Estimation Methods For the Upper Tail of the Wealth Distribution*

Simon J. Toussaint†

First Version: November 23, 2020
Current Version: December 13, 2020

Abstract

This paper investigates methods to estimate the upper tail of the wealth distribution. I compare data types and estimation methods using data from the Netherlands for the period 1993–2018, exploiting the unique availability of multiple types of data for this context. In addition to comparing the existing methods of OLS regression, Maximum Likelihood, and Generalized Pareto interpolation, I develop a new method to combine data from several sources. This method, called Robust Pareto Regression, combines local estimates of wealth concentration from individual data sources, and uses fixed effects methods to correct for the heterogeneity across data sources and years. Several conclusions emerge: (i) No data source on its own accurately captures the top tail, meaning that all sources need to be adjusted or combined to estimate top wealth. (ii) Combining surveys with rich lists is highly sensitive to the quality of the underlying data sources; generalized Pareto interpolation partly addresses this concern, but straight Pareto regression and Maximum Likelihood methods do not. (iii) Robust Pareto Regression is preferable to existing methods, since it more adequately adjusts for data heterogeneity, and easily shows trends over time.

JEL Classification: D30, D31, E01, E21, H20, H26, N30
Keywords: Wealth inequality; wealth distribution; Pareto distribution.

*I thank my advisors, Coen Teulings and Bas van Bavel, for many helpful suggestions and support. I further thank Wiemer Salverda for helpful suggestions; Sander Schimmelpenninck and Caroline Meijer for their help with accessing the *Quote* data; Clara Martínez-Toledano and Thomas Blanchet for help with the `gpinter` code and making the data consistent with National Accounts totals. Mistakes and errors in this paper can of course only be attributed to me.

†Utrecht University School of Economics, Kriekenpitplein 21-22, 3284 EC, the Netherlands. E-mail: s.j.toussaint@uu.nl

1 Introduction

This paper discusses methods to estimate the upper tail of the wealth distribution. There are three types of data sources from which the upper tail can be estimated: Fiscal data, deriving from a wealth tax (e.g., Fagereng et al. 2020), inheritance tax (Kopczuk and Saez 2004), or from the capitalized values of the income tax (Saez and Zucman 2016); household surveys of wealth (e.g., Kuhn, Schularick, and Steins 2020); and rich lists such as the *Forbes List of Billionaires* (Gomez 2019). None of these sources perfectly captures the upper tail. Fiscal data is subject to tax avoidance and evasion (Alstadsæter, Johannesen, and Zucman 2019), and typically measures wealth components such as closely held businesses at book value rather than market value (Albers, Bartels, and Schularick 2020). Rich households are significantly less likely to respond to household surveys, which biases survey data (Vermeulen 2018). Finally, rich lists only measure the very top of the distribution; moreover, these rankings have difficulty to measure opaque wealth components such as liabilities and savings accounts.

Because none of these data sources perfectly measures the top of the distribution, we have to make assumptions to estimate wealth concentration. Since top wealth is approximately Pareto distributed (Pareto 1896), researchers often exploit convenient properties of this functional form to estimate the key parameters that govern this distribution, which can then be used to draw conclusions about wealth concentration. These parameters can be estimated using either linear regression (Gabaix and Ibragimov 2011), Maximum Likelihood (Vermeulen 2018), or generalized Pareto interpolation techniques (Blanchet, Fournier, and Piketty 2017).

Data limitations often mean that these data sources and methods are applied in isolation, which precludes a systematic comparison of strengths and weaknesses of different (combinations of) data sources and methodologies. This is where this paper comes in. I use data from several data sources to trace the top tail of the Dutch wealth distribution from 1993 until 2018. A unique feature of the Dutch context is that we not only have access to different kinds of data, but that there are multiple sources for each data type. Specifically, the period 1993–2018 is (partly or fully) covered by fiscal data (from several different tax regimes), two household surveys, two rich lists, and National Accounts data. Hence, we can critically compare different methodologies for the years where data sources overlap.

This systematic comparison yields several conclusions. First, each data source on its own is insufficient to estimate top wealth. For the surveys and rich lists, this may not be surprising; but even the 2011–2018 fiscal series needs to be adjusted to capture top wealth. In particular, Statistics Netherlands estimates closely-held assets at book value instead of market value. Second, the combination of different data sources, while preferable to analyzing individual sources, is only valid when the underlying heterogeneity in quality and scope of data sources is addressed adequately. Existing research that combines data sources is limited to combinations of surveys with rich lists, using one of the three methods described above. While these combinations are preferable to either surveys or rich lists above, my results show that the results of these combinations are highly dependent on *ad hoc* assumptions, such as the exact location of the threshold of the distribution, as well as the exact functional form of the distribution. Generalized Pareto interpolation, which is the most flexible of the three methods, is the best approach in addressing these issues, yet even that method fails to systematically account for heterogeneity in data sources; nor does it capture trends over time.

As my major contribution, therefore, I develop a novel estimation method for the upper tail, using fixed effects regression to average across different data sources and years. Essentially, I first compute many 'local' estimates of top wealth, using only a single data source at a time. These local estimates, computed using any of the three methods described above, then serve as new data points in a regression framework. Since each local estimate delivers information on key properties of the upper tail, this aggregative method in effect delivers an average over all data sources and years. Moreover, we can explicitly allow for heterogeneity in data sources and years, by using a rich set of fixed effects. These fixed effects not only serve to make the estimation more reliable, but are informative in themselves, since they represent averaged differences between data sources and over time. Therefore, this method explicitly quantifies the differences in data quality and other sources of heterogeneity; moreover, the method also directly shows trends over time. This new method, which I term Robust Pareto Regression, is easy to compute and can be adapted to many differing settings.

Applying Robust Pareto Regression to the Dutch context, I show that heterogeneity between data sources is substantial. Taking fiscal data as a benchmark, survey data result in far lower estimates of wealth concentration. Interestingly, the fiscal data itself also results in estimates that are implausibly low. In particular, the estimated threshold of the Pareto distribution is at least an order of magnitude lower than is typically found in the literature. As a consequence, the inclusion of the rich lists is of essential importance to accurately capture top wealth. As a second result, I show that the threshold of the Pareto distribution has increased from the 1990s until the mid-2000s, after which it has remained roughly constant. This, combined with only a modest decrease in the thickness of the upper tail, leads to the conclusion that wealth concentration in the Netherlands has increased since the 1990s.

This paper contributes to several strands of literature. First, my results add to the increasing empirical evidence on the long–run dynamics of wealth inequality (Piketty 2014, Roine and Waldenström 2015, Zucman 2019). Most studies have to rely on one type of data source or method, such as surveys (Kuhn, Schularick, and Steins 2020), surveys combined with rich lists (Vermeulen 2018), or fiscal data (Saez and Zucman 2016, Smith et al. 2019). In contrast, my results rest on a combination of different data sources and techniques. Closest to my paper is the study by Albers, Bartels, and Schularick (2020), who also use different data sources and methods to trace the long–run distribution of wealth in Germany. Compared to their study, my contribution is to explicitly compare and contrast different methods, which is possible thanks to the overlapping data sources available for this time period.

Second, my paper belongs to the small but growing literature which studies wealth inequality in the Dutch context. The seminal studies in this regard are Wilterdink (1984) and Soltow and van Zanden (1998); more recent additions are van Bavel and Salverda (2014), Salverda et al. (2013), and van Bavel and Frankema (2017). Closest to my paper is van Bavel and Frankema (2017), who explicitly compare all available data sources mentioned here, with the exception of the *SEP* survey. Compared to them, my paper adds to our understanding of differences between data sources by explicitly comparing different existing methods. Moreover, the Robust Pareto Regression framework directly exposes differences across data sources in an econometrically robust way.

Finally, this paper contributes to a fast-growing literature which studies different methodologies to estimate top wealth or income shares. Examples include the Pareto-regression approach (Gabaix and Ibragimov 2011), Maximum Likelihood methods (Vermeulen 2018), and the generalized Pareto interpolation method (Blanchet, Fournier, and Piketty 2017). The first major contribution of this paper is that it explicitly studies the consequences of different approaches for estimating wealth concentration, exploiting the availability of multiple data sources. Second, I contribute to this literature by introducing a new method, the Robust Pareto Regression.

The rest of this paper is structured as follows. Section 2 discusses the data used, starting with the wealth concepts and concluding with the various sources for the 1990S and beyond. Section 3 discusses the theoretical and methodological framework used in the literature to estimate the upper tail on the basis of (combinations of) these sources. In Section 4, I contrast existing insights and methods, which I achieve by first looking at individual data sources, and then by looking at common strategies to combine surveys and rich lists. In Section 5, I present the Robust Pareto Regression framework, and use this framework to draw econometrically valid conclusions about differences between data sources and trends over time. Section 6 concludes.

2 Wealth Concepts and Data

This paper draws on four types of data: National Accounts data, household surveys, income and wealth tax data, and rich lists. I start by discussing the macro-data from the National Accounts, which serve as a reference point for the wealth concept I employ in this paper. Next, I describe the three other sources, which I use for estimates on the tail of the wealth distribution.

2.1 Household Wealth: Concepts and Data

The wealth series employed in this paper largely follow the guidelines prescribed by the Distributional National Accounts project, which underpins the World Inequality Database (WID). Distributional National Accounts use the available household balance sheets from the System of National Accounts to create a concept of total household wealth, and subsequently allocate this wealth to households over the wealth distribution. This method ensures that 100% of household wealth is accounted for. The difficulty lies in 'translating' total household wealth from National Accounts to the wealth distribution derived from sources such as surveys or fiscal data. Typically, these distributional sources do not correspond to the National Accounts, both in terms of total aggregate wealth as well as the distribution of wealth components. To alleviate this issue, I follow the National Accounts to create a consistent series of household wealth, but rely on the distributional sources for the precise composition and distribution of household wealth. Moreover, while the WID reports wealth on an individual basis, I report wealth at the household level, as is more common in the broader literature.

Household wealth, calculated in this way, follows the definition spelled out in the System of National Accounts, which is the total value of assets minus liabilities. Assets include all financial and non-financial assets over which ownership rights can be enforced and which provide economic benefits to their owners. This definition includes most major wealth components, including housing, real estate, savings accounts, stocks and bonds, and life insurance and pension funds which can be accessed and capitalized by their households. Notably, the definition excludes non-privately saved pension wealth. For the context of the Netherlands, this is a significant omission because the market value of pension funds far exceeds 100% of national income. Several authors argue that, due to its size and its potential substitution effects with private savings, these pension funds should be included in Dutch wealth estimates (Caminada, Goudswaard, and Knoef 2014; see van Bavel and Frankema 2017 for a discussion). In this paper, I will remain consistent with international practice, and exclude pension wealth from the wealth concept.

Statistics Netherlands provides household financial balance sheets since 1990, and non-financial balance sheets since 1995. Hence, for the years 1993 and 1994, our total wealth concept is incomplete. This may affect the estimates provided in this paper for those two years.

2.2 Distributional Data

For the past 30 years, we possess a variety of sources that each provide estimates of the Dutch wealth distribution. Two of these sources – the *SEP* survey and the *Forbes* rich list – start in 1987; two sources – the *DNB* survey and the Statistics Netherlands fiscal series – start in 1993; and the *Quote* rich list starts in 1997. To provide the most consistent possible comparison, I limit the start of my time series to 1993. Post-1993, all mentioned data sources are available, allowing for rich comparisons. In addition, the data from Statistics Netherlands is indispensable to correct severe biases in the wealth surveys, as I will discuss in the next section; hence, we gain only limited insights from the *SEP* survey pre-1993.

2.2.1 Statistics Netherlands, 1993 – 2000 and 2002 – 2018

For the years 1993 – 2000, Statistics Netherlands started publishing disaggregated statistics about the wealth distribution. These statistics were constructed on the basis of a representative sample drawn from the Dutch tax registers, and adjusted using the *SEP* survey. Hence, we have estimates of the size and distribution of wealth components across the distribution. Since this sample is based on fiscal data, it may suffer from tax avoidance and tax evasion.

After the tax code changed in 2001, Statistics Netherlands – henceforth also referred to as *CBS* – did not resume publishing wealth inequality estimates until 2006; however, some data is already available since 2002[1]. This second series of fiscal data combines data from the several *boxen*: Tax 'boxes' that have different rates for different sources of income. For its wealth inequality statistics, Statistics Netherlands takes the cadastral value of real estate minus mortgage debt from Box 1; the value of private business equity from Box 2; and the net value of financial capital and deposits from Box 3. From 2006 until 2011, these statistics were reported on basis of a representative sample; after 2011, statistics are based on the universe of tax files. For 2006 – 2018, Statistics Netherlands reports the top 1% wealth share; for 2011 – 2018, they also report the top 0.1% share.

The freely available statistics from Statistics Netherlands are quite detailed, yet remain highly aggregated. However, they have also constructed several custom-made datasets on the wealth distribution, which have been provided to me. Two of these data sets divide the wealth distribution in centiles, and provide disaggregations of the distribution of each wealth component for each centile. Additionally, a third dataset provides similar information for the top 0.1% percentile for the 2011 – 2018 period. This detailed information can be used to make the estimation and correction of the official inequality statistics more precise.

One major issue with the new tax code is Box 2, which taxes income for owners of closely held businesses. Since the tax rate in this box is lower than the income tax in Box 1, the box is prone to tax arbitrage and tax delay (Jacobs 2015). Of more direct relevance for this paper is the fact that capital in Box 2 is difficult to estimate. Statistics Netherlands uses the book value of this wealth component, which it estimates at around €200 billion in 2017, while a recent government report arrives at double the figure by estimating the market value (Ministerie van Financiën 2020). Using this recent report, I will correct Statistics Netherlands' results since 1993; in particular, I will double the value of unlisted business wealth for all reported years. In Section 4, I discuss the implications of this correction.

1. The data from 2002 – 2005 is similar to the data from 2006 – 2010, since it was also based on a sample from the fiscal data. However, debts and checking accounts were incompletely registered, making this data less reliable, particularly for the bottom of the distribution.

2.2.2 Household Surveys

The first survey I use is the *Sociaal-Economisch Panelonderzoek* or Socio-Economic Panel Survey (SEP), which was initiated in 1984 by Statistics Netherlands, and ran until 2002. It first began asking questions about assets and liabilities in 1987, and continued doing so until its termination. The survey had a sample size of around 5,000 households. Its coverage of wealth components is quite comprehensive, with the notable exception of business wealth for self-employed individuals.

As a second survey, I use the *DNB Household Survey* (DNB), which was initiated in 1993, and is still active. This household survey has a significantly lower sample size, of around 2,000. Otherwise, it is very similar in structure and methodology to the SEP. Both household surveys ask respondents various questions about their assets and liabilities, which we can use to construct total wealth (in the survey), as well as the quantiles of wealth.

2.2.3 Rich Lists

There are two rich lsits available for the Netherlands. The first rich list is the *Forbes* List of Billionaires. This list covers all dollar billionaires worldwide since 1987, and is still active. Since its inception until 1987, there were typically two or three Dutch billionaires present on this list. Including these data points vastly improves our estimate of wealth inequality, as I will explain in section 4. Using *Forbes* carries the advantage of consistency: Using one rich list since 1987 makes the results of the different household surveys more comparable. Nevertheless, it is preferable to have a larger sample of the rich. Therefore, I will use the *Quote 500* as well. This rich list provides the 500 richest Dutch families and households for each year since 1997. Although this rich list has a broader coverage compared to *Forbes*, it has the disadvantage that its unit of observation has not remained consistent: After 2012, *Quote* began publishing a separate list for the richest families; moreover, in the years 2014 – 2016, it also published a separate list for expatriated Dutch individuals. To keep the information from the *Quote* as consistent as possible, I combine the separate lists into a single rich list for the relevant years[2].

These rich lists are based on publicly available information, such as Chamber of Commerce records and stock market data. These are combined with estimates about non-public wealth components, to arrive at an estimate of total wealth; hence, these lists present a valuable addition to traditional sources, although they have difficulty to estimate non-public wealth components, such as debts, bank deposits and offshore accounts.

One conceptual issue with rich lists is the unit of observation. *Forbes* principally includes individuals; *Quote* includes both individuals and families (although in separate lists from 2012 onward). Both units are not identical to households, the unit of observation in surveys. Individuals is closer, since individuals' wealth will typically be the bulk of that particular household's wealth. Families are more difficult to reconcile, since the size of the family is often unknown. For most practical purposes, this difference in unit of observation does not matter, since we mainly use rich lists to provide estimates on the properties of the right tail, rather than to arrive at precise household-by-household breakdowns. Nevertheless, the differences in unit of analysis is a source of heterogeneity between data sources that needs to be kept in mind when estimating the upper tail.

2. As a robustness check, I have run my analyses on the list of individuals only, excluding the lists of families and expatriates. The resulting top wealth shares are slightly lower in levels, but the trends are qualitatively identical.

3 Theoretical and Methodological Framework

With the data sources described in the previous section, we attempt to estimate the top tail of the wealth distribution over time. Although each source separately reveals interesting patterns, we can only approach the true right tail by combining or adjusting sources. I begin by describing the methodological framework, discussing Pareto interpolations and generalized Pareto interpolation. Next, I discuss how these estimation methods have been applied to estimate top wealth, using either single data sources or combinations of different sources.

Since Pareto (1896), it is known that the distribution of wealth has a heavy upper tail. For the upper tail, this exact Pareto distribution is defined as

$$\mathbb{P}(W > X_0) = \left(\frac{W}{X_0}\right)^{-\xi} \quad \text{for } W \in [X_0, \infty) \text{ and } X_0,\ \xi > 0 \tag{1}$$

where W refers to (total) wealth as defined in section 2, and X_0 is the threshold above which the Pareto distribution holds, which is also referred to as the scale parameter. $\mathbb{P}(W > X_0)$, therefore, is the probability of wealth level W exceeding threshold X_0 being found in the wealth distribution. The magnitude of this probability, and hence the thickness of the right tail, is governed by the exponent ξ, also known as the Pareto coefficient. A lower value of ξ means a thicker tail, and hence higher wealth inequality. Pareto distributions have several interesting characteristics (Gabaix 2009). For instance, Pareto distributions feature *fractal* inequality: The ratio of the wealth share held by the richest 0.1% to the share of the richest 1%, for instance, is equal to the ratio of the top 0.01% share to the 0.1% share, the 0.001% share to the 0.01% share, and so on. In particular, the conditional mean for the distribution above wealth level W_i equals

$$\mathbb{E}(W > W_i \mid W_i > X_0) = b \cdot W_i \tag{2}$$

where $b \equiv \xi/(\xi - 1)$ is the inverted Pareto coefficient, which is constant for an exact Pareto distribution (Atkinson, Piketty, and Saez 2011). For instance $b = 2$ means that the average wealth above €1 million equals €2 million, above €2 million equals €4 million, and so on. As a final useful fact, note that we can use the coefficient ξ to recover top wealth shares, using the fact that

$$S(p) = (100/p)^{1/\xi - 1} \tag{3}$$

where $S(p)$ refers to the share held by fractile p (Jones 2015). For instance, for $\xi = 1.5$, the top 1% share of wealth is approximately 22%. Note that this conversion only works if ξ is strictly greater than 1; for any values below 1, the wealth distribution is explosive, and top wealth shares approach infinity.

We can operationalize this distributional assumption as follows. Order all n observations by size of wealth, such that $W_{(1)} \geq ... \geq W_{(n)}$, with n the cutoff point corresponding to X_0, i.e., the smallest value in the upper tail for which the Pareto distribution holds. Then we can take logs of both rank and size to obtain a simple regression equation

$$\ln(r) = \underbrace{\ln(n) + \xi \ln(X_0)}_{= \text{ Constant } C} - \xi \ln(W) \tag{4}$$

with r referring to the rank of the observation. This equation delivers an estimate for ξ, which we can use to reconstruct the top wealth share. Gabaix and Ibragimov (2011) argue that equation 4 is prone to small-sample bias, and hence propose a correction to the left-hand side $\ln(r - 1/2)$, which removes the bias up to first order. Equation 4 and its adjusted form by Gabaix and Ibragimov (2011) are both discretized versions of the distribution; for quasi-continuous data, we can use $\ln \mathbb{P}(W > X_0)$ instead of $\ln(r)$. Then, the term $\ln(n)$ also disappears, which means that we can directly estimate both X_0 and ξ; this is a fact of which I will make extensive use in developing Robust Pareto Regression in Section 5.

An alternative method to OLS to estimate ξ is Maximum Likelihood – the so-called Hill estimator – which results in (Vermeulen 2018)

$$\hat{\xi}_{\text{ML}} = \left[\ln \mathbb{E}(W > W_i \mid W_i > X_0) - \ln(X_0)\right]^{-1} \tag{5}$$

Note that the Maximum Likelihood estimator 5 can be thought of as the sample analogue to equation 2, with the difference that ξ is not estimated as $b/(b-1)$ but rather as the inverse of $\ln b$.

Although strict Pareto interpolation – either using OLS or Maximum Likelihood – is simple and widely applied (see e.g., Vermeulen 2018), it is not without criticism. The main assumption is that wealth follows an exact Pareto distribution, which may be too stringent for real-world data. Hence, Blanchet, Fournier, and Piketty (2017) propose to estimate wealth shares with *generalized Pareto curves*, which they define as the curve of inverted Pareto coefficients $b(p)$, where $p \in [0, 1]$ is the rank, and $b(p) \coloneqq \xi(p)/(\xi(p) - 1)$ is the ratio between average wealth above rank p and the p-th threshold $W(p)$:

$$b(p) = \frac{\mathbb{E}[W > W(p) \mid W(p) > X_0]}{W(p)} \tag{6}$$

For an exact Pareto distribution, $b(p)$ is constant and equal to the b derived in equation 2; however, Blanchet, Fournier, and Piketty (2017) argue that the coefficient of $b(p)$ varies between quantiles. This more general estimation method takes as input information on the location of fractiles p, corresponding thresholds $W(p)$ and bracket averages $\mathbb{E}(W|W(p+1) > W > W(p))$. Using these inputs, the authors interpolate all other values of the distribution via fifth-degree splines, with the parameters for the splines chosen by a linear programming algorithm[3]. Comparing both methods, Albers, Bartels, and Schularick (2020) find that strict Pareto methods result in a higher inequality estimate than the generalized interpolation method for Germany.

Note that the Maximum Likelihood estimator 5 and the generalized Pareto interpolation method 6 are different ways of summarizing the same information, namely the average wealth above a threshold. To see this, observe that $X_0 = Q(p)$ for the percentile corresponding with the minimum quantile X_0. The main difference is that Maximum Likelihood log-transforms the inverse coefficient b before inverting it, whereas the generalized Pareto interpolation method does not. Therefore, we should expect that the two methods yield similar estimates for ξ for a given X_0, at least for sample sizes large enough. Indeed, this is shown in Table 1, which compares the two methods for the *Forbes List of Billionaires*. Although this dataset is hampered by small observations, we can still conclude that the two methods are highly comparable. One notable difference between the two methods concerns the fact that a logarithmic transformation compresses the support of the distribution, and hence reduces the influence of outliers on the estimate of ξ.

The three estimation methods discussed above are all commonly used in the literature; as of yet, no explicit discussion exists of the advantages or disadvantages of particular methods compared to others. The methods are typically applied in isolation, to single data sets. Examples include Bach, Thiemann, and Zucco (2019) (straight OLS regression), Jacobs, Jongen, and Zoutman (2013) (Maximum Likelihood), and Garbinti,

3. For a full description of the methodology, see Blanchet, Fournier, and Piketty (2017), section 3.

Table 1: Comparison of Generalized Pareto Interpolation and Maximum Likelihood Using *Forbes*

		Estimator	
Year	**N**	**gpinter**	**ML-Hill**
1993	3	1.50	1.44
1994	3	1.69	1.80
1995	3	2.71	3.74
1996	3	4.57	6.55
1997	4	2.30	2.59
1998	2	2.53	3.98
1999	4	5.93	7.38
2000	3	2.25	2.64
2001	2	1.54	1.91
2002	2	1.48	1.78
2003	4	1.94	2.30
2004	4	1.95	2.23
2005	4	1.74	1.72
2006	4	1.62	1.48
2007	4	1.42	1.21
2008	5	1.71	1.79
2009	3	2.20	2.87
2010	5	1.62	1.62
2011	6	1.64	1.62
2012	6	1.73	1.88
2013	6	1.65	1.90
2014	7	1.58	1.59
2015	9	1.50	1.38
2016	9	1.42	1.23
2017	10	1.55	1.50
2018	9	1.55	1.60

Note: `gpinter` estimates are calculated by a direct computation of $b = \mathbb{E}[W|W > X_0]/X_0$, where X_0 is the minimum value of the *Forbes* list; the reported estimate is then computed as $\xi = b/(b-1)$. The estimates for `ML-Hill` are estimated using equation 5.

Goupille-Lebret, and Piketty (2020) (generalized Pareto interpolation). In addition, several authors use one of these methods to *combine* different data sources. This is commonly done in the context of household surveys.

Household surveys represent a sample drawn from the wider population. If the variables measured in the sample are normally distributed in the population, a reasonably large sample size will ensure that the sample estimates are representative. However, if variables follow a heavy-tailed distribution – such as wealth – particular care is needed to ensure representativeness. In a simple sample, rich individuals are significantly less likely to act as participants. This results in a downward bias in parameter estimates, known as *differential unit non-response bias*. For this reason, rich lists represent an essential source of information, because they contain estimates of the very top of the distribution that are otherwise missing from survey data. Examples of studies which combine surveys with rich lists include Vermeulen (2018) (OLS and Maximum Likelihood), Blanchet (2016), and Albers, Bartels, and Schularick (2020) (generalized Pareto interpolation).

The main challenge when combining surveys and rich lists is in estimating the gap between the maximum value of the survey and the minimal value of the rich list. When estimating this gap with OLS or Maximum Likelihood, this is quite straightforward, since we can simply pool the data from the rich list and the data from the survey and apply these methods as before. Generalized Pareto interpolation, however, we need to make additional assumptions, since the bracket average of the gap between the survey and the rich list is missing and must be estimated separately. The *R* software package `gpinter`, which I use to compute generalized Pareto interpolation, offers the possibility to apply this method without bracket average, in effect fitting a generalized Pareto distribution based on the fractiles p and thresholds $W(p)$ alone. The resulting estimate of the bracket average can then be combined with all known bracket averages to re-estimate the generalized Pareto interpolation method.

4 Individual Data Sources And Simple Combinations

In this section, I compare the various different data sources and methodologies discussed in the previous two sections. First, I discuss the adjustments to household surveys. I begin by analyzing the need to calibrate surveys with fiscal data; then, I compare and contrast three different strategies to combine surveys with rich lists – OLS, Maximum Likelihood, and generalized Pareto interpolation. Second, I discuss the adjustments to fiscal data based on correcting for the value of closely held wealth. Although all these adjustments help in improving estimates of the top tail, we will see that they are imperfect as well.

4.1 Adjusting Surveys

It is well known that simple random survey designs have difficulty to capture the right tail of the wealth distribution, due to differential unit non-response bias. What is new in the Dutch context, however, is the fact that the two available surveys also have difficulty to capture the *left tail*: Both surveys provide too few negative values. Moreover, the summary statistics (Tables 2 and 3) reveal that the moments of the survey distributions vary strongly from year to year. This is a particular issue for the *DNB*, which has the additional problem of a low sample size. If we compare these summary statistics to information on the wealth distribution from Statistics Netherlands (summarized in Table A.1 in the Online Appendix), it becomes immediately apparent that the surveys are widely unrepresentative both at the bottom and the top of the distribution.

Table 2: Summary Statistics, *Sociaal-Economisch Panelonderzoek*

Year	N	P05	P25	Median	Mean	P75	P95	SD
1993	5,185	-7,390	1,000	18,924	78,375	103,160	337,544	153,069
1994	5,187	-7,855	1,620	27,115	93,857	130,092	376,986	166,818
1995	5,109	-6,494	1,750	29,314	105,538	150,394	417,982	183,865
1996	5,179	-9,154	3,700	45,030	122,706	173,500	459,669	211,211
1997	5,049	-9,456	4,200	56,046	133,328	196,331	500,000	210,612
1998	4,963	-7,712	5,700	72,500	165,412	238,722	578,900	308,438
1999	5,022	-8,362	6,549	90,618	186,216	273,550	671,072	272,335
2000	5,007	-8,198	7,940	115,415	227,133	332,575	812,450	350,137
2001	4,851	-9,000	9,206	139,324	258,803	378,236	900,965	370,382
2002	4,413	-4,226	4,991	70,067	127,246	181,580	439,114	192,172

Note: Values in constant 2002 EUR.

Table 3: Summary Statistics, *DNB Household Survey*

Year	N	P05	P25	Median	Mean	P75	P95	SD
1993	2,694	-2,169	18,052	111,326	188,748	259,961	645,853	272,210
1994	2,834	-10,500	7,830	39,975	105,494	129,026	445,807	199,843
1995	2,764	-2,494	22,968	124,759	195,533	269,016	672,073	269,632
1996	2,517	-1,298	29,490	141,000	211,517	294,709	692,802	270,794
1997	2,150	-4,085	20,150	145,837	260,828	302,245	725,100	1,941,956
1998	1,754	-6,761	2,218	108,867	191,493	274,964	659,891	335,641
1999	1,526	-9,373	2,484	50,810	144,239	212,421	549,590	238,029
2000	1,524	-10,377	51	40,025	191,842	274,670	770,585	371,124
2001	1,641	-10,939	7,647	89,930	239,896	330,064	900,289	424,469
2002	1,638	-14,254	13,512	139,924	290,004	407,686	1,064,733	446,884
2003	1,764	-6,838	8,924	85,460	141,083	208,683	475,560	197,488
2004	1,731	-9,465	10,130	85,667	145,277	215,348	478,703	214,188
2005	1,785	-5,741	9,400	88,991	155,566	223,089	502,448	230,352
2006	1,725	-5,525	10,350	87,285	161,300	236,300	541,685	236,693
2007	1,635	-5,898	13,480	103,766	176,043	255,089	571,854	269,268
2008	1,519	-5,008	16,405	110,400	183,223	273,114	605,021	262,989
2009	1,529	-3,977	13,250	93,783	174,006	257,130	576,399	253,898
2010	1,681	-5,085	16,802	131,119	193,482	290,371	611,710	261,872
2011	1,594	-4,578	20,761	152,154	209,756	306,738	638,725	271,845
2012	1,739	-11,254	13,368	132,325	193,545	281,836	590,770	297,806
2013	1,810	-20,000	5,638	79,358	149,838	231,400	528,468	271,591
2014	1,826	-20,324	8,406	68,202	140,844	219,692	513,058	257,296
2015	1,935	-15,767	11,476	102,355	166,884	249,711	568,463	273,051
2016	1,961	-19,300	10,500	103,092	173,198	250,150	569,801	265,438
2017	2,038	-16,207	7,961	85,053	164,015	241,512	569,801	250,975
2018	1,888	-16,443	8,197	85,972	161,749	242,509	561,208	242,777

Note: Values until 2002 in NLG; from 2002 in EUR. All values are in nominal terms.

As a result of this unrepresentativeness, existing strategies to derive top wealth shares from survey data fall short. Most papers take the survey data at face value until the upper percentiles, and make assumptions to correct the data above said upper percentiles. Albers, Bartels, and Schularick (2020), for instance, use a

German survey, which they take as given until the 99th percentile, and make Pareto-based corrections above the 99th percentile, following Bartels and Metzing (2019). The summary statistics for the Dutch surveys reveal, however, that this strategy must be carefully evaluated. If surveys oversample the tails – as is the case in the German context – these problems will be alleviated; simple random surveys such as the *DNB* and *SEP*, however, need to be calibrated in order to produce reliable results.

To this end, I use available quantiles from Statistics Netherlands to calibrate the quantiles for the surveys. This means, for instance, that I use the fiscal series for the 'true' values of the 25th percentile, median, 75th percentile, and so on. Using these calibrated percentiles, I can carry out all strategies to combine surveys with rich lists. Had we not calibrated these surveys, the computed bracket averages would be incorrect, which might systematically bias the top wealth shares. Calibrating survey values with fiscal data will improve the representativeness of the surveys. One consequence of the necessity for calibration is that we have to discard values of the *SEP* and *Forbes* before 1993, since that is the first year for which Statistics Netherlands publishes distributional data.

Once calibrated in this way, I combine the surveys with the two rich lists, *Forbes* and *Quote*. As discussed in section 3, there are various ways to combine surveys and rich lists. In Table 4, I compare three methods: OLS regressions, Maximum Likelihood, and generalized Pareto interpolation. I do so for both the *SEP* and *DNB*, both combined with *Forbes*[4]. For all methods, I calculate the local Pareto coefficient above the threshold of €1 million. This threshold value has often been used as a threshold for the Pareto distribution (Vermeulen 2018); moreover, it roughly corresponds to the 99th percentile in most years, which is another threshold often chosen for the Pareto distribution (e.g., Albers, Bartels, and Schularick 2020). For the generalized Pareto interpolation method, which delivers a range of Pareto coefficients, I report the average Pareto coefficient above €1 million.

The results reported in Table 4 reveal that OLS regressions result in very low Pareto coefficients. In particular, for every year and for both surveys, the regression method results in a coefficient $\xi < 1$. Such low coefficients indicate an explosive top wealth share which approaches infinity, which is both theoretically impossible and practically unhelpful. This problem is not driven by the sample size of the survey, since both the results for the *DNB* and the *SEP* suffer from this problem. We can conclude that using simple regressions to estimate top wealth shares results in implausible low Pareto coefficients. By contrast, Maximum Likelihood methods fare slightly better: In many years, the reported coefficients at least exceed 1. However, these results are highly sensitive to the underlying survey data, as the large variance of coefficients indicates. The ML-Hill coefficient fluctuates between 0.2 and 1.2 for the *DNB*, within the span of three years. A comparison with the estimates for the *SEP* show that this variability is a result of the sample size of the *DNB*; indeed, the Maximum Likelihood coefficients for the *SEP* are much more stable, and mostly above 1. Finally, the generalized Pareto method performs far better than both other methods. The reported coefficients driven by this method are both more stable, and more plausible; all values are moreover above 1. A comparison between the generalized Pareto coefficients of the two surveys also shows that this method is highly robust to the underlying sample size[5].

In summary, after comparing these three commonly used methods, I conclude that the generalized Pareto interpolation performs best. Moreover, this method results in lower top wealth shares than the other two methods; this finding corroborates the conclusion by Albers, Bartels, and Schularick (2020) for Germany. However, close inspection of the underlying data reveals that even generalized Pareto methods must be

4. The results are qualitatively identical when the surveys are combined with *Quote*.
5. Although this stability has likely benefited from the calibration with fiscal data.

Table 4: Estimated Pareto Coefficient ξ, Surveys + Rich Lists

	DNB + Forbes			**SEP + Forbes**		
Year	**OLS**	**ML-Hill**	**gpinter**	**OLS**	**ML-Hill**	**gpinter**
1993	0.42	1.31	1.85	0.31	0.72	1.37
1994	0.30	0.70	2.05	0.37	1.10	1.38
1995	0.41	1.19	1.36	0.39	1.05	1.35
1996	0.41	1.36	1.29	0.43	1.36	1.29
1997	0.36	0.96	1.27	0.38	1.09	1.27
1998	0.43	1.27	1.34	0.62	1.84	1.35
1999	0.25	0.52	1.27	0.51	1.67	1.27
2000	0.38	1.02	1.33	0.66	1.92	1.32
2001	0.54	1.45	1.43	0.92	2.29	1.42
2002	0.70	1.77	1.48	0.44	1.49	1.48
2003	0.29	0.63	1.52			
2004	0.30	0.62	1.47			
2005	0.31	0.71	1.45			
2006	0.30	0.67	1.47			
2007	0.34	0.78	1.54			
2008	0.29	0.56	1.46			
2009	0.35	0.84	1.50			
2010	0.31	0.62	1.49			
2011	0.31	0.65	1.45			
2012	0.32	0.61	1.46			
2013	0.27	0.51	1.44			
2014	0.25	0.42	1.41			
2015	0.27	0.47	1.41			
2016	0.29	0.55	1.42			
2017	0.27	0.48	1.41			
2018	0.25	0.41	1.39			

Note: OLS refers to regressions of the form of equation 4; ML-Hill refers to the equivalent estimate by the Hill estimator; gpinter refers to generalized Pareto interpolation. All coefficients are reported for a threshold of €1 million.

viewed with skepticism. This is illustrated in Figure 1, where I have plotted all values of *SEP* above €1 million, as well as all values of *Forbes*, for the year 2002. The regression line drawn between the two groups of data corresponds to the estimate of ξ reported in Table 4 for 2002. It is immediately clear that this line fits the two data sources poorly. Maximum Likelihood methods, which also deliver a single estimate for ξ, will not fare significantly better. Crucially, however, generalized Pareto methods also fall short to capture the heterogeneity that we see. The survey data has a steep slope, indicating that the survey predicts a thin tail. The two data points from *Forbes*, meanwhile, have a slope that is much less steep. Generalized Pareto interpolation would attempt to fit a polynomial on both data sources, as well as the gap in between, which best fits the data. This would result in a line for the gap that is approximately flat. This is an unlikely result, which seems driven by the inherent differences between data sources. This may be specific to the Dutch context, since surveys in other countries often have better coverage at the top due to oversampling (Vermeulen 2018). Nevertheless, estimating the large gap between the top of the survey and the bottom of the rich list with a single OLS fit seems problematic even in this context, given that this does not acknowledge the inherent heterogeneity in these data sources.

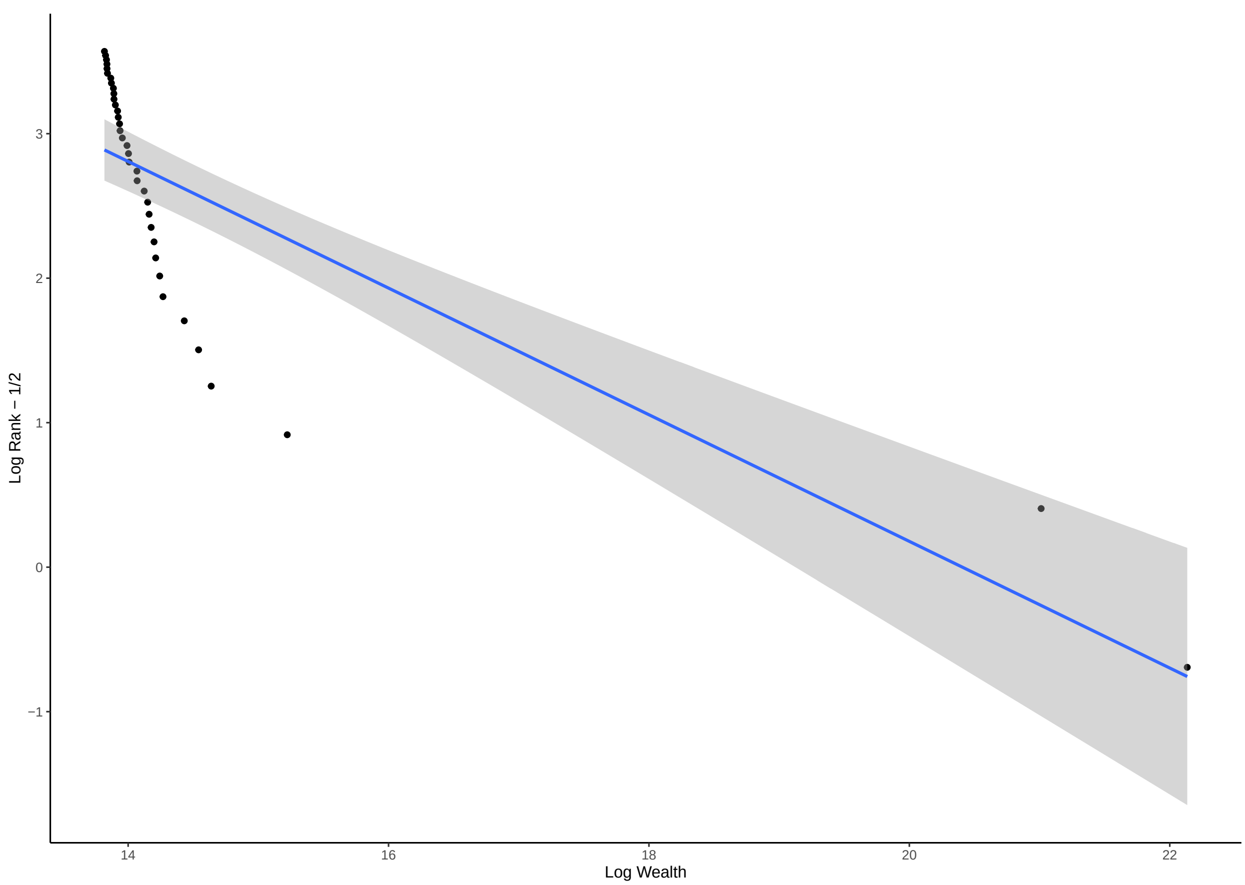

Figure 1: Combining *SEP* with *Forbes*, 2002.

4.2 Adjusting Fiscal Data

As discussed in Section 2, even the fiscal data compiled by *Statistics Netherlands* has trouble measuring certain wealth components. Over time, the coverage of liabilities has improved, as can be glimpsed from the summary statistics in Table A.1, where we observe that after 2010, the bottom percentiles are adjusted downward. This adjustment is mostly due to better coverage of small debts, although the economic crisis doubtlessly also played a major role. Nevertheless, these statistics do not include the value of closely held business at market value, but at book value, as discussed in Section 2. To correct for this, I use the report by the Ministry of Finance, which re-estimates the size of business wealth (*aanmerkelijk belang*, lit. "significant ownership"). This wealth component comprises shares of more than 5% in unlisted limited liability companies. For 2017, the report estimates the total value of *aanmerkelijk belang* at €400 billion, double the figure given by Statistics Netherlands (Ministerie van Financiën 2020). Based on this estimate, I double the value of unlisted business equity across the distribution, for all years (1993 – 2018). Of course, it is far from certain whether the market value of this wealth component was always double the book value estimated by Statistics Netherlands; nevertheless, given the available information, this represents the most natural assumption. Moreover, there are several reasons to have confidence in this exercise. First, the report mentions that the number of *directeur-grootaandeelhouders*, director-majority shareholders, has remained stable over the years, around 250,000. Since directors-majority shareholders represent the bulk of this wealth component, this suggests that the structure of this component has remained quite stable over time. However, the total number of individuals with *aanmerkelijk belang* has been revised upward from the original statistics, and has been increasing over time. This suggests that the official statistics miss many households with small values of unlisted business equity. This is partly addressed, since this exercise doubles wealth along the distribution. Since this wealth component is predominantly concentrated among the wealthiest 1%[6], doubling *aanmerkelijk belang* primarily affects top wealth shares. These and other issues are discussed in Toussaint et al. (2020).

To correct for the underestimation of closely held business wealth in fiscal data, I double the value of this wealth component along the distribution. Since this wealth component is predominantly owned by the richest 1%, this exercise leaves the values of lower quantiles largely unaffected. It does, however, significantly adjust top wealth shares upward. Figure 2 shows that correcting for closely held businesses results in top 1% wealth shares that are adjusted upward about 5–8 percentage points. The increase of the top 0.1% share is of similar magnitude . This is a significant increase, and begs the question whether this adjustment is valid. As discussed in section 2, we know that this adjustment is approximately valid for the years around 2017. For earlier years, the exact ratio of market value to book value for *aanmerkelijk belang* will likely not be exactly equal to 2. Unfortunately, other common strategies to estimate this wealth component, such as multiplying corporate profits with discounted p/e-ratios (Albers, Bartels, and Schularick 2020), are not feasible in the context of the Netherlands due to data availability.

Nevertheless, there are several reasons to be confident in the rough order of magnitude of the new series. First, as discussed above, most owners of this wealth component are so-called director-majority shareholders of unlisted corporations. The number and makeup of director-majority shareholders has not changed significantly over time, suggesting that no dramatic changes have taken place. As a further robustness check to validate my doubling exercise, I use the detailed distributional data on the top 0.1% share provided by Statistics Netherlands. In this group, I double unlisted business wealth, and then use generalized Pareto techniques to estimate the threshold values and average wealth of the richest 500 (the top 0.65% of the top 0.1%). As described in Blanchet, Fournier, and Piketty (2017), this amounts to fitting a Generalized Pareto Distribution to the values beyond the final threshold; this fitted distribution can then be used to estimate

6. According to the report by the Ministry of Finance, as well as the detailed data on the wealth distribution provided by Statistics Netherlands, the richest 1% own about 75% of this wealth component, and the richest 0.1% about 50%.

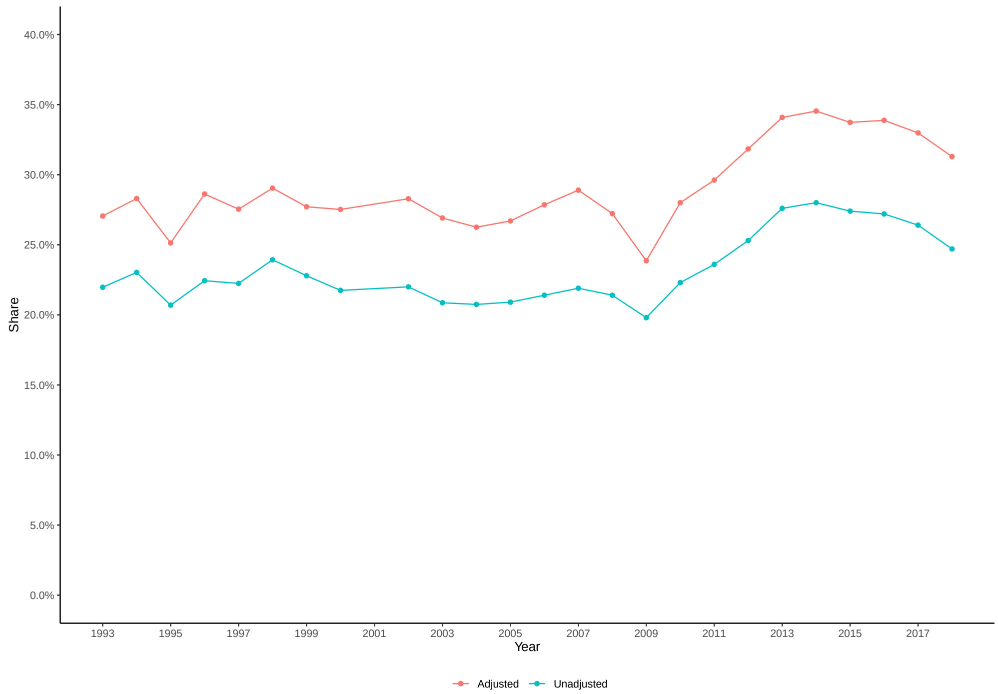

Figure 2: Adjusting Fiscal Data for Closely Held Wealth, Top 1% Share

the threshold and average wealth at the fractile where the *Quote* list begins. I compare these predicted values with the actual values reported by the Quote 500. The results are collected in Table 5. Note that for this exercise, I have not consolidated the *Quote* lists of list of individuals, families, and expatriates, as I do for the wealth shares. By directly comparing the extrapolated estimates from adjusted fiscal data with non-expatriated individuals, the units of observation are the most similar.

Table 5: Comparing results with *Quote 500*

Year	**Threshold Value**		**Average Wealth**	
	Predicted	**Actual**	**Predicted**	**Actual**
2011	37	55	188	287
2012	38	50	193	213
2013	38	54	162	226
2014*	39	35	170	167
2015*	40	41	171	184
2016*†	29	25	112	117
2017	45	71	207	292
2018	47	80	195	326

Note: Values in millions of EUR. Extrapolation based on distributional statistics of top 0.1% percentile provided by Statistics Netherlands. Before 2012, the *Quote 500* covered both families and individuals; after 2012, only individuals. In years with a *, expatriated Dutch individuals are excluded from the list. In 2016 (†), the list contained 1,000 individuals instead of 500; this explains the lower threshold value.

From the table, we can conclude that our exercise matches the threshold value quite well, especially in the years (2014—2016) where Quote excluded expatriated Dutch individuals. In those years, the rich list matches most closely with fiscal data, which is encouraging. The average wealth is underestimated quite severely, however. This is an indication that wealth within the richest 0.0065% of the population is distributed even more unequally than in the top 0.1%. This is confirmed by inspecting Pareto coefficient ξ. For the richest 0.0065%, $\xi \approx 1.2$, much lower than the Pareto coefficients found for the 99th percentile reported in Table 4.

Because of the many methodological changes in the construction of the *Quote 500*, it is difficult to use the rich list as an independent source to evaluate the dynamics of top wealth. Therefore, this list should not be interpreted on its own, but only in relation to the estimated values based on the adjusted fiscal data.

5 Robust Pareto Regression

The previous section demonstrated that individual data sources are insufficient to estimate top wealth. Adjusting surveys by calibrating them to fiscal data partially addresses this concern, as does correcting fiscal data for unlisted wealth. Nevertheless, these results are strongly dependent on the specific functional assumptions made. For instance, it might be that a different choice of threshold would alter the results in Table 4; likewise, if the ratio of corrected unlisted wealth to uncorrected unlisted wealth is not equal to 2, corrected fiscal data might look rather different. What we would like is a way to combine all available information from different data sources in a consistent manner, while correcting for the underlying heterogeneity of data sources and methods. In this section, I develop a new methodology to do just that. The aim of this method, which I term *Robust Pareto Regression*, is to estimate averaged parameters of the Pareto distribution, as well as their development over time and between datasets. First, I develop the theoretical framework, and show how I use this to obtain robust estimates for the Pareto parameters. Next, I discuss how I implement this method for the Dutch data sources, and present the results.

5.1 Theoretical Framework and Identification

Recall from Section 3 that a Pareto distribution is governed by two parameters: The Pareto coefficient ξ, and the scale parameter X_0. Our aim is to robustly estimate these parameters by combining individual estimates from different years and data sources. To that end, I present some useful properties of the Pareto functional form. In particular, letting lowercase variables denote logs, we have that if a stochastic variable X is Pareto distributed, then x is exponentially distributed:

$$\mathbb{P}\left(w > x | w > x_0\right) \equiv \mathrm{P}_{kjt} = e^{-\xi(x-x_0)} \tag{7}$$

$$\mathbb{E}\left(w | w > x\right) \equiv \mathrm{E}_{wx} = x + \xi^{-1} \tag{8}$$

We use these relationships to derive linear relationships for the log-transformed variables across datasets j and years t. For each j and t, using one of the three methods described in Section 3, we can estimate the local Pareto parameters for different thresholds w_{kjt}, where k indexes wealth classes or thresholds. For instance, for $k =$ €1 million, we can use data from the data source $j = DNB$, using either linear regression or Maximum Likelihood[7]. We use these local estimates as new data points for our robust Pareto regression.

7. Recall that the Maximum Likelihood estimator and the generalized Pareto interpolation method are intimately related; therefore, we have cannot use both methods simultaneously for a data source. In practice, since the household surveys are very

We have K_{jt} datapoints ($k = 1, ..., K_{jt}$) of the form $\{\mathrm{P}_{kjt}, w_{kjt}, \mathrm{E}_{kjt}\}$, where w_{kjt} is the lower bound for w of wealth in class k, P_{kjt} is the fraction of individuals having a higher wealth than the lower bound, and where E_{kjt} is the conditional mean of the log-transformed variables, defined as

$$\mathrm{E}_{kjt} \equiv \mathbb{E}\left(w | w > w_{kjt}\right)$$

Using equations 7 and 8, we can derive the following linear relationships:

$$\begin{aligned} \ln \mathrm{P}_{kjt} &= -\xi\left(w_{kjt} - x_0\right) = -\xi w_{kjt} + \chi, && (9) \\ \left(\mathrm{E}_{kjt} - w_{kjt}\right)^{-1} &= \xi && (10) \end{aligned}$$

where $\chi \equiv \xi x_0$ is the effective lower bound of the full Pareto distribution. These equations should look familiar, since they are the aggregative analogues of equations 4 and 5, which we have used for each wealth class k, data source j and year t as inputs for these aggregative relationships. Since each data point is a triple representing an estimate of $w_{kjt}, \ln \mathrm{P}_{kjt}$, and E_{kjt}, we can plug in all these estimates into a large regression. Then, in regression equation 9, the coefficient on $-w_{kjt}$ identifies ξ, and the intercept identifies χ. Likewise, since we know both endogenous terms on the left-hand side of equation 10, if we regress these values on a constant, we have identified ξ as well.

Since there may be underlying heterogeneity between years and data sources, running pooled regressions as described above would not be valid. Indeed, we have seen in previous sections that different data sources yield radically different estimates of top wealth. We correct for this by incorporating both year and data fixed effects into the identifying equations:

$$\ln \mathrm{P}_{kjt} = -\xi_1 w_{kjt} - \sum_{j=2}^{5} \xi_j w_{kjt} \times \delta_j + \delta_j - \sum_{t=1994}^{2018} \xi_t w_{kjt} \times \gamma_t + \gamma_t + \varepsilon_{kjt} \quad (11)$$

$$\left(\mathrm{E}_{kjt} - w_{kjt}\right)^{-1} = \delta_j + \gamma_t + \varepsilon_{kjt} \quad (12)$$

where δ_j and γ_t represent data and year fixed effects, and ε_{kjt} is an idiosyncratic error term. Since using dummies for all years and data sources would result in perfect multicollinearity, we have to set one of the data sources and one of the years to 0. This is why we include the term $-w_{kjt}$ beside the interaction terms, since ξ_1, the coefficient on $-w_{kjt}$, identifies the value of ξ for the data source and year we have set to 0. Likewise, the value of the intercept of this regression identifies the value of χ for the data source and year set to 0. Using these benchmark values for ξ and χ, we can derive values for differences in ξ across data sources and years – identified by the coefficients for the interaction terms $-w_{kjt} \times \delta_j$ and $-w_{kjt} \times \gamma_t$ – and differences in χ – identified by the coefficients on the fixed effects δ_j and γ_t. As a robustness check for this method, we can run separate regressions using equation 12. For these equations, the coefficients on the fixed effects identify ξ, not χ. We can compare the estimated ξ from equation 11 with the estimate from 12; if the fixed effects have fully absorbed all underlying heterogeneity, the two equations should result in roughly the same value for ξ.

thin-tailed, generalized Pareto interpolation estimates for top wealth based solely on the survey are highly unreliable anyway, so there is little loss in their exclusion.

5.2 Implementation and Results

For each data source, I use the methods developed in Section 3 to construct the local estimates of the Pareto coefficients. For the household surveys *DNB* and *SEP*, I use increasing thresholds $w_k = \{500,000;$ 1,000,000; 2,000,000; ...$\}$, with the condition that at there are at least 10 households at or above the threshold[8]; at each threshold, I use OLS and Maximum Likelihood to estimate the coefficients. For *CBS*, we have data on two thresholds, $w_k = \{500,000; 1,000,000\}$, for which the relevant information is computed using generalized Pareto interpolation. For the rich lists, we take the threshold w_k to be the minimal value in the list, and compute all relevant variables directly[9]. This results in 252 observations for 26 years, over the 5 datasets. To get a sense of the data I have aggregated, we can turn to Figure 3, which plots the values of $\ln \mathrm{P}_{kjt}$ against w for the year 2002. For this year, we have multiple values of all five data sources, making it the ideal year to illustrate the method.

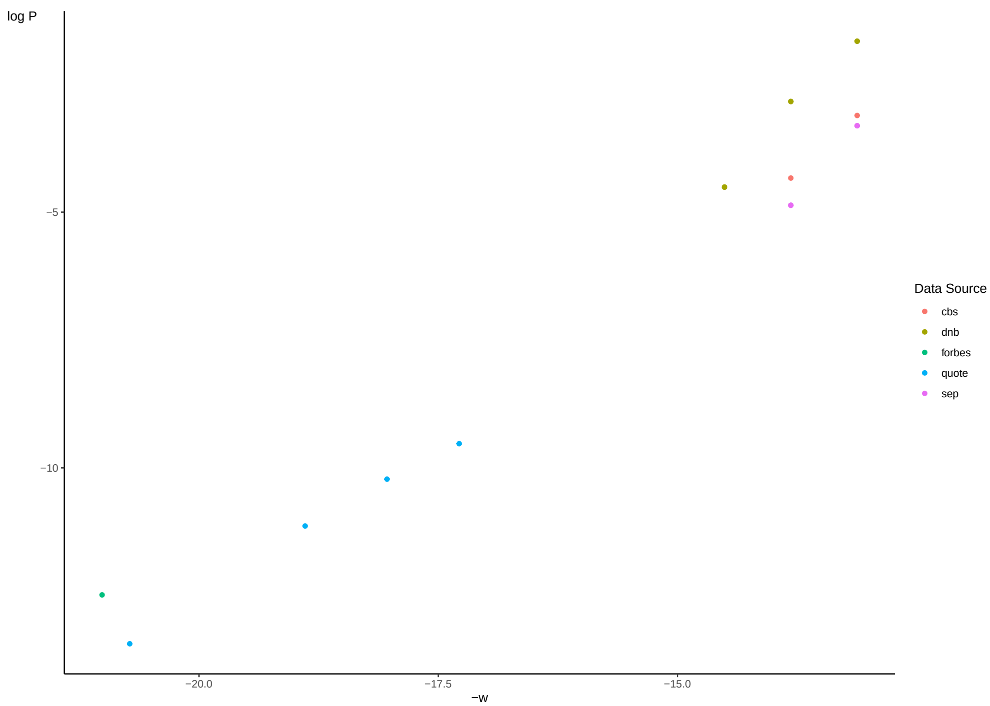

Figure 3: Different Estimates of $-w_{kjt}$ and $\ln \mathrm{P}_{kjt}$ for the Year 2002

We see that there are different thresholds w_{kjt}, varying from approximately 13 log points to 21. The estimates, colored by data source, align remarkably well in a linear relationship. In our general regression, the slope of the regression line through the data points belonging to *CBS* represents the estimate of ξ for 2002, since we have taken *CBS* as the baseline. The intercept of this regression line represents the estimated average value of χ. The year fixed effect for 2002 captures the difference between 2002's value for χ and 1993's, and the interaction effect between the year fixed effect and $-w$ does the same for ξ. Similarly, the data fixed effects and their interactions represent the deviations of regression slopes and intercepts from the regression slope for estimates taken from *CBS*. Specifically, the value of the interaction of a data fixed effect

8. All values are expressed in EUR; to convert values from data sources in other currencies to EUR – i.e., USD for *Forbes* and NLG for *DNB* and *Quote* before 2002 – I use the end-of-year exchange rate.

9. The *Quote 500* has published tabulated data for the years 2001 – 2008, i.e., information on the wealth held by the top 10, top 100, and top 250. These wealth classes represent additional thresholds w_k, which I also include in my estimates.

with $-w$ represents the difference in slope, and the value of the data fixed effect represents the difference in intercepts, multiplied by the estimate of ξ for that data source. This can be seen in Figure 4.

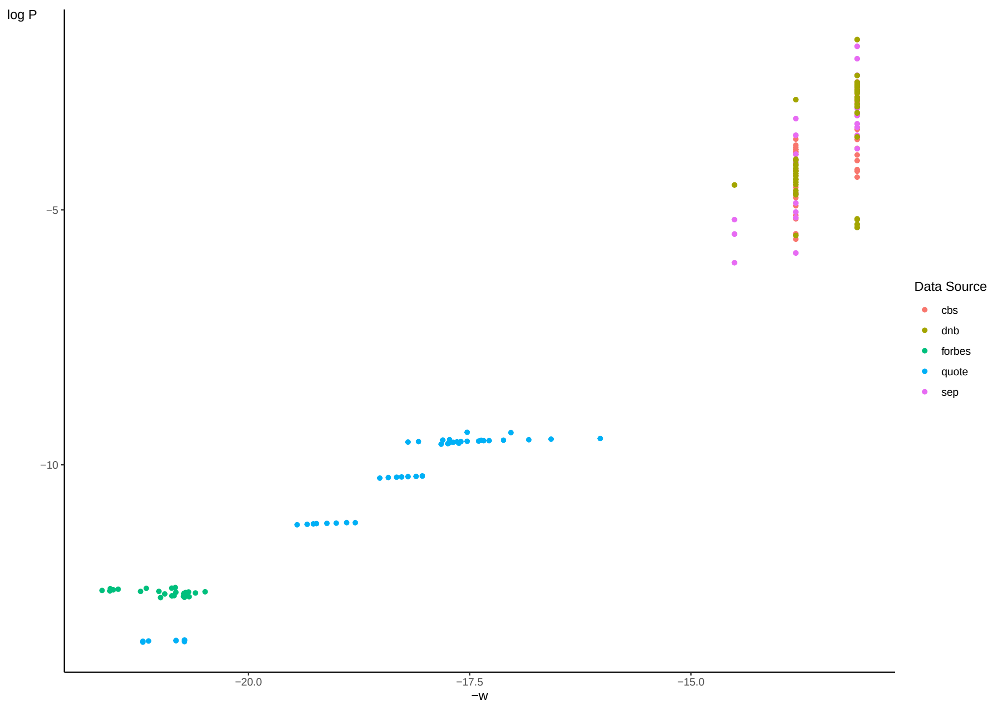

Figure 4: Different Estimates of $-w_{kjt}$ and $\ln \mathrm{P}_{kjt}$ per Data Source

Table 6 reports the results. To avoid multicollinearity, I drop the values for *CBS* and the year 1993. This means that the intercept for models (1)–(4), combined with the coefficient for $-w$, can be interpreted as the estimated average value of $\ln \mathrm{P}_{kjt}$ for the data source *CBS* in the year 1993. All other coefficients should be interpreted *relative* to this benchmark. I report heteroskedasticity-robust standard errors; using clustered standard errors at the data–year level does not materially alter any of my results.

Table 6: Robust Pareto Regressions

	Dependent variable:				
	$\ln \mathrm{P}_{kjt}$				$(\mathrm{E} - w_{kjt})^{-1}$
	(1)	(2)	(3)	(4)	(5)
Constant	18.655***	18.446***	19.496***	16.048***	1.991***
	(2.914)	(2.383)	(2.617)	(2.357)	(0.207)
$-w$	1.724***	1.724***	1.791***	1.554***	
	(0.216)	(0.177)	(0.195)	(0.174)	
$-w \times$ DNB	−0.141	0.071	0.041	0.096	
	(0.266)	(0.219)	(0.240)	(0.204)	
$-w \times$ Forbes	−2.658***	−2.773***	−3.222***	−2.632***	
	(0.382)	(0.326)	(0.382)	(0.560)	
$-w \times$ Quote	−0.695***	−0.615***	−0.581***	−0.636***	
	(0.224)	(0.184)	(0.202)	(0.174)	
$-w \times$ SEP	0.384	0.601***	0.528**	0.656***	
	(0.271)	(0.225)	(0.247)	(0.210)	
DNB	−2.024	0.807	0.408	1.139	1.389***
	(3.581)	(2.945)	(3.238)	(2.745)	(0.114)
Forbes	−51.615***	−54.018***	−63.437***	−51.041***	0.298**
	(7.207)	(6.214)	(7.378)	(11.434)	(0.148)
Quote	−11.114***	−9.895***	−9.338***	−10.148***	−0.286**
	(3.110)	(2.551)	(2.816)	(2.447)	(0.133)
SEP	5.719	8.662***	7.628**	9.384***	1.043***
	(3.672)	(3.035)	(3.336)	(2.843)	(0.145)
Year Trend	0.062***				
	(0.005)				
Year FE			✓	✓	✓
$-w \times$ Year FE		✓		✓	
Observations	252	252	252	252	252
R^2	0.980	0.988	0.985	0.991	0.645
Adjusted R^2	0.979	0.986	0.983	0.988	0.598
Residual Std. Error	0.530 (df = 241)	0.433 (df = 217)	0.476 (df = 217)	0.403 (df = 192)	0.624 (df = 222)

We observe that across models (1) – (4), the results are remarkably similar. Using the intercept and the coefficient for $-w$, we can identify ξ and χ for the fixed effects which serve as baseline, namely the data source *CBS* in 1993. If we take model (4), we get a value of $\xi = 1.554$. The effective lower bound is $\chi = 16.048$. The interaction effects between $-w$ and the data sources show differences in the estimated value of ξ across data sources, relative to *CBS*. These differences are also plotted in Figure 5. We observe that the rich lists *Quote* and *Forbes* both have far lower average values for ξ. The surveys differ from each other; *DNB* is not distinguishably different from *CBS*, but *SEP* clearly results in higher values of ξ and therefore a thinner tail.

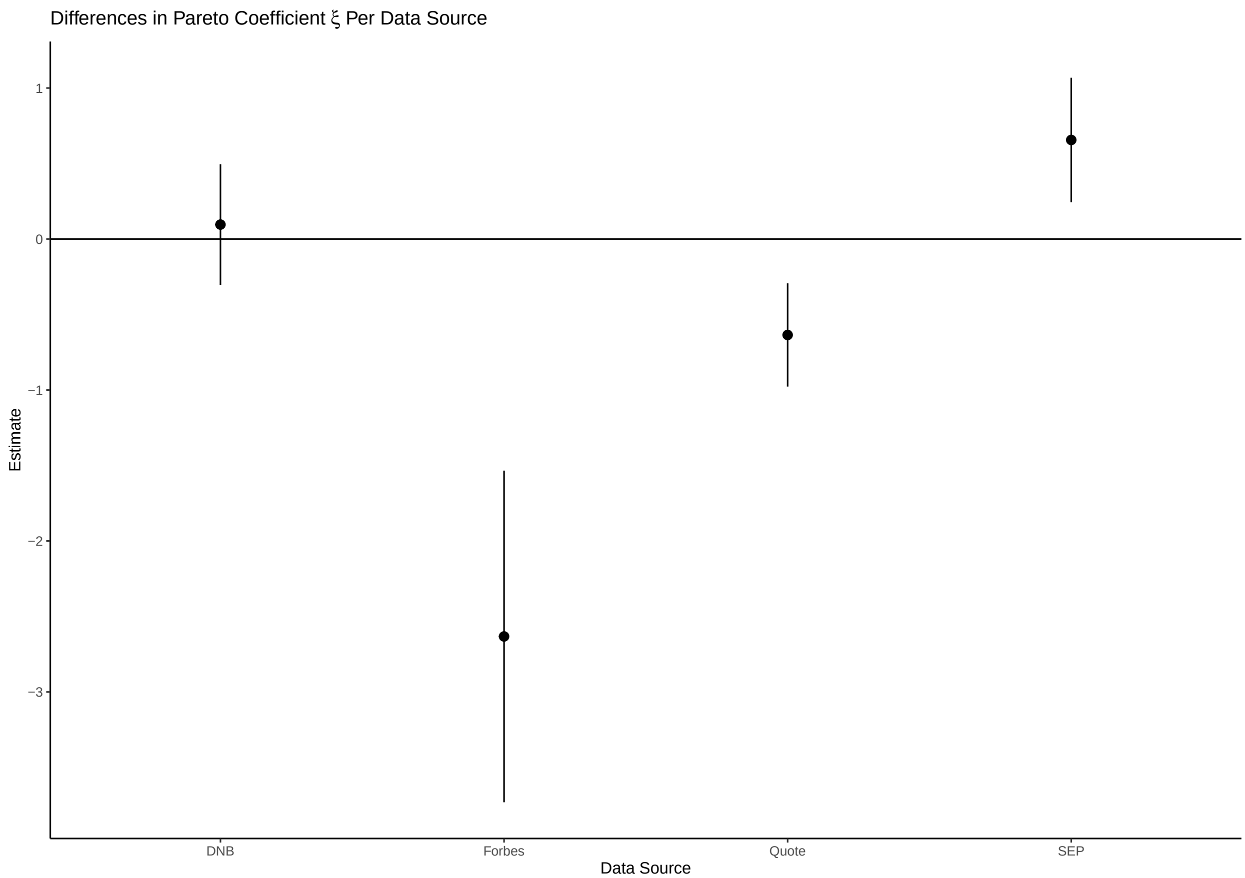

Figure 5: Differences in Pareto Coefficient ξ Across Data Sources

The differences between data sources also matter for the average value of χ. Figure A.1 (in the online Appendix) reports these differences. It is important to note that the large negative coefficient for *Forbes* is likely driven by its sample size; we have only a single estimate for this data source per year; hence, its fixed effect is very sensitive to the underlying values. The rich list *Quote*, which is similar to *Forbes* yet provides multiple estimates for multiple years, fares better. Both rich lists predict *lower* values for χ than *CBS*, which may look curious at first glance; we would expect the rich lists to find higher values for the threshold than fiscal data. This curiosity is explained by the fact that the reported values for ξ and χ are both relative to *CBS*, and are both negative; hence, if we transform the lower bound χ back into the minimum threshold X_0, we find that X_0 is larger for the rich lists than for *CBS*. Clearly, the number of individual estimates per data source matters for the confidence we have in these point estimates. Nevertheless, we can conclude based on these figures that solely relying on *CBS* for wealth estimates results in Pareto parameters – and hence, top wealth shares – that are implausibly low. Moreover, the household surveys are of relatively little value to estimate top wealth. *SEP*, in particular, delivers very high values for ξ and very low estimates for X_0.

Next, we turn to trends over time. The annual trends in ξ – taken from model (4) – are plotted in Figure 6, and the trends in χ are plotted in Figure A.2, in the online appendix. All estimates are relative to 1993. We see a clear increase in ξ, and a similar corresponding increase in χ, after the 1990s. Since χ is a function of ξ, the two developments are not independent; notice, however, that the magnitude of χ's increase more than the offsets the simultaneous decline in ξ. Hence, we can safely conclude that both the Pareto coefficient ξ – but very slightly – and the effective scale parameter χ have increased. Both facts are evidence of an increase in wealth concentration over time, since the threshold has increased and the thickness has only decreased slightly. Both results also point out, however, that after the turn of the century, the Pareto parameters have remained relatively constant on average. This does not necessarily mean that top wealth shares have not increased over time, but only that the average thickness of the right tail has remained constant, as has the location of its threshold.

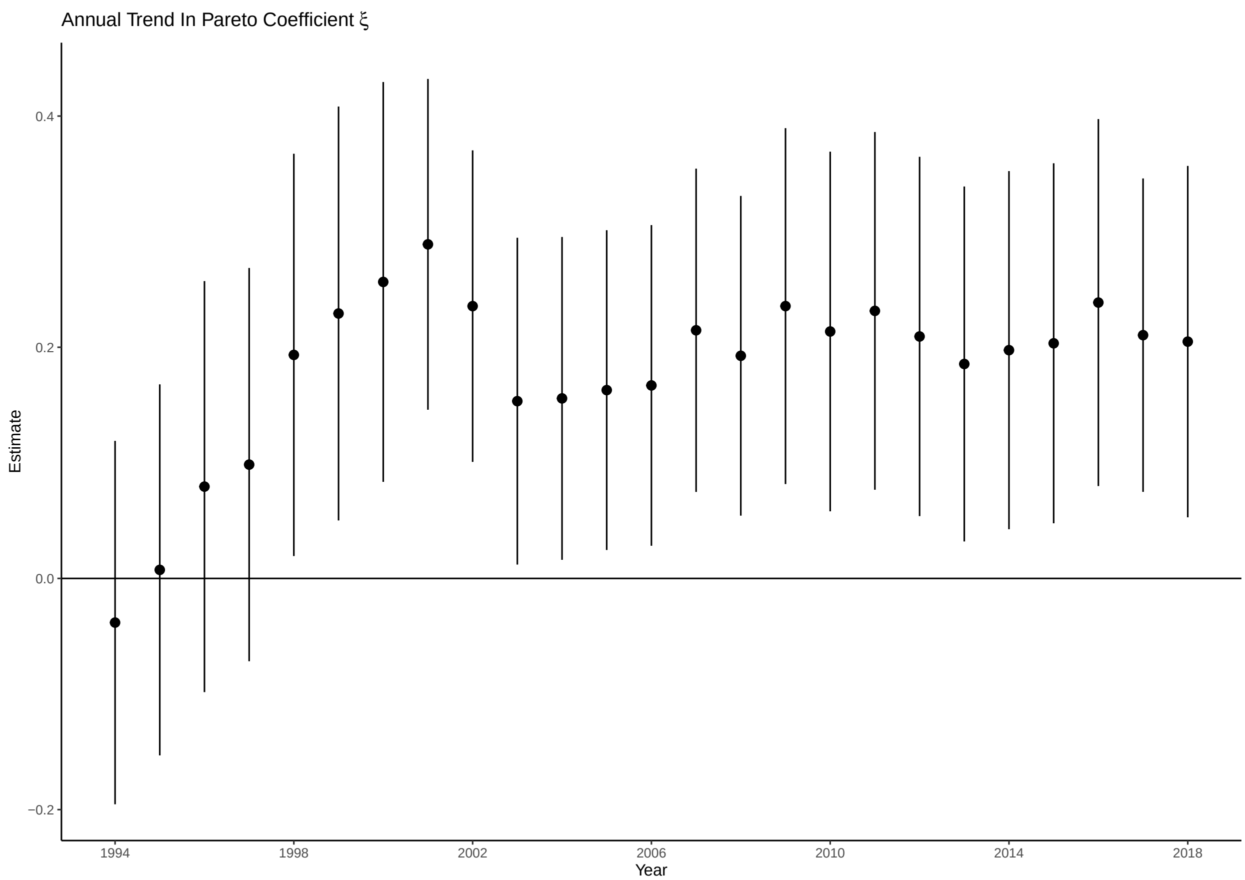

Figure 6: Yearly Trends in ξ

Finally, we can compare the results between equations 11 and 12. Since the regressions come from the same underlying data, the estimates for ξ should not be too different. Equation 12 is estimated in model (5) in Table 6. Note that the interpretation of the coefficients is different to those of models (1)–(4); in model (5), the year and data fixed effects identify ξ, not χ, and χ is not identified. If we compare the intercept for fit

(4) – which represents the average value of ξ for *CBS* in 1993 – with the coefficient for $-w$ in models (1)–(4), we observe that the two models indeed deliver roughly similar estimates for ξ. Notably, however, the sign and magnitude of the coefficients for the data fixed effects differ somewhat from the relevant coefficients in specification (1)–(4). If we compare the coefficient for *DNB* in model (5) with that of $-w \times$ DNB in models (1)–(4), we notice that model (5) estimates a significantly larger ξ compared to *CBS*, whereas the first three specifications find no significant difference. Moreover, the sign for *Forbes* flips from negative to positive, and its difference from *CBS* is reduced by an order of magnitude. This suggests that solely identifying ξ from fixed effects results in significantly different outcomes. If we compare the overall performance of the models, we have reason to prefer the first four specifications to model (5), and model (4) in particular. The R^2 much lower in model (5), and the residual standard error is much higher. Both suggest that the model fit is significantly worse in the final specification. We conclude that estimating the parameters based on equation 11 is a more robust approach.

One limitation of my approach is that it requires a large number of estimators relative to the available data points. In particular, specification (4) only has 192 degrees of freedom to estimate all parameters, which is not very extensive. Hence, these results should be viewed with some uncertainty; with more observations and hence more degrees of freedom, we could be more certain about these results.

6 Conclusion

Measuring the top tail of the wealth distribution is a difficult endeavor. In this paper, I have exploited many data sources and many differing methodologies to provide as robust as possible an answer to this question, using data from the Netherlands from 1993 to 2018. Drawing conclusions is difficult due to the fragmented nature of the data, and the inherent imperfections in each source. Nevertheless, by combining all available information in a consistent manner, we can conclude several things. First, all data sources need to be adjusted in some sense to come closer to the true level of wealth concentration. Surveys need to be calibrated to fiscal values to correct under-reporting of the left and right tail, and need to be combined with rich lists to capture top wealth. Fiscal data needs to be corrected for the market value of closely held assets, in order to more closely capture the right tail.

A second conclusion is that data sources on their own, even when calibrated, fall short of robustly gauging wealth concentration. I have shown that existing methods to combine surveys and rich lists go some way in addressing this gap, but often result in impossible results and are highly sensitive to the quality of the underlying data sources. Of the existing methods, generalized Pareto interpolation provides the most valid results; however, only combining surveys and rich lists still leaves a lot to be desired.

Therefore, my third conclusion has been that my newly developed method of Robust Pareto Regression is to be preferred to any of the existing methods. This method is simple to operate, and adequately adjusts Pareto estimates for the underlying heterogeneity across data sources and years. Using this method, I have shown that wealth concentration has modestly increased over time in the Netherlands. Furthermore, the method clearly shows that rich lists are essential to fully capture top wealth, since other data sources – surveys in particular – deliver implausibly low values for both ξ and χ.

The methodology behind Robust Pareto Regression is simple and easy to implement in different settings where data heterogeneity is a source of concern. This can be at the national level, like I have done; however, cross-country comparisons are also a promising implementation, since we only need to add country fixed effects and their interactions to the estimating equations. The method performs best when there are a sufficient number of observations to compensate for the many dummy variables. Because rich lists typically only have very few brackets that can serve as data points, the total number of observations is mainly determined by the granularity of the available fiscal data, and the quality of the household surveys. The more wealth brackets are available for fiscal data at high levels of wealth, the better; similarly, the number of households in the survey that have wealth above high thresholds also determines the number of data points that surveys contribute. Many countries, such as the United States and Germany, have surveys which oversample rich households; these countries should have little trouble to implement Robust Pareto Regression.

My results should be seen as a first step, rather than a definitive conclusion. Several issues need to be resolved in future research to improve upon my study. First, this paper does not analyze the various existing survey-enhancement methods; future work can compare these methods to the results already obtained here. Second, these results need to be corrected for tax evasion, which is estimated at 4% of household wealth in the Netherlands (Alstadsæter, Johannesen, and Zucman 2018). Although the distributional implications of offshore wealth are unclear, if they are similar to the results found for Scandinavia, top wealth shares can be expected to increase even further[10].

10. See Leenders et al. (2020) for an initial analysis of tax evasion in the Netherlands.

References

Albers, Thilo N. H., Charlotte Bartels, and Moritz Schularick. 2020. "The Distribution of Wealth in Germany, 1895–2018." *ECONtribute Report.*

Alstadsæter, Annette, Niels Johannesen, and Gabriel Zucman. 2018. "Who owns the wealth in tax havens? Macro evidence and implications for global inequality." *Journal of Public Economics* 162:89–100.

———. 2019. "Tax Evasion and Inequality." *American Economic Review* 109 (6): 2073–2103.

Atkinson, Anthony B., Thomas Piketty, and Emmanuel Saez. 2011. "Top Incomes in the Long Run of History." *Journal of Economic Literature* 49 (1): 3–71.

Bach, Stefan, Andreas Thiemann, and Aline Zucco. 2019. "Looking for the missing rich: Tracing the top tail of the wealth distribution." *International Tax and Public Finance* 26 (6): 1234–1258.

Bartels, Charlotte, and Maria Metzing. 2019. "An integrated approach for a top-corrected income distribution." *Journal of Economic Inequality* 17 (2): 125–143.

Van Bavel, Bas J. P., and Ewout Frankema. 2017. "Wealth Inequality in the Netherlands, c. 1950-2015. The Paradox of a Northern European Welfare State." *Tijdschrift voor Sociale en Economische Geschiedenis/ The Low Countries Journal of Social and Economic History* 14 (2): 29–62.

Van Bavel, Bas J. P., and Wiemer Salverda. 2014. "Vermogensongelijkheid in Nederland." *Economisch Statistische Berichten* 99 (4688): 2012–2015.

Blanchet, Thomas. 2016. "Wealth inequality in Europe and in the United States: estimations from surveys, national accounts and wealth rankings." Master Thesis, Paris School of Economics.

Blanchet, Thomas, Juliette Fournier, and Thomas Piketty. 2017. "Generalized Pareto Curves: Theory and Applications." *WID Working Paper* 2017/3.

Caminada, Koen, Kees Goudswaard, and Marike Knoef. 2014. "Vermogen in Nederland gelijker verdeeld sinds eind negentiende eeuw." *MeJudice.*

Fagereng, Andreas, Luigi Guiso, Davide Malacrino, and Luigi Pistaferri. 2020. "Heterogeneity and Persistence in Returns to Wealth." *Econometrica* 88 (1): 115–170.

Gabaix, Xavier. 2009. "Power Laws in Economics and Finance." *Annual Review of Economics* 1:255–294.

Gabaix, Xavier, and Rustam Ibragimov. 2011. "Rank - 1/2: A simple way to improve the ols estimation of tail exponents." *Journal of Business and Economic Statistics* 29 (1): 24–39.

Garbinti, Bertrand, Jonathan Goupille-Lebret, and Thomas Piketty. 2020. "Accounting for Wealth Inequality Dynamics: Methods, Estimates and Simulations for France." *Journal of the European Economic Association,* Forthcoming.

Gomez, Matthieu. 2019. "Displacement and the Rise in Top Wealth Inequality." *Mimeo.*

Jacobs, Bas. 2015. "Belastingen op Kapitaalinkomen in Nederland." *Tijdschrift voor Openbare Financiën* 47 (1): 24–48.

Jacobs, Bas, Egbert Jongen, and Floris T. Zoutman. 2013. "Meer over de top." *CPB Achtergronddocument.*

Jones, Charles I. 2015. "Pareto and Piketty: The Macroeconomics of Top Income and Wealth Inequality." *Journal of Economic Perspectives* 29 (1): 29–46.

Kopczuk, Wojciech, and Emmanuel Saez. 2004. "Top Wealth Shares in the United States, 1916–2000: Evidence from Estate Tax Returns." *National Tax Journal* 57 (2): 445–487.

Kuhn, Moritz, Moritz Schularick, and Ulrike I. Steins. 2020. "Income and Wealth Inequality in America , 1949 – 2016." *Journal of Political Economy* 128 (9): 3469–3519.

Leenders, Wouter, Arjan Lejour, Simon Rabaté, and Maarten van 't Riet. 2020. "Offshore Tax Evasion and Wealth Inequality: Evidence from a Tax Amnesty in the Netherlands." *CPB Discussion Paper.*

Ministerie van Financiën. 2020. *Bouwstenen voor een beter belastingstelsel: Inkomen uit aanmerkelijk belang.* Technical report. Dutch Ministry of Finance.

Pareto, Vilfredo. 1896. *Cours d'Économie Politique.* Librairie Droz.

Piketty, Thomas. 2014. *Capital in the Twenty-First Century.* Cambridge, MA: The Belknap Press.

Roine, Jesper, and Daniel Waldenström. 2015. "Long run trends in the distribution of income and wealth." In *Handbook of Income Distribution,* edited by Anthony B. Atkinson and François Bourguignon, vol. 2A, 469–592.

Saez, Emmanuel, and Gabriel Zucman. 2016. "Wealth Inequality in the United States since 1913: Evidence from Capitalized Income Tax Data." *Quarterly Journal of Economics* 131 (2): 519–578.

Salverda, Wiemer, Christina Haas, Marloes de Graaf-Zijl, Bram Lancee, Natascha Notten, and Tahnee Ooms. 2013. *Growing inequalities and their impacts in the Netherlands.* Technical report.

Smith, Matthew, Danny Yagan, Owen M. Zidar, and Eric Zwick. 2019. "Capitalists in the Twenty-First Century." *Quarterly Journal of Economics* 134 (4): 1675–1745.

Soltow, Lee, and Jan Luiten van Zanden. 1998. *Income and wealth inequality in the Netherlands, 16th-20th century.* Amsterdam: Het Spinhuis.

Toussaint, Simon J., Bas J. P. van Bavel, Wiemer Salverda, and Coen N. Teulings. 2020. "Nederlandse vermogens schever verdeeld dan gedacht." *Economisch Statistische Berichten* 105 (4789): 438–441.

Vermeulen, Philip. 2018. "How Fat is the Top Tail of the Wealth Distribution?" *Review of Income and Wealth* 64 (2): 357–387.

Wilterdink, Nico. 1984. *Vermogensverhoudingen in Nederland: Ontwikkelingen sinds de negentiende eeuw.* Amsterdam: Synopsis.

Zucman, Gabriel. 2019. "Global Wealth Inequality." *Annual Review of Economics* 11 (1): 109–138.

A Additional Figures and Tables

Table A.1: Summary Statistics, Statistics Netherlands

Year	Households	P05	P25	Median	Mean	P75	P95
1993	6368	-4	1	14	57	60	220
1994	6445	-5	1	16	62	67	235
1995	6469	-4	1	18	64	74	242
1996	6518	-5	2	20	72	80	261
1997	6581	-5	2	22	78	89	281
1998	6656	-8	1	20	84	96	304
1999	6745	-8	2	23	93	110	341
2000	6801	-10	2	25	108	130	386
2001							
2002	6862	0	4	33	133	161	468
2003	6919	0	3	37	141	176	502
2004	6948	-1	3	35	141	177	505
2005	6990	-1	3	39	150	188	530
2006	7025	-4	2	35	160	195	586
2007	7082	0	3	43	173	211	626
2008	7146	0	3	47	185	225	664
2009	7215	-8	3	42	179	217	652
2010	7281	-22	2	34	170	201	638
2011	7348	-30	1	33	165	196	620
2012	7412	-40	0	26	154	179	595
2013	7468	-58	0	17	138	154	561
2014	7496	-57	0	17	139	153	564
2015	7569	-52	0	20	145	160	576
2016	7623	-45	0	22	153	168	597
2017	7695	-35	1	28	165	182	627
2018	7755	-27	1	38	182	203	676

Note: All values in thousands of EUR.

Table A.2: Summary Statistics, *Forbes List of Billionaires* and *Quote 500*

	Forbes			**Quote**		
Year	**N**	**Threshold**	**Average Wealth**	**N**	**Threshold**	**Average Wealth**
1993	3	1,940	5,830			
1994	3	2,080	5,110			
1995	3	2,820	4,460			
1996	3	4,500	5,450			
1997	4	4,040	7,130	500	20	191
1998	2	4,350	7,190	500	35	258
1999	4	5,050	6,070	500	45	312
2000	3	4,740	8,530	500	60	363
2001	2	3,500	10,000	500	75	363
2002	2	1,330	4,100	500	32	166
2003	4	792	1,640	500	36	181
2004	4	954	1,960	500	41	195
2005	4	1,100	2,600	500	44	221
2006	4	987	2,580	500	46	243
2007	4	883	2,990	500	48	273
2008	5	1,010	2,430	500	50	290
2009	3	1,250	2,290	500	45	254
2010	5	973	2,540	500	51	271
2011	6	1,000	2,570	500	55	286
2012	6	985	2,330	550	50	274
2013	6	1,120	2,860	550	54	288
2014	7	1,150	3,130	550	35	238
2015	9	1,010	3,030	650	41	393
2016	9	949	3,230	650	55	300
2017	10	1,000	2,820	550	71	370
2018	9	1,310	3,710	550	80	402

Note: Values until 2002 in millions of NLG; from 2002 onward in millions of EUR. For converting USD values to NLG (and EUR), I use the end-of-year exchange rate. The values for *Quote* from 2012 onward are calculated over the consolidated lists, which includes both individuals and families; in 2016 and 2017, it also includes the list of expatriated Dutch individuals. In all other years, expatriated individuals were already included in the main list.

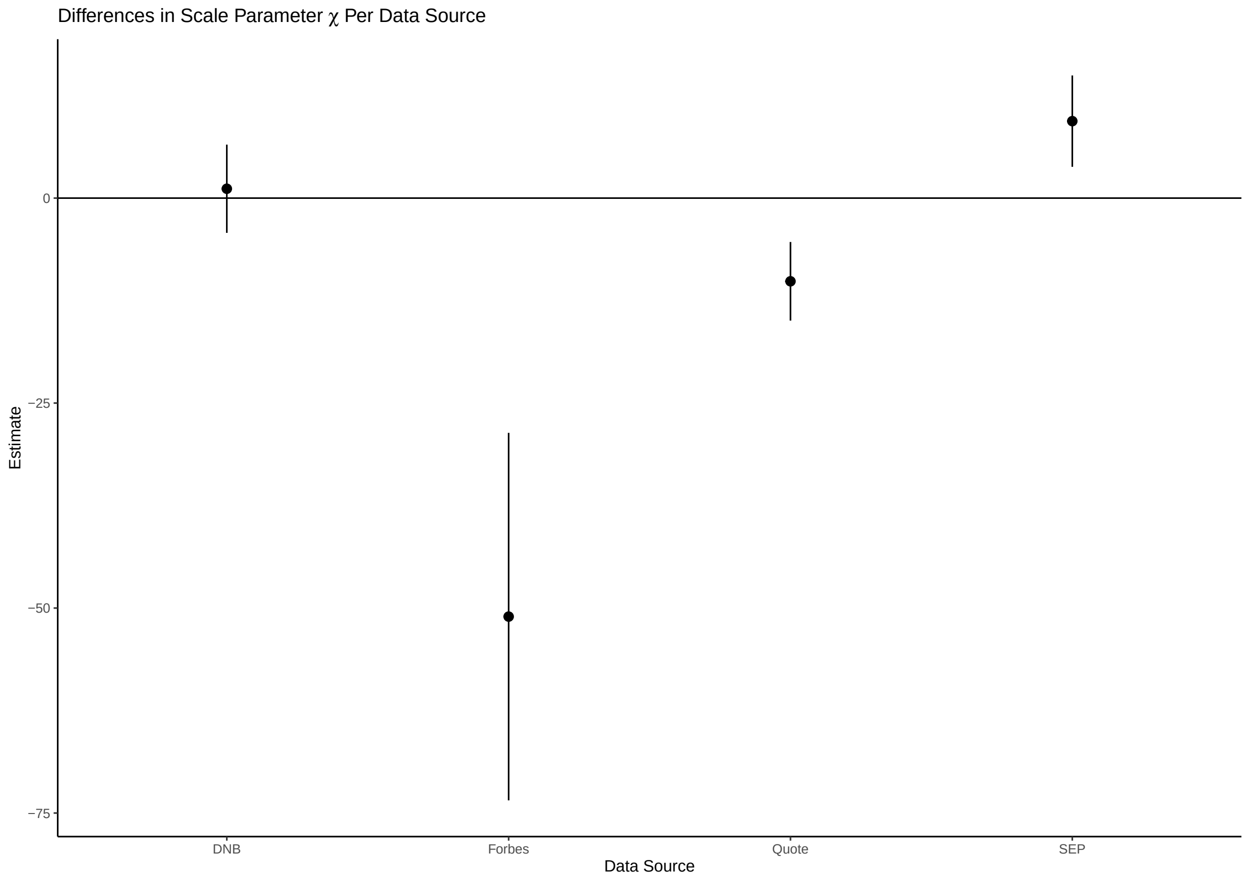

Figure A.1: Average Differences In χ Between Data Sources

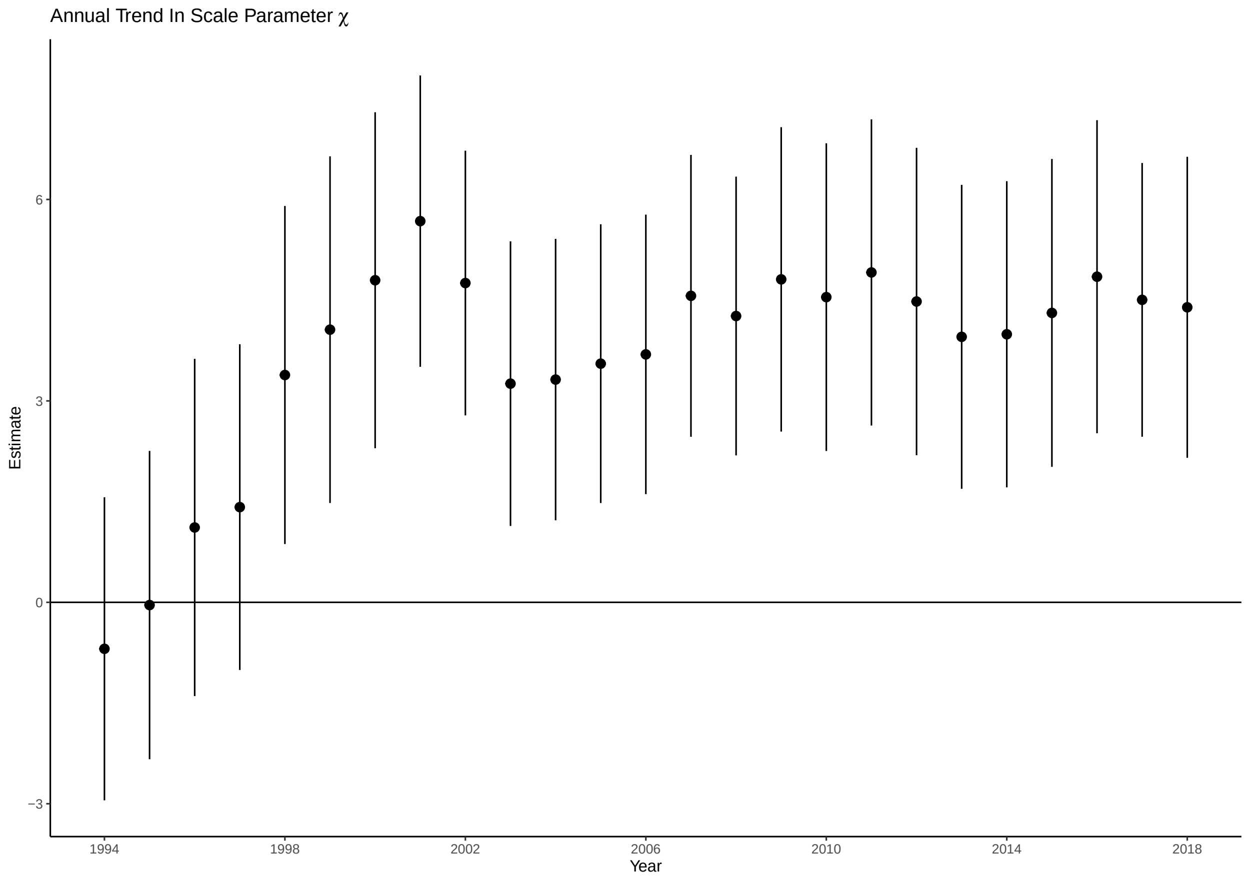

Figure A.2: Yearly Trends in χ