a suitable guide to the estimation of the scale and shape parameter, in our case the presence of five different and autonomous implicates leads to an additional complication, namely how to synchronously identify a good fit across all five implicates. In the application of Wald's well known maxi–min model (Wald, 1945) we found a satisfying answer to the latter concern: the maxi–min model posits that in the face of different alternatives with uncertain consequences, one should rank these alternatives on the basis of their worst-case consequences, which in our case corresponds to the worst fit across all implicates, and choose that option where the worst case is at least as good as all other alternatives. The maxi–min principle introduces a certain degree of conservatism to the chosen estimation results by focusing on the relatively worst fits and nullifying the impact of single exceptionally well-fitting subsamples across all implicates. In detail, we first choose the maximal test statistics (i.e., the worst fit) across implicates for each sample size and then identify the minimum of these test statistics (i.e., the best fit) across all sample sizes. By applying this procedure we find that the threshold value of the 78th percentile proves to be the most suitable candidate for providing a statistically reliable estimate for the true value of m_i.[11] Notwithstanding the potential problems arising from this complication to match the approach by Clauset *et al.* (2009) to the specific structure of our dataset, the interval around the result is characterized by small test statistics as well as stable alpha parameters across implicates, which increases our confidence in the reliability of our proposed approach. The final estimation results across all implicates are collected in Table 1.

[11]While this result at first sight seems to be determined solely by implicate 1, which shows the worst fit throughout most of the estimation range, removing this implicate would leave the 79th percentile as the most suitable candidate for providing a statistically reliable estimate for the true value of m_i.

TABLE 1
Estimating the Pareto Distribution for the Austrian Wealth Distribution: Results

Implicate	$\hat{\alpha}_i$	$\hat{m}_i$
#1	1.28808	€281,242
#2	1.14815	€287,809
#3	1.3332	€289,811
#4	1.24881	€293,161
#5	1.36649	€288,422
Average	1.276946	€288,089

So far we have established what we deem to be a non-arbitrary technique of fitting a Pareto distribution to the upper tail of the Austrian HFCS sample. However, even if this upper tail does not follow a Pareto distribution, the according parameters could still be estimated without noticing the mistake. Therefore, we rigorously test the hypothesis that our data are actually drawn from a Pareto distribution prior to using these estimates for correcting the HFCS sample.

3.3. *Testing the Pareto Hypothesis*

In the previous subsection we elaborated on how to find reliable distribution parameters. However, it remains to be shown that the estimated distributions truly represent the data. On first sight this might seem superfluous, since the p-values based on the Cramer von Mises tests would provide an immediate answer to the question of whether the data within a given subsample are statistically different from the estimated distribution or not. However, those standard p-values are derived under the assumption that the distribution against which the data are tested is perfectly known, whereas in our case the distribution to test against is just an estimation. As a result, the standard p-values are not suitable for clarifying this issue.

In this context we again follow Clauset *et al.* (2009), who suggest comparing the goodness of fit of the original data and its estimated distribution with the goodness of fit of newly created data vectors based on the original data as well as the estimated distribution. While these new data vectors are created by means of a bootstrap—that is, repeated random drawing from the estimated distribution (above $\hat{m}_i$) and the original data (below $\hat{m}_i$)—the general idea is to test the goodness of fit of the original estimation against the goodness of fit of a series of estimations based on these newly generated data vectors, where the data for top-wealth households (i.e., all households above $\hat{m}_i$) is already known to truly follow a Pareto distribution. If the goodness of fit of the original estimation is not significantly worse than the goodness of fit of the estimations based on the newly generated data vectors, there is good reason to believe that the estimated distribution adequately represents the underlying data.

Following this strategy, we create $B = 10{,}000$ synthetic datasets (X_{ib} where $b = 1 \ldots B$) for each implicate by drawing a number x_{ij} with probability t_i/n from the previously estimated distribution with the parameters $\hat{\alpha}_i$ and $\hat{m}_i$, where $n = 2380$ is the number of total observations and t_i is the number of observations above $\hat{m}_i$. With probability $1 - t_i/n$ we pick a random element x_{ij} from the original dataset below $\hat{m}_i$. Repeating this process from $j = 1 \ldots n$ yields a synthetic dataset with 2380 observations where all elements above $\hat{m}_i$ are drawn from the originally estimated distribution. For each implicate we use these 10,000 data sets to compute an artificial p-value (p_i) for the hypothesis test that the original data follow a Pareto distribution with $\hat{\alpha}_i$ and $\hat{m}_i$ more closely than the synthetic datasets follow their estimated distributions. Thus, we want to define a p-value for testing the null hypothesis, that the HFCS data truly follow a Pareto distribution above our estimated scale parameter, against the alternative hypothesis that it does not follow a Pareto distribution. In order to obtain this p-value, we repeat all the steps from Section 3.2 for each of these synthetic datasets: scale parameters $\hat{m}_{ib}$ and shape parameters $\hat{\alpha}_{ib}$ are estimated as described above and the corresponding CvM test statistics ts_{ib} are computed. Since the synthetic datasets truly follow a Pareto distribution above $\hat{m}_{ib}$, these test statistics are capturing only random variations but no systematical differences between the synthetic data and the estimated distribution. Thus whenever the condition

$$ts_i \leq ts_{ib} \tag{2}$$

holds, the difference between the original data and the original estimation is actually smaller than or equal to the difference between the synthetic data vectors and their respective estimated distributions which is purely due to random variation. By counting the instances where (2) holds (denoted by c_i) we obtain the p-value (p) for our hypothesis after averaging over all implicates, i.e.

$$p = \frac{1}{5}\sum_{i=1}^{5} c_i / B. \tag{3}$$

The interpretation of this artificial p-value is pretty standard, namely that below the 10 percent level the null hypothesis is rejected since the difference between the HFCS data and the estimated Pareto distribution is significantly greater than the differences due to random variation in our synthetic datasets. However, if enough synthetic test statistics are larger than ts_i, the difference between the actual data and the estimated Pareto distribution is mostly smaller than or equal to pure random variation and thus the null hypothesis cannot be rejected. Unfortunately the results for each single implicate are partially idiosyncratic and far from consistent across all implicates.

Table 2 indicates that the Pareto distribution is a plausible model for implicates 2, 4, and 5, and that it is strongly rejected for implicate 1 and weakly for implicate 3. On average, however, the hypothesis still holds. We focus on this latter result since the variability expressed by the single implicates is due to imputing

TABLE 2
p-Values Across All Implicates

Implicate	#1	#2	#3	#4	#5	Mean
p-value	0.019	0.2776	0.0983	0.5421	0.1781	0.223

missing data based on a series of different statistical models. Thus, only an average across those implicates seems to be a justifiable criterion, since the different implicates have to be interpreted conjointly to appropriately consider the variability between implicates.

3.4. *Correcting the Data*

After identifying $\hat{\alpha}_i$ and $\hat{m}_i$ we use this information to correct for the "missing rich" by removing all observations exceeding €4 million in net wealth from the original dataset. We choose this €4 million cut-off point because the frequency of observations starts to markedly decline beyond this level of net wealth. Since patterns slightly differ across implicates, and to verify the robustness of our results, we also included two scenarios with €3 million and €5 million cut-off points respectively (see Figure 5). These variations had only a minor impact on the final results (see footnote 12).

In the standard case all eliminated observations are part of the 100th percentile (except for the second implicate, where the 100th percentile starts at €4.6 million) and represent between 11,374 and 44,081 households depicted by 8 to 30 observations depending on the specific implicate under consideration. This treatment implies that we assume that the alleged non-observation bias affects this group of households, and instead suggest relying on the estimated Pareto distribution for observations above €4 million in net wealth.

To determine how many households should be added to the sample based on the estimated Pareto distribution, we look at the number of households (HH_i) with net wealth holdings above $\hat{m}_i$ and below $\mu \equiv$ €4 million according to the HFCS dataset. HH_i varies between 785,924 for $i = 2$ and 817,418 for $i = 1$. By drawing on the properties of the underlying probability distribution function we compute the number of households above €4 million (H_i) by:

$$H_i = HH_i \frac{1 - P_i(\mu)}{P_i(\mu)}. \tag{4}$$

H_i varies between 22,982 for $i = 5$ and 40,251 for $i = 2$. This approach ensures that the correction for rich households only depends on high quality observations from the HFCS data. Given H_i, one can derive the wealth x_i for each household above μ by exploiting the fact that

$$1 - P_i(x_i) = \Pr(X_i > x_i) = \left(\frac{m_i}{x_i}\right)^{\alpha_i} \equiv \frac{H_{x_i}}{HH_i + H_i}.$$

Rearranging terms gives

$$x_i = m_i \left(\frac{HH_i + H_i}{H_{x_i}}\right)^{1/\alpha_i}, \tag{5}$$

where H_{x_i} is the number of households reporting a net wealth of at least x_i or, put differently, the rank of a given household. By applying (5) consecutively we

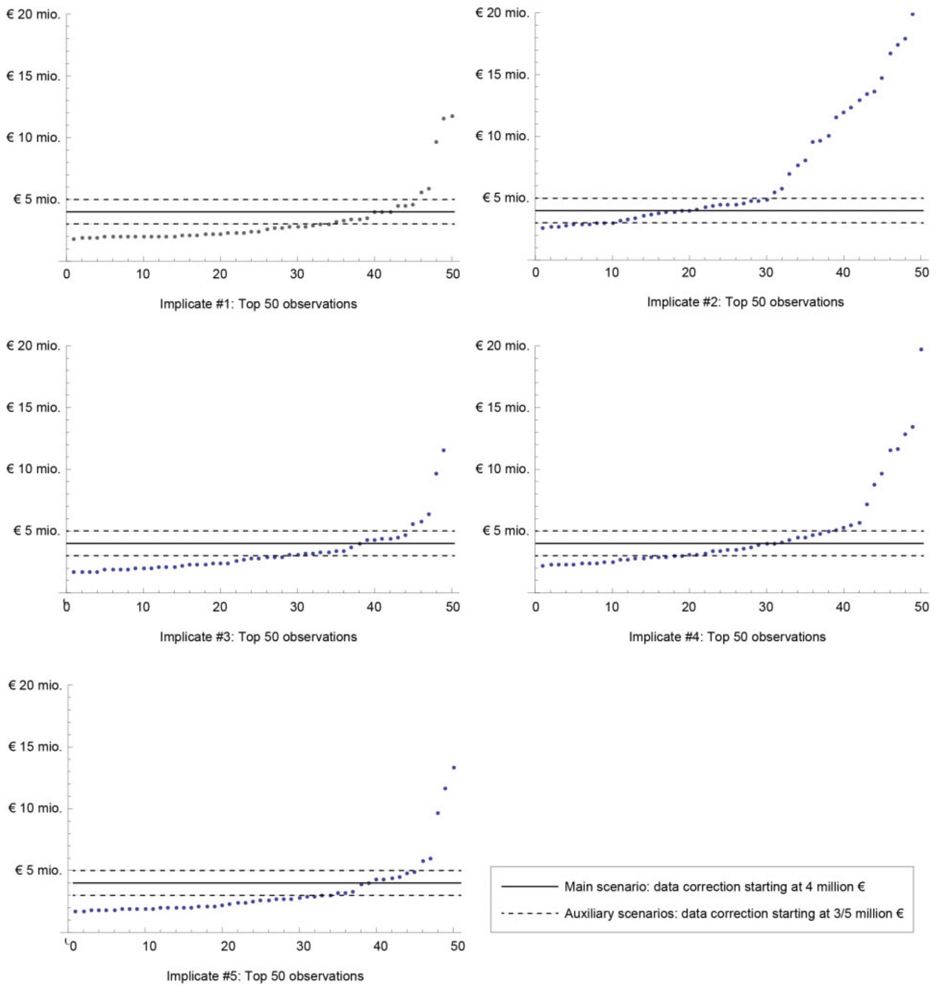

Figure 5. Plots of the Richest 50 Households in Each Implicate and the Cut-Off Points

generate new observations for net wealth above μ. It is important to note, however, that we limit net wealth by €1 billion in our application (specifically: any observation above that value was set equal to €1 billion). This truncation of the newly generated sample is motivated by our preference for conservative estimates as well as a certain modesty regarding the possibility to correctly estimate net wealth for a tiny group of people at the far end of the distribution. However, we will return to this issue when comparing our final results with a journalist list on Austrian millionaires; see Section 4).

Finally, we have to slightly adapt the sample weights, since the number of households added and the number of households removed from the sample differ. The net change, which is the difference between the number of households above €4 million according to the original HFCS sample and HH_i varies between +16,280

TABLE 3

Austrian's Richest 5% According to the Corrected Data

Percentile	Total Net Wealth (in billion €)	Average Net Wealth per Household (in million €)	Share of Total Net Wealth	Culminated Share of Total Net Wealth
96	40.8	1.1	3.2%	58.53%
	(38.8)	(1.0)	(3.9%)	(47.6%)
97	50.3	1.3	4.0%	
	(48.7)	(1.3)	(4.9%)	
98	66.7	1.8	5.2%	
	(65.5)	(1.7)	(6.6%)	
99	101.2	2.7	7.9%	
	(94.1)	(2.5)	(9.3%)	
100	497.3	13.4	38.2%	
	(237)	(6.4)	(22.9%)	
Total Sample	1278	0.339	100%	100%
	(1000)	(0.265)		

Note: Values using original HFCS data in parentheses.

($i = 1$) and −3828 ($i = 2$). In the latter case, the original HFCS dataset reported a higher number of households above μ compared to the estimated Pareto distribution. Thus in this case the weights of the remaining observations below the estimated scale parameter need not be reduced, but increased. In either situation the alteration of sampling weights is done proportionally to the total number of households less the observations above $\hat{m}_i$. For example, in relation to the number of households below $\hat{m}_i$ the net change in the first implicate equals 0.55 percent of total weights. As a result the weights for observations below the scale parameter are reduced by 0.55 percent. On average (across the implicates) the weights are reduced by 0.21 percent. After having corrected our survey data for the "missing rich" we may now contrast the estimates derived from the corrected data with those obtained from the original HFCS data.

4. The Impact of the Missing Rich

We summarize the findings obtained in this way in Table 3, which also provides estimates derived from the original HFCS in parentheses. Table 3 indicates that the total wealth of the richest percentile grows by more than 100 percent, namely from €237 billion to €497 billion. Hence, total wealth also increases significantly from roughly €1000 billion to about €1278 billion, representing an increase of 28 percent in the estimate for total wealth due to the proposed data correction.[12] In this context the decrease in the share of percentiles 99–96 is due to the correction of the number of households richer than €4 million.

[12]These results are insensitive with respect to the cut-off point, above which HFCS observations are replaced by the estimated Pareto distribution. If we change this value from €4 million to either €3 million or €5 million, the resulting total wealth estimates are €1266 and €1284 billion, respectively. The share of the 100th percentile changes from 38.2 to 38.6 percent in case of €3 million cut-off and remains unchanged when a €5 million cut-off is used.

Correspondingly, the share of wealth held by selected population groups changes significantly, with the most remarkable change in the share of the richest percentile, which increases from 22.9 to 38.2 percent. The share of the poorest 50 percent of the wealth distribution, on the other hand, decreases from 2.8 to 2.2 percent.

In sum our estimations suggest that the size of wealth omitted when obtaining estimates on top or total wealth from survey data is indeed significant. This bias has two potential sources: a non-observation bias inherent in survey designs (see Section 2), and an additional and mostly unobservable bias stemming from differential non-response. While our procedure is not able to separate these effects, it allows for addressing the total bias arising from this constellation. In our specific application the estimate for total wealth changed by roughly a quarter, although the correction of the data affected less than 1 percent of the underlying population. Consequently, one has to conclude that the implications of these biases are far from trivial and can hardly be ignored when dealing with top wealth data. However, since the problem of differential non-response is hard to verify in practice, we focus our discussion on the problem of non-observation.

The most interesting questions about the results presented above is how closely the corrected data fits the actual wealth distribution. Due to the lack of detailed tax data on wealth in Austria, which would be a first choice when looking for a benchmark to compare our results against, we chose a different approach. We use Journalists' rich lists as an exogenous source of information to assess the validity of our results. Up until now we assumed maximal wealth to be equal to €1 billion when correcting the HFCS data. Practically, this implied that net wealth of all households exceeding €1 billion according to the estimated Pareto distributions were set to €1 billion. Although this restriction might lead to an underestimation of total wealth, we nonetheless imposed it for reasons of statistical conservatism. However, even though we do not want our estimates to rely on the upper end of the distribution's tail, we will now relax this restriction in order to validate our general strategy.

In doing so we first compared the estimated number of households with net wealth greater than €1 billion as implied by our estimated distributions with the available media information. The latter varies considerably between years (and it is unclear whether these variations are due to actual changes in wealth or just to changes in journalists' informational status), from 19 for 2010, to 24 for 2011 (the year of the HFCS survey, see Table A1 in part III of the online Appendix), to 30 as reported in 2013. Moreover, it does not distinguish between households and family clans: some of the entries can be decomposed into several households. Accounting for this can increase or decrease the number of billionaires, since some clan fortunes are large enough to make its individual members billionaires and others are not. In order to get a rough understanding of this, we tried to decompose this list into individual households (see Table A2 in part III of the online Appendix). In doing so we used information available in the media to assess the number of members (households) of these family clans and divided the fortune equally among them. In our case this adjustment does not lead to a change in the number of billionaires, which stays constant at 24 in 2011, since some families drop out while others split into several billionaires. Our own calculations from the

TABLE 4
Total Wealth of the Richest Households (in million €) and Deviations (%)

	Top 30	Top 60
Original data	430	860
Pareto estimate	87,536	110,077
Trend list	92,160	109,350
Trend list, adjusted	78,295	97,201
Deviation original data from Trend list	−99.5%	−99.2%
Deviation original data from adjusted Trend list	−99.5%	−99.1%
Deviation Pareto from Trend list	−5.0%	0.7%
Deviation Pareto from adjusted Trend list	11.8%	13.2%

Pareto function point to 30 billionaires and thus are well in line with the figures reported by journalists.

Probably more interesting than the number of billionaires is the total volume of wealth at the top as reported by journalists and as predicted by the Pareto distribution. The results of such a comparison are shown in Table 4. Specifically, the total net wealth of the richest 30 households according to the (unrestricted) Pareto distribution is 5 percent lower than net wealth reported by the top 30 entries of the journalist list. If we look at the upper 60 entries, we see that the Pareto distribution comes even closer to just 0.7 percent deviation. For comparison, total wealth from the original sample data deviates from total wealth reported by journalists by −99.5 percent and −99.2 percent, respectively.[13] However, as we have already argued, the journalist list has the disadvantage that it does not distinguish between clans and single households. If we compare the Pareto estimates to the list where we tried to take this fact into account, using the Pareto distribution to estimate the richest 30 households leads to an overestimation of 11.8 percent (Top 30) and 13.2 percent (Top 60), respectively. Overall, this comparison shows that our estimates are closely in line with non-sample information about wealth holdings in Austria and thus suggests a good performance of our estimation strategy.

In what follows we compare our results with the existing literature and compute upper and lower bounds for our estimates, thereby scrutinizing the adequacy and plausibility of our methodological setup as well as our results.

5. Robustness Checks

Given the sharp increase in net wealth holdings in the 100th percentile due to the data correction we suggest in Section 3, the reader may be interested in the uncertainty of our estimates or be skeptical about our results anyway since they depend entirely on the statistical Pareto model. One way to assess the validity of our estimates is to compare them to the results of similar studies based on the same dataset. Two such studies became available very recently (Eckerstorfer *et al.* 2014; Vermeulen 2014) and their results are broadly comparable to ours: Eckerstorfer *et al.* (2014), for instance, estimate the share of the top 1 percent of wealth-holders

[13]Since the sample does not contain weights as small as 60, the total wealth of the top 60 (top 30) households is equal to the average wealth of the wealthiest household across all implicates (€14.3 million) multiplied by 60 (30).

to be 39.7 percent, whereas the estimates provided by Vermeulen (2014) range from 30 percent to 41 percent. Our estimate of 38.2 percent, hence seems to be well in line with the results obtained from broadly comparable studies.

In addition to this comparison, we can validate the robustness of our results by providing upper and lower bounds of estimated net wealth with respect to data variability across subsamples by means of a bootstrap. Due to the complex survey design of the HFCS, which involves stratified sampling as well as multiple imputation to correct for item-non-response, one is confronted with serious complications when trying to compute confidence intervals reflecting the uncertainty of the estimation process. While current literature offers procedures to compute confidence intervals with either multiply imputed data (Rubin, 1987) or data from complex surveys (Rao and Wu, 1988; Rao *et al.*, 1992; Kolenikov, 2010), there is, according to our knowledge, no contribution which shows how to construct appropriate confidence intervals when multiple imputation as well as a complex survey design are used in the data collection process. Therefore, we implemented an approach to validate the robustness of our estimates with regard to sampling variation and suggest focusing only on the uncertainty arising from the variability of the original data. In doing so, we apply a bootstrap in order to test the robustness of our results with respect to random resampling. Even though we cannot express the uncertainty of the estimation process itself this way, we are still able to demonstrate the robustness of our results due to potential outliers and irregularities within certain subsets of the original HFCS sample. However, it is important to bear in mind that the bounds reported below do not serve as direct substitutes for traditional confidence intervals.

The bootstrap procedure for computing an upper and lower bound of $\hat{\alpha}_i$ involves the construction of $U = 1000$ random samples consisting of $n_U = 2/3 \cdot n \approx 1587$ observations randomly picked from the original HFCS data set for each implicate. Then we re-estimate $\hat{\alpha}_i$ for each random sample. After ordering them in ascending order, the 26th estimate of α_i is identified as the lower and the 975th as the upper bound of $\hat{\alpha}_i$. By repeating the data correction procedure described in Section 3.3 for the upper and lower bound $\hat{\alpha}_i$'s, we obtain new estimates for the distribution of net wealth. The results are reported in Table 5 (already averaged across implicates). As one can see, the upper bound of net wealth within the 100th percentile deviates approximately by €135 billion from the point estimate while the lower bound deviates by €110 billion, indicating that the €237 billion reported in the original HFCS data are very likely to be downward biased.

TABLE 5

Upper and Lower Estimation Bounds

Lower Bound	Point Estimate	Upper Bound
Pareto's alpha		
$\alpha = 1.35$	$\alpha = 1.28$	$\alpha = 1.21$
Wealth attributed to the richest percentile		
€387 billion	€497 billion	€633 billion

6. Conclusion

In this paper we tried to correct for the underrepresentation of the wealthiest households by means of a novel approach based on Clauset *et al.* (2009). There are several conceptual advantages of this approach in comparison to former contributions. First, it allows for correcting for the downward bias inherent in survey data without resorting to alternative data sources on top wealth households. Second, the method can be justified simply by referring to a non-observation bias, which naturally arises if small samples are drawn from a skewed distribution. Finally, and most importantly, we employed a Cramer–von-Mises test instead of graphical evidence or ad hoc assumptions to determine a suitable scale parameter for the Pareto distribution.

In order to illustrate the capability as well as the specific application of our approach, we focus on the Austrian case. Applying our procedure significantly influences final estimation results: estimated aggregate wealth increases from about €1000 billion to €1278 billion, where the increase is mainly due to the increase of wealth within the highest percentile (wealth within this percentile increases by 110 percent). Amongst other things it follows that the richest 10 percent of Austrian households possess 69.3 percent of total net wealth instead of the 61 percent that follow from the original HFCS data. The change in the share of the richest percentile is even more remarkable: it increases from 22.9 percent (HFCS) to 38.2 percent. When we compare our results to a detailed list of Austrian billionaires published by Austrian media, we find that our non-arbitrary approach of fitting a Pareto distribution is very well in line with non-sample evidence and also closely fits the data.

Finally, we address the validity of our results by computing upper and lower bounds of wealth estimates based on a bootstrap procedure. Especially the fact that the lower bound of our estimates for top wealth is still higher than implied by the original data indicates that the underrepresentation of wealthy individuals cannot be explained by potential irregularities and outliers in the sample and, hence, is a robust finding. A natural limitation of our approach is the unknown underlying distribution of wealth due to a lack of available tax data. Even in cases where data on wealth taxes were available it is still questionable to what extent it represents the actual distribution, since the observed tax base might differ from actual household wealth, for example due to tax exemptions as well as tax evasion and avoidance behavior (as discussed in Section 1). Nevertheless it could be an interesting task to compile a high quality tax dataset and test our method against it. Future research might also refine the estimation and testing procedures presented above with respect to the use of survey weights.

References

Albacete, N., P. Lindner, K. Wagner, and S. Zottel, "Household Finance and Consumption Survey des Eurosystems 2010: Methodische Grundlagen für Österreich," *Geldpolitik & Wirtschaft*, 3/2012 Addendum, 2012.

Alvaredo, F. and E. Saez, "Income and Wealth Concentration in Spain in a Historical and Fiscal Perspective," *Journal of the European Economic Association*, 7, 1140–67, 2009.

Avery, R. B., G. E. Elliehausen, and A. B. Kennickell, "Measuring Wealth with Survey Data: An Evaluation of the 1983 Survey of Consumer Finances," *Review of Income and Wealth*, 34, 339–69, 1986.

Bach, S. and M. Beznoska, "Aufkommens- und Verteilungswirkung einer Wiederbelebung der Vermögenssteuer," DIW: Politikberatung Kompakt 68, Berlin, 2011.

Bach, S., M. Beznoska, and V. Steiner, "Aufkommens- und Verteilungswirkungen einer Grünen Vermögensabgabe," DIW: Politikberatung Kompakt 59, Berlin, 2010.

———, "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, 2014.

Clauset, C., C. R. Shalizi, and M. E. J. Newman, "Power Law Distributions in Empirical Data," *SIAM Review*, 51, 661–703, 2009.

Cowell, F. A., "Inequality Among the Wealthy," Centre for Analysis of Social Exclusion at the London School of Economics, 2011.

Dagum, C., "Wealth Distribution Models: Analysis and Applications," *Statistica*, 66, 235–68, 2006.

Davies, J. B. and A. F. Shorrocks, "The Distribution of Wealth," in A. B. Atkinson and F. Bourguignon (eds), *Handbook of Income Distribution*, Vol. 1, Elsevier, Oxford, 605–75, 2000.

Durán-Cabré, J. M. and A. Esteller-Moré, "Tax Data for Wealth Concentration Analysis: An Application to Spanish Wealth Tax," *Review of Income and Wealth*, 56, 620–31, 2010.

ECB, "The Eurosystem Household Finance and Consumption Survey—Methodological Report for the First Wave," Statistics Paper Series, No. 1/April, 2013.

Eckerstorfer, P., V. Steiner, and F. Wakolbinger, "Steuerreformvorschläge in der Diskussion—Eine Mikrosimulationsanalyse der Aufkommens- und Verteilungswirkungen für Österreich," *Wirtschaft und Gesellschaft*, 40, 203–33, 2014.

Feenberg, D. R. and J. M. Poterba, "Income Inequality and the Incomes of Very High-Income Taxpayers: Evidence from Tax Returns," *Tax Policy and the Economy*, 7, 145–77, 1993.

Greene, W., *Econometric Analysis*, 7th edn, Pearson, Essex, 2012.

Hill, B., "A Simple General Approach to Inference About the Tail of a Distribution," *Annals of Statistics*, 3, 1163–74, 1975.

Hoeller, P., I. Joumard, D. Bloch, and M. Pisu, "Less Income Inequality and More Growth—Are They Compatible? Part 1: Mapping Income Inequality Across the OECD," OECD Working Paper No. 924, 2012.

Kennickell, A. and D. McManus, "Sampling for Household Financial Characteristics Using Frame Information on Past Income," Proceedings of the Survey Research Methods Section, American Statistical Association, Vol. 1, 88–97, 1993.

Kolenikov, S., "Resampling Variance Estimation for Complex Survey Data," *Stata Journal*, 10, 165–99, 2010.

Kopczuk, W. and E. Saez, "Top Wealth Shares in the United States, 1916–2000: Evidence from Estate Tax Returns," NBER Working Paper 10399, 2004.

Piketty, T., "Income Inequality in France, 1901–1998," *Journal of Political Economy*, 111, 1004–42, 2003.

———, *Capital in the 21st Century*, Harvard University Press, Cambridge, MA, 2014.

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

Piketty, T. and G. Zucman, "Capital is Back: Wealth-Income Ratios in Rich Countries 1700–2010," *Quarterly Journal of Economics*, 129, 1155–210, 2014.

Rao, J. N. K. and C. F. J. Wu, "Resampling Inference with Complex Survey Data," *Journal of the American Statistical Association*, 83, 231–41, 1988.

Rao, J. N. K., C. F. J. Wu, and K. Yue, "Some Recent Work on Resampling Methods for Complex Surveys," *Survey Methodology*, 18, 209–17, 1992.

Rubin, D. B., *Multiple Imputation for Nonresponse in Surveys*, J. Wiley & Sons, New York, 1987.

Singer, E., "Introduction: Nonresponse Bias in Household Surveys," *Public Opinion Quarterly*, 70, 637–45, 2006.

Vermeulen, P., "How Fat is the Tail of the Wealth Distribution?" ECB Working Paper No. 1692, July 2014.

Wald, A., "Statistical Decision Functions Which Minimize the Maximum Risk," *Annals of Mathematics*, 46, 265–80, 1945.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Part I: Non-observation bias Monte Carlo
Part II: Correcting for the missing rich
Part III: List of the richest Austrians