the review of income and wealth

Review of Income and Wealth
Series 0, Number 0, 2023
DOI: 10.1111/roiw.12657

ESTIMATING THE DISTRIBUTION OF HOUSEHOLD WEALTH: METHODS FOR ADJUSTING SURVEY DATA ESTIMATES USING NATIONAL ACCOUNTS AND RICH LIST DATA

BY MICHELE CANTARELLA*

Laboratory for the Analysis of compleX Economic Systems, Scuola IMT Alti Studi Lucca
Centre for Consumer Society Research, University of Helsinki
"Marco Biagi" Department of Economics, Università degli Studi di Modena e Reggio Emilia

ANDREA NERI

Economic and Financial Statistics Department, Banca d'Italia

AND

MARIA GIOVANNA RANALLI

Department of Political Science, Università degli Studi di Perugia

Income and wealth surveys are usually affected by unit non-response and reporting errors, which contribute to a mismatch with macroeconomic figures from national accounts. In this paper, we develop a novel allocation method to address these two issues simultaneously, when only limited external information is available. The proposed approach combines information from a power law (Pareto) model with imputation procedures based on calibration. We apply the proposed simultaneous approach to produce distributional indicators for four Euro-Area countries that are consistent with their national accounts. In particular, we use data from the 2014 Finnish, French, German, and Italian Household Finance and Consumption Survey and rich list data from Forbes or national press sources, along with household sector aggregates from national accounts, as auxiliary sources of information. A bootstrap procedure is also applied to evaluate the precision of the final estimates.

JEL Codes: D31, E01, E21

Keywords: household finance and consumption survey, non-response, pareto distribution, reporting error, survey calibration

1. INTRODUCTION

In the past few decades, several economic downturns have hit the global economy starting with the subprime mortgage crisis that originated in the USA in 2007 up to the most recent crisis caused by the COVID-19 outbreak. These events have increased the demand for timely, coherent, and consistent distributional

Note: The authors wish to thank all members of the HFCN network and of the EG-LMM and EG-DFA working groups for their useful comments and suggestions.

*Correspondence to: Michele Cantarella, Laboratory for the Analysis of compleX Economic Systems, Scuola IMT Alti Studi Lucca, Lucca, Italy (michele.cantarella@imtlucca.it).

information relating to household wealth that has become an important instrument of protection against unexpected events. Such information is crucial for policy analysis and microsimulation. It has been used, for example, to assess how much debt is concentrated in the hands of financially vulnerable households (see, for instance, Ampudia et al., 2016; Michelangeli and Rampazzi 2016), or to estimate the aggregate consumption response to wealth shocks (Christelis *et al.*, 2021; Arrondel *et al.*, 2019; Paiella 2007; Guiso *et al.*, 2005; Mumtaz and Theophilopoulou 2020; Colciago *et al.*, 2019; Casiraghi *et al.*, 2018; Coibion *et al.*, 2017).

This paper proposes a method to produce distributional indicators of household wealth that combine the distributional information coming from surveys with the aggregate statistics from national accounts. Sample surveys are generally one important tool to collect distributional information on household wealth. Most national central banks (NCB) and national statistical institutes (NSI) in the Euro-Area (and guest countries soon to join the Euro) conduct the Eurosystem Household Finance and Consumption Survey (HFCS) that collects harmonized household-level data on households' finances and consumption (Eurosystem Household Finance and Consumption Network 2009). This is the most important and comparable source of distributional information on such topics. The second source of information on household wealth comes from national accounts, which record a country's stock of assets (both financial and non-financial) and liabilities over time. National accounts are not error free, yet they are the most comparable and widely used aggregate information relating to household wealth. This is the reason why they represent the best benchmark for our exercise.

In theory, because the HFCS was designed to be representative of all households, aggregating these micro-data should correspond to the macro aggregates. In practice, however, there are significant differences: aggregate totals based on surveys are often substantially below those found in national accounts. Before using the distributional information from survey data, it is therefore crucial to explain and possibly eliminate the differences between the two sources of information.[1]

There are several reasons for the differences (Expert Group on Linking Macro and Micro data 2020). On the survey side, two relevant issues are unit non-response and reporting errors. There is substantial evidence that a household's decision on whether to participate in the survey is not random. In particular, the wealthier the household, the more difficult it becomes to contact and to persuade it to participate in the survey, to the point that, after an upper bound of wealth is reached, the probability of sampling approaches zero (Chakraborty *et al.*, 2019; Kennickell 2008; Kennickell 2019; Vermeulen 2018). The severity of this problem strongly depends on the availability of auxiliary information that can be used to select rich households (design stage) and to adjust ex-post to compensate for their lower probability to be interviewed (estimation stage, Ranalli *et al.*, 2023). Only a few countries are in a position to apply an effective adjustment for such an issue. On the contrary, item non-response is limited in the HFCS, and it is dealt with using multiple imputations (Eurosystem Household Finance and Consumption Network 2020).

[1]In 2015, the European System of Central Banks (ESCB) established an expert group with the aim of comparing and harmonizing macro data (i.e. national accounts/financial accounts) and micro data (i.e. the HFCS) on wealth.

Because these households own a large share of total wealth, their underrepresentation in the final sample is likely to result in a biased picture of wealth distribution. Moreover, wealth surveys generally include both complex and sensitive items. It follows that respondents are not always able or even willing to report the correct amount of wealth they hold. Like non-response, reporting error is not random and it differs across population subgroups and portfolio items (D'Alessio and Neri 2015; Ranalli and Neri 2011). It follows that differential non-response and reporting error can affect the representativity of a survey in terms of wealth distribution even when administrative records are used for sampling and calibration.

The ideal solution for overcoming these problems would be to find an external reference point at the top to reconstruct the tail of the distribution, a strategy often followed by both top-wealth and top-income adjustments. For top-wealth adjustments, some have linked survey data with administrative records (such as tax records or credit registers, as in Blanchet *et al.*, 2022b; Garbinti *et al.*, 2018; Garbinti *et al.*, 2020). Alternative approaches to data linkage are directly based on the use of wealth (tax) records (Alvaredo and Saez 2009; Atkinson 2016) or capital income information from tax records to construct wealth estimates assuming certain rates of return on assets (Saez and Zucman 2016). Similar approaches have been followed by the literature on top-income adjustments (see Jenkins 2022, for an overview of the literature), which benefit from access to plentiful sources of administrative tax data.

Unfortunately, these data sources are sparser for wealth, and when such administrative records exist and are not limited in scope, they are not usually available for confidentiality reasons. The result is that, in many national contexts, survey data are often the only official source of information available to those seeking to reconstruct the distribution of wealth; therefore, top-wealth adjustments cannot always rely on the same adjustments. By contrast, top-income adjustments are easier to replicate in new national contexts because tax data are more readily available.

For that reason, the recent literature has developed methods to re-estimate the wealth distribution after combining survey data with the limited external information publicly available, such as aggregate figures from national accounts or lists of rich individuals' total wealth. This is usually achieved by fitting a Pareto distribution to the combined data.

Adjustments based on Pareto models have gained popularity in the literature. Within the top-incomes literature, Jenkins (2017) estimated income inequality in the UK by fitting generalized Pareto and Pareto distributions to tax data to derive top-tail inequality information and combine it with survey data. More recently, Blanchet *et al.* (2022a) make use of generalized Pareto interpolation on a combination of survey, income tax, and national accounts data to reconstruct an income distribution series in Europe and compare it with the USA. Similarly, Carranza *et al.* (2022) showed how these methodologies can be applied to other income surveys such as the European Union Statistics on Income and Living Conditions (EU-SILC), underlining the importance of using external data for calibrating surveys suffering from the missing rich problem.

These methods have proven useful for reconstructing the wealth distribution too. Vermeulen (2018); Vermeulen (2016) uses Forbes annual World's Billionaires List in combination with wealth surveys to estimate the total wealth held by rich households. He shows that the use of such lists increases the quality of the results (compared to estimating a Pareto model from survey data alone). Similar methods based on fitting a Pareto distribution on rich list enhanced data have been used since 2010 in the Credit Suisse Global Wealth Report (Shorrocks *et al.*, 2021).

Building on these approaches, Chakraborty *et al.* (2019), Waltl (2021), and Waltl and Chakraborty (2022) extended the analysis by benchmarking survey results to the national accounts. It is important to highlight that these studies do not set themselves to adjust survey data through imputation or calibration, as rich lists are only used to aid in the estimation process.

A set of common assumptions for all these studies remain that the number of households in the tail is either held fixed or can be retrieved from an external source, and that unit non-response of wealthy households is the only reason for the micro/macro gap.

Few other studies have taken the step of enhancing survey data with information from the estimated Pareto distribution, following a "replacing and re-weighting" approach. Bach *et al.* (2019) implemented these methodologies to impute rich list data to wealth surveys, and the aforementioned research from Blanchet *et al.* (2022b) combines survey and tax data to fit a Pareto distribution and re-calibrate the weights from survey data. These strategies have also been followed by some statistical offices: to our knowledge, the UK Department for Work and Pensions has used this strategy since 1992, and the UK Office for National Statistics introduced it in 2021 (backdating to years from 2001). These methods stress the importance of considering not only the amount of wealth (or income) at the top but also the number (and relative size) of households in this group (Bourguignon 2018).

This paper adds to the literature that has attempted to produce distributional indicators of wealth consistent with the national accounts, by addressing these limitations and proposing a methodology that draws on existing and well-established methods. We offer two separate methodologies for the correction of non-response and reporting error that can be used in conjunction. This work contributes to this literature in four ways.

First, whereas some of the aforementioned studies focus only on the missing part of the tail, assuming existing survey observations a representative, we claim that differential non-response also affects the representativeness of existing survey observations, which in turn affects estimates for the total number of households in the Pareto tail, and their total wealth. Every study that relies on internal information to obtain the number of households in the tail will suffer from this limitation. As mentioned earlier, few studies have considered this issue by re-estimating the size of the top-wealth group. We add to this literature by offering a simple, yet novel, approach to producing adjusted estimates for the number of households in the Pareto tail in the presence of truncation and/or differential non-response. We then propose a correction for differential non-response that accounts for the missing rich and re-weights observed survey households. This correction is not a substitute for the imputation (Bach *et al.*, 2019) or simulation (Waltl 2021) procedures developed in the literature, but rather, it complements them by allowing for the correction of non-response bias among existing survey observations.

Second, while existing papers focused only on non-response at the tail of the distribution, we present a methodology that allows us to correct for reporting error as well. Dealing with both aspects simultaneously is important, even when the research purpose is to estimate the share of total wealth held by wealthy households. Indeed, some rich households may misreport their true wealth and therefore could be misclassified in the adjustment process. One advantage of our approach is that it enables us to compute distributional indicators that refer to "non-rich" households, such as those relating to financial vulnerability.

Our third contribution is that even if we apply well-established methods (such as Pareto modeling, imputation, and calibration), we show how to combine and use them in a single framework and how to evaluate the precision of the results. As corrections for reporting error have to rely on specific assumptions about reporting behavior, our framework allows for these adjustments to be performed on top of non-response correction methods that are instead always appropriate. In particular, we propose the use of survey calibration as an imputation method for reporting error, which generalizes and encompasses proportional allocation as a particular case. In fact, in proportional allocation, adjustments are made by separately scaling each wealth component up (or sometimes down) to match the national account totals, so that different scaling factors are used for different wealth components, yet the factors are the same for all households. In the framework we propose, it is possible to use this approach or to extend it to allow for a different scaling factor for each household. The approach is general enough to make the user choose whether this (possibly different) scaling factor has to be different for each wealth component or can be common to all (or to a subset of) wealth components.

The fourth and final contribution is to produce a modified and readily usable data set in which survey values have been adjusted for the aforementioned quality issues and, by construction, the totals add up to the national accounts. While the existing papers focused mainly on methods to estimate total wealth held at the top of the distribution, our adjusted data set can be used for estimating any distributional indicator that may be of interest.

This paper is structured as follows. Section 2 describes the data sources used in our application. Section 3 presents the Pareto approach (Subsection 3.1), calibration (Subsection 3.2), and the methodology we propose to combine the two (Subsections 3.3,3.4, and 3.5). Section 4 describes the tools used to assess the properties of the proposed methods. Section 5.1 describes how the method applies to our data, while Section 5.2 discusses the results and the main findings of the application. Finally, Section 6 concludes.

2. Data

This paper uses the HFCS and two sources of auxiliary information, namely, the national accounts, which include both financial and non-financial accounts, and rich list data from several sources.

The HFCS is a joint project of NCB and statistical institutes of the Eurosystem and of some EU countries that are yet to join the Euro. The survey collects detailed household-level data on various aspects of household balance sheets and related economic and demographic variables, including income, private pensions, employment, and measures of consumption. The HFCS is conducted using a decentralized approach. A group of experts from the European Central Bank (ECB) and the NCB (the Household Finance and Consumption Network, HFCN) coordinates the project, to maximize the between-country comparability of the final data. However, differences between national surveys remain large, as we discuss below.

We used the second wave of HFCS (2014), and we restricted our analysis to four countries: Italy, France, Germany, and Finland. This choice was motivated by two considerations. First, rich lists and non-financial accounts are available for this subset of countries. Second, these surveys present methodological differences that can be used to evaluate our method. For example, some countries oversample rich households using individual tax records (as in the French and Finnish surveys) or using the information at the regional level (as in the German one), while others do not oversample (as in the Italian case). Moreover, in some instances, the survey is linked to administrative data (as in the Finnish one). In both cases of oversampling and use of administrative records, we should expect the adjustment method to have a lower effect.

Our variable of interest is household net wealth defined as the sum of deposits, bonds, shares, mutual funds, money owed to the household, the value of insurance policies and pension funds, business wealth, and housing wealth, minus debts. Because item non-response is limited in the HFCS and not all surveys make full use of multiple imputations (such as the Italian survey, for which all five replicates are identical), we have decided to use only one data set for each country (the first one).

The second source of information was provided by national accounts. The financial component (financial accounts) is produced by NCBs and reports the total financial assets and liabilities held by households, classified by financial instrument and in order of liquidity based on the original maturity and negotiability (cash, deposits, insurance, and pension instruments). Non-financial accounts are produced by NSIs and contain the total value of dwellings, other buildings and structures, and land owned by households. Even if national account figures may suffer from quality issues and be based on different concepts and definitions from those used in the survey, they are still widely considered superior to survey data.

Rich lists were our third source of information. They have already been used in the literature to adjust for missing rich households (Vermeulen 2018; Waltl and Chakraborty 2022). Their use may generate concerns because the methodology adopted to populate them is often obscure, and usually only figures for net worth are provided, with no financial instrument breakdown. Some studies have tried to resolve these issues using other types of Pareto adjustments (Blanchet *et al.*, 2021; Waltl 2021). Others (such as Schröder *et al.*, 2019) have also explored new ways of sampling high-wealth individuals with adequate precision. However, these methods can be used only in specific instances when information on these households exists and is easily accessible. When these sources are not available, rich lists remain a reliable alternative, and evidence from Waltl (2021) indicates that, after rich lists have been integrated in the estimation of the Pareto tail, there might be little difference between the wealth estimated by the various Pareto adjustments. Nothing prevents our methodology being adapted in contexts for which registry data on top fortunes are available to researchers or survey providers. We preferred to work with rich lists as these are more readily available to researchers.

TABLE 1
Rich Lists and Survey Data

	Rich List				HFCS	
	Size	Max/Min w_i		% NA	Size	Max w_i
Country	(1)	(2)	(3)	(4)	(5)	(6)
IT	39	11,049.09	958.805	1.06%	8,156	13.62
FR	116	20,636.80	183.662	2.47%	12,035	230.36
DE	478	31,746.92	256.023	5.55%	4,461	48.15
FI	96	2,350.57	9.770	1.82%	11,030	51.51

Notes: Cols. (1) and (5): sample size of rich lists and HFCS. Cols. (2) and (3): maximum and minimum gross wealth w_i in the rich list (millions of EURO). Col. (4): wealth in the rich list as a percentage of national accounts (NA). Wealth figures adjusted for liabilities and comparable portfolio compositions, in accordance with Expert Group on Linking Macro and Micro data (2020); Chakraborty *et al.* (2019). Col. (6): maximum wealth in the HFCS (millions of EURO).

In our case, when applicable, we used the same rich lists as those in Waltl (2021). These are all produced by national press sources and include the 2014 edition of Challenges' "Les 500 plus grandes fortunes de France" for France, Manager Magazin's list for Germany and Arvopaperi's list for Finland. These local lists were not available for Italy, and we extracted the richest Italian households from "Forbes 2014 World's Billionaires List" instead. Ideally, longer rich lists such as the ones produced by national magazines are to be preferred: the list of Italian billionaires, coming from the global Forbes' rich list, is shorter and presumably less precise than the other lists we have used. These factors will affect the precision of our estimates, so we expect more variation from the Italian survey estimates. Nonetheless, what matters for our estimations is that the rank of the richest individuals is preserved, so that the richest, the second richest, the third richest individuals (and so on until the list is over) are captured.

We also adjusted the rich list data by the debts and portfolio composition based on portfolio shares from top-wealth observations in the HFCS.[2] In this way, estimates for portfolio compositions among the top fortunes can be obtained, and rich list data can be fully integrated with the HFCS for estimation purposes.

Table 1 provides an overview of the coverage of fortunes in the HFCS survey and the rich lists. It shows that while rich lists can be short, the households contained in them can account alone for up to 5 percent of the total wealth of a country. The gap after the truncation point of the survey can be quite large. Surveys with the weakest oversampling methods (such as the Italian and German ones) fall short of even capturing household in the bottom part of the rich list. Even in surveys featuring oversampling (such the French and Finnish ones) the wealthiest households appear to be distant from the richest fortunes on record.

[2]This is a simplifying assumption. An improvement over this form of portfolio allocation is offered by the approach used in Waltl and Chakraborty (2022).

3. Methodology

Let w be household net wealth, w_i the net wealth for each individual household i in the population, for $i = 1, \dots, N$, where N is the population size, and $t(w) = \sum_{i=1}^{N} w_i$ is the population total to be estimated using survey data. Let $\mathcal{S}$ be the set of units selected in the sample and $\mathcal{S}_0 \subseteq \mathcal{S}$ be the final set of respondents. Let $\hat{t}(w) = \sum_{i \in \mathcal{S}_0} d_i w_i$ be the Horvitz–Thompson estimator, where d_i is a sampling weight adjusted to account for smaller sample size of respondents; that is, it is given by the inverse of the inclusion probability of household i grossed up so that $\sum_{i \in \mathcal{S}_0} d_i = \hat{N}$, where $\hat{N}$ is an estimate of the population size N.

Because of unit non-response and reporting error, the expected value of the Horvitz–Thompson estimator $\hat{t}(w)$ is generally lower than $t(w)$. Unit non-response occurs when some households refuse to participate in the survey. If this decision is related to household wealth (i.e. wealthier households are more difficult to enlist in the survey than others), then the sample of respondents $\mathcal{S}_0$ may not adequately represent the upper tail of the distribution. Reporting error occurs when the information collected in the survey w_i is different from the true unknown value w_i^*, for $i \in \mathcal{S}_0$. The error term $(w_i^* - w_i)$ may depend on many factors such as the difficulty of respondents to recall the required information or their unwillingness to report their true wealth.

Our methodology to address these issues is based on techniques that are well established in the literature. We used the Pareto distribution to reassess the wealth held by the richest households and the size of their population (Subsection 3.1). Once these population parameters are estimated, the calibration approach (Subsection 3.2), which is commonly used in survey sampling to deal with design weight adjustments, is used to compensate for unit non-response of wealthy households (Subsection 3.3). We define the combination of these two approaches as the *non-response* adjustments.

Reporting error might still persist, so we offer reporting error adjustments that can be run in conjunction with non-response corrections. Our proposed methods for the correction of reporting error involve using the calibration approach again to adjust reported values, instead of weights.

Correction methods for non-response and reporting error might or might not be dependent on each other depending on the assumptions made about the nature of reporting error. In Subsection 3.5, we offer a simplified procedure that provides the adjustment in a single step, which assumes that non-response and reporting error are independent of each other. Should non-response and reporting error depend on each other, we offer in Subsection 3.4 a framework for implementing them simultaneously in an iterative process.

In any case, the final product of the methodology is an adjusted survey data set with total estimates of net wealth, real assets, financial assets, and liabilities that match the aggregate figures in the national accounts balance sheet. This data set can be used to compute several distributional indicators of interest.

Before applying the method, we reclassified some definitions of wealth items used in the survey data to remove as many of the conceptual differences with national accounts as possible (see for instance Expert Group on Linking Macro and Micro data 2020; Chakraborty *et al.*, 2019). In particular, we removed the wealth held by non-profit institutions serving households (NPISHs) from national accounts totals, and we focused only on the items with the highest level of comparability.

3.1. *Pareto Tail Estimation*

Following from the literature, the proposed method is based on the fundamental assumption that the wealth distribution follows a power law.

Wealth is Pareto distributed if, above a certain wealth threshold $w_0 > 0$, the complementary cumulative distribution function (CCDF) of wealth is approximated by a power law, which, for $w_i \geq w_0$, can be expressed as:

$$P(W \leq w_i) = 1 - (w_0/w_i)^{\alpha}, \tag{1}$$

where the parameter $\alpha \in \mathbb{R}^+$ indicates the shape of the tail. The lower the value of α, the fatter the tail, and the more concentrated the wealth is.

The method described in Vermeulen (2016); Vermeulen (2018) has been developed precisely to overcome these issues. The first step of the adjustment consists in selecting the threshold w_0. Previous research has often adopted the arbitrary threshold of £1 million, which has been validated using the properties of Van der Wijk's law (see Waltl and Chakraborty 2022, for more detail). Van der Wijk's statistics for each country are shown in the Online Appendix, Figure 9, suggesting that a Pareto distribution should be supported at the 1 million threshold for all countries. In addition, in line with the findings from Waltl and Chakraborty (2022), Van der Wijk's statistics stabilizes at much lower values of wealth, suggesting that lower threshold can also be supported.

However, differential non-response and truncation can affect Van der Wijk's law in significant ways. In Online Appendix B, we show precisely that if the threshold and truncation point are too close to each other, then Van der Wijk's law cannot be invoked, and the issue only worsens in the presence of differential non-response. Figure 10 exemplifies this issue for the Italian case, showing how the distance between the threshold and the truncation point might be too small for robust inference. Waltl and Chakraborty (2022) also encountered this issue when setting the threshold at higher levels (2 million) for Austria and Germany. Nonetheless, in Online Appendix B, we also discuss that Van der Wijk's statistics can still stabilize if the threshold is sufficiently far away from the truncation point.

We then adopted the 1 million threshold, but also produced robustness estimates with a lower 500k threshold for Italy. Furthermore, drawing from the mean excess function, we developed a method for the estimation of a minimum threshold for which a Pareto distribution is supported, which we have detailed in Online Appendix A. Results with the estimated thresholds are also presented in the Online Appendix. The estimated thresholds, along with the 500k threshold, are far enough from the truncation point that the mean excess function is still approximately linear to wealth (as shown in Figure 10, in the Online Appendix) even in the presence of truncation and non-response.

Moving to the estimation of the parameters of the Pareto tail, define $\mathcal{S}_T = \{i \in \mathcal{S}_0, \text{ s.t. } w_i \geq w_0\}$ as the subsample of m_T respondents whose wealth is above the threshold. Let $\mathcal{S}_R$ be the set of population units in the rich list. Assume that for all units in $\mathcal{S}_R$, w_i is larger than the maximum value observed in $\mathcal{S}_T$.[3] We appended $\mathcal{S}_R$ and $\mathcal{S}_T$ creating a new set $\mathcal{S}_I = \mathcal{S}_T \cup \mathcal{S}_R$ of dimension m_I. In our approach, the imputed set $\mathcal{S}_I$ is used only for the estimation of α, and afterwards all adjustments were only applied to the original sample of respondents $\mathcal{S}_0$.

Let $\overline{D}_I$ be the average survey weight of all units in $\mathcal{S}_I$, that is, $\overline{D}_I = \sum_{i \in \mathcal{S}_I} d_i / m_I$, where we consider $d_i = 1$ for $i \in \mathcal{S}_R$. Now, let $\mathcal{S}_j$ be the subset of $\mathcal{S}_I$ with the j-th wealthiest households, for $j = 1, \dots, m_I$. That is, $\mathcal{S}_1$ is made of the wealthiest household, $\mathcal{S}_2$ is made of the two wealthiest households, and so forth, so that $\mathcal{S}_{m_I} = \mathcal{S}_I$. Then, $\overline{D}_j$ is the average weight of the j wealthiest households, i.e. $\overline{D}_j = \sum_{i \in \mathcal{S}_j} d_i / j$. Linear estimates for α can then be obtained through the following least squares specification (see also Gabaix and Ibragimov 2011):

$$\ln[(j - 1/2)\overline{D}_j / \overline{D}_I] = C - \alpha \ln(w_j), \tag{2}$$

where $C = \ln(m) + \alpha \ln(w_0)$ and w_j is the net wealth of the j-th wealthiest household.

Using Monte Carlo simulation, Vermeulen (2018) has shown that this method can produce estimates that approximate the true population parameter, while accounting for the increase in non-response probability as wealth increases. This means that information on the Pareto parameter can be retrieved even when survey weights decay progressively until the probability of non-response reaches 1. The method has been replicated by Waltl and Chakraborty (2022) and Waltl (2021), who showed that this estimator can produce unbiased and consistent estimates of α even when information on top tail observations is obtained from commonly available rich lists.

The second step of the adjustment consists of estimating the total wealth in the top tail $t(w; top)$ by multiplying the estimate of the total number of rich households from the survey, $\hat{N}_T = \sum_{i \in \mathcal{S}_T} d_i$, by the mean of the estimated Pareto distribution, given by $\alpha w_0 / (\alpha - 1)$, for $\alpha > 1$. The adjustment proposed by Vermeulen (2016; 2018) stops here: after this estimate is produced, no further adjustment is made to the survey. As our intention was to use this information to adjust survey data, this estimate could be used to calibrate the sampling weights of rich households in the survey to the total estimated wealth in the Pareto tail.

However, this approach would assume that the sample estimate of N_T (the total number of households whose wealth exceeds w_0) is unbiased. We believe that this assumption should be relaxed. After all, as discussed earlier, the presumption of differential non-response is based on the assumption that richer households are progressively less likely to be survey respondents. The absence of these households leads not only to the missing tail but also to a smooth underestimation of survey weights as wealth increases. These imbalances will be reflected in final survey weights even if some kind of post-stratification re-weighting is performed, and even in the presence of oversampling.[4] As these households are all located in the higher part of the

[3]We then removed from the rich lists all observations displaying lower wealth than what the maximum wealth observed in the survey. This is an extremely rare occurrence, and only few observations from the Finnish rich list were removed.

[4]Unless the oversampling strategy covers the full Pareto tail.

Pareto tail, the undersampling of even a small number of households could have a sizeable impact on the amount of estimated wealth in the tail.

To the best of our knowledge, of all studies based on Pareto adjustments, only the method from Blanchet *et al.* (2022b) implicitly considers this issue by calibrating survey data so that it reflected the "true" density observed in tax records. While Blanchet *et al.* (2022b) focus on the income distribution and, as such, have easier access to tax data with much larger support for the Pareto tail, we did not have access to such information and a large gap between the survey and the rich list often remains. This implies that, even in the absence of differential non-response in the observed part of the Pareto tail of the wealth distribution, we would need to estimate the total number of households in the Pareto tail. We then propose a novel method for the estimation of the number of missing rich households and their wealth.

Recall that w_t is the truncation point above which there are no rich households in the sample. Let us focus here on $\mathcal{S}_T$ only and let $\mathcal{S}_j$ be now the subset of $\mathcal{S}_T$ with the j-th wealthiest households, for $j = 1, \dots, m_T$. That is, $\mathcal{S}_1$ is made of the wealthiest household in $\mathcal{S}_T$, $w_{m_T} = w_t$, and $\mathcal{S}_{m_T} = \mathcal{S}_T$. Following from the Glivenko–Cantelli theorem, because of truncation, the empirical cumulative distribution function resulting from this sample is unlike the theoretical distribution implied by the Pareto adjustment. In particular, the following relation does not hold:

$$\frac{\hat{N}_T - \hat{N}_{j-1}}{\hat{N}_T} - \left[1 - \left(\frac{w_0}{w_j}\right)^{\alpha}\right] \approx 0, \tag{3}$$

where $\hat{N}_{j-1} = \sum_{i \in \mathcal{S}_{j-1}} d_i$, that is, the sum of weights for all households with wealth greater than w_j, so that $\hat{N}_T - \hat{N}_{j-1}$ is the survey estimate of the number of households in the population whose wealth is between w_0 and w_j. This relation means that the empirical CCDF will always suffer from bias equal to or larger than zero because units whose wealth exceeds w_t are unobserved. Substituting $\hat{N}_T$ in the denominator with the unknown N_T in equation (3) and rearranging it leads to the following equation:

$$N_T \approx \frac{\hat{N}_T - \hat{N}_{j-1}}{1 - (w_0/w_j)^{\alpha}}. \tag{4}$$

Analytically, the estimate from equation (4) should be the same for each $j \in \mathcal{S}_T$. In practice, with empirical data, variability in survey weights will affect the estimate of the number of households in the tail. Because of differential non-response, this becomes a particularly thorny problem when weight quality can deteriorate as the observed wealth gets closer to the truncation point w_t, where the maximum wealth in the survey sample $\mathcal{S}_T$ is recorded. The results can then be improved by estimating N_T for each value of wealth of top tail observations and then getting the average. That is:

$$\tilde{N}_T = \frac{1}{m_T} \sum_{j=1}^{m_T} \frac{\hat{N}_T - \hat{N}_{j-1}}{1 - (w_0/w_j)^{\alpha}}. \tag{5}$$

The intuition behind this method stems from the idea that, due to differential non-response, the probability of non-response increases along the Pareto distribution. If non-differential non-response is negligible, then $\tilde{N}_T$ should be comparable for each $j \in S_T$. In the presence of differential non-response, observations closer to the threshold offer a more accurate estimate $\tilde{N}_T$ and are weighted more simply because richer observations are more likely to be missing. Further improvements can be obtained by weighting observations by their rank, or by establishing regions of support along the Pareto tail in which the value of $\tilde{N}_T$ is assumed to be correct.

An estimator of the number of missing, unobserved, households after the truncation point (the *missing tail*, henceforth) can be computed as $\hat{N}_t = \tilde{N}_T(w_0/w_t)^\alpha$. To account for these missing households, the total number of observable households can be estimated as

$$\hat{N}_{obs} = \tilde{N}_T[1 - (w_0/w_t)^\alpha]. \tag{6}$$

Finally, the total wealth in the top tail $\hat{t}(w; top)$ can be estimated by the product of the estimated number of households and the Pareto mean as

$$\hat{t}(w; top) = \tilde{N}_T \frac{\alpha w_0}{(\alpha - 1)}. \tag{7}$$

Wealth in the missing part of the tail can similarly be computed as

$$\hat{t}(w; miss) = \tilde{N}_t \alpha w_t/(\alpha - 1) \tag{8}$$

by setting the new threshold at the truncation point w_t. Note that this is possible because the Pareto shape parameter does not change along the Pareto distribution.

3.2. *Calibration*

Calibration is a method whose aim is to correct the sampling weights d_i through re-weighting methods while keeping the individual responses w_i unchanged (Deville and Särndal 1992; Särndal 2007). In the literature, this approach is mainly used: (1) to force consistency in certain survey estimates with known population quantities; (2) to reduce non-sampling errors such as non-response errors and coverage errors; and (3) to improve the precision of estimates (Haziza *et al.*, 2017).

Calibration is achieved through the following optimization problem for finding a new set of weights $d_i^* = d_i a_i$:

$$\min_{d_i^*} \sum_{i \in S_0} G(d_i^*; d_i) \quad s.t. \quad t(z) = \sum_{i \in S_0} d_i^* z_i, \tag{9}$$

where $G(d_i^*; d_i)$ is a distance function between the basic design weights and the new calibrated weights, z_i is the value on unit i taken by a (possibly) vector valued auxiliary variable z, and $t(z)$ are the benchmark constraints, that is, the known vector of population totals or counts of the calibration variables z. The adjustment factors a_is are a function of the z_is, and they are computed so that final weights meet benchmark constraints, $t(z)$, while, at the same time, being kept as close as possible to the initial ones. Closeness can be defined by means of several distance functions (see table 1 in Deville and Särndal 1992), the most common being the one that resembles a chi-square statistic,

$$G(d_i^*; d_i) = \frac{(d_i^* - d_i)^2}{d_i c_i}, \tag{10}$$

where c_i are known constants, the role of which will be discussed in more detail later, for which an analytical solution always exists.

The final output is a single new set of weights to be used for all variables. The magnitude of the adjustment factors and therefore the variability of the final set of weights is a function of a number of constraints, that is, the length of $t(z)$, and the imbalance (the difference between the Horvitz–Thompson estimate and the population total of z). Very variable weights hinder the quality of final estimates for sub-populations and for variables that are not involved in the calibration procedure. For these reasons, weights are usually required to meet range restrictions such as to be positive and/or within a chosen range. This can be achieved by carefully choosing and tuning the distance function $G(\cdot)$.

The method was originally proposed to improve the efficiency of the estimators and to ensure coherence with population information, but then it was also largely applied to adjust for non-response (Särndal and Lundström 2005). For example, Little and Vartivarian (2005) showed that if the variables used to construct the weights are associated both with non-participation and with the variable of interest, the bias and the variance of the estimator are reduced.

The main problem with the use of household balance sheet data in re-weighting methods is that wealth is generally skewed and concentrated in the hands of a small group of the population that has both low propensity to participate in the survey and different socio-demographic characteristics from the average population.

3.3. *Adjusting for Non-response: Pareto-Calibration*

We begin by exploiting the information obtained after fitting a Pareto distribution, as in Subsection 3.1, to adjust the wealth distribution in the survey for differential non-response using the calibration methods described in Subsection 3.2. We used w_0, $\hat{\alpha}$ and equation (5) to estimate the total number of observable households over the threshold $\hat{N}_{obs}$ (see equation 6) and their total net wealth by $\hat{t}(w; obs) = \hat{t}(w; top) - \hat{t}(w; miss)$, where $\hat{t}(w; miss) = \tilde{N}_T \alpha w_t / (\alpha - 1)$. We then derived $\hat{N}_{bot} = N_{bot} - (\tilde{N}_T - N_{obs})$ as the adjusted number of households below the threshold w_0. We then calibrated the sampling weights from sample S_0 using the following constraints:

$$t(z) = (\hat{t}(w; obs), \hat{N}_{obs}, \hat{N}_{bot}, \hat{t}(y; bot), t(x)), \tag{11}$$

where $\hat{t}(y; bot)$ is a vector of Horvitz–Thompson estimators decomposing the initial wealth of observations below the threshold into their corresponding portfolio items, and $t(x)$ is a vector of population counts (scaled to $\hat{N}_{bot}$) for demographic characteristics. To allow for the calibration among households with zero wealth, wealth is expressed in normal terms (and so will be for the rest of the paper).

The main intuition behind this method is the fact that the effect of a unitary change in the weights over the total of wealth in the sample tends to be zero in the non-Pareto part of the survey, but is large in the Pareto tail.

After calibrating survey data according to these parameters, we obtained non-response adjusted weights d_i^*'s. The combined approach of fitting the Pareto tail, re-estimating the number of households, and then calibrating the survey weights will be referred to as "Pareto-calibration" (or, for brevity, "P-C") from now on. It is worth noting that our Pareto-calibration approach is also intended to ensure that the transition between the Pareto tail and the rest of the distribution remains smooth. Given the calibration constraints, wealthier households will see the largest increase in their weights, whereas the weights of households near the threshold will remain mostly unchanged. This will be shown later in Figure 1.

This calibration procedure is standard and has been used to a similar extent in Blanchet *et al.* (2022b). The main difference with the work from Blanchet *et al.* (2022b) relates to how information on the density of households is obtained because, due to the nature of rich lists, there is barely an overlap between the data sources that we use. The final steps involving calibration alone are comparable.

Should the survey be suffering from differential non-response issues only, this step might be sufficient to fill the gap with the financial accounts. However, this is not always the case: provided that we have a good approximation of wealth distribution in the tail, the remaining differences in coverage between the estimate obtained in equation (7) and the national accounts will then be left to reporting error.

3.4. *Adjusting for Non-response and Reporting Error: Simultaneous Approach*

To correct for reporting error, we combined the adjustment for differential non-response described in Subsection 3.3 with the following procedure. First, we ran the Pareto-calibration adjustment, as described earlier. Let d_i^*, for $i \in S_0$, be the final weight from the non-response adjustment procedure. Next, we ran a calibration procedure as in equation (9) in which (1) the d_i^*'s are now the starting weights and (2) the set of benchmark constraints $t(z)$ are given by the vector of macro aggregates of wealth items y. The adjustment factor a_i, for $i \in S_0$, obtained by this procedure is such that

$$\sum_{i \in S_0} d_i^* a_i y_i = t(y). \tag{12}$$

We applied this adjustment factor directly to the variables of interest so that

$$y_i^* = a_i y_i, \tag{13}$$

that is, we choose to adjust the values of observations for the components of wealth, rather than their weights. In this way, we avoided the possibility of large adjustments affecting weights that are used to compute estimates for all variables in the survey, and not only for wealth components. This approach shares similar traits with reverse calibration introduced by Chambers and Ren (2004) to deal with outlier-robust imputation.

It is worth noting that the method allows the user to choose between options that range from using one single calibration that includes all the constraints to using a different calibration for each constraint. In the first case, every household has a different adjustment factor a_i that depends on all the values of y. The underlying assumption is that reporting issues are household-specific but equal for different wealth components. In the second instance, the method would produce adjustment factors that are both household and item specific. Proportional allocation, which consists of allocating the gap by multiplying each component of y_i by the corresponding inverse of the item-specific coverage ratio, can be embedded in our approach as it can be seen as a particular case of univariate calibration. In fact, if we focus on a single item, y_1, the adjustment factor used by proportional allocation can be obtained as the solution to a univariate calibration procedure in which (1) the starting weights are again the d_i^*s, (2) there is only one benchmark constraint $\sum_{i \in S_0} d_i^* a_i y_{i1} = t(y_1)$, and (3) the distance function $G(\cdot)$ is chi-squared as in (10) with constants $c_i = 1/y_{i1}$. The proof has been omitted for brevity, but it is close in spirit to Example 1 in Deville and Särndal (1992).

This equivalence sheds some light on the role of the constants c_i's in the distance function (9). In univariate calibration, if they are chosen to be the inverse of the variable in the constraint, then the adjustment factors are shrunk toward a common value for all households as in proportional allocation. On the contrary, if they are set to be constant, the adjustment factors would be roughly proportional to the values of the item. For this reason, in the proposed multivariate calibration for imputation, we have set the constants to possibly depend on the wealth of the household, that is,

$$c_i = \left(\frac{1}{w_i}\right)^{\tau}, \tag{14}$$

where $\tau \geq 0$ can be seen as a shrinkage factor: larger values provide adjustment factors that are more uniform across households, while values toward 0 provide adjustment factors with a higher variability and correlation with w_i.[5]

Note that the adjustment factors a_i may be very variable because we are using a multivariate calibration approach, and the extent of reporting error can be considerable. This is particularly relevant when using the chi-squared distance function for which the adjustment factors can even take negative values. For this reason, we recommend the use of alternative distance functions, such as the raking (Case 2 in Deville and Särndal 1992) for which positive adjustments are ensured, or the range-restricted version of the chi-squared and of the raking distance functions (Cases 6 and 7 in Deville and Särndal 1992), for which adjustments are bounded to be in a pre-specified interval.

To account for the missing wealthy households, we add a single observation with weight $\hat{N}_t$ and wealth $\hat{t}(w; miss)/\hat{N}_t$. This observation's portfolio was also allocated using portfolio shares in the Pareto tail of the distribution.

At the end of the multivariate calibration, the gap is filled. However, the distribution of w_i has changed, because its components have changed. The Pareto tail parameter and its threshold might have changed. Therefore, we might need to find

[5]For this work, we set $\tau = 1$. Future research might seek to retrieve information on τ using external data where no misreporting behavior is present.

a new Pareto threshold (our solution is to rescale the threshold w_0 by the average adjustment factor) and repeat the Pareto-calibration procedure again. This requires an iterative procedure that alternates between the non-response and reporting error corrections. The Pareto-calibration and multivariate calibration steps are then iterated until convergence. Convergence has been set on the parameter α of the Pareto distribution: if the estimated values in two consecutive steps differ by less than a small predefined threshold, the procedure stops.

3.5. *A Special Case: Single-Iteration Approach*

If one is willing to assume that (1) the relative reporting error is independent of the observed wealth, at least among the very rich, converging in probability to a constant ζ, i.e. $(w^* - w)/w \xrightarrow{p} \zeta$, so that, on average, the unobserved "true" total wealth will be given by $\hat{w}_i^* = \zeta w_i$, and (2) that the degree of relative reporting error between survey and rich list is also similar, then the non-response and reporting corrections can be performed into a single step. Should these assumptions hold, survey wealth would still be Pareto distributed with tail parameter α after adjusting for reporting error. It follows that total wealth in the survey would scale up to $\sum_{i \in S_0} \zeta d_i^* w_i$, and the Pareto CCDF would turn into $F_\alpha(\zeta w_i) = 1 - (\zeta w_0 / \zeta w_i)^\alpha$. Simplifying this last formula and updating equation (7) for reporting error, we obtain the following estimate for total wealth:

$$\zeta \hat{t}(w) = \zeta \left(\frac{\alpha w_0}{(\alpha - 1)} \hat{N}_{obs} + \sum_{i \in S_0} d_i^* w_i \right). \tag{15}$$

This implies that our estimate for α does not depend on the scaling of the variables. In this case, the coefficient for the Pareto-adjusted coverage ratio, given the national accounts total wealth, as in $\zeta = t(w)/\hat{t}(w)$, will yield the scalar to which to re-allocate reported survey wealth. To account for the missing wealth, wealth should be scaled to $\zeta(\hat{t}(w) - \hat{t}(w, miss))$, which, after Pareto-calibration, simplifies to $\zeta \sum_{i \in S_0} d_i^* w_i$.

If the Pareto shape parameter is unaffected by the rescaling, the iterative procedure is no longer needed. The adjustment for reporting error and for non-response at the tail of the distribution can be run independently of each other.

Ideally, if the aforementioned assumptions hold, whatever the adjustment method for reporting error is used, the final data should still be Pareto distributed among rich households. We can then adopt a single-iteration approach by employing either proportional allocation or a non-iterative version of the multivariate calibration method described in Subsection 3.4.

As will be discussed later, the robustness of these estimates strongly depends on the level of discrepancy in terms of reporting error between survey and rich list data. The assumption (2), in fact, does not hold in all contexts.

4. Assessment of the Method

The ideal approach for assessing the quality of the results would be to compare them with an external benchmark, for instance, from highly reliable administrative records. Without such auxiliary information, we assess the method in two ways. First, we assess the robustness of our results by comparing them with other estimators based on different assumptions. Second, we assess the precision of our results by estimating their variability.

We compute five alternative estimators. All adjustments incorporate $\hat{t}(w; miss)$, the wealth held by the missing tail, into the estimate, which is computed using the formula (8) under the procedure described in Subsection 3.1. The estimators are as follows:

- "Survey & missing tail" (B.+Tail). The results are produced using the unadjusted survey data, and an estimation of the total wealth held by rich households with zero probability of being in the survey (missing tail).
- "Pareto-calibration" (P-C). Survey data are adjusted with the Pareto-calibration model. Survey weights are calibrated using constraints from equation (11) and the total wealth of the missing tail is included in the estimate.
- "Pareto-calibration, proportional allocation" (P-C, P.A.). This method adds a correction for reporting error based on proportional allocation (as in Fesseau and Mattonetti 2013) to the previous one. This is a naive method based on the assumption that reporting error is equal across households and that it only depends on the financial instrument.
- "Pareto-calibration, single-iteration approach" (P-C, S.I.). In this method, the correction for reporting error is only iterated once, as described in Subsection 3.5
- "Pareto-calibration, simultaneous approach" (P-C, SIM.). In this method, the correction for non-response and reporting error are iterated until convergence, as described in Subsection 3.4.

Variance estimation in our methodology has two main components. The first is the sampling variance, which indicates the variability introduced by choosing a sample instead of enumerating the whole population, assuming that the information collected in the survey is otherwise entirely correct. The second source of variability arises from the imputation process and can be attributed to the fact that the methodology for filling the gap can produce several different plausible imputed data sets. To estimate the overall variability, we use the Rao–Wu rescaled bootstrap weights released with HFCS data. In particular, the replicate samples are drawn independently and with replacement in each stratum. The number of units drawn in each stratum of size n_h is set to $n_h - 1$. The final estimation weight for each observation is then rescaled by a factor $n_h/(n_h - 1)$, and multiplied by the frequency of the observation in the replicate sample (Eurosystem Household Finance and Consumption Network 2020). For each of the so-obtained 1,000 sets of bootstrap weights, we replicated all the methods previously described. In each replication, the parameters of the Pareto distribution were re-estimated, introducing additional variability. We then obtain the mean and standard deviation from all successful simulations[6] to evaluate the robustness of our methods and derive a measure of their variability.

Only the weights of the survey were re-sampled. The rich list should be held fixed, and the weight of observations in the rich list was held to equal 1. This was

[6]A simulation is flagged as unsuccessful and discarded whenever a calibration procedure fails because of lack of convergence under the chosen restraints.

done for two reasons: (1) because the procedure is intended to estimate the variability of each adjustment method as the survey sample changes and (2) because it is vital for the rich list to preserve the correct rank of the observations to estimate the Pareto distribution. In other words, re-sampling the rich list would lead to the same problem of "missing wealth" that Pareto adjustments are intended to address because the very richest would be less likely to be re-sampled.

5. Results

The method described in the previous sections has been applied to the second 2014 wave of the HFCS. This section describes both the methodological and the economic results.

5.1. *Model Estimation*

We begin with the Pareto-calibration model, hence by fitting the Pareto distribution and adjusting the survey accordingly. Figure 1 shows the estimates of the parameter of the Pareto distribution. In particular, α indicates the Pareto shape parameter estimated by imputing the rich list, while θ refers to the estimation results with survey data only. The figure also illustrates the outcome of the Pareto-calibration process, showing the empirical CCDF on a log-log scale before and after the adjustment.

Can calibration account for smoothly declining response rates without affecting the nature of the overall distribution? Further evidence in this direction is offered by Figure 2, showing the change in weights in the Pareto tail before and after the Pareto-calibration method is employed. Notably, the figure shows that calibration approach maximizes the change in weights for observations furthest from the threshold, and this is especially true for surveys that do not feature a proper sampling of rich households. While some small discontinuities could still emerge, Figures 1 and 2 together show that the method is preferable to a uniform rescaling approach as not only the Pareto tail parameter remains relatively unaffected, but also the deviation from the original weights is minimized for most observations.[7]

Table 2 shows coverage ratios between survey wealth estimates and financial accounts. Column (1) shows initial coverage ratios, while Column (2) displays the coverage ratio after Pareto-calibration for the observed part of the survey only (so that only $\hat{t}(w; obs)$ from the Pareto tail is included), and Column (3) adds to the previous estimate total wealth after truncation ($\hat{t}(w; miss)$). Columns from (4) to (7) focus on the number of households in the Pareto tail. Column (4) displays the original number of households in the survey tail. Column (5) shows the estimated number of households in the Pareto tail, while the corresponding confidence interval of this estimate is shown in Column (6). Column (7) reports the number of "missing rich" right of the truncation point.

[7]For the full distribution, see Figures 3–6 in the Online Appendix, which show the empirical probability density distribution before and after the Pareto-calibration step. Wealth is expressed in log terms for visualization purposes only. These density plots were produced with a Gaussian smoothing kernel whose bandwidth was estimated using Silverman (2018)'s (Silverman (2018)) "rule of thumb."

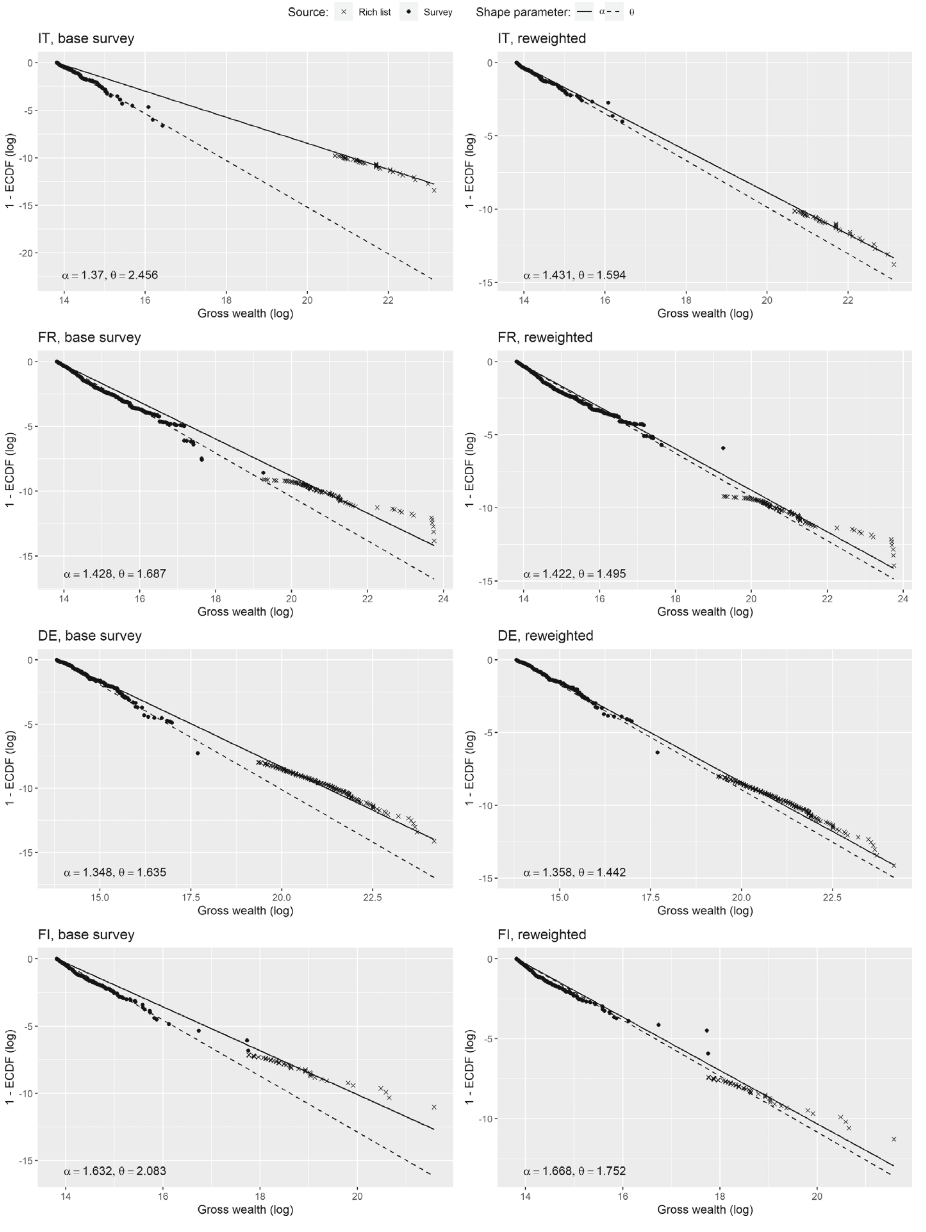

FIGURE 1. Pareto Tail Re-weighting.

Notes: Empirical cumulative distribution functions (log scale) for survey wealth distributions in the Pareto tail. Re-weighting achieved using the Pareto-calibration method, using the calibration benchmarks from equation (11). θ parameters estimated using survey data only, α estimated using Vermeulen's Silverman (2018) regression method with imputed rich list.

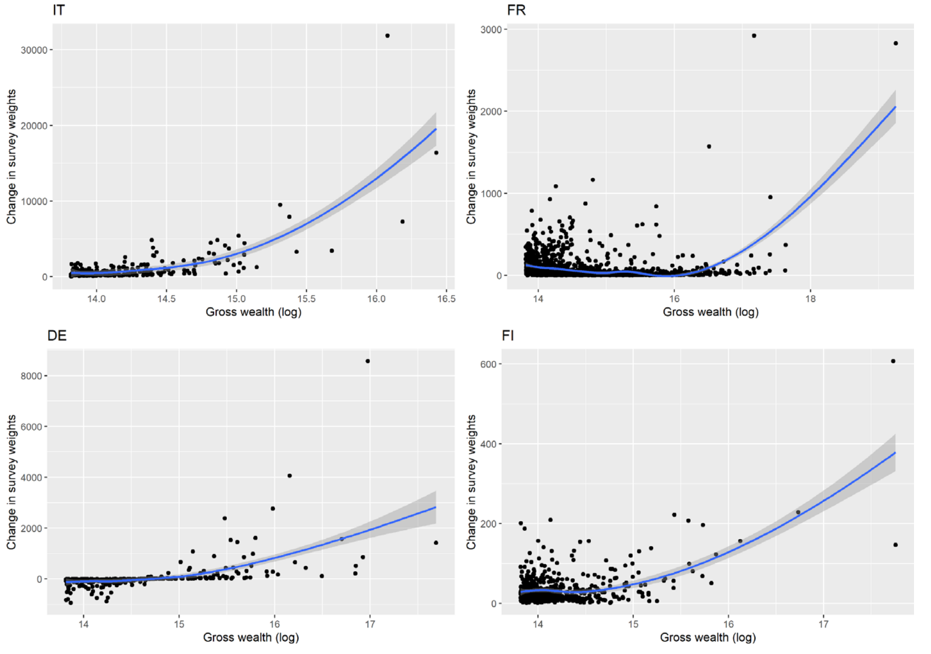

FIGURE 2. Pareto Tail Re-weighting.

Notes: Change in weights in the Pareto tail after the Pareto-calibration method, using the calibration benchmarks from equation (11), and superimposed local regression fit. Threshold set at EUR 1 million. [Colour figure can be viewed at wileyonlinelibrary.com].

TABLE 2
Pareto Adjustments for Gross Wealth

	Coverage Ratios			Estimated Tail Households			
Country	Base (1)	P-C (2)	+ Missing Tail (3)	Base (4)	Estimated (5)	95% C.I. (6)	Missing (7)
IT	0.553	0.662	0.801	681,305	1,095,932	$\pm$ 4009.422	27,395.55
FR	0.673	0.772	0.806	1,001,778	1,145,012	$\pm$ 190.461	473.00
DE	0.827	0.870	0.997	1,350,286	1,419,375	$\pm$ 3016.263	7399.215
FI	0.917	1.035	1.061	60,794	87,290	$\pm$ 145.599	127.250
IT ($w_0 = 500k$)	0.553	0.685	0.787	2,493,748	2,972,082	$\pm$ 9865.057	23,799.14

Notes: Estimated coverage ratios (Total survey wealth/Total wealth in the National Accounts) and number of households in the tail. Cols. (1) and (4): base survey estimates for coverage ratios and no. of households. Col. (2): coverage ratios after Pareto-calibration, calibrating the households to the estimated number shown in col. (5) (confidence intervals in col. 6). Col. (3): coverage ratios after adding the missing tail from Col. (7) (estimated number of households after the truncation point) to the Pareto-calibration estimate.

Overall, these figures suggest that the proposed Pareto-calibration approach can produce substantial improvements in survey coverage, especially in the absence of auxiliary information. In the case of Finland and Germany, the discrepancies between micro and macro figures virtually disappear after calibrating survey weights and accounting for the unobservable households. These results suggest that corrections for differential non-response can solve the missing wealth problem alone in cases in which reporting error is limited: the fact that the Finnish survey relies on registry data, in which reporting error is virtually absent, and that the coverage gap is completely filled after calibrating the survey weights, provides evidence in support of this argument.

Moreover, it should be noted that the improvements in coverage between columns (1) and (2) are much larger than the improvements that could be noted between columns (2) and (3). This makes a strong case in favor of our Pareto-calibration approach, meaning that simply addressing for the decay in survey weights in the observed part of the distribution can already address the largest part of the non-response problem.

Furthermore, having re-estimated the number of households in the Pareto tail of the survey, our method also shows substantial improvements in coverage over the grossing up methods already explored in the literature, and suggests that adjustments for non-response should also focus on correcting the number of households estimated to be in the Pareto tail, rather than only the wealth contained in it.

Coverage is also significantly improved for Italy and France, but the persistence of a mismatch between survey data and financial accounts points to the presence of reporting error. In the case of the Italian survey, a case could be for the 1 million empirical threshold being too high. The change in the Pareto α coefficient before and after the Pareto-calibration is significant and suggests that the threshold is placed at a point where survey weights are already suffering from decay due to non-response. Considering that the Italian survey is already known for suffering from both non-response and reporting error, it is not unreasonable to believe that the threshold could be placed lower in the distribution. This is evident from a visual inspection of the wealth distribution, as can be seen from Figures 3 and 7.

As discussed, the 1 million threshold might be a generous one. Using our threshold selection method based on the mean excess function (detailed in Online Appendix A), we find a lower boundary for the threshold at EUR 310,084 for Italy, 567,378 for France, 254,000 for Germany, and 880,806 for Finland. For these thresholds, the Pareto distribution is supported and the distorting effect of non-response and truncation is also minimized, as discussed in Online Appendix B with regard to the properties of Van der Wijk's law under truncation. This suggests that observations at 1 million threshold might already be, at best, a generous one (as discussed already in Waltl and Chakraborty 2022) and, at worst, an inadequate one for countries like Italy. In the Online Appendix, Figures 7 and 8 show the distributional changes for Italy with a EUR 500,000 threshold: these distributional changes now appear much more reasonable, with no significant change in the estimated Pareto parameter.

After dealing with the issue of differential non-response in the tail of the distribution, we further adjusted for reporting error along the whole distribution. Depending on the assumption on the nature of reporting error, some methodologies might be more or less appropriate. In this section, we only detail the estimation procedure for the simultaneous approach, as the other proposed approaches, being far simpler, lead immediately to distributional results.

Regarding the simultaneous approach, for multivariate calibration $t(y)$ we used the financial instruments with high conceptual comparability between survey and financial accounts as benchmark constraints—namely, deposits, bonds, shares, funds, insurance products, and liabilities—following from the comparability scale provided by Expert Group on Linking Macro and Micro data (2020). The resulting adjustment factors are then applied to financial instruments with lower comparability—business and housing wealth—which should ensure that the adjustment will not be biased by the presence of instruments with low comparability, assuming that reporting error is comparable across similar financial instruments.

We then iterated the simultaneous approach until convergence. Convergence has been set on the parameter α of the Pareto distribution: if the estimated value in two consecutive steps differs by less than a small predefined threshold,[8] the procedure stops. Convergence is usually achieved in a limited number of steps (between 1 and 3 in the application at hand).

Table 3 shows the average values of the adjustment factors a_i's (as well as coefficients of variation) as a function of gross wealth percentiles at the end of the iterative procedure for the four countries. That is, these are the overall adjustment of the survey variables at the end of the procedure obtained as the ratio between the final imputed values and the ones from the original survey.

5.2. *Distributional Results*

Table 4 shows the distributional results indicating the proportion of net wealth held by the top 1, 5, 10, and 20 percentiles, along with the bottom 50 percent. Gini inequality indices are also presented in Column (6), while Column (7) provides the estimated Pareto tail parameter α given the data. These figures have been reproduced under each allocation method. The bootstrap-based standard deviation is reported in parentheses for each estimate.

The row denoted with "Base Survey" presents distributional figures from the unadjusted HFCS data.

For all other rows corresponding to the alternative estimators discussed in Section 4, we also included an adjustment for the unobserved part of the Pareto tail as presented in Subsection 3.1. To do this, these missing households were imputed as a single observation whose weight and wealth were respectively equal to the estimated number of unobserved households and the estimated average wealth in the unobserved Pareto tail.

The rows denoted by the "B.+Tail" ("Survey & missing tail") method display estimates produced using the unadjusted survey data, and the missing tail households.

Survey weights were then adjusted using the proposed Pareto-calibration method to produce the figures shown in the set of rows "P-C" ("Pareto-calibration"). This method produces a further increase in inequality, even if the magnitude is smaller than the previous step.

[8]In the current application, this tolerance was set at 0.05.

TABLE 3
Simultaneous Approach: Mean and Coefficient of Variation of Overall Adjustment Factors a_i

	Percentile									
Country	0.10 (1)	0.20 (2)	0.30 (3)	0.40 (4)	0.50 (5)	0.60 (6)	0.70 (7)	0.80 (8)	0.90 (9)	1.00 (10)
IT	1.278	1.742	2.098	1.209	1.093	1.083	1.116	1.139	1.181	1.288
	(0.289)	(0.295)	(0.274)	(0.102)	(0.024)	(0.018)	(0.019)	(0.025)	(0.022)	(0.088)
FR	1.037	1.587	2.403	1.778	1.228	1.131	1.122	1.125	1.132	1.215
	(0.107)	(0.241)	(0.162)	(0.351)	(0.059)	(0.007)	(0.005)	(0.007)	(0.009)	(0.132)
DE	1.162	1.368	1.703	1.876	1.713	1.406	1.212	1.144	1.134	1.121
	(0.169)	(0.126)	(0.076)	(0.048)	(0.080)	(0.095)	(0.028)	(0.011)	(0.009)	(0.020)
FI	0.942	1.078	1.334	1.101	1.040	1.027	1.032	1.039	1.047	1.047
	(0.087)	(0.095)	(0.075)	(0.047)	(0.005)	(0.003)	(0.005)	(0.002)	(0.003)	(0.029)
IT (500k)	1.271	1.578	1.632	1.132	1.072	1.072	1.108	1.123	1.169	1.292
	(0.136)	(0.154)	(0.225)	(0.049)	(0.017)	(0.016)	(0.014)	(0.025)	(0.020)	(0.114)

Notes: Final multivariate calibration adjustment factors a_i, equations (12) and (13), from the multivariate calibration approach for imputation as a function of gross wealth percentiles.

TABLE 4
WEALTH DISTRIBUTION ESTIMATES

	Wealth Shares							
Method	Top 1% (1)	Top 5% (2)	Top 10% (3)	Top 20% (4)	Bot 50% (5)	Gini (6)	Tail α (7)	S.r. (8)
	Italy							
Base survey	0.112	0.292	0.423	0.597	0.103	0.597	1.860	–
	(0.074)	(0.032)	(0.021)	(0.013)	(0.039)	(0.012)	(0.075)	
B.+Tail	0.284	0.431	0.537	0.677	0.083	0.676	1.370	1.000
	(0.020)	(0.017)	(0.015)	(0.012)	(0.004)	(0.011)	(0.016)	
P-C	0.323	0.482	0.584	0.712	0.075	0.707	1.431	0.996
	(0.016)	(0.016)	(0.016)	(0.012)	(0.004)	(0.011)	(0.016)	
P-C, P.A.	0.342	0.498	0.597	0.721	0.073	0.716	1.475	0.996
	(0.026)	(0.022)	(0.020)	(0.015)	(0.005)	(0.014)	(0.011)	
P-C, S.I.	0.400	0.556	0.649	0.762	0.064	0.712	1.370	0.996
	(0.123)	(0.096)	(0.077)	(0.052)	(0.014)	(0.052)	(0.016)	
P-C, SIM.	0.290	0.465	0.580	0.716	0.074	0.701	1.475	0.996
	(0.019)	(0.017)	(0.016)	(0.013)	(0.005)	(0.011)	(0.019)	
	France							
Base survey	0.168	0.352	0.485	0.655	0.073	0.655	1.768	–
	(0.089)	(0.036)	(0.022)	(0.012)	(0.040)	(0.011)	(0.052)	
B.+Tail	0.216	0.391	0.518	0.679	0.066	0.679	1.428	1.000
	(0.019)	(0.015)	(0.012)	(0.009)	(0.003)	(0.009)	(0.014)	
P-C	0.292	0.455	0.571	0.715	0.059	0.713	1.422	1.000
	(0.010)	(0.009)	(0.008)	(0.006)	(0.002)	(0.006)	(0.014)	
P-C, P.A.	0.295	0.458	0.572	0.716	0.058	0.715	1.448	1.000
	(0.011)	(0.009)	(0.008)	(0.006)	(0.002)	(0.006)	(0.019)	
P-C, S.I.	0.297	0.463	0.576	0.715	0.070	0.705	1.428	0.998
	(0.025)	(0.020)	(0.016)	(0.012)	(0.004)	(0.015)	(0.014)	
P-C, SIM.	0.295	0.461	0.575	0.714	0.071	0.705	1.432	1.000
	(0.016)	(0.014)	(0.013)	(0.010)	(0.003)	(0.006)	(0.017)	
	Germany							
Base survey	0.219	0.439	0.573	0.742	0.030	0.739	1.575	–
	(0.111)	(0.050)	(0.031)	(0.015)	(0.073)	(0.014)	(0.159)	
B.+Tail	0.326	0.517	0.632	0.778	0.026	0.775	1.348	0.981
	(0.019)	(0.014)	(0.012)	(0.008)	(0.002)	(0.007)	(0.019)	
P-C	0.345	0.537	0.648	0.787	0.025	0.784	1.358	0.952
	(0.019)	(0.015)	(0.013)	(0.009)	(0.002)	(0.008)	(0.019)	
P-C, P.A.	0.342	0.535	0.647	0.786	0.025	0.784	1.359	0.952
	(0.022)	(0.017)	(0.014)	(0.009)	(0.002)	(0.008)	(0.012)	
P-C, S.I.	0.348	0.524	0.634	0.772	0.035	0.759	1.348	0.941
	(0.059)	(0.043)	(0.034)	(0.021)	(0.004)	(0.028)	(0.019)	
P-C, SIM.	0.323	0.506	0.621	0.764	0.037	0.759	1.378	0.940
	(0.016)	(0.012)	(0.012)	(0.009)	(0.003)	(0.009)	(0.015)	

Table 4
Continued

Method	Wealth Shares Top 1% (1)	Top 5% (2)	Top 10% (3)	Top 20% (4)	Bot 50% (5)	Gini (6)	Tail α (7)	S.r. (8)
	Finland							
Base survey	0.120	0.285	0.416	0.595	0.101	0.596	2.162	–
	(0.050)	(0.019)	(0.012)	(0.007)	(0.025)	(0.007)	(0.069)	
B.+Tail	0.144	0.305	0.432	0.606	0.098	0.608	1.632	1.000
	(0.007)	(0.006)	(0.005)	(0.004)	(0.003)	(0.004)	(0.021)	
P-C	0.205	0.363	0.483	0.643	0.089	0.642	1.668	0.993
	(0.008)	(0.008)	(0.007)	(0.006)	(0.002)	(0.006)	(0.021)	
P-C, P.A.	0.209	0.367	0.486	0.645	0.089	0.644	1.649	0.993
	(0.009)	(0.009)	(0.008)	(0.006)	(0.003)	(0.006)	(0.019)	
P-C, S.I.	0.202	0.362	0.481	0.642	0.093	0.622	1.632	0.989
	(0.040)	(0.033)	(0.027)	(0.019)	(0.005)	(0.010)	(0.021)	
P-C, SIM.	0.174	0.344	0.476	0.647	0.087	0.640	1.767	0.992
	(0.008)	(0.008)	(0.008)	(0.007)	(0.003)	(0.006)	(0.029)	

Notes: Wealth shares (cols. 1–5) and inequality indexes (cols. 6 and 7) computed using the different estimators described in Section 4. Pareto threshold is set at net wealth of EUR 1 million. Standard deviation reported in parenthesis. Figures estimated using different adjustments for the HFCS data, and accounting for the unobserved part of the Pareto tail. Survey weights are replaced by bootstrap weights. Success rates (col. 8) report the observed probability of convergence for calibration.

In the set of rows "P-C, P.A." ("Par-cal, proportional allocation"), portfolio items are scaled proportionally to the financial accounts aggregates, after the Pareto adjustment.

The final sets of rows show the results obtained when combining the Pareto-calibration method with the multivariate calibration approach, either in a single iteration ("P-C, S.I.," "Par-cal, single-iteration approach") or through an iterative process ("P-C, SIM.," "simultaneous approach").

As expected, all the results point to the fact that inequality is underestimated in survey data. The increase in the estimated degree of inequality is proportional to the severity of both non-response among wealthy households and reporting error problems. For instance, the Gini index increases by 10 points in Italy when considering the simultaneous approach. For the other countries, the increase ranges from two points in Germany to five points in France.

A second result is that the Pareto non-response adjustment has a larger influence in determining the final inequality statistics compared to the reporting error adjustments. This should be expected. The magnitude of the increase depends on the severity of the non-response problem. For surveys, such as the Italian and German ones, in which truncation bias is particularly pronounced, the sole inclusion of these unobserved households increases the proportion of wealth held by the top 1 percent households by at least 10 percentage points, respectively. This increase is much less pronounced for the French and Finnish surveys, where the truncation is more modest, thanks to oversampling.

After the Pareto-calibration step, the change in distributional estimates from the base survey can still be significant even for countries featuring oversampling.

The top 1 percent shares increase by around 10 percent points for France and Finland, and around 20 for Germany and Italy. These results are perfectly in line with our expectations and result from the recalculation of survey weights in the Pareto tail. Even when administrative records are used for oversampling, only the section of the wealth distribution that is oversampled will not suffer from non-response, even in the best-case scenario. The change in estimates from the Survey & Missing tail step then suggests that differential non-response can still affect a sizeable section of the Pareto tail. As discussed earlier, performing this non-response adjustment is often sufficient to completely fill the gap with the National Accounts, as is the case for Finland and Germany.

Another important result is that in some cases, the simultaneous approach produces different results compared to the single-iteration approach. Therefore, we do not find a strong support for the assumption that the relative error converges in probability to a constant, as discussed in Section 3.5.

This is the case for the Italian survey, in which the single-iteration approach leads to inflated and highly variable results. As discussed, we believe this also to be connected to the choice of a Pareto threshold. Results obtained under a lower threshold (provided in Table 5 in the Online Appendix) are closer to the ones produced by the simultaneous approach (at least for some indicators such as the Gini index).

Finally, as to variance estimation, the adjustment methods generally produced a decrease in the reliability of the results in the highest percentiles. This was expected as the sample size grows smaller as we move upwards in the distribution of wealth. What matters is that the final coefficients of variation are not very different from those based on the base survey data, and often estimates get more precise at the higher wealth percentile groups. The length and quality of the rich list will also affect the variability of estimates: note that estimates for the Italian survey (whose adjustment is using a shorter, and lower quality, rich list) feature a much higher standard deviation in their higher percentiles than in the other surveys. Compared to other methods, the simultaneous approach produced the least variability in the wealth estimates.

It is worth stressing that our results show a higher level of wealth inequality compared to those obtained by Vermeulen (2018) but are broadly comparable to the estimates for financial assets shown in Vermeulen (2016). The increase in top inequality figures, including in the Pareto-calibration step, was expected and welcomed. As we discussed earlier, the reassessment of the number of households suggests that the portion of wealth held in the Pareto tail is also larger than previously estimated. Our results suggest that, even in the presence of oversampling and linked survey-registry data, the number of households in the tail can be underestimated, and this can lead to the underestimation of top fortunes under conventional Pareto adjustment procedures.

It is important to recall that all adjustments for reporting error are performed on top of the Pareto-calibration method, and have to rely on specific assumptions about under-reporting behavior. From this point of view, the question on which method is the most appropriate will have a different answer on a case-by-case basis. Is there a reason to believe that reporting error increases in relative terms along the wealth distribution? For countries like Finland, this assumption is clearly incorrect because of the linkage with administrative data. In cases like these, estimates of the simultaneous approach can be expected to be incorrect, while proportional allocation would be more appropriate. In all cases, however, we believe that the Pareto-calibration approach constitutes a valid baseline after which adjustments for reporting error can be applied.

At least, with regard to the non-response Pareto-calibration adjustment, we can compare our results with other studies which have dealt with reconstructing the wealth distribution using other methodologies and alternative data sources. Garbinti *et al.* (2020) use tax data to reconstruct wealth distribution series in France. Estimates for the wealth held by the top 1 percent and the top 10 percent in 2012 have been estimated at around 25 percent and 57 percent, respectively. These estimates are quite close to ours. At the Pareto-calibration step (Table 4), we estimated 29 percent for the top 1 percent and 57 percent for the top 10 percent. These results are valuable, as they suggest that, even with limited data sources, our Pareto-calibration method produces estimates similar to studies based on better data sources, with tolerable (and expected) deviations only for the top 1 percent estimates, which are likely attributable to quality differences between the rich list and the survey data. Furthermore, using the mean excess method for the detection of the Pareto threshold (see Online Appendix A, Table 6), we found even closer estimates (27.4 percent and 55.6 percent) at the Pareto-calibrationstep.

Evidence from other countries is spottier, but some information from Italy can be inferred from Acciari and Morelli (2020), who reconstructed the distribution of inheritances in Italy using inheritance tax data. From their estimation, in 2012, wealth held by decedent top 1 percent and top 10 percent wealthiest households in the population amounted to around 20 percent and 53 percent of the total share. While deceased households are only a small sample of the general population and estimates are biased by the long-running linkage between wealth and life expectancy, it is valuable to note that these estimates are not very far from ours. In particular, in our results under the mean excess method for the detection of the Pareto threshold (Online Appendix A, Table 6), top shares for these wealth groups were estimated at 26 percent and 54 percent.

Full results with the thresholds estimated with the mean excess function method are discussed in Online Appendix A and shown in Table 6.

6. Conclusions

In this paper, we have shown how a combination of well-established methodologies for the fitting of a Pareto distribution and the calibration of survey data can be used to correct for non-response and misreporting when only limited external information is available.

Our main conclusions may be summarized as follows. The consistency between micro and macro data is essential if we are to obtain a more reliable picture of household wealth distribution. Inequality estimates from survey data understate the population parameters, depending on the severity of both non-response and reporting error.

We found that the proposed adjustment for non-response is appropriate even when the survey makes use of oversampling and/or linked survey-registry data. In some cases, such as Finland or Germany, this adjustment alone is sufficient to fill the gap with the national accounts. Adjustments for reporting error, if needed, can then be applied on top of these non-response adjustments, depending on the severity of the reporting error and assumptions about its nature. After assessing the precision of our results, we found that reiterating the non-response adjustment with a multivariate reporting adjustment can allow for differential reporting error while also producing the lowest increase in variability.

Finally, as a by-product of our methodology, we obtained an adjusted micro-data set which allows to compute many distributional indicators in addition to the ones shown in this paper.

Additional work is still needed for the refinement of the methodology we propose. For example, the estimation of the number of wealthy households could be further validated and improved using alternatives to rich lists (such as tax records) or by applying additional methods (such as the Type II Pareto or the Estate Multiplier Method). In addition, the correction of reporting error could be further improved by enriching the auxiliary granular vector with more granular external information (such as administrative records). The development of diagnostic tools to help the user with the choice of the tuning parameters involved in the procedure is also on the checklist.

It is worth stressing that our method could easily be adapted in case additional external (possibly aggregate) information is available (other than national accounts). Indeed, both the non-response and the reporting error adjustments could be enhanced with external information and be run separately when needed.

References

Acciari, P. and S. Morelli, "Wealth transfers and net wealth at death: Evidence from the Italian inheritance tax records 1995–2016," *Measuring Distribution and Mobility of Income and Wealth*, University of Chicago Press, Cambridge, MA, 175–204, 2020.

Alvaredo, F. and E. Saez, "Income and wealth concentration in Spain from a historical and fiscal perspective," *Journal of the European Economic Association*, 7(5), 1140–67, 2009.

Ampudia, M., H. van Vlokhoven and D. Żochowski, "Financial fragility of Euro Area households," *Journal of Financial Stability*, 27, 250–62, 2016.

Arrondel, L., P. Lamarche and F. Savignac, "Does inequality matter for the consumption-wealth channel? Empirical evidence," *European Economic Review*, 111, 139–65, 2019.

Atkinson, A. B., "Pareto and the upper tail of the income distribution in the UK: 1799 to the present," *Economica*, 84(334), 129–56, 2016.

Bach, S., A. Thiemann and A. Zucco, "Looking for the missing rich: Tracing the top tail of the wealth distribution," *International Tax and Public Finance*, 26(6), 1234–58, 2019.

Blanchet, T., L. Chancel and A. Gethin, "Why is Europe more equal than the United States?", *American Economic Journal: Applied Economics*, 14, 480–518, 2022a.

Blanchet, T., I. Flores and M. Morgan, "The weight of the rich: Improving surveys using tax data," *The Journal of Economic Inequality*, 20(1), 119–50, 2022b.

Blanchet, T., J. Fournier and T. Piketty, "Generalized Pareto curves: Theory and applications," *Review of Income and Wealth*, 68(1), 263–88, 2021.

Bourguignon, F., "Simple adjustments of observed distributions for missing income and missing people," *The Journal of Economic Inequality*, 16(2), 171–88, 2018.

Carranza, R., M. Morgan and B. Nolan, "Top income adjustments and inequality: An investigation of the EU-silc," *Review of Income and Wealth*, Advance online publication, 2022.

Casiraghi, M., E. Gaiotti, L. Rodano and A. Secchi, "A "reverse Robin Hood"? The distributional implications of non-standard monetary policy for Italian households," *Journal of International Money and Finance*, 85, 215–35, 2018.
Chakraborty, R., I. Kavonius, S. Perez-Duarte and P. Vermeulen, "Is the top tail of the wealth distribution the missing link between the Household Finance and Consumption Survey and National Accounts?" *Journal of Official Statistics*, 35, 31–65, 2019.
Chambers, R. L. and Ren, R. (2004). Outlier robust imputation of survey data. The Proceedings of the American Statistical Association, Boston, MA, 3336–3344.
Christelis, D., D. Georgarakos, T. Jappelli, L. Pistaferri and M. van Rooij, "Heterogeneous wealth effects," *European Economic Review*, 137, 103805, 2021.
Coibion, O., Y. Gorodnichenko, L. Kueng and J. Silvia, "Innocent bystanders? Monetary policy and inequality," *Journal of Monetary Economics*, 88, 70–89, 2017.
Colciago, A., A. Samarina and J. de Haan, "Central bank policies and income and wealth inequality: A survey," *Journal of Economic Surveys*, 33(4), 1199–231, 2019.
D'Alessio, G. and Neri, A. (2015). Income and wealth sample estimates consistent with macro aggregates: Some experiments. Questioni di Economia e Finanza (Occasional Papers) 272, Bank of Italy, Economic Research and International Relations Area, Rome, Italy.
Deville, J.-C. and C.-E. Särndal, "Calibration estimators in survey sampling," *Journal of the American Statistical Association*, 87(418), 376–82, 1992.
Eurosystem Household Finance and Consumption Network. Survey data on household finance and consumption research: Summary and policy use. ECB Occasional Paper Series No 100, European Central Bank, Frankfurt, Germany, 2009.
Eurosystem Household Finance and Consumption Network. The household finance and consumption survey: Methodological report for the 2017 wave. ECB Statistics Paper Series No 35, European Central Bank, Frankfurt, Germany, 2020.
Expert Group on Linking Macro and Micro data. Understanding household wealth: Linking macro and micro data to produce distributional financial accounts. ECB Statistics Paper Series No 37, European Central Bank, Frankfurt, Germany, 2020.
Fesseau, M. and Mattonetti, M. L. Distributional measures across household groups in a National Accounts framework. OECD Statistics Working Papers No. 2013/08, Organisation for Economic Co-Operation and Development (OECD), Paris, France, 2013.
Gabaix, X. and R. Ibragimov, "Rank-1 / 2: A simple way to improve the OLS estimation of tail exponents," *Journal of Business & Economic Statistics*, 29(1), 24–39, 2011.
Garbinti, B., J. Goupille-Lebret and T. Piketty, "Income inequality in France, 1900–2014: Evidence from Distributional National Accounts (DINA)," *Journal of Public Economics*, 162, 63–77, 2018.
Garbinti, B., J. Goupille-Lebret and T. Piketty, "Accounting for wealth-inequality dynamics: Methods, estimates, and simulations for France," *Journal of the European Economic Association*, 19(1), 620–63, 2020.
Guiso, L., M. Paiella and I. Visco, Do capital gains affect consumption? Estimates of wealth effects from Italian households' behavior, Essays in Memory of Albert Ando, Long-run Growth and Short-run Stabilization, 2005.
Haziza, D. and J. F. Beaumont, "Construction of weights in surveys: A review," *Statistical Science*, 32(2), 206–26, 2017.
Jenkins, S. P., "Pareto models, top incomes and recent trends in UK income inequality," *Economica*, 84(334), 261–89, 2017.
Jenkins, S. P., "Top-income adjustments and official statistics on income distribution: The case of the UK," *The Journal of Economic Inequality*, 20(1), 151–68, 2022.
Kennickell, A., "The role of over-sampling of the wealthy in the Survey of Consumer Finances," *Irving Fisher Committee Bulletin*, 28, 403-408, 2008.
Kennickell, A., "The tail that wags: Differences in effective right tail coverage and estimates of wealth inequality," *The Journal of Economic Inequality*, 17, 443–59, 2019.
Little, R. J. and S. Vartivarian, "Does weighting for nonresponse increase the variance of survey means?," *Survey Methodology*, 2005.
Michelangeli, V. and C. Rampazzi, Indicators of financial vulnerability: A household level study. Questioni di Economia e Finanza (Occasional Papers) 369, Bank of Italy, Economic Research and International Relations Area, Rome, Italy, 2016.
Mumtaz, H. and A. Theophilopoulou, "Monetary policy and wealth inequality over the great recession in the UK," *An Empirical Analysis. European Economic Review*, 130, 103598, 2020.
Paiella, M., "Does wealth affect consumption? Evidence for Italy," *Journal of Macroeconomics*, 29(1), 189–205, 2007.
Ranalli, M. G., A. Matei and A. Neri, "Generalised calibration with latent variables for the treatment of unit nonresponse in sample surveys," *Statistical Methods & Applications*, 32, 169–195, 2023.

Ranalli, M. G. and A. Neri, "To misreport or not to report?, The case of the Italian survey on household income and wealth," *Statistics in Transition New Series*, 12(2), 281–300, 2011.
Saez, E. and G. Zucman, "Wealth inequality in the United States since 1913: Evidence from capitalized income tax data," *The Quarterly Journal of Economics*, 131, 519–78, 2016.
Särndal, C.-E., "The calibration approach in survey theory and practice," *Survey Methodology*, 33(2), 99–199, 2007.
Särndal, C.-E. and S. Lundström, *Estimation in Surveys with Nonresponse*, John Wiley & Sons, Ltd, Hoboken, NJ, 2005.
Schröder, C., C. Bartels, M. M. Grabka, J. König, M. Kroh and R. Siegers, "A novel sampling strategy for surveying high net-worth individuals—a pretest application using the Socio-Economic Panel," *Review of Income and Wealth*, 66(4), 825–49, 2019.
Shorrocks, A., J. Davies and R. Lluberas, *The global wealth report 2021*, Technical Report 11, Credit Suisse Research Institute, Zürich, Switzerland, 2021.
Silverman, B. W., *Density estimation for statistics and data analysis*, Routledge, Oxfordshire, England, UK, 2018.
Vermeulen, P., "Estimating the top tail of the wealth distribution," *American Economic Review*, 106(5), 646–50, 2016.
Vermeulen, P., "How fat is the top tail of the wealth distribution?" *Review of Income and Wealth*, 64(2), 357–87, 2018.
Waltl, S. R., "Wealth inequality: A hybrid approach toward multidimensional distributional National Accounts in Europe," *Review of Income and Wealth*, 68(1), 74–108, 2021.
Waltl, S. R. and R. Chakraborty, "Missing the wealthy in the HFCS: Micro problems with macro implications," *The Journal of Economic Inequality*, 20(1), 169–203, 2022.

Supporting Information

Additional supporting information may be found in the online version of this article at the publisher's web site:

Appendix S1: Supporting Information.