Looking for the missing rich: tracing the top tail of the wealth distribution

Stefan Bach[1,2] · **Andreas Thiemann**[3] · **Aline Zucco**[1]

Abstract

We analyse the top tail of the wealth distribution in France, Germany, and Spain using the first and second waves of the Household Finance and Consumption Survey (HFCS). Since top wealth is likely to be under-represented in household surveys, we integrate big fortunes from rich lists, estimate a Pareto distribution, and impute the missing rich. In addition to the *Forbes* list, we rely on national rich lists since they represent a broader base of the big fortunes in those countries. As a result, the top 1% wealth share increases notably for the three selected countries after imputing the top wealth. We find that national rich lists can improve the estimation of the Pareto coefficient in particular when the list of national USD billionaires is short.

Keywords Wealth distribution · Missing rich · Pareto distribution · HFCS

JEL Classification D31 · C46 · C81

1 Introduction

Rising inequality in income and wealth is gaining increased attention in both public and academic debate. The widespread discussion triggered by Piketty's (2014) book, *Capital in the Twenty-First Century*, focuses on concentration at the top and

---

✉ Andreas Thiemann
   andreas.thiemann@ec.europa.eu

   Stefan Bach
   sbach@diw.de

   Aline Zucco
   azucco@diw.de

[1] German Institute for Economic Research (DIW Berlin), Berlin, Germany

[2] University of Potsdam, Potsdam, Germany

[3] European Commission, Joint Research Centre (JRC), Seville, Spain

the underlying trends in modern capitalism. Economists and policy makers alike are aware of increasing heterogeneity in income and wealth, along with its consequences for financial stability, savings and investment, employment, growth, and social cohesion. Against the backdrop of tax systems being less progressive over the last decades and rising public debt following the 2008 financial crisis, some politicians are considering higher taxes on capital income and top wealth (Förster et al. 2014; Saez and Zucman 2019). However, the lack of precise information about wealth concentration at the top of the distribution makes it difficult to assess the economic impact of such tax reforms.

In this paper, we rely on an approach suggested by Vermeulen (2018), which combines household survey data with rich lists to jointly estimate a Pareto distribution for the top tail to adjust the top wealth distribution in France, Germany, and Spain. In particular, our main contribution is to investigate how the use of national rich lists instead of the *Forbes* list improves top tail estimations. We derive an adjusted wealth distribution to better account for the wealth at the very top, which wealth surveys do usually not cover properly. Hence, we also provide new estimates of top shares of household wealth for France, Germany, and Spain.

We base the analysis on the Eurosystem's Household Finance and Consumption Survey (HFCS) comparisons (European Central Bank 2013), a household survey conducted in most Eurozone countries. The comprehensive information on wealth distributions facilitates the international comparisons. For instance, the data reveal that Germany has one of the most unequal wealth distributions in Europe.

However, with respect to the top wealth distribution, household surveys have inherent, crucial drawbacks: non-response and under-reporting (Vermeulen 2016a, 2018). Generally, personal wealth is considerably more concentrated than income and it is difficult to capture the top wealth distribution through small-scale voluntary surveys. The potential non-observation bias, i.e. the lack of reliability due to small samples, can be partly reduced by oversampling rich households. Moreover, non-response bias is probable as response rates tend to decrease with high income and wealth, especially at the top (Vermeulen 2018). The bias because of under-reporting is striking when comparing survey data with national accounts (Vermeulen 2016a; Chakraborty et al. 2018; Chakraborty and Waltl 2018).

One viable solution to better capture the missing rich is estimating the top wealth concentration by relying on functional form assumptions on the shape of the top tail distribution. Traditionally, these estimations assume a Pareto distribution, which approximates well the top tail of income and wealth (Davies and Shorrocks 2000), in some cases the models rely on more complex functional forms (Clauset et al. 2009; Burkhauser et al. 2012; Brzezinski 2014). However, the problem of biased wealth concentration remains if wealthy households are substantially under-represented in survey data.

Administrative tax data generally cover also wealthy taxpayers. Yet, few countries levy a recurrent wealth tax. Estate tax records are used to infer top wealth by mortality multipliers (Kopczuk and Saez 2004; Alvaredo et al. 2016) for which researchers must deal with differential mortality. The capitalization of capital income tax records (Saez and Zucman 2016) raises complex issues to assess proper discount rates, in particular with respect to risk premia. In general, tax record data could be heavily flawed by

explicit tax privileges, tax avoidance, tax evasion, favorable valuation procedures, and outdated valuations. Alstadsæter et al. (2019) show for three Scandinavian countries that offshore tax evasion is highly concentrated among the rich, relying on a unique data set of administrative wealth data matched to leaked data from offshore financial institutions.

A potential way of dealing with the "missing rich" is integrating different data sources. A promising approach combines household survey data on wealth with rich lists of big fortunes to jointly estimate a Pareto distribution for the top tail of wealth (see e.g. Davies 1993 for Canada, Bach et al. 2014 for Germany, and Eckerstorfer et al. 2016 for Austria).

Vermeulen (2018) augments the US Survey of Consumer Finances (SCF), the UK Wealth and Assets Survey and the HFCS data with the *Forbes* list in order to show the potential under-representation of top wealth in the survey data for the USA, the UK, and several Eurozone countries. He finds that differential non-response seems to be rather high in some Eurozone countries, particularly Germany. This leads to an underestimation of the top wealth shares when the estimation is exclusively based on survey data without extreme tail observations.

In this paper, while relying on Vermeulen (2018), we extend it along two dimensions. First, we investigate how replacing the *Forbes* list by more detailed country-specific wealth rankings affects the top tail estimation. In doing so, we experiment with different lengths and specifications of national lists of the richest persons or families, provided by the media. In contrast to the *Forbes* list, national rich lists typically cover a larger share of the very wealthy. We find that national rich lists can provide a real value added, particularly in countries where only few dollar billionaires make it onto the *Forbes* list, such as Spain. Second, we compare the results based on the first and second waves (separate surveys) of the HFCS to draw conclusions regarding wealth dynamics within a country.

Based on our preferred specification, relying on national rich lists, we find the following: For Germany, the top wealth imputation leads to an increase of the top 1% wealth share from 24 to 34% in the first wave and from 24 to 35% in the second wave. For France (first wave) and Spain, we find smaller effects of the wealth imputation since rich households are better represented in the survey data. The Spanish top 1% wealth share increases by 8 (6) percentage points to 23% (22%) in the first (second) wave of the HFCS. In France, the top 1% wealth share increases from 18 to 25% in the first wave. In the second wave, however, the top 1% owns 31% of total wealth after the top wealth imputation, which is 12 percentage points more than in the original HFCS.

The remainder of the paper proceeds as follows: Sect. 2 describes the data, and Sect. 3 explains the estimation and imputation of top wealth. Section 4 discusses the results of the top wealth imputation on the wealth distribution, while the last Sect. 5 concludes.

2 Data

In this paper, we use different types of data: household survey data, namely the HFCS, national rich lists, and the *Forbes* list.

Table 1 Response behaviour and oversampling in the first and second waves of the HFCS. *Source:* European Central Bank (2013, 2016a)

Countries	Gross sample size	Net sample size	Response rate (%)	Oversampling rate[a]
First wave				
France	24,289	15,006	69	129
Germany	20,501	3565	19	117
Spain	11,782	6197	57	192
Second wave				
France	20,272	12,035	65	132
Germany[b]	16,221	4461	29	141
Spain[b]	13,442	6106	48	234

[a] Effective oversampling rate of the top 10%, in percent: $\frac{(S90-0.1)}{0.1}$, where S90 is the share of sample households in the wealthiest 10%

[b] Countries with panel component. Response rate including panel

2.1 HFCS

The HFCS is a decentralized household survey focusing on Eurozone countries, conducted by national central banks or statistical offices. The HFCS aims at collecting information about income, consumption, and assets of households. We use the first and second waves, which were collected between 2008 and 2011 (European Central Bank 2013, p. 8) and between 2011 and 2015 (European Central Bank 2016a, p. 4), respectively. While the HFCS intends to oversample wealthy households to address potential non-observation bias, the selection criteria applied in the oversampling process differ across countries (European Central Bank 2013, p. 9).

Table 1 shows that the response behaviour varies substantially across countries and waves. The effective oversampling rate describes to what extent the ratio of the top 10% is oversampled compared to its share of the population (European Central Bank 2013, p. 36). To address item non-response, i.e. participants who refuse or are unable to answer certain questions, the HFCS applies a multiple imputation approach (European Central Bank 2013, p. 39). Throughout the paper, we base our results on all five implicates.[1]

Although the HFCS is compiled in a harmonized way, it still relies on decentralized country surveys, making cross-country comparisons difficult. Further, the response rate varies across countries and waves. For instance, in Spain, 57% of contacted households participated in the first wave while this number dropped to 48% in the second wave. In contrast to Germany and Spain, French households are obliged to participate if sampled (European Central Bank 2013, p. 41). Furthermore, Germany and Spain exclude the homeless and institutionalized populations, while France only excludes the latter (European Central Bank 2013, p. 33). That is, including or excluding homeless persons, thus the poorest part of the population, can have an important impact on the

---

[1] Implicates are the set of imputed values for each missing value. The distance between the values of the five implicates in the HFCS reflects the inherent imputation uncertainty (European Central Bank 2016a). In the second wave, the French data do not contain implicates but one single data set.

wealth concentration. For our purpose, however, differential success in oversampling of the top 10% is the major challenge.

The oversampling of the rich in Germany uses geographical information on high-income municipalities, whereas in France and Spain it relies on net wealth information from fiscal sources. Moreover, the timing and duration of fieldwork differs notably between these three countries.[2]

The HFCS collects information on households' assets and liabilities in detail. Net wealth is measured as the sum of real estate properties, business properties, financial assets, corporate shares, and the main household assets, including cars, deducting liabilities. Household net wealth does not include claims to social security or occupational and private pensions, or health care plans. It is based on self-assessed property valuations of the survey respondents. There is no evidence that respondent self-assessment leads to a systematic bias. Critics argue that it is particularly difficult to measure liabilities accurately in surveys. We address this argument by re-doing the top tail adjustment using a gross wealth concept. The results (provided in the Online Appendix), however, are relatively similar to those that rely on net wealth.

2.2 Rich lists

Since the 1980s, business media and researchers have provided lists that rank the large fortunes held by the super-rich. We use the *World's billionaires list* of Forbes (2014) and national lists of the richest persons or families of the selected countries. We select the annual rich lists such that we match the fieldwork period in the two waves of the HFCS.

The reliability of these lists is contentious since the data are not consistently collected, instead using a variety of sources, e.g. public registers, financial markets, business media, and interviews of wealthy individuals themselves. In addition, the rich lists are not compiled relying on one single approach, but combine various methods. The completeness of these lists is questionable, especially with regard to smaller fortunes, which are often dominated by non-quoted corporate shares, which makes it more difficult to assess their precise value. Hence, the precision of wealth estimates seems to be inversely correlated with the rank, which shows the heaping of rich list entries at round numbers, e.g. at 300, 400 or 500 million euros. Further, the editor removed some individuals from the German *manager magazin* list because of privacy claims, which adds selectivity.

Furthermore, a rich list may over- or underestimate wealth concentration. On the one hand, it is probable that the number of households, which are listed as one family in the rich list, is underestimated, which leads to an overestimation of wealth concentration. Moreover, in many cases, rich lists report wealth for entire entrepreneurial

---

[2] In the first wave, Spanish households were interviewed between November 2008 and July 2009, and in France between October 2009 and February 2010. In Germany, the fieldwork period was from September 2010 to July 2011. These temporal differences persist throughout the second wave of the HFCS: While the survey was then conducted between October 2011 and April 2012 in Spain, the interviews of German and French households were about two years later (in Germany between April and November of 2014 and in France between October 2014 and February 2015) (Tiefensee and Grabka 2016; European Central Bank 2016a).

"families" that may consist of several households. In particular, many successful "*German Mittelstand*" firms, if not major enterprises, on the German *manager magazin* list have been family-owned for generations. Thus, using publicly available information on the number of shareholders of the respective family-owned firms, we correct the German national list (see below). Moreover, we remove households that are obviously residing abroad. This phenomenon is also likely to be present in the French and Spanish national rich lists, thus leading to an overestimation of the top wealth concentration. However, as we do not have the necessary information for this adjustment, we neglect this issue in the French and Spanish rich lists. Table 6 in the Appendix tests alternative specifications of national rich lists. The sensitivity test shows that alternative specifications would generally lead to similar results.

On the other hand, rich lists presumably ignore private assets and liabilities, which may result in the underestimation of wealth concentration, as top wealth households have typically real estate and financial portfolios. In some cases, however, corporate investments might be leveraged by private debt, even though this could have unfavourable tax consequences in the countries analysed in this paper. When comparing estate tax files and the *Forbes* list, US Internal Revenue Service (IRS) researchers find that the list overestimates net worth by approximately 50% (Raub et al. 2010). The main reasons for this inconsistency are valuation difficulties and tax exemptions as well as family relations (individuals vs. couples) and other structural differences. The German *manager magazin* includes valuables and real estate, while the Spanish *El Mundo* list does not. These methodological differences might influence the results in the respective countries and should be kept in mind when comparing results across countries. Despite these caveats, rich lists are a rare source that offers information on the super-rich and their wealth.

manager magazin (Germany)

Each year, the *manager magazin* publishes a wealth ranking of the richest persons or families in Germany. From 2000 to 2009, the magazine ranked the 300 wealthiest Germans (and their wealth); since 2010, the 500 richest. The incompleteness and selectivity of the rich list tends to increase when ranks are lower since there is scarce information for households holding non-quoted firms or other assets. Therefore, we only use the top 200 from the German list. As mentioned above, in many cases a family comprises more than one household. We use publically available information on the number of shareholders to correct the respective observations: First, we split the wealth of rich list observations into several households according to the anecdotal evidence, partly provided by the editor of the list. Then, we re-rank all observations of this new list and consider the new top 200 as the corrected rich list.

However, measurement errors might clearly remain since there is often sparse information on the ownership structure in the financial accounts and other company disclosures. Generally, German "*Mittelstand*" entrepreneurs are rather reluctant to provide information on their financial affairs and anxious to keep capital markets and external investors out of their firms. For lower-ranked families, we generally assume four households per family. We also remove obvious non-resident households from the list (Table 2).

Table 2 Summary statistics of the national rich lists in France, Germany, and Spain. *Source:* Manager magazin (2011, 2014), Challenges (2010, 2015) and El Mundo (2009, 2012) and Forbes (2009, 2010, 2011, 2012, 2014, 2015); own calculations

Country	Rich list	N	Mean in billion euros	SD in billion euros	Min in billion euros	Max in billion euros
First wave						
Germany	Manager Magazin 200 (households-adjusted)	200	1.36	1.85	0.50	17
	Manager Magazin 200 (original families)	200	1.91	2.29	0.55	17
	Forbes (2011)	52	3.53	3.47	0.82	19
France	Challenges 200	200	1.08	2.60	0.16	23
	Forbes (2010)	11	5.47	6.35	0.81	20
Spain	El Mundo	74	1.49	2.06	0.50	16
	Forbes (2009)	12	2.36	3.77	0.78	14
Second wave						
Germany	Manager Magazin 200 (households-adjusted)	200	1.78	2.18	0.60	15
	Manager Magazin 200 (original families)	200	2.46	3.45	0.70	31
	Forbes (2014)	85	3.47	3.55	0.74	18
France	Challenges 200	200	1.97	4.28	0.35	31
	Forbes (2015)	43	5.10	7.40	0.97	35
Spain	El Mundo	117	1.11	3.66	0.19	39
	Forbes (2012)	16	3.36	6.77	0.76	29

The corrected manager magazin rich list adjusts the entries by the likely number of households per entry

Challenges (France)

Since 1996, *Challenges* magazine annually publishes a ranking of the 500 richest households in France. A team of journalists constructs and updates a large database that enables them to estimate the wealth of these households. It relies on various sources of information: Public data on share ownership and accounts, investigations on the ownership structure of unlisted companies, professional publications, seminars, award ceremonies, and surveys that are sent to rich households directly (Treguier 2012). Similar to the German case, we only use the 200 richest entries from the French list to avoid the risk of lower precision at lower ranks.

El Mundo (Spain)

The Spanish national rich list has been compiled since 2006 by the Spanish newspaper *El Mundo*, and it provides two separate wealth rankings of the wealthiest families or individuals. While the first list of the top 50 (top 100 in 2012) "visible fortunes" relies on public information about the ownership structure of listed companies on the stock market, the list of the top 50 (top 100 in 2012) "estimated fortunes" reflects the

estimated share value of unlisted companies. The estimation uses information about purchase-sales of shares, venture capital investments, and direct estimates of fortunes. The joint list for 2009, used in this paper, consists of the top 50 visible fortunes and the top 27 estimated fortunes, where the last entry from the latter list reports the same net wealth as the 50th person from the first list. For the second wave, we use the joint list of 2012, compiled in the same way. It contains 100 visible fortunes and 17 estimated fortunes. Hence, the final list represents the 74 and the 117 richest Spanish individuals (El Mundo 2009, 2012) in the first and second waves, respectively.

Forbes (Global)

Similar to the national rich lists, *Forbes* journalists compile international information on big fortunes in creating the wealth ranking (Forbes 2014). To get a spot among the wealthy on the *Forbes* list, personal net wealth has to exceed one billion US dollars. Compared to the national lists, the *Forbes* list might be more reliable as it focuses on the super-rich, for whom information is easier to collect. Moreover, many billionaires cooperate with the editors. However, the list may remain incomplete and selective in comparison with national lists. We match the respective *Forbes* billionaire lists with the latest year of the survey: We use the *Forbes* list 2011 and 2014 for Germany, 2010 and 2015 for France, and 2009 and 2012 for Spain. For our analysis, we recalculate the wealth in euros.[3]

3 Methodology of estimation and imputation of the top wealth distribution

In our paper, we investigate how the use of different rich lists affects the estimation of the Pareto coefficients for the top tail wealth distribution. Further, we adjust the wealth distributions for France, Germany, and Spain, replacing the survey top tail by the estimated one. In this section, first, we briefly sketch the theoretical background of our approach. Secondly, we estimate the Pareto coefficients for each country, investigating how replacing the *Forbes* list by a national rich list affects the results. We make use of these newly estimated Pareto coefficients and replace the top tail of the survey-based wealth distribution with synthetic households that follow the Pareto distribution.[4]

3.1 Theoretical background

For any level of net wealth w, at least as high as $w_{\min}$, the Pareto density function is defined as

$$f(w) = \frac{\alpha w_{\min}^{\alpha}}{w^{\alpha+1}}. \tag{1}$$

---

[3] The exchange rate (1 EUR in USD) corresponds to the date of the "snapshot" of the *Forbes* Billionaires Lists: 1.2823 USD (13/02/09, ES), 1.3572 (12/02/10, FR) and 1.344 (14/02/11, DE) for the first wave and 1.3169 (14/02/12, ES), 1.3573 (12/02/14, DE) and 1.13821 (13/02/15, FR) for the second wave.

[4] We refer interested readers to Dalitz (2016), Vermeulen (2018), Cowell (2011), Gabaix (2009), Gabaix and Ibragimov (2012), Clauset et al. (2009), Kleiber and Kotz (2003), Chakraborty and Waltl (2018).

The corresponding complementary cumulative distribution function is defined on the interval $[w_{\min}, \infty)$ and for $\alpha > 0$ as

$$P(W > w) = \left(\frac{w_{\min}}{w}\right)^{\alpha}. \tag{2}$$

The coefficient α, also called tail index, determines the "fatness" of the top tail. The lower α, the fatter the tail and the more concentrated is the wealth distribution.

Empirically, a finite sample approximately follows a Pareto distribution if

$$\frac{i}{n} \cong \left(\frac{w_{\min}}{w_i}\right)^{\alpha}; \; \forall w_i. \tag{3}$$

w_i denotes wealth of household i. Note that households are ordered decreasingly by their wealth. Therefore, i indicates also the rank of a household, which implies that w_1 is net wealth of the richest household, w_2 of the second richest, and so forth. The poorest household with $w_n > w_{\min}$ has the lowest rank, n, which is equal to the sum of top tail households. After re-arranging, we obtain

$$ln(i) \cong C - \alpha ln(w_i) \tag{4}$$

with $C = ln(n) + \alpha ln(w_{\min})$.

Vermeulen (2018) describes how this rank-wealth-relationship can be adapted to survey weights. Furthermore, Gabaix and Ibragimov (2012) show that the log-log-rank-size estimates are biased in finite samples and suggest to shift the rank by 0.5. We obtain the following relationship

$$ln\left(\left(i - \frac{1}{2}\right)\frac{\overline{N_{fi}}}{\bar{N}}\right) = C^* - \alpha ln(w_i) \tag{5}$$

The richest household $i = 1$ has a survey weight of N_1, the second richest household a weight of N_2, and so forth. Further, $\bar{N} = \frac{\sum_{j=1}^{n} N_j}{n} = \frac{N}{n}$ reports the average survey weight of all observations, N being the sum of survey weights in the top tail, $\overline{N_{fi}} = \frac{\sum_{j=1}^{i} N_j}{i}$, the average weight of the first i observations, and $C^* = ln(\frac{\bar{N}}{N}) + \alpha ln(w_{\min})$.[5] Throughout the paper, we estimate the Pareto α according to Eq. 5 using ordinary least squares (OLS).[6] To estimate standard errors, we follow the HFCN (European Central Bank 2016b, p.58) making use of the 5 implicates and the first 200 replicate weights according to the Rubin's rule.

---

[5] Vermeulen (2018, Online Appendix, p. 17) shows how to derive this formula.

[6] Vermeulen (2018) proposes also a pseudo-maximum likelihood (PML) estimator that also takes into account the complex survey weights of the HFCS. However, testing the performance of the OLS and PML estimators in the presence of differential non-response, he prefers the OLS estimator (Krenek and Schratzenstaller 2017).
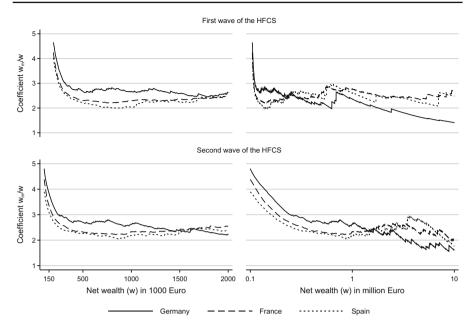
Fig. 1 Ratio mean wealth above w, divided by w, w_m/w. *Source:* First and second waves of the HFCS; own calculations

3.2 Estimation of the Pareto coefficient

The estimation of the Pareto coefficient α depends on $w_{\min}$ and the rich list that we add to the HFCS sample. To test how the rich list choice affects the result, we compare estimates based on different HFCS samples that are composed of the HFCS and combined with either a national rich list or the *Forbes World's billionaires list*. When setting $w_{\min}$, we face a trade-off as a high value leaves us with fewer observations to estimate α, whereas a low $w_{\min}$ risks that not all observations actually follow a Pareto distribution. To obtain a proper cut-off point, we exploit the distinctive property of the Pareto distribution: The average wealth, w_m, above any wealth threshold w is a constant multiple of that threshold, also known as "van der Wijk's law" (see Cowell (2011); Embrechts et al. (1997)). Based on the HFCS data, we plot $\frac{w_m}{w}$ for wealth thresholds above 100,000 euros, exemplary for the first implicate in Fig. 1, given in linear scale up to 2 million euros and in log scale up to 10 million euros.

The graphs suggest that household wealth above 500,000 euros, which is around the 90th percentile in the respective countries, is Pareto-distributed.[7] Therefore, we set the cut-off point of the Pareto distribution to 0.5 million euros. To choose the

---

[7] Eckerstorfer et al. (2016) propose a sophisticated method to derive the cut-off point above which wealth follows a Pareto distribution. They suggest identifying suitable parameter combinations of maximum likelihood estimates and goodness-of-fit tests. Dalitz (2016) and Krenek and Schratzenstaller (2017) use the Kolmogorov–Smirnov (K–S) criterion to identify the $w_{\min}$ that fits best to the empirical distribution. The K–S test compares alternative top tail distributions to the empirical one to determine the optimal lower bound. While it provides a quantitative decision criterion, the K–S test still relies on the empirical top tail distribution recorded in the HFCS.

optimal combination of w_{min} and rich list, we follow Vermeulen (2018) and test several minimum wealth thresholds: 0.5, 1, and 2 million euros. To assess also how α estimates differ by type and length of a rich list, we compare several combined samples of wealthy HFCS households and the top 100, 200 and 300 entries of each national rich list as well as the entire *Forbes* list. For Spain, we rely on the full rich list, which implies 74 and 117 observations in the first and second wave, respectively. Table 3 shows the estimated α coefficients for combinations of w_{min} and different subsamples for each country and wave.[8]

First, we replicate the corresponding α estimates of the first wave of Vermeulen (2018, Online Appendix, Table A1).[9] All these point estimates (except one) are identical with Vermeulen (2018) and most standard errors are also identical; for a few standard errors that deviate, differences are very small. In general, the results confirm that estimations combining survey data with rich lists generally lead to smaller α estimates and, therefore, higher wealth concentration. Relying purely on wealthy observations of the HFCS would underestimate wealth inequality.

Comparing estimates of samples that include rich list observations (national or *Forbes*) by w_{min} shows that they are rather sensitive in France and Spain, but not in Germany. Experimenting with the number of observations from the national rich list (top 100, 200, or 300) shows that α estimates are very similar in Germany, but also in France. For instance, α varies between 1.33 and 1.35 when relying on the French *Challenges* list, based on w_{min} of 1 million euros and the second wave.

Next, we analyse how estimates differ depending on whether we use a sample that includes observations from a national rich list or from the *Forbes* list. Replacing *Forbes* list entries with the ones from national rich lists reduces α estimates in France and Spain, indicating higher wealth concentration. In France, for instance, considering the first wave and w_{min} of 0.5 million euros, α decreases from 1.73 (*Forbes*) to 1.55 (*Challenges* top 200). In the second wave, however, *Forbes* and *Challenges* estimates are somewhat more similar, which potentially could be explained by the higher number of USD billionaires, who made it on the Forbes list in 2015 (43) compared to 2010 (11). The discrepancy between estimates based on the different rich lists is also visible in Spain. Taking the second wave as an example, α falls from 1.71 (*Forbes*) to 1.58 (*El Mundo*) when setting w_{min} to 0.5 million euros. In Germany, however, replacing the *manager magazin* list by the *Forbes* list hardly makes any difference for the α estimates, independent of wave or w_{min}. A potential explanation is the rather high number of USD billionaires who made it on the German *Forbes* list. These findings suggest that, in particular, when the *Forbes* list of national USD billionaires is short, national rich lists can improve the estimation of the Pareto coefficient, given that both lists are of similar quality. However, when the country-specific *Forbes* list is sufficiently large, the value added of replacing it with a national rich list is rather small. Table 6, in the Appendix, provides estimation results when we experiment with alternative rich list specifications. The results suggest that splitting rich list observations into several households has the largest effect.

---

[8] The Online Appendix illustrates the different samples graphically for France, Germany, and Spain in the first and second waves.

[9] Table 5 (in the Appendix) provides the top 1% shares that correspond to Table 3.

Table 3 Estimated α-coefficients for different subsamples. *Source*: HFCS (both waves), Manager magazin (2011, 2014); Challenges (2010, 2015); El Mundo (2009, 2012) and Forbes (2011, 2014, 2010, 2015, 2009, 2012); own calculations

Country, wave	Excluding rich list			Including rich list National rich list Top 100			Top 200			Top 300			Forbes		
	w_{min} (in euro)														
	≥ 2 M	≥ 1 M	≥ .5 M	≥ 2 M	≥ 1 M	≥ .5 M	≥ 2 M	≥ 1 M	≥ .5 M	≥ 2 M	≥ 1 M	≥ .5 M	≥ 2 M	≥ 1 M	≥ .5 M
Germany, 1st	1.87	1.64	1.54	1.39	1.40	1.41	1.38	1.39	1.40	1.38	1.40	1.40	1.38	1.39	1.40
	(0.39)	(0.23)	(0.14)	(0.04)	(0.02)	(0.01)	(0.04)	(0.02)	(0.01)	(0.04)	(0.02)	(0.02)	(0.04)	(0.02)	(0.01)
Germany, 2nd	1.81	1.59	1.50	1.35	1.37	1.38	1.36	1.37	1.38	1.36	1.38	1.39	1.34	1.36	1.37
	(0.35)	(0.20)	(0.12)	(0.03)	(0.02)	(0.01)	(0.04)	(0.02)	(0.01)	(0.04)	(0.02)	(0.01)	(0.03)	(0.02)	(0.01)
France, 1st	1.67	1.78	1.83	1.37	1.45	1.54	1.40	1.48	1.55	1.43	1.50	1.57	1.50	1.63	1.73
	(0.15)	(0.09)	(0.06)	(0.02)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.03)	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)
France, 2nd	1.51	1.63	1.73	1.26	1.33	1.44	1.28	1.34	1.43	1.30	1.35	1.44	1.32	1.41	1.53
	(0.19)	(0.13)	(0.08)	(0.02)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.02)	(0.02)	(0.03)
Spain, 1st	1.67	1.76	1.87	1.36	1.45	1.57	–	–	–	–	–	–	1.59	1.68	1.80
	(0.13)	(0.11)	(0.08)	(0.03)	(0.02)	(0.02)	–	–	–	–	–	–	(0.05)	(0.05)	(0.05)
Spain, 2nd	1.72	1.75	1.79	1.43	1.51	1.58	–	–	–	–	–	–	1.57	1.65	1.71
	(0.11)	(0.09)	(0.08)	(0.04)	(0.02)	(0.02)	–	–	–	–	–	–	(0.05)	(0.04)	(0.05)

All estimates rely on the regression method (OLS), as described in Sect. 3.1. Bootstrap standard errors are reported in brackets. National rich list in Germany: manager magazin, in France: Challenges, and in Spain: El Mundo. The Top 100 in Spain contains 74 observations in the first, and 117 in the second wave. Table 5 (in the Appendix) provides the corresponding top 1% wealth shares after the top tail imputation

The Online Appendix provides graphs that illustrate the wealth distribution of the top tail for France, Germany, and Spain, distinguished by the type of rich list and the specific cut-off points w_{min}. The graphs plot the ccdf (Eq. 2) as well as the empirical and the estimated Pareto distributions for different combinations of rich lists and w_{min}. Comparing the plots for the top 300, top 200, and the *Forbes* rich list shows that the top 200 provides a good fit to the Pareto lines for Germany and France, including HFCS and national rich list. Therefore, we choose the top 200 households of the corresponding rich lists for Germany and France as the baseline specification. When choosing the rich list sample, we face a trade-off between efficiency and precision. On the one hand, larger rich lists increase the risk of heaping at round numbers, which reflects that wealth ranking estimates are less reliable. On the other hand we aim at using as much information from the rich list as possible and, thus, prefer the top 200 over the top 100 rich list. For Spain, we use the entire *El Mundo* list.

3.3 Imputation of the missing rich households

There is a large gap between the richest household in the German part of the HFCS (in both waves) and the poorest household in the national rich list (Online Appendix, section 1). The same is true for France and Spain; however, the gap is smaller than that of Germany. This suggests that the HFCS in France and in Spain better represents the top tail than it does in Germany. To fill the gap of the missing rich, we create synthetic observations according to the Pareto density function of the respective α. Moreover, we replace HFCS households above w_{min} and reallocate their weights to "synthetic households". According to the discrete Pareto distribution, we assign population weights to each imputed household such that total weights of the new households, i.e. those owning wealth of at least w_{min}, match the total weights of the corresponding households in the original HFCS that are replaced. The end of the top tail, we replace with rich list observations.[10] Figure 2 plots the adjusted tail wealth distribution for Germany (second wave) as an example. The joint tail wealth distributions for the three countries are plotted in the Online Appendix.

4 Results: impact of adjusting for the missing top wealth on the wealth distribution

In this section, we analyse the impact of adjusting for the missing rich on the wealth distribution. In doing so, we rely on the integrated data sets, composed of households from the HFCS, from the imputation, and from the corresponding national rich lists. First, we compare our results, referring to the first wave of the HFCS, with those of Vermeulen (2018). Our results are different by construction, as we build an adjusted wealth distribution based on newly created households, or taken from the rich list, whereas Vermeulen (2018) calculates the implied top shares. Considering the top 1%

---

[10] Table 7 (in the Appendix) shows how the tail wealth changes, when we do not replace the very end by rich list observations, but fully impute top tail wealth. Moreover, in the Online Appendix we describe in detail how we represent the continuous Pareto function by a discrete version, represented by the synthetic households.
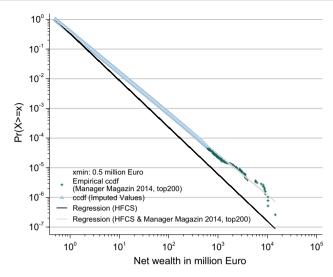
Fig. 2 Adjusted tail wealth distribution, Germany-Second wave of the HFCS. *Source:* HFCS. (Second wave), Manager magazin (2014); own calculations

wealth share, he reports 19, 34, and 16% for France, Germany, and Spain, respectively, considering, for instance, w_{min} of 0.5 million euros and the sample including the Forbes list. According to our estimates, the corresponding top 1% shares are 25, 34, and 23% in France, Germany, and Spain. Despite the methodological differences, the estimates are identical in Germany and somewhat higher in Spain and France (see Fig. 6).

Figure 3 shows how the adjustment of the HFCS for the missing rich affects the household net wealth distribution in Germany for the first and second waves.[11] The left plot focuses on the first wave of the HFCS and compares the wealth distribution based on the original HFCS to the one including the top tail adjusted sample. The plot shows that wealth is strongly concentrated, regardless of whether we include the rich list information or not. Using the original HFCS reveals that the richest decile owns almost 60% of total wealth, whereas the bottom half owns merely 3%. Moreover, wealth is further concentrated within the last decile, as the top 1% owns almost a quarter of total net wealth. The black bars show the wealth distribution after the top tail adjustment: Both wealth concentration and the total net wealth increase substantially. Total net wealth increases by about 1400 billion euros to 9134 billion euros (+18%) in the first wave. The share of household net wealth, held by the top decile, increases by more than 6 percentage points to 65%, while the share of the richest 1% climbs by 9 percentage points to 34%. The imputation mainly affects the top 0.1%, therefore leading to an increase of 13 percentage points to 17%.

Using the second wave of the HFCS, we observe a similar pattern (right plot). The wealth share of the top decile increases by 6 percentage points to about 66%, while the top 1% share rises from 24 to 35% due to the top tail adjustment. Similarly to the first wave, wealth of the richest 0.1% increases by 12 percentage points. We illustrate

---

[11] The Online Appendix provides additional tables with the underlying distributional results for each country and wave.

Share of total wealth in %

(a) First wave

HFCS (total net wealth: €7 743 bn; Gini: 0.74)

HFCS + imputed top wealth (total net wealth: €9 134 bn; Gini: 0.78)

(b) Second wave

HFCS (total net wealth: €8 500 bn; Gini: 0.75)

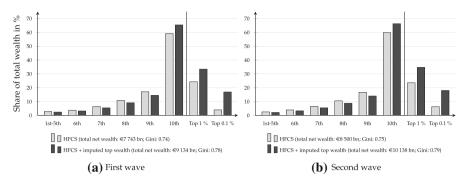HFCS + imputed top wealth (total net wealth: €10 138 bn; Gini: 0.79)

Fig. 3 The distribution of household net wealth in Germany. *Source:* HFCS (both waves), Manager magazin (2011), Manager magazin (2014), own calculations

the impact of the top tail adjustment on wealth concentration by comparing the Gini coefficient, which is a common inequality measure. A higher Gini coefficient reflects a more unequal distribution. The increase in the Gini coefficient by 4 percentage points to 0.78 (0.79) in the first (second) wave after the top tail adjustment confirms the higher wealth inequality.[12] Remarkably, the top wealth concentration seems to be fairly constant between the first and second waves, even though total net wealth increased by 1000 billion euros (+ 11%) to 10,138 billion euros.

Figure 4 illustrates the impact of adjusting for the missing rich on the French household net wealth distribution for both waves. In France, wealth is also strongly concentrated: The lower half owns 5.4% (6.3%), while the top 1% holds about 18% (19%) of total net wealth, based on the first (second) wave of the original HFCS.

Adjusting the top tail of the French net wealth distribution increases total wealth by 820 billion (+ 13%) to 7320 billion euros in the first wave, and by more than 2000 billion (+ 30%) to 9061 billion euros in the second wave. The top 1 % wealth share increases by 7 (12) percentage points in the first (second) wave as a result of the top tail adjustment. Wealth inequality, expressed by the Gini coefficient, increases by about 0.03 points in the first wave and about 0.07 points in the second due to the top tail adjustment.

Total net wealth increased from the first wave (2009/2010) to the second (2014/2015) by about 24%, or 1740 billion euros, a remarkable increase compared with Germany. A potential explanation is the long gap between the two waves. Wealth concentration measured by the Gini coefficient has increased from 0.71 to 0.74 across the two waves.

Figure 5 shows how the adjustment for the missing rich changes the net wealth distribution in Spain. First, we focus on the wealth distribution obtained from the original HFCS. The poorer half of all households owns 13% (12%) of total net wealth based on the first (second) wave. The richest decile, however, holds 2152 billion euros (43.4%) in the first and 2173 billion euros (45.6%) in the second wave. The net wealth

---

[12] We calculate the Gini coefficients by replacing zero or negative net wealth by one euro; however, smaller positive values do not affect the results. In Germany, the share of households holding zero or negative net wealth is 5% in the first wave and 6% in the second wave (in France: 3% in the first and 2% in the second wave; in Spain: 2% in the first and second waves).
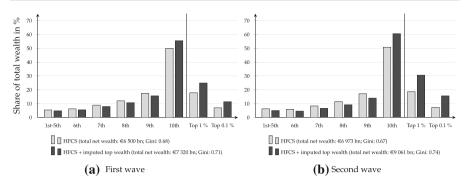
Fig. 4 The distribution of household net wealth in France. *Source:* HFCS (both waves), Challenges (2010), Challenges (2015), own calculations
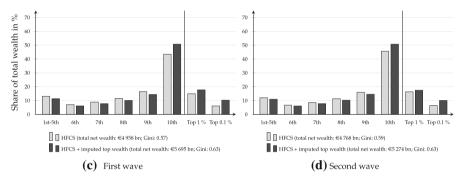

Fig. 5 The distribution of household net wealth in Spain. *Source:* HFCS (both waves), El Mundo (2009), El Mundo (2012), own calculations

share of the top 1% is 14.9% (16.3%) in the first (second) wave of the HFCS. Net wealth—in absolute terms—is slightly lower in the second wave (4770 billion euros) relative to the first (4960 billion euros). A potential explanation is the global recession that hit Spain in 2009, shortly after the first wave was conducted. In the aftermath of the economic crisis, the worth of business assets and real estate decreased substantially, thus resulting in a overall reduction of wealth. The top tail adjustment increases total net wealth by 15% (+ 737 billion euros) in the first wave and 11% (+ 506 billion euros) in the second. Hence, the top tail adjustment affects the wealth distribution only moderately compared to Germany. Furthermore, the top 1% wealth share increases by about 6–8 percentage points in both waves due to the imputation of the top tail. Wealth inequality, measured by the Gini coefficient, increases after the top tail adjustment by 0.05 points in the first wave and by 0.04 in the second wave. However, overall net wealth inequality is still notably lower than in France or in Germany.

To test the robustness of the results, Fig. 6 shows how the wealth shares held by the top 1% of households depend on the choice of the sample and $w_{\min}$. Table 5 in the Appendix reports the underlying data. The plot shows that combining the HFCS with additional information from rich lists, the corresponding imputed top 1% share increases for all three countries and waves. Moreover, we test whether the estimated
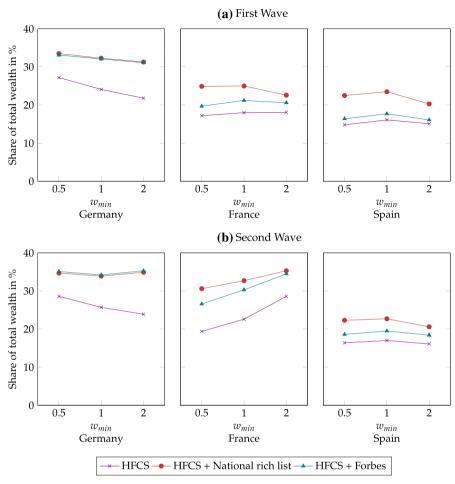
Fig. 6 Net wealth share of top 1% when replacing the top tail by different samples and w_{min}. *Note:* The plot reports top 1% wealth shares when replacing the top tail by the imputed one, as described in Sect. 3. The shares of "HFCS" refer to a replacement of the top tail when the Pareto α is estimated based on a pure HFCS sample without adding a rich list. *Source:* HFCS (both waves), Manager magazin (2011), Manager magazin (2014), Challenges (2010), Challenges (2015), El Mundo (2009), El Mundo (2012), own calculations

share of top 1% is sensitive to the choice of w_{min}. We find that the choice of w_{min} has only a minor impact on the calculated shares. Further, the top 1% shares estimated based on a combined sample of HFCS and national rich list or the *Forbes* list are almost identical in Germany and somewhat different in France and Spain. For the latter two, relying on a sample that uses information from national rich lists increases wealth concentration compared to using the *Forbes* list.

Finally, we compare our estimations with macroeconomic wealth aggregates from the national and financial accounts for the household sector.[13] Table 4 shows aggregate

_____

[13] The Online Appendix provides an overview of total assets and liabilities for France, Germany and Spain in the corresponding time period.

Table 4 Comparison of total net wealth between national accounts and top tail adjusted net wealth. *Source:* National accounts: France: INSEE, national accounts; Banque de France and European Central Bank, financial accounts; Germany: Federal Statistical Office, national accounts; Deutsche Bundesbank, financial accounts; Spain: Banco de España and European Central Bank, financial accounts; estimation on real estate assets by Banco de España. Adjusted HFCS based on the HFCS (first and second waves) and Manager Magazin (2011, 2014), Challenges (2010,2015) and El Mundo (2009, 2012), own calculations

Countries	National accounts in billion	Adjusted HFCS	Difference in %
First wave			
Germany	7969	9134	15
France	9463	7320	− 23
Spain	6394	5695	− 11
Second wave			
Germany	9355	10,138	8
France	9964	9061	− 9
Spain	5804	5247	− 9

The 'adjusted HFCS' column reports total net wealth after the top tail adjustment, described in Sect. 3

net wealth, after deducting currency, non-life insurance technical reserves, and pension entitlements. National accounts-based net wealth is adjusted by removing items, not included in the HFCS net wealth measure, to ensure comparability with the survey-based wealth definition.

In the case of Germany, the adjusted household net wealth aggregate reported in national and financial accounts of 7969 billion euros (2010) even falls short of our estimation for total personal net wealth of 9134 billion euros (including imputed top wealth) in the first wave. In the second wave, the gap between the aggregated data and the estimation is smaller (National accounts: 9355 billion euros; Estimation: 10,138 billion euros). However, German financial accounts presumably underestimate unlisted corporate shares and other equity by at least 1000 billion euros since there is no reliable data on financial or tax accounts data of the "*German Mittelstand*" and many family-owned major enterprises.[14] In contrast, the personal net wealth aggregate for France reported in national and financial accounts is much higher than our estimate (9463 compared to 7320 billion euros, in the first wave, and 9964 compared to 9061 billion euros, in the second wave). Likewise, in Spain, the household net wealth aggregate in macroeconomic statistics considerably exceeds our estimates in both waves (first wave: 6394 compared to 5695 billion euros; second wave: 5,804 compared to 5247 billion euros). The remarkable underestimation of household net wealth in France and Spain compared to the macro-aggregates suggests that total (top) wealth is still under-represented. Yet, national and financial accounts of household wealth might be flawed by estimation risks, in particular with respect to non-financial assets, corporate shares in non-quoted firms, and financial assets abroad. Thus, the differences between the

---

[14] According to the financial accounts, corporations in Germany had a "net worth" of 2,990 billion euros in 2014 (Federal Statistical Office 2018). Since companies cannot be owned by themselves, this wealth would be fully attributed to other sectors (households, non-profit institutions, government, and rest of the world). This is not possible due to a lack of data.

national and financial accounts and results from household surveys should be analysed in detail for the different components of household wealth and liabilities (Chakraborty and Waltl 2018; Chakraborty et al. 2018).

5 Conclusion

In this paper, we analyze the top tail of the wealth distribution by constructing an integrated micro-database for France, Germany, and Spain that better represents the top wealth concentration. Following Vermeulen (2018), we use the first and second waves of the HFCS and combine it with national rich lists and the *Forbes* list in these three countries to estimate a joint Pareto distribution for the wealth top tail. Further, we compare how estimates differ depending on whether we use a sample that is based on the *Forbes* list or a national rich list. We conclude that national rich lists can improve the estimation of the Pareto coefficient, especially when the list of national billionaires is short. However, when the *Forbes* list is sufficiently large, then, the value added of replacing it with a national rich list is rather small. For Germany, using the *Forbes* list or the national list makes no difference. For France and Spain, replacing the Forbes list with a national list yields substantially higher estimates. For instance, the top 1% wealth share is up to 8 percentage points higher in Spain when replacing the *Forbes* by the *El mundo* list. Following the top wealth imputation, the top percentile share of household wealth in Germany jumps up from 24 to 34% in the first wave and from 24 to 35% in the second. For France, wealth concentration increases by 7 percentage points in the first wave. In the second wave, however, the top 1% wealth share almost doubles from 19 to 31% due to the top tail adjustment. For Spain, we find that the imputation has only a smaller effect since rich households are better captured in their surveys.

The data in our analysis refer to the period between 2008 and 2011, for the first wave, and to the period between 2011 and 2015, for the second wave of the HFCS. Historically low interest rates adversely affect fixed-income securities such as bank deposits, bonds, and pension plans, while increasing the market valuation of investments such as real estate, businesses, and corporate shares. As the latter dominates top wealth strata, the wealth distribution might have further concentrated, at least in Germany. Counter-factual microsimulation analyses could shed light on the distributional impact involved (Domanski et al. 2016). Moreover, our integrated database could be used for the analyses of redistribution policies, for instance, wealth taxation[15] or programs to promote housing ownership and capital formation. It must be mentioned that our findings should be interpreted with some caution.

Uncertainty emerges from the estimation strategy of the top wealth concentration, which relies on the Pareto distribution, and from measurement errors in household wealth in both the HFCS and the rich lists. In particular, the reliability of rich lists is

---

[15] Bach and Thiemann (2016b) rely on the integrated database to simulate the tax revenue of reviving the recurrent wealth tax in Germany. Not surprisingly, the results show that tax revenues are substantially higher if the integrated database is used instead of the original HFCS data. Moreover, based on the integrated database, Bach and Thiemann (2016a) simulate future estates and inheritances using static ageing procedures and estimate future tax revenue and distributional effects of estate taxation scenarios.

contentious. We suppose that these wealth rankings rather under-report the very top wealth concentration with respect to some selectivity in favour of corporate wealth and against other assets, such as real estate properties and financial portfolios. It is difficult to evaluate the self-assessed property valuations of the survey respondents or the valuations of properties collected in the rich lists. However, we have no evidence of systematic biases in this respect. Regardless, these issues indicate substantial need for research. Tax files from wealth and estate taxation or disclosed financial statements of large family-owned corporations, foundations, or trusts are an important source for further top wealth research. Sampling design, survey strategy, and field work of voluntary household surveys might be improved to better collect data from the wealthy strata of the population.

A Appendix

A. 1 Sensitivity of estimating the top 1% share of household net wealth

Table 5 provides the underlying numbers of Fig. 6. The table also reports the top 1% shares of household net wealth that correspond the α estimates, shown in Table 3.

A.2 Sensitivity of national rich lists

We follow Chakraborty and Waltl (2018, p. 51) and repeat their sensitivity tests of national rich lists on the top tail estimation by re-arranging rich list entries in different ways. We perform these tests for the top 200 observations (or the full Spanish list), because lower-ranked observations appear to be less precise (see Sect. 2.2, Table 6).

We compare the estimate based on the top 200 wealthy of the respective national rich lists (0) to specifications, where we restrict the length of rich list further (1–3). Specification (4) drops the richest observation, while (5) and (6) split the rich list observations into 2 or 4 households. Finally, specification (7) randomly splits rich list observations into 1, 2, 3 or 4 households, with equal shares. The results, shown here, are the average of repeating the random assignment 1,000 times.

Comparing the specifications (1–2) to the estimate based on the top 200 of the national rich list (0) shows that the results are not strongly affected by the length of rich list. When only using the top 10% (3), however, the estimates become less stable,

Table 5 Net wealth share of the top 1% for different sample choices and w_{min}. *Source*: HFCS (both waves),Manager magazin (2011, 2014); Challenges (2010, 2015); El Mundo (2009, 2012) and Forbes (2011, 2012) own calculations

Country, wave	Excluding rich list w_{min} (in euro)			Including rich list National rich list Top 100			Top 200			Top 300			Forbes		
	≥ 2 M	≥ 1 M	≥ .5 M	≥ 2 M	≥ 1 M	≥ .5 M	≥ 2 M	≥ 1 M	≥ .5 M	≥ 2 M	≥ 1 M	≥ .5 M	≥ 2 M	≥ 1 M	≥ .5 M
Germany, 1st	21.8	24.1	27.2	31.1	32.0	33.2	31.3	32.3	33.5	30.8	31.7	32.8	31.1	32.1	33.1
Germany, 2nd	23.9	25.7	28.6	35.1	34.0	34.8	34.9	33.9	34.7	34.3	33.2	34.0	35.3	34.2	35.1
France, 1st	18.1	18.0	17.2	23.7	26.2	25.6	22.6	25.0	24.9	21.9	24.1	24.1	20.6	21.2	19.7
France, 2nd	28.6	22.6	19.4	36.8	33.5	30.4	35.3	32.7	30.6	34.1	31.7	30.1	34.5	30.3	26.6
Spain, 1st	15.1	16.1	14.8	20.3	23.5	22.5	–	–	–	–	–	–	16.1	17.7	16.4
Spain, 2nd	16.1	17.0	16.4	20.6	22.7	22.3	–	–	–	–	–	–	18.4	19.5	18.6

All underlying Pareto estimates rely on the regression method (OLS), as described in Sect. 3.1. The HFCS top tail is replaced an synthetic households that follow the estimated Pareto distribution. The very end is replaced by the corresponding rich list (exempt for HFCS sample (excluding rich list)). National rich list in Germany: manager magazin, in France: Challenges, and in Spain: El Mundo. The Top 100 in Spain contains 74 observations in the first, and 117 in the second wave

Table 6 Sensitivity of national rich lists. *Source:* HFCS (both waves), Manager magazin (2011, 2014), Challenges (2010, 2015); El Mundo (2009, 2012), own calculations

	Germany			France			Spain		
	Obs.	α	Tail wealth	Obs.	α	Tail wealth	Obs.	α	Tail wealth
First wave									
(0) Top 200 rich list	200	1.339	6435	200	1.548	4077	74	1.569	3172
(1) Top 50% of all observations	100	1.332	6545	103	1.538	4127	38	1.628	2980
(2) Top 25% of all observations	50	1.340	6432	54	1.555	4045	19	1.695	2805
(3) Top 10% of all observations	20	1.378	5945	20	1.626	3751	8	1.767	2648
(4) Excluding the top observation	199	1.343	6387	199	1.555	4047	73	1.579	3137
(5) Split into 2 households	400	1.372	6017	400	1.570	3975	148	1.535	3299
(6) Split into 4 households	800	1.412	5591	800	1.548	4078	296	1.527	3332
(7) Random split	501	1.399	5713	501	1.562	4014	186	1.530	3318
Second wave									
(0) Top 200 rich list	200	1.326	7471	200	1.456	5527	117	1.577	2800
(1) Top 50% of all observations	103	1.320	7565	100	1.444	5630	59	1.594	2749
(2) Top 25% of all observations	52	1.333	7344	52	1.472	5395	32	1.640	2625
(3) Top 10% of all observations	20	1.360	6930	20	1.542	4922	12	1.700	2487
(4) Excluding the top observation	199	1.329	7408	199	1.437	5692	116	1.587	2768
(5) Split into 2 households	400	1.355	7009	400	1.437	5691	234	1.588	2766
(6) Split into 4 households	800	1.392	6522	800	1.459	5502	468	1.613	2696
(7) Random split	501	1.382	6643	501	1.457	5516	293	1.606	2713

The table shows the sensitivity tests based on the top 200 entries of the German list (*manager magazin*), the French list (*Challenges*), and on the Spanish full list (*El Mundo*). α is estimated, setting $w_{\min}$ to 0.5 million euros. The tail wealth is calculated according to the following formula: $Tail\ wealth = N_{tt} * w_{\min} * \left(\frac{\alpha}{\alpha - 1} \right)$, where α is the Pareto coefficient, $w_{\min}$ the minimum wealth threshold of the Pareto distribution, and N_{tt} the weighted sum of households in the Pareto tail (Vermeulen 2016b, p. 15)

Table 7 Total tail wealth by imputation approach. *Source:* HFCS (both waves), Manager magazin (2011; 2014), Challenges (2010, 2015); El Mundo (2009, 2012), own calculations

| | Total tail wealth (in billion euros) | | | | |
	A) w_{min} up to rich list	B) Rich list	C)=A)+B)	D) Full imputation	Diff. (%)
First wave					
Germany	5373	273	5646	5744	− 1.7
France	3905	216	4120	4077	1.1
Spain	3109	110	3219	3172	1.5
Second wave					
Germany	6225	355	6580	6679	−1.5
France	5399	393	5792	5738	0.9
Spain	2709	130	2839	2800	1.4

The table shows the sensitivity tests based on the top 200 entries of the German list (*manager magazin*), the French list (*Challenges*), and on the Spanish full list (*El Mundo*). α is estimated, setting w_{min} to 0.5 million euros. Tail wealth in A) is calculated according to $\int_{w_{min}}^{rl_{low}} N_{tt} w \frac{\alpha w_{min}^{\alpha}}{w^{\alpha+1}} dw$, rl_{low} being net wealth of the poorest rich list observation and N_{tt} the weighted sum of households in the Pareto tail. The tail wealth in D) is calculated according to the following formula: $Tail\ wealth = N_{tt} w_{min}(\frac{\alpha}{\alpha - 1})$, where α is the Pareto coefficient, w_{min} the minimum wealth threshold of the Pareto distribution (Vermeulen 2016b, p. 15)

especially in Spain where the rich list is shorter than in France or Germany. Excluding the top observation (4) hardly affects the top tail estimation, even in Spain.[16]

When we split rich list observations into 2, 3 or randomly into 1–4 households (5–7), the estimates tend to be rather different compared with those based on the full list (0). Hence, having better information about how wealth of rich list observations is distributed across several households could improve the precision of estimates substantially. Therefore, we have adjusted the national rich list for Germany based on the available anecdotal evidence, partly provided by the editor of the list.

A.3 Pareto tail imputation sensitivity

In this paper, we replace HFCS households in the top tail of the wealth distribution by imputed synthetic households that follow a Pareto distribution and rich list observations at the top. Hence, the maximum household net wealth is determined by net wealth of the richest household of each rich list. The Pareto distribution, however, is defined on the interval between the wealth threshold, w_{min}, and infinity.

Table 7 compares the impact of the top tail adjustment on total tail wealth of two alternatives: first, the top tail adjustment as performed in this paper, replacing HFCS houseolds in the top tail by imputed synthetic households and rich list observations at the top; second, replacing the HFCS top tail by a truly Pareto-distributed top tail up to infinity. Column A) shows the total imputed tail wealth between w_{min}, here set to 0.5 million euros, and the lowest wealth level of the corresponding national rich list.

---

[16] Amancio Ortega Gaona, is estimated to have a net worth of 16.4 billion euros (39 billion euros) in 2009 (2012), whereas the wealth of the second richest amounts to 6 billion euros in 2009 or 2012.

B) reports total net wealth of the corresponding national rich lists, where we rely on the top 200 observations (full list in case of Spain), and C) reports the sum of A) and B), which is total tail wealth based on the approach of the paper. Column D) shows total tail wealth when replacing the top tail by a true Pareto distribution (up to infinity). The results for Germany show that a full imputation according to the Pareto function would lead to about 1.5–1.7% higher total tail wealth than the approach adopted in this paper. In France and Spain, however, applying the full imputation would yield slightly lower total tail wealth. In general, however, both imputation approaches yield relatively similar results.

References

Alstadsæter, A., Johannesen, N., & Zucman, G. (2019). Tax evasion and inequality. *American Economic Review*, *109*(6), 2073–2103.

Alvaredo, F., Atkinson, A., Chancel, L., Piketty, T., Saez, E., & Zucman, G. (2016). Distributional national accounts (DINA) guidelines: Concepts and methods used in WID.world. Working Paper Series 2016/1.

Bach, S., Beznoska, M., & Steiner, V. (2014). A wealth tax on the rich to bring down public debt? Revenue and distributional effects of a capital levy in Germany. *Fiscal Studies*, *35*(1), 67–89.

Bach, S., & Thiemann, A. (2016a). Inheritance tax revenue low despite surge in inheritances. *DIW Economic Bulletin*, *4+5*, 41–48.

Bach, S., & Thiemann, A. (2016b). Reviving Germany's wealth tax creates high revenue. *DIW Economic Bulletin*, *4*(5), 50–59.

Brzezinski, M. (2014). Do wealth distributions follow power laws? Evidence from "rich lists". *Physica A: Statistical Mechanics and its Applications*, *406*(15), 155–162.

Burkhauser, R. V., Shuaizhang, F., Jenkins, S. P., & Larrimore, J. (2012). Recent trends in top income shares in the United States: Reconciling estimates from March CPS and IRS tax return data. *The Review of Economics and Statistics*, *94*(2), 371–388.

Chakraborty, R., Kavonius, I. K., Pérez-Duarte, S., & Vermeulen, P. (2018). Is the top tail of the wealth distribution the missing link between the household finance and consumption survey and national accounts? European Central Bank, Working Paper Series, 2187.

Chakraborty, R., & Waltl, S. R. (2018). Missing the wealthy in the HFCS: Micro problems with macro implications. European Central Bank, Working Paper Series, 2163.

Challenges. (2010). Les 500 plus grandes fortunes professionnelles de France. Challenges 220.

Challenges. (2015). Les 500 plus grandes fortunes professionnelles de France. Challenges 441.

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

Cowell, F. (2011). *Measuring inequality*. Oxford: Oxford University Press.

Dalitz, C. (2016). Estimating wealth distribution: Top tail and inequality. Technischer Bericht Nr. 2016-01, Hochschule Niederrhein.

Davies, J. (1993). The distribution of wealth in Canda. *Research in Economic Inequality*, *4*, 159–180.

Davies, J. B., & Shorrocks, A. F. (2000). Chapter 11: The distribution of wealth. In A. Atkinson & F. Bourguignon (Eds.), *Handbook of income* distribution, (vol 1, pp. 605–675). Amsterdam: Elsevier.

Domanski, D., Scatigna, M., & Zabai, A. (2016). Wealth inequality and monetary policy. *BIS Quarterly Review*, *1*, 45–64.

Eckerstorfer, P., Halak, J., Kapeller, J., Schütz, B., Springholz, F., & Wildauer, R. (2016). Correcting for the missing rich: An application to wealth survey data. *Review of Income and Wealth*, *62*(4), 605–627.

El Mundo. (2009). Los 100 mas ricos de España. El mundo magazine 532.

El Mundo. (2012). Los 200 Ricos de España. El mundo magazine 691.

Embrechts, P., Mikosch, T., & Klüppelberg, C. (1997). *Modelling extremal events: For insurance and finance*. Berlin: Springer.

European Central Bank. (2013). The Eurosystem household finance and consumption survey. Methodological report for the first wave. Statistical Paper Series 1, European Central Bank.

European Central Bank. (2016a). The household finance and consumption survey: Methodological report for the second wave. Statistics Paper Series 17, European Central Bank.

European Central Bank. (2016b). The household finance and consumption survey: Methodological report for the second wave. Statistical Paper Series 17, European Central Bank. European Central Bank-Statistical Paper Series, 1.

Federal Statistical Office. (2018). Balance sheets for institutional sectors and the total economy 1999–2017. Wiesbaden.

Forbes. (2009). The World's Billionaires 2009. The World's Billionaires 2009.

Forbes. (2010). The World's Billionaires 2010. The World's Billionaires 2010.

Forbes. (2011). The World's Billionaires 2011. The World's Billionaires 2011.

Forbes. (2012). The World's Billionaires 2012. The World's Billionaires 2012.

Forbes. (2014). The World's Billionaires 2014. The World's Billionaires 2014.

Forbes. (2015). The World's Billionaires 2015. The World's Billionaires 2015.

Förster, M., Llena-Nozal, A., & Nafilyan, V. (2014). Trends in top incomes and their taxation in OECD Countries. *OECD Social, Employment and Migration Working Paper*, *159*, 1–93.

Gabaix, X. (2009). Power laws in economics and finance. *Annual Review of Economics*, *1*(1), 255–294.

Gabaix, X., & Ibragimov, (2012). A simple way to improve the ols estimation of tail exponents. *Journal of Business and Economic Statistics*, *29*(1), 24–39.

Kleiber, C., & Kotz, S. (2003). *Statistical size distribution in economics and actuarial sciences*. London: Wiley.

Kopczuk, W., & Saez, E. (2004). Top wealth shares in the United States, 1916–2000: Evidence from estate tax returns. *National Tax Journal*, *57*((2, part2)), 445–488.

Krenek, A., & Schratzenstaller, M. (2017). Sustainability-oriented Future EU Funding: A European Net Wealth Tax. Working Paper-Series 10, FairTax.

Manager magazin. (2011). Die 500 reichsten Deutschen (The 500 richest Germans). manager magazin spezial, Oktober 2011.

Manager magazin. (2014). Die 500 reichsten Deutschen (The 500 richest Germans). manager magazin spezial, Oktober 2014.

Piketty, T. (2014). *Capital in the twenty-first century*. Harvard: Harvard University Press.

Raub, B., Johnson, B., & &Newcomb, J. (2010). A comparison of wealth estimates for America's wealthiest descendants using tax data and data from the Forbes 400. In 103rd Annual conference on taxation national tax association proceedings.

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

Saez, E., Zucman, G. (2019). Progressive wealth taxation, unpublished conference paper.

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

Treguier, E. (2012). Comment? Value-t-on leur patrimoine? http://www.challenges.fr/entreprise/20120711.CHA8798/comment-evalue-t-on-leur-patrimoine.html.

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

Vermeulen, P. (2016b). Estimating the top tail of the wealth distribution. Working Paper Series 1907, European Central Bank.

Vermeulen, P. (2018). How fat is the top tail of the wealth distribution? *Review of Income and Wealth*, *64*(2), 357–387.