Technische Universität München

Department of Mathematics

Contributing to estimating the distribution of household wealth

Solving under-reporting via optimization problems with invariant Gini coefficient

Master Thesis

by

Dennis Ohlwerter

Supervisor:	Prof. Dr. Matthias Scherer
Advisor:	Dr. Janina Engel
Submission Date:	February 14th, 2021

I hereby declare that this thesis is my own work and that no other sources have been used except those clearly indicated and referenced.

Munich, February 14th, 2021

Abstract

In this thesis, we address the macro-economic problem that aggregated micro data from the Household Finance and Consumption Survey (HFCS) of the ECB does not match the macro data from national accounts (NtlA) statistics. Earlier studies have already identified that extremely wealthy households are generally underrepresented in household surveys. In Chapter 2, we present a method proposed by Vermeulen ([22]) to overcome this underrepresentation of very rich households. He suggests to combine the HFCS data with rich lists, enabling an estimation of the upper part of the wealth distribution via a Pareto distribution. After estimating this distribution, it is possible to sample further households of the upper part of the wealth distribution. To justify this method we additionally provide an overview of some wealth accumulation processes that lead to a Pareto distribution for wealth in the long term. Though some parts of the discrepancies between HFCS and NtlA statistics are explained by the missing wealthy, Chakraborty and Waltl ([5]) find that large parts are still unexplained. Therefore, we provide a solution to close the persisting gaps via an optimization problem that aims at preserving the level of inequality. Finally, we compare the findings with an approach that uses a multivariate calibration via a small case study covering household wealth of Germany.

Zusammenfassung

In dieser Masterarbeit befassen wir uns mit dem makroökonomischen Problem, dass aggregierte Mikrodaten aus der Household Finance and Consumption Survey (HFCS) der EZB nicht mit den Makrodaten aus der Statistik der National Accounts (NtlA) übereinstimmen. Frühere Untersuchungen haben bereits gezeigt, dass extrem wohlhabende Haushalte in Haushaltsbefragungen generell unterrepräsentiert sind. In Kapitel 2 stellen wir eine von Vermeulen ([22]) entwickelte Methode vor, um dieser Unterrepräsentation von sehr reichen Haushalten entgegenzuwirken. Er schlägt vor, die HFCS Daten mit Reichenlisten zu kombinieren, um so den oberen Teil der Vermögensverteilung durch eine Pareto Verteilung zu schätzen. Nach der Schätzung dieser Verteilung ist es möglich, weitere Haushalte für den oberen Teil der Vermögensverteilung zu erzeugen. Um diese Methode zu rechtfertigen, geben wir zusätzlich einen Überblick über einige Prozesse der Vermögensakkumulation, die langfristig zu einer Pareto Verteilung des Vermögens führen. Obwohl ein Teil der Differenzen zwischen der HFCS und der NtlA Statistik durch die sogenannten „missing wealthy“ erklärt werden können, stellen Chakraborty und Waltl ([5]) fest, dass große Teile noch unerklärt sind. Daher schlagen wir vor, die verbleibenden Lücken durch Lösung eines passenden Optimierungsproblems zu schließen, das insbesondere darauf abzielt, das Niveau der Ungleichheit zu erhalten. Abschließend vergleichen wir die Ergebnisse anhand einer kleinen Fallstudie über das Vermögen der privaten Haushalte in Deutschland mit einem Ansatz, der eine multivariate Kalibrierung verwendet.

Contents

1 Introduction

Macroeconomic data and indicators are frequently used to assess the current state of economies. However, it is not straightforward to derive accurate conclusions from available data about income and wealth dynamics, especially in terms of wealth distributions. Analysing and comparing distributions of wealth is important not only to assess the impact of policies from governments and central banks but also for research purposes and information for the public.

Our starting point is the Household Finance and Consumption Survey (HFCS) and national accounts (NtlA) statistics published by the ECB and national central banks.

The HFCS is conducted at the national level and provides household-level data on assets, liabilities, income and consumption along with related economic and demographic variables. Therefore, it does not only provide insights into the financial situation of households but also on their economic behaviour. These aspects can have major implications for the development of the respective economies. The Deutsche Bundesbank emphasizes[1], that central banks need micro-level information since *"aggregate data are deemed insufficient"* and micro-data *"opens up the possibility of understanding structural relationships"*.
The intended survey frequency is three years and is conducted by the Household Finance and Consumption Network (HFCN) which consists of *"statisticians and economists from the ECB, the national central banks of the Eurosystem and a number of national statistical institutes"*[2]. For example, the results of the third wave were published in 2020, with data of over 91,000 households from 19 EU countries as well as Croatia, Hungary and Poland being collected from 2016 to 2019. According to the ECB[3], *"the HFCS questionnaire consists of two main parts:*

1. *Questions relating to the household as a whole, including questions on real assets and their financing, other liabilities and credit constraints, private businesses, financial assets, intergenerational transfers and gifts, and consumption and saving;*

2. *Questions relating to individual household members, covering demographics (for all household members), employment, future pension entitlements and income (for household members aged 16 and over)."*

A comprehensive overview of general aspects of the HFCS and the third wave in detail is given by ([7]). The data set is often used in research. An application of the first two waves can be found in ([6]) with a focus on deriving wealth inequality from the data. In particular, several inequality measures are calculated and Costa and Pérez-Duarte ([6]) are analyzing the evolution and trends of wealth inequality derived from wave 1 and 2.

It turns out that aggregates (financial and non-financial instrument holdings of the household sector such as deposits and housing wealth) from the HFCS are usually lower than

[1] https://www.bundesbank.de/en/bundesbank/research/panel-on-household-finances/about-the-phf/about-the-phf-617320

[2] https://www.ecb.europa.eu/pub/economic-research/research-networks/html/researcher_hfcn.en.html

[3] https://www.ecb.europa.eu/stats/ecb_surveys/hfcs/html/index.en.html

figures from national accounts. Therefore, researchers as well as authorities, including the ECB, are investigating the reasons causing this macro-micro gap. One reason for the observed discrepancies in the aggregates is called "the missing wealthy" ([5]): As the participation of super-rich households in wealth-related surveys is very unlikely, the HFCS does not adequately capture the very top of the wealth distribution, and therefore the wealthiest households are underrepresented.

As stated by Vermeulen ([22]), *"[h]ousehold surveys are widely believed to suffer from various degrees of non-response and differential non-response"*. This non-response is particularly pronounced in very rich cohorts since it is often harder to contact them due to their lack of time or their reluctance to reveal sensitive information. *"But if non-responding households are having higher wealth in some systematic way"*, as emphasised by Vermeulen ([22]), *"wealth estimates will be biased downwards, particularly estimates of wealth at the top of the distribution"*. One way how survey analysts deal with that problem is oversampling the wealthy and thus the sample weights in the survey can be adjusted to tackle the problem of non-response. But Vermeulen ([22]) notes, that not all wealth surveys oversample the rich. Therefore, he proposes a method to overcome the underrepresentation of very rich households: The HFCS data is combined with rich lists (e.g. Forbes World's billionaires data or country-specific lists like the ranking of Germany's wealthiest persons provided by the "Manager Magazine") enabling an estimation of the upper part of the wealth distribution via a Pareto distribution. We will have a closer look at that in Chapter 2. The fact that the upper tail of the wealth distribution is Paretian was already empirically observed in many papers, e.g. for the Forbes 400 in ([12, 13]) or for the 100 wealthiest Canadians in ([18]). To provide a more economic explanation of this phenomenon, in Section 2.2, this thesis presents an overview of some theoretical models that corroborate these empirical findings. Nevertheless, Chakraborty and Waltl ([5]) examined that *"the missing wealthy do not explain large parts of the macro-micro gap for highly comparable instruments (liabilities, bonds, deposits and mutual funds) [...] still leaving significant parts unexplained"*. Therefore, we provide a solution to close the persisting gaps via reasonable optimization problems.

The remainder of the thesis is structured as follows. Chapter 2 provides some information on the method that Vermeulen ([22]) proposed to estimate the Pareto distribution. In addition, a newly developed approach by ECB staff members to sample additional households from the estimated Pareto distribution is described. To justify this method, Chapter 2 also provides an overview of some wealth accumulation processes that lead to a Pareto distribution for wealth in the long term. In Chapter 3, we seek to find a sound economic approach to close the remaining discrepancies between the aggregates of the HFCS data and the NtlA statistics. In Chapter 4, the findings of Chapter 3 will be compared to the approach that is currently used (multivariate calibration, see ([19])) and analysed via a small case study covering household wealth of Germany. Chapter 5 concludes.

2 Modelling the top of the wealth distribution with a Pareto distribution

The purpose of this section is twofold. First, we give an overview of how the Pareto distribution for the top wealth distribution is estimated. Second, we provide reasons for a natural appearance of a Pareto distribution for wealth accumulation. In particular, this section presents two of the most intuitive approaches that explain why the upper tail of the wealth distribution follows a Pareto law.

2.1 Modelling the Pareto distribution and sampling additional households

We believe in the (common) assumption that household wealth X beyond a certain threshold w_{min} of a country follows the Pareto distribution F, defined by,

$$F(x) = P(X \leq x) = \begin{cases} 1 - \left(\frac{w_{min}}{x}\right)^{\alpha}, & \text{for } x \geq w_{min}, \\ 0, & \text{for } x < w_{min}. \end{cases} \tag{1}$$

The shape parameter α is called the tail index (or Pareto index when dealing with wealth distribution) and determines the heaviness of the tail. From Equation (1), we can see that the lower α, the heavier the tail, i.e., the more concentrated wealth is. The second parameter that determines the Pareto distribution is the scale parameter w_{min}.

2.1.1 Estimating the Pareto index

There do also exist estimation procedures for w_{min}, but we are more interested in estimating the Pareto index α and take w_{min} as given. For example, Vermeulen ([22]) proposes $w_{min} = 1$ million euros. We want to derive an estimator for α. Assume that there is a finite number of n households, each of them with wealth at or above w_{min}. Further, assume that the sample is ordered by increasing wealth, i.e. $x_1 \geq x_2 \geq \ldots \geq x_n \geq w_{min}$, and that each household is assigned with a rank, i.e., the rank of the household with wealth x_i is i. Since our sample follows a Pareto distribution by assumption, we replace the probability $P(X > x)$ by the empirical frequency $\frac{i}{n}$. Therefore, we get the relationship

$$\frac{i}{n} \cong \left(\frac{w_{min}}{x_i}\right)^{\alpha}. \tag{2}$$

Taking logarithms on both sides of Equation (2), we get

$$\ln\left(\frac{i}{n}\right) \cong \alpha \ln\left(\frac{w_{min}}{x_i}\right), \tag{3}$$

or equivalently

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

with $C = \ln(n) + \alpha \ln(w_{min})$. If w_{min} is known, α can then be estimated with linear regression without the constant term C.

Since the HFCS has a more complex survey design and the households have different survey weights, Vermeulen ([23]) adopts this method taking into account the weights of the sample points. Therefore, we consider a sample with different weighted households. In this case, the total sum of weights equals N. Again, we rank the sample households according to wealth. For example, the wealthiest household has wealth x_1 and a survey weight of N_1, and the second wealthiest household has wealth x_2 and survey weight of N_2, and so on. We replace $\frac{i}{n}$ with $\frac{N_1+N_2+...+N_i}{N}$ in (3) to get

$$\ln\left(\frac{N_1 + N_2 + \ldots + N_i}{N}\right) \cong \alpha \ln\left(\frac{x_{min}}{x_i}\right) \tag{5}$$

which equals to

$$\ln\left(i\frac{(N_1 + N_2 + \ldots + N_i)}{i}\frac{1}{N}\right) \cong \alpha \ln\left(\frac{w_{min}}{x_i}\right). \tag{6}$$

We now define $\bar{N} := \frac{\sum_{j=1}^{n} N_j}{n}$ as the average weight of a sample point and $\bar{N}_{fi} := \frac{\sum_{j=1}^{i} N_j}{i}$ as the average weight of the first i sample points. Then we have

$$\ln\left(i\frac{\bar{N}_{fi}}{\bar{N}}\frac{\bar{N}}{N}\right) \cong \alpha \ln\left(\frac{w_{min}}{x_i}\right), \tag{7}$$

leading to the regression

$$\ln\left(i\frac{\bar{N}_{fi}}{\bar{N}}\right) = C - \alpha \ln\left(x_i\right), \tag{8}$$

with $C = -\ln\left(\frac{\bar{N}}{N}\right) + \alpha \ln\left(w_{min}\right)$. We note that this regression is almost identical to the regression for samples with equal survey weights. But in this case, the rank of the sample observation i is weighted by the ratio of the average weight of the first i observations to the average weight of all observations.

2.1.2 Estimating the number of additional households[4]

The main purpose of estimating a Pareto distribution for the top wealth distribution is to overcome the problem of non-response of wealthy households in the HFCS. Therefore, the next step is to sample additional households from the interval that is not covered by the HFCS and neither by a rich list. We assume that the Pareto distribution was estimated from the HFCS data combined with the relevant rich lists, e.g., the ranking of Germany's wealthiest persons provided by "Manager Magazine". The rich lists are very important because, without them, we underestimate the heaviness of the tail.

[4] This subsection is derived from internal R code and notes developed by ECB staff members.

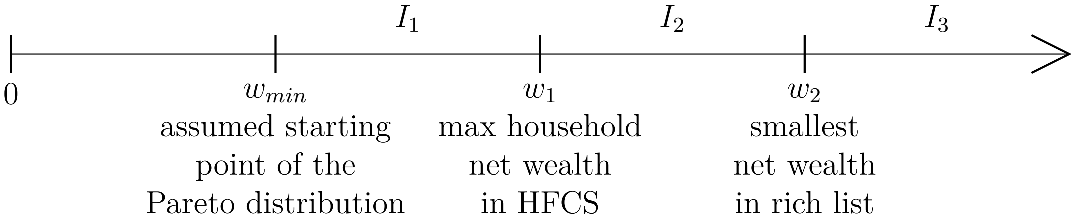
I_1
I_2
I_3
0
w_{min}
assumed starting point of the Pareto distribution
w_1
max household net wealth in HFCS
w_2
smallest net wealth in rich list

Figure 1: Illustration of the split of the support of the Pareto distribution into three distinct intervals.

First of all, we split the support of the Pareto distribution, i.e. the interval $[w_{min}, \infty[$, into three intervals:

$$\begin{aligned} I_1 &:= \Big[w_{min}, w_1\Big], \\ I_2 &:= \Big]w_1, w_2\Big[, \\ I_3 &:= \Big[w_2, \infty\Big[, \end{aligned} \tag{9}$$

for some $w_1, w_2 \in \mathbb{R}$ with $w_{min} < w_1 < w_2$.
More precisely, regarding our data, we choose

$$\begin{aligned} w_1 &:= \max\Big\{ \text{ household wealth in HFCS } \Big\} \\ w_2 &:= \min\Big\{ \text{ wealth in rich list } \Big\}. \end{aligned} \tag{10}$$

In other words,

$$\begin{aligned} I_1 &= \Big[\text{ all HFCS households with wealth of at least } w_{min}\Big], \\ I_2 &= \Big] \text{ gap between HFCS observations and rich list } \Big[, \\ I_3 &= \Big[\text{ from the observations of the rich list until infinity}\Big[. \end{aligned} \tag{11}$$

An illustration of this split is given in Figure 1.

Furthermore, we will denote the random variables of the number of households in interval I_i with M_i respectively for $i \in \{1, 2, 3\}$. The total number of households in all intervals is denoted by M and given by

$$M := M_1 + M_2 + M_3. \tag{12}$$

Note that we actually know values of M_1 and M_3 by using the data (HFCS and rich list). The number of households m_1 (value of M_1) in interval I_1 is simply the sum of the weights of the households in the HFCS with wealth of at least w_{min}. Likewise, the number of households m_3 (value of M_3) in interval I_3 is given by the number of households in the rich list. What is unknown is the number of households M_2 in interval I_2.

We assume that (X_i) are independent and identically distributed random variables with distribution function given in (1) and that the random variable M is independent of (X_i).

We need to estimate the number of households M_2 that fall in interval I_2. We observe that M_2 in interval I_2 can be expressed in terms of (X_i) by

$$M_2 = \sum_{i=1}^{M} \mathbb{1}_{\{X_i \in I_2\}}. \tag{13}$$

In the following, we will derive an expression for $\mathbb{E}[M_2]$ and use this for estimating M_2. Note that the expected number of households in I_1 is given by

$$\mathbb{E}[M_1] = \mathbb{E}\left[\sum_{i=1}^{M} \mathbb{1}_{\{X_i \in I_1\}}\right]. \tag{14}$$

Using Wald's lemma yields

$$\begin{aligned}
\mathbb{E}[M_1] &= \mathbb{E}\left[\sum_{i=1}^{M} \mathbb{1}_{\{X_i \in I_1\}}\right] \\
&= \mathbb{E}[M]\,\mathbb{E}\left[\mathbb{1}_{\{X_1 \in I_1\}}\right], \quad \text{by Wald's lemma, } (X_i) \text{ i.i.d.} \\
&= \mathbb{E}[M]\,P\Big(X_1 \in [w_{min}, w_1]\Big) \\
&= \mathbb{E}[M]\left[P\left(X_1 \leq w_1\right)\right] \\
&= \mathbb{E}[M]\left[1 - \left(\frac{w_{min}}{w_1}\right)^{\alpha}\right] \\
&= \mathbb{E}[M]\,\frac{w_1^{\alpha} - w_{min}^{\alpha}}{w_1^{\alpha}}.
\end{aligned} \tag{15}$$

We can now reformulate Equation (15) to

$$\text{Equation (15)} \quad \Leftrightarrow \quad \mathbb{E}[M] = \mathbb{E}[M_1]\,\frac{w_1^{\alpha}}{w_1^{\alpha} - w_{min}^{\alpha}}. \tag{16}$$

Likewise, it is straightforward to calculate the expected number of households in I_2,

$$\begin{aligned}
\mathbb{E}[M_2] &= \mathbb{E}\left[\sum_{i=1}^{M} \mathbb{1}_{\{X_i \in I_2\}}\right] \\
&= \mathbb{E}[M]\,P\Big(X_1 \in\,]w_1, w_2[\Big) \\
&= \mathbb{E}[M]\left[\left(1 - \left(\frac{w_{min}}{w_2}\right)^{\alpha}\right) - \left(1 - \left(\frac{w_{min}}{w_1}\right)^{\alpha}\right)\right] \\
&= \mathbb{E}[M]\left[\left(\frac{w_{min}}{w_1}\right)^{\alpha} - \left(\frac{w_{min}}{w_2}\right)^{\alpha}\right].
\end{aligned} \tag{17}$$

We insert Equation (16) into Equation (17) and get

$$\begin{aligned}
\mathbb{E}[M_2] &= \mathbb{E}[M]\left[\left(\frac{w_{min}}{w_1}\right)^{\alpha} - \left(\frac{w_{min}}{w_2}\right)^{\alpha}\right] \\
&= \mathbb{E}[M_1]\,\frac{w_1^{\alpha}}{w_1^{\alpha} - w_{min}^{\alpha}}\left[\left(\frac{w_{min}}{w_1}\right)^{\alpha} - \left(\frac{w_{min}}{w_2}\right)^{\alpha}\right] \\
&= \mathbb{E}[M_1]\,\frac{w_1^{\alpha}}{w_1^{\alpha} - w_{min}^{\alpha}}\left[\frac{w_{min}^{\alpha} w_2^{\alpha} - w_{min}^{\alpha} w_1^{\alpha}}{w_1^{\alpha} w_2^{\alpha}}\right] \\
&= \mathbb{E}[M_1]\,\frac{w_{min}^{\alpha} w_2^{\alpha} - w_{min}^{\alpha} w_1^{\alpha}}{w_1^{\alpha} w_2^{\alpha} - w_{min}^{\alpha} w_2^{\alpha}}.
\end{aligned} \tag{18}$$

Therefore, we are going to use $\hat{m}_2 := \mathbb{E}[M_2]$ to estimate the unknown number of households m_2. Since we can estimate $\mathbb{E}[M_1]$ via the observed number of households in I_1, namely m_1, we get

$$\hat{m}_2 = m_1 \frac{w_{min}^{\alpha} w_2^{\alpha} - w_{min}^{\alpha} w_1^{\alpha}}{w_1^{\alpha} w_2^{\alpha} - w_{min}^{\alpha} w_2^{\alpha}}. \tag{19}$$

2.1.3 Sampling according to a given Pareto distribution[5]

Once we have estimated the number of unobserved households in interval I_2, we can add this amount of additional households by randomly sampling them according to the given Pareto distribution via inverse transform sampling. This process is described in the following.

Let $U \sim \text{Unif}\left[F(w_1), F(w_2)\right]$. Applying the inverse of the Pareto distribution $F^{-1}(U)$ yields random variables following the Pareto distribution conditioned on falling into I_2. More precisely,

$$\left[F(w_1), F(w_2)\right] = \left[1 - \left(\frac{w_{min}}{w_1}\right)^{\alpha}, 1 - \left(\frac{w_{min}}{w_2}\right)^{\alpha}\right] \tag{20}$$

and for $x \geq w_{min}$ the inverse of the Pareto distribution function is given by

$$\begin{aligned} \underbrace{F(x)}_{=:y} &= 1 - \left(\frac{w_{min}}{x}\right)^{\alpha} \\ \Leftrightarrow \qquad x &= \frac{w_{min}}{(1-y)^{1/\alpha}} \\ \Rightarrow \quad F^{-1}(y) &= \frac{w_{min}}{(1-y)^{1/\alpha}}. \end{aligned} \tag{21}$$

2.2 Wealth distribution models

In Section 2.1, we assumed that the household wealth for the top of the wealth distribution follows a Pareto law. Now, we seek to answer the question of why this assumption is reasonable. Therefore, we will get to know some theoretical wealth accumulation processes that have been proposed in the past and explain why the upper tail of wealth distribution follows a Pareto law.

2.2.1 Wealth accumulation process with homogeneous investment talent

Levy and Levy ([15]) show that general wealth accumulation processes with homogeneous investment talent lead to a Pareto wealth distribution. The three main assumptions of the proposed model are:

1. Wealth accumulation follows a stochastic multiplicative process,
2. a lower bound of wealth w_{min} exists,

[5] This subsection is derived from internal R code and notes developed by ECB staff members.

3. a homogeneous talent (of all actors) for wealth accumulation.

Therefore, the reason for inequality is rather chance than some sort of individual investment ability. By Levy and Levy ([15]), the process of wealth accumulation can be generally formulated as

$$X_{t+1}(i) = \lambda_t(i)X_t(i) + S_t(i) - C_t(i) \tag{22}$$

where $X_t(i) \in \mathbb{R}_{\geq x_{min}}$ denotes the amount of wealth of the i-th individual at date $t \in \mathbb{N}$, $\lambda_t(i) \in \mathbb{R}$ is the return on the wealth at date t, and the amount of salary as well as consumption at time t of the i-th individual is given by $S_t(i)$ and $C_t(i)$.

For very rich individuals, wealth is mostly driven by return on investments rather than salary or consumption. Since we are interested in the evolution of wealth for very rich households, we neglect the income and consumption component and obtain

$$X_{t+1}(i) = \lambda_t(i)X_t(i). \tag{23}$$

Due to our assumption of homogeneous investment talent, all individuals draw randomly from the same distribution, called $g(\lambda)$, which determines the returns on wealth. Historical long-term returns on investments (e.g. stock market) suggest that it seems reasonable to assume a positive drift of λ, i.e. $E[\lambda] > 1$.
With the assumptions made above, it is possible to prove the following theorem:

Theorem 1. ([15]) *For any nondegenerate initial wealth distribution and nontrivial return distribution, the wealth accumulation process with positive drift given by Equation* (23) *with a lower bound leads to convergence of the normalized wealth distribution to the Pareto distribution.*

Proof. We define the normalized wealth of the i-th individual at time t as $x_t(i) = \frac{X_t(i)}{\sum_j X_t(j)}$, i.e. the individual's wealth as a fraction of the total wealth. The normalized wealth accumulation process is given by

$$x_{t+1}(i) = \frac{X_{t+1}(i)}{\sum_j X_{t+1}(j)} = \frac{X_t(i)}{\sum_j X_t(j)}\lambda_t(i)\frac{\sum_j X_t(j)}{\sum_j X_{t+1}(j)} =: \tilde{\lambda}_t(i)x_t(i), \tag{24}$$

where we define $\tilde{\lambda}_t(i) := \lambda_t(i)\frac{\sum_j X_t(j)}{\sum_j X_{t+1}(j)}$. The cumulative normalized wealth distribution $F(x, t+1)$ at time $t+1$ can be expressed as

$$F(x, t+1) = \int_0^\infty F\left(\frac{x}{\tilde{\lambda}}, t\right) g(\tilde{\lambda})\mathrm{d}\tilde{\lambda}. \tag{25}$$

The authors note that the distribution $F(w, t)$ undergoes a continuous smoothing process regardless of the starting point $F(x, 0)$. Due to the assumption of a lower bound on wealth, they point out that this process is analogous to diffusion towards a barrier and it can be shown that $F(x, t)$ converges to a stationary distribution, namely $F(x)$. According to Equation (25), the stationary distribution can be written as

$$F(x) = \int_0^\infty F\left(\frac{x}{\tilde{\lambda}}\right) g(\tilde{\lambda})\mathrm{d}\tilde{\lambda}. \tag{26}$$

Differentiating with respect to x gives

$$f(x) = \int_0^\infty f\left(\frac{x}{\tilde{\lambda}}\right)\frac{1}{\tilde{\lambda}}g(\tilde{\lambda})\mathrm{d}\tilde{\lambda}. \tag{27}$$

We know that the Pareto density function is given by

$$f(x) = \frac{\alpha w_{min}^{\alpha}}{x^{\alpha+1}} \qquad \text{for } \alpha > 0 \text{ and } x \geq w_{min} > 0. \tag{28}$$

Inserting this density into equation (27) yields

$$\begin{aligned}\frac{\alpha w_{min}^{\alpha}}{x^{\alpha+1}} &= \int_0^\infty \frac{\alpha w_{min}^{\alpha}}{(x/\tilde{\lambda})^{\alpha+1}}\frac{1}{\tilde{\lambda}}g(\tilde{\lambda})\mathrm{d}\tilde{\lambda} \\ &= \frac{\alpha w_{min}^{\alpha}}{x^{\alpha+1}}\int_0^\infty \tilde{\lambda}^{\alpha}g(\tilde{\lambda})\mathrm{d}\tilde{\lambda}.\end{aligned} \tag{29}$$

This means that for α solving $\int_0^\infty \tilde{\lambda}^{\alpha}g(\tilde{\lambda})\mathrm{d}\tilde{\lambda} = 1$, the stationary distribution is Paretian. Note that the actual amount of wealth at any given time is the normalized wealth times a constant, i.e. $X = Cx$. A transformation of the probability density function shows

$$h(X) = f(x)\frac{\partial x}{\partial X} = f\left(\frac{X}{C}\right)\frac{1}{C} = \frac{\alpha(w_{min}C)^{\alpha}}{X^{\alpha+1}}. \tag{30}$$

We conclude that not only the normalized wealth follows a Pareto distribution, but also X itself. □

The authors also try to answer the question of whether the Pareto wealth distribution still holds with non-homogeneous investment talent. Indeed, they give numerical evidence that a certain degree of investment talent differential is possible. Interestingly, in this case, wealth inequality is not only a result of luck (randomness of the wealth accumulation process) but also a consequence of different investment skills.

2.2.2 Exponential wealth accumulation with stable population

The second model that we are investigating was proposed by Jones ([9, 10]). A similar, slightly more complex model has already been introduced by Wold and Whittle ([24]). In this model ([9, 10]), one assumes that the wealth x of an individual at time t is given by the following equation

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = (r - \tau - \alpha)x(t), \tag{31}$$

where $r \in \mathbb{R}$ is the interest rate, $\tau \in [0, 1]$ is a tax wealth and $\alpha \in [0, 1]$ is assumed to be the constant consumption rate.

If $x_{t-a}(0)$ denotes the initial wealth of a newborn at time $t-a$, the wealth of an individual of age a at time t is given by

$$x_t(a) = x_{t-a}(0)e^{(r-\tau-\alpha)a}. \tag{32}$$

Furthermore, the model assumes that the concept of a stable population holds. This concept was first introduced by Alfred J. Lotka ([16]). The two main assumptions are a constant mortality and age structure. In our case, the population growth is given by

$$N_t = N_0 e^{\bar{n}t}, \tag{33}$$

where N_0 is the initial number of individuals and $\bar{n}$ denotes the growth rate.
Assuming that death follows a Poisson process with arrival rate $\bar{d}$, one can show that under the assumption of a constant age structure, the stationary distribution for an individual's age is given by

$$P(\text{Age} > a) = e^{-(\bar{n}+\bar{d})a}. \tag{34}$$

Note that the rate $\bar{n} + \bar{d}$ can be interpreted as (constant) birth rate $\bar{b}$.

In addition, the aggregated capital at time t of the population is denoted by K_t. Thus, the average capital k_t is given by $k_t = \frac{K_t}{N_t}$. Considering that the capital of people dying at time t is $\bar{d}K_t$ and the number of newborns at time t is $(\bar{n} + \bar{d})N_t$, we can calculate the average wealth inherited by a newborn as

$$x_t(0) = \frac{\bar{d}K_t}{(\bar{n} + \bar{d})N_t} =: \bar{x}k_t, \tag{35}$$

with $\bar{x} := \frac{\bar{d}}{\bar{n}+\bar{d}}$.
It is also assumed that the economy is in steady-state, i.e., the capital per person is growing at a constant rate g. Thus, the capital per person can also be expressed as

$$k_t = k_0 e^{gt}.$$

The amount of wealth of a person of age a at time t inherited at their birth can then be written as

$$x_{t-a}(0) = \bar{x}k_{t-a} = \bar{x}k_0 e^{g(t-a)} = \bar{x}k_t e^{-ga}. \tag{36}$$

We can insert this expression into (32) and obtain the cross-section wealth at time t

$$x_t(a) = \bar{x}k_t e^{(r-g-\tau-\alpha)a}. \tag{37}$$

Now, we invert Equation (37) to obtain the age at which a person's wealth exhibits a certain threshold x:

$$a(x) = \frac{1}{r - g - \tau - \alpha} \log\left(\frac{x}{\bar{x}k_t}\right). \tag{38}$$

By using relation (34), we can derive an expression for the distribution of wealth:

$$\begin{aligned} P(\text{Wealth} > x) &= P(\text{Age} > a(x)) \\ &= e^{-(\bar{n}+\bar{d})a(x)} \\ &= \left(\frac{x}{\bar{x}k_t}\right)^{-\frac{\bar{n}+\bar{d}}{r-g-\tau-\alpha}}. \end{aligned} \tag{39}$$

It is important to note that we have not yet established for which cohorts our wealth accumulation model as described in Equation (31) holds. But as in the first model, this process is typically only accurate for very wealthy households. The fact that consumption can be described as a fraction of wealth does not properly apply for average households that typically make their living by earning a salary rather than collecting interests (or capital gains). In this case, a lump sum or a fraction of salary that describes the consumption part of the wealth accumulation process would be more appropriate. Furthermore, the wealth tax τ, if existent, usually applies only for wealth above a high threshold.

Jones ([10]) emphasizes that the main mechanism leading to a Pareto distribution can be summarized in one sentence: *"exponential growth that occurs for an exponentially distributed amount of time leads to a Pareto distribution"*.
He also expounds some interesting implications that can be derived from Equation (39). First, the main source of wealth inequality is the term $r - g$, and the higher the difference between the interest rate r and the exogenous growth rate g, the more wealth inequality increases. He also notes that a higher wealth tax τ will lower wealth inequality.

2.2.3 Further models

The models above constitute just a small fraction of what has been proposed on that topic. Further interesting approaches are briefly described in this section.

Generalized Lotka-Volterra model Interestingly, many econophysicists have come up with theories developed by physicists and adopted them to wealth inequality. A comprehensive overview of that research can be found in ([4]). One example is the generalized Lotka-Volterra model for wealth distribution ([4, 20, 21]): A total number of $N \in \mathbb{N}$ agents redistribute their wealth according to the following multiplicative random process:

$$x_{i,t+1} = (1 + \xi_t)\, x_{i,t} + \frac{a}{N} \sum_j x_{j,t} - c \sum_j x_{i,t} x_{j,t} \tag{40}$$

where $x_{i,t}$ is the wealth of agent i at date t and ξ_t is chosen randomly from a positive set which has a variance V. Economically, the first term on the right-hand side introduces some stochastic return on wealth held by agent i, the second term can be interpreted as a redistribution (e.g. social welfare) at each time step to ensure that the wealth of all agents is always positive. The latter part of Equation (40) controls the overall growth of the system and ensures that both external limiting factors (e.g. finite amount of resources) and market effects like competition are included. Richmond and Salomon ([20, 21]) have already performed an analysis of this model in the early 2000s. Inter alia, they have shown that the stable distribution in generalized Lotka-Volterra models for wealth distribution follows a Pareto law.

Yard-Sale model The last model we are introducing is the so-called Yard-Sale model ([8]). It was analysed by Boghosian ([1, 2]). The model assumes that wealth distribution is a result of wealth transfer between economic agents whose transaction size is proportional to the wealth of the less wealthy agent. We define β as the fraction of the wealth of the less wealthy that is transferred in such a process. Therefore, the wealth that changes hands in a transaction can be written as

$$\Delta\left(x, x', r\right) = \beta z \min\left(x, x'\right) = \beta z \left(x \mathbf{1}_{\{x'-x \geq 0\}} + x' \mathbf{1}_{\{x-x' \geq 0\}}\right) \tag{41}$$

where z is a random variable equal to $+1$ or -1 with equal probability and x, x' denote the wealth of the first and second agent, respectively.
After the transaction, the wealth of the agents is

$$\begin{aligned} x_{\text{new}} &= x + \Delta\left(x, x', r\right), \\ x'_{\text{new}} &= x' - \Delta\left(x, x', r\right). \end{aligned}$$

Boghosian ([1, 2]) gives numerical evidence that the Yard-Sale model results in a distribution where most of the wealth is possessed by one single agent. But in case of introducing some sort of wealth redistribution like a wealth tax, i.e., each agent gains

$$\Delta_r(x) = \tau\left(\frac{X}{N} - x\right) \tag{42}$$

after redistribution, where τ denotes the tax rate, X the total amount of wealth and N the number of agents, Boghosian shows by establishing the Fokker-Planck equation of that model that the steady-state solution of the wealth distribution exhibits an approximate power law at large x.

All presented models provide a foundation of what has been detected by empirical observations in research: The wealth distribution of wealthy people follows a Pareto law. Thus, it is reasonable to estimate the wealth of residents that are not adequately captured by the HFCS via a Pareto distribution, as we have done in Section 2.1.

3 Closing the remaining discrepancies with an optimization approach

As mentioned in Chapter 1, Chakraborty and Waltl ([5]) examined that the procedure described in Section 2.1 leaves significant parts of the observed wealth gap between national accounts and HFCS data unexplained. Therefore, the main goal of this chapter is to close the remaining discrepancies with an optimization approach. In the following, we assume that the HFCS data has already undergone the procedure proposed in Section 2.1, i.e., the HFCS data have already been adjusted for the missing wealthy via a Pareto distribution. Therefore, let $\boldsymbol{d} = (d_1, \ldots, d_n) \in \mathbb{R}^n_{>0}$ denote the weights of the $n \in \mathbb{N}$ households participating in the HFCS (in a given country) after the Pareto fitting including the observations from the rich list and synthetically sampled households where needed. Moreover, let $\boldsymbol{x} = (x_1, \ldots, x_n)$ denote the wealth (of all observed households in increasing order or the holdings in a certain instrument like deposits, bonds, shares, etc.). Furthermore, we introduce an adjustment coefficient vector $\boldsymbol{a} = (a_1, \ldots, a_n)$. First of all, we want to ensure that the adjusted HFCS aggregates match the NtlA figures (denoted by F), i.e., $\sum_{i=1}^{n} d_i a_i x_i = F$ should hold. Since we do not have further information that some households should be more affected of wealth allocation than others, we try to allocate the additional wealth in a homogeneous way. The most convenient way to do this is introducing a suitable optimization function. That is, we want to minimize a certain function f under the constraint that the adjusted instrument holdings match the national accounts F, i.e.,

$$\begin{aligned} &\min_{\boldsymbol{a} \in \mathbb{R}^n_{>0}} f(\boldsymbol{a}, \boldsymbol{d}, \boldsymbol{x}) \\ &\sum_{i=1}^{n} d_i a_i x_i = F. \end{aligned} \tag{43}$$

Note that the constraint can also be restricted to a certain instrument like deposits.
Now, the first step is to choose an appropriate function for f. Since we do not have further information on how to allocate the wealth, the objective function should modify the extended HFCS data (including the added wealthy household) as little as possible. In contrast to the Pareto fitting, where one seeks to add wealthy households to the data, we assume for our optimization that the wealth inequality resulting from the HFCS after the Pareto fitting is correct. More precisely, we see no additional reasons that would justify a further modification of wealth inequality. Thus, the objective function will be related to an inequality measure. Particularly, the difference of the wealth inequality before and after adjustments with the factors a_i should be as low as possible. There are many possible choices for such an inequality measure, including the Gini coefficient, the Atkinson index or the Generalised Entropy indices. An overview is given in ([6]). Here, we use the Gini coefficient for several reasons: First of all, it is the most common inequality measure. Furthermore, its definition is easy to interpret, and, as we will see later, the empirical version of the Gini coefficient has some useful mathematical properties (especially linearity).

3.1 Gini coefficient[6]

To understand the Gini coefficient it is first important to introduce the so-called Lorenz curve, which is a graphical representation of the distribution of wealth. The Lorenz curve is a plot of the cumulative distribution of wealth versus the cumulative distribution of the population. Formally, assume that for $0 \leq q \leq 1$ the quantile function Q of the wealth distribution function F is defined as

$$Q(F, q) := \inf\{y \mid F(y) \geq q\} := y_q, \tag{44}$$

and the cumulative wealth function C is defined as

$$C(F, q) := \int_{-\infty}^{Q(F,q)} y \mathrm{d}F(y). \tag{45}$$

Then, the Lorenz curve is given by

$$L(F, q) = \frac{C(F, q)}{\mu(F)}, \tag{46}$$

where $\mu(F)$ denotes the mean of the distribution F, see also ([6]). It follows from the definition that the Lorenz curve is equal to the 45-degree line in case of equality, i.e., if every agent possesses the same wealth.

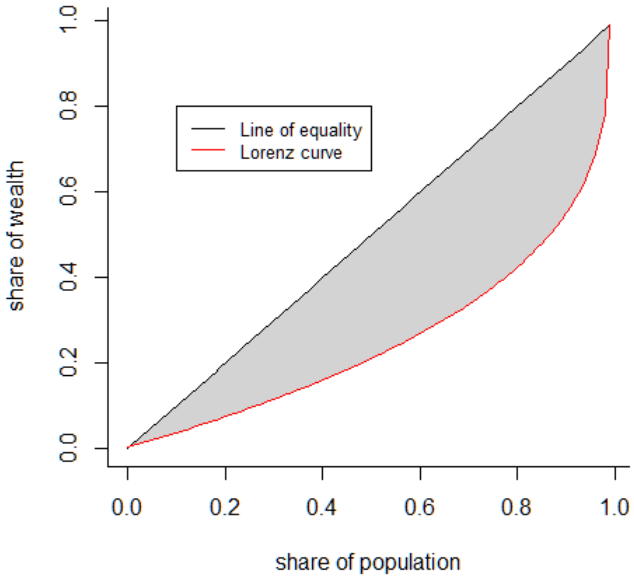

Figure 2: Lorenz curve for a Pareto distribution. The shaded area is called area of concentration.

The Gini coefficient as a measure of inequality, named after the Italian statistician and

[6] The definitions are adopted from ([6]).

sociologist Corrado Gini, is closely linked to the Lorenz curve. We call the area between the 45-degree line and the Lorenz curve the area of concentration. The Gini coefficient is defined as the ratio of the area of concentration to the maximum possible concentration area, i.e.

$$G = \frac{0.5 - \int_0^1 L(F,q)\mathrm{d}q}{0.5} = 1 - 2\int_0^1 L(F,q)\mathrm{d}q. \tag{47}$$

Since $0 \leq \int_0^1 L(F,q)\mathrm{d}q \leq 0.5$, $G \in [0,1]$. In case of perfect equality, i.e., if the Lorenz curve is equal to the Line of equality, the Gini coefficient is given by $G = 0$. In case of perfect inequality, i.e., if one person possesses all wealth, the Lorenz curve is given by $L(F,q) = 0$ for all $q \in [0,1)$ and $L(F,q) = 1$ for $q = 1$. In this case, the Gini coefficient calculates as $G = 1$.

In order to work with HFCS data, we need a weighted empirical version of the Gini coefficient. Therefore, recall that $\boldsymbol{d} = (d_1,\dots,d_n) \in \mathbb{R}^n_{>0}$ denotes the vector of weights and $\boldsymbol{x} = (x_1,\dots,x_n)$ denotes the vector of wealth in increasing order of the n households participating in the HFCS (in a given country) after the Pareto fitting. We define

$$W_k := \frac{\sum_{\ell=1}^k d_\ell}{\sum_{\ell=1}^n d_\ell},\ W_0 = 0 \quad \text{and} \quad X_k := \frac{\sum_{\ell=1}^k d_\ell x_\ell}{\sum_{\ell=1}^n d_\ell x_\ell},\ X_0 = 0, \tag{48}$$

as the cumulative share of population and the cumulative share of wealth, respectively. The Gini coefficient is then defined as

$$G := 1 - \sum_{k=1}^n (W_k - W_{k-1})(X_k + X_{k-1}). \tag{49}$$

3.2 Definition of the problem that has to be solved

We have already established the fact that we would like to minimize the difference of the Gini coefficient before and after matching the HFCS data with the NtlA figures. Therefore, the objective function f could look like

$$f(\boldsymbol{a},\boldsymbol{d},\boldsymbol{x}) = (G^* - G(\boldsymbol{a},\boldsymbol{d},\boldsymbol{x}))^2, \tag{50}$$

or

$$f(\boldsymbol{a},\boldsymbol{d},\boldsymbol{x}) = |G^* - G(\boldsymbol{a},\boldsymbol{d},\boldsymbol{x})|, \tag{51}$$

where G^* denotes the Gini coefficient after the Pareto fitting and

$$\begin{aligned} G(\boldsymbol{a},\boldsymbol{d},\boldsymbol{x}) &= 1 - \sum_{k=1}^n (W_k - W_{k-1})(X_k + X_{k-1}) \\ &= 1 - \left[\sum_{k=1}^n (W_k - W_{k-1}) \frac{2\sum_{\ell=1}^{k-1} a_\ell d_\ell x_\ell + a_k d_k x_k}{F}\right] \end{aligned} \tag{52}$$

with $W_k = \frac{\sum_{\ell=1}^k d_\ell}{\sum_{\ell=1}^n d_\ell}$, $W_0 = 0$ and $X_k = \frac{\sum_{\ell=1}^k a_\ell d_l x_\ell}{\sum_{\ell=1}^n a_\ell d_\ell x_\ell}, X_0 = 0$. However, we observe that the Gini coefficient is invariant under multiplication with a constant. Considering our constraint in (43), it is evident that one solution is given by

$$a_1 = \dots = a_n = \frac{F}{\sum_{i=1}^n d_i x_i}. \tag{53}$$

Consequently, our interest lies in investigating the solution space of the following system of equations:

$$G^* - G(\boldsymbol{a}, \boldsymbol{d}, \boldsymbol{x}) = 0, \tag{54}$$

$$\sum_{i=1}^{n} d_i a_i x_i = F. \tag{55}$$

For economic reasons, we want to ensure that the wealth order of the included households does not change. Therefore, we further introduce additional constraints. The wealth order is preserved if $a_i x_i \leq a_{i+1} x_{i+1}$ holds for all $i \in \{1, \ldots, n-1\}$. Since the aggregated sums of the HFCS are usually lower than financial account figures, as stated by Chakraborty and Waltl ([5]), we have to allocate additional wealth. Hence, we want to ensure that each household has at least the same amount of wealth after adjustment. Therefore, we add the constraint $a_i \geq 1$ for all $i \in \{1, \ldots, n\}$. Nevertheless, this usually leads to an infinite solution space. To see this, consider the following example:
Assume that $n = 4$, $\boldsymbol{x} = (1, 3, 5, 7)$, $\boldsymbol{d} = (2, 4, 3, 3)$, $\boldsymbol{W} = (\frac{2}{12}, \frac{6}{12}, \frac{9}{12}, 1)$ and $F = 100$. We want to solve the problem

$$\begin{cases} \text{solve} \quad G^* - G(\boldsymbol{a}, \boldsymbol{d}, \boldsymbol{x}) = 0 \\ \text{subject to} \\ \quad -a_i \leq -1 \quad \forall i, \\ \quad\quad a_1 \leq 3a_2 \leq 5a_3 \leq 7a_4, \\ \quad 2a_1 + 12a_2 + 15a_3 + 21a_4 = 100. \end{cases} \tag{56}$$

A straightforward calculation shows that $a_3 = \frac{568 - 38a_1 - 156a_2}{90}$ and $a_4 = \frac{16 + 13a_1 + 42a_2}{63} \geq 1$. To ensure that $a_3 \geq 1$, the inequality $a_1 \leq \frac{478 - 156a_2}{38}$ has to hold. Now, the inequalities that have been introduced to preserve the wealth order can be rewritten as

$$5a_3 \leq 7a_4 \Leftrightarrow a_1 \geq \frac{67}{8} - \frac{30}{8} a_2, \tag{57}$$

$$3a_2 \leq 5a_3 \Leftrightarrow a_1 \leq \frac{284}{19} - \frac{105}{19} a_2, \tag{58}$$

$$a_1 \leq 3a_2. \tag{59}$$

A graphical presentation of the solution space in this example is given by Figure 3. Indeed, the solution space is infinite.

As a consequence, we now try to confine the solution space but still want to ensure that the wealth order does not change. We, therefore, replace the constraints $a_i x_i \leq a_{i+1} x_{i+1}$ for all $i \in \{1, \ldots, n-1\}$ by the stronger constraints $a_i \leq a_{i+1}$ for all $i \in \{1, \ldots, n-1\}$. Thus, the problem we want to solve equates to

$$\begin{cases} \text{solve} \quad G^* - G(\boldsymbol{a}, \boldsymbol{d}, \boldsymbol{x}) = 0 \\ \text{subject to} \\ \quad\quad -a_i \leq -1 \, \forall i, \\ \quad a_i - a_{i+1} \leq 0 \quad \forall i \in \{1, \ldots, n-1\}, \\ \quad \sum_{i=1}^{n} d_i a_i x_i = F. \end{cases} \tag{60}$$

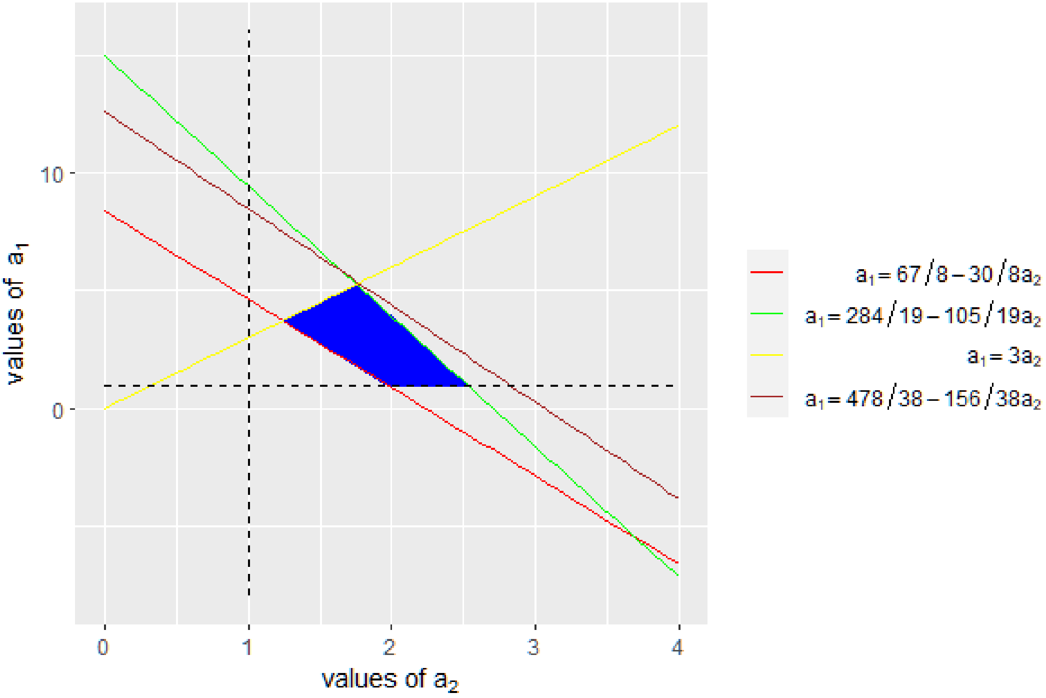

Figure 3: The blue area shows the solution space of Problem (56) for the values a_1 and a_2. The adjustment coefficients a_3 and a_4 can be derived by the upper expressions.

With the constraints in Problem (60), we will show in the following, that the solution given in Equation (53) is unique.

Theorem 2. Assuming that $\sum_{i=1}^{n} d_i x_i \leq F$, the unique solution to the optimization problem defined in Problem (60) is given by

$$a_1 = \cdots = a_n = \frac{F}{\sum_{i=1}^{n} d_i x_i}. \tag{61}$$

Proof. The strategy to prove the theorem will be as follows:
First of all, we observe that there are two equation, namely $G^* - G(\boldsymbol{a}, \boldsymbol{d}, \boldsymbol{x}) = 0$ and $\sum_{i=1}^{n} d_i a_i x_i = F$, each of them having an $n-1$ dimensional solution space (hyperplane). We will intersect these hyperplanes in order to get expressions for a_{n-1} and a_n dependent on $a_1, \ldots, a_{n-2}$ (Appendix, Lemma 1). By applying the inequality constraints in (60), we will first show that $a_1 \leq \frac{F}{\sum_{i=1}^{n} d_i x_i}$ (Appendix, Lemma 2). Then, by induction, we will conclude that $a_m \leq \frac{F}{\sum_{i=1}^{n} d_i x_i}$ for all $m \in \{1, \ldots, n-2\}$ (Appendix, Lemma 3).
Having proved that $a_m \leq \frac{F}{\sum_{i=1}^{n} d_i x_i}$ holds for all $m \in \{1, \ldots, n-2\}$, we will show step by step (starting with $m = n-2$) that $a_m \geq \frac{F}{\sum_{i=1}^{n} d_i x_i}$ and conclude that $a_m = \frac{F}{\sum_{i=1}^{n} d_i x_i}$ for all $m \in \{1, \ldots, n-2\}$ (Appendix, Lemma 4). Inserting these values into the expressions we have established for a_{n-1} and a_n, we can conclude that the solution $a_1 = \cdots = a_n = \frac{F}{\sum_{i=1}^{n} d_i x_i}$ is, indeed, unique. □

3.3 Lower and upper bound for the Gini coefficient after adjustment

Besides the solution derived in the previous section, it is also of practical interest to derive the lower and upper bound of inequality that can be implied via the additional amount of wealth that needs to be allocated (or subtracted in the case of overcoverage). Our focus lies on the case of undercoverage, i.e. $\sum_{i=1}^{n} d_i x_i < F$, because it is the common case. In the case of overcoverage, the results can be derived similarly. We will use the results of this section when we perform a small case study in Chapter 4. Having a lower and upper bound is interesting, because it provides a range of inequality and we can not only detect where the Gini coefficient resulting from the proportional adjustment is located but also compare it to Gini coefficients obtained by applying other methods as we will see in the next chapter.

3.3.1 Upper bound

Let us first take a look at the upper bound, in the case of undercoverage, i.e., we want to maximize the Gini coefficient such that

(i) all households keep at least their current instrument holdings,

(ii) the households' ranking in terms of their instrument holdings is preserved and

(iii) the NtlA total is matched.

This equates to the following problem:

$$\begin{cases} \max\limits_{\boldsymbol{a}\in\mathbb{R}^n} G(\boldsymbol{a},\boldsymbol{d},\boldsymbol{x}) \\ \text{subject to} \\ \quad -a_i \leq -1\,\forall i, \\ \quad a_i x_i - a_{i+1}x_{i+1} \leq 0\,\forall i \in \{1,\dots,n-1\}, \\ \quad \sum\limits_{i=1}^{n} d_i a_i x_i = F. \end{cases} \tag{62}$$

Let's start with a more general optimization problem, where we relax the constraints, ignoring the constraint of preserving the households' ranking. We will later see that the solution to this relaxed problem, which is easier to solve, also constitutes a solution to the OP of Problem (62).

$$\begin{cases} \max\limits_{\boldsymbol{a}\in\mathbb{R}^n} G(\boldsymbol{a},\boldsymbol{d},\boldsymbol{x}) \\ \text{subject to} \\ \quad -a_i \leq -1\,\forall i, \\ \quad \sum\limits_{i=1}^{n} d_i a_i x_i = F. \end{cases} \tag{63}$$

Recall the definition of the Gini coefficient:

$$G = 1 - \sum_{k=1}^{n} (W_k - W_{k-1})(X_k + X_{k-1}),$$

with

$$W_k = \frac{\sum_{\ell=1}^{k} d_\ell}{\sum_{\ell=1}^{n} d_\ell}, W_0 = 0 \quad \text{and} \quad X_k = \frac{\sum_{\ell=1}^{k} a_\ell d_\ell x_\ell}{\sum_{\ell=1}^{n} a_\ell d_\ell x_\ell}, X_0 = 0.$$

Note that the Gini coefficient can be written as

$$\begin{aligned}
G(\mathbf{a}, \mathbf{d}, \mathbf{x}) &= 1 - \left[\sum_{k=1}^{n} \underbrace{(W_k - W_{k-1})}_{\Delta W_k} \frac{2\sum_{\ell=1}^{k-1} a_\ell d_\ell x_\ell + a_k d_k x_k}{F}\right] \\
&= 1 - \left[\sum_{k=1}^{n} \Delta W_k \frac{2\sum_{\ell=1}^{k-1} a_\ell d_\ell x_\ell + a_k d_k x_k}{F}\right] \\
&= 1 - \frac{1}{F} \sum_{i=1}^{n} a_i \left(\Delta W_i d_i x_i + \sum_{k=i+1}^{n} 2\Delta W_k d_i x_i\right) \text{ (change the order of summation)} \\
&= 1 - \frac{1}{F} \sum_{i=1}^{n} a_i d_i x_i \left(\Delta W_i + \sum_{k=i+1}^{n} 2\Delta W_k\right) \\
&= 1 - \frac{1}{F} \sum_{i=1}^{n} a_i d_i x_i \left(\Delta W_i + 2(\underbrace{W_n}_{=1} - W_i)\right) \\
&= 1 - \frac{1}{F} \sum_{i=1}^{n} a_i d_i x_i \left(2 - W_i - W_{i-1}\right).
\end{aligned} \tag{64}$$

Hence, (63) is equivalent to

$$\begin{cases}
\min_{\boldsymbol{a} \in \mathbb{R}^n} \sum_{i=1}^{n} a_i d_i x_i \left(2 - W_i - W_{i-1}\right) \\
\text{subject to} \\
\quad - a_i \leq -1 \, \forall i, \\
\quad \sum_{i=1}^{n} d_i a_i x_i = F.
\end{cases} \tag{65}$$

The Lagrange function is given by

$$L(a, \lambda, \nu) = \sum_{i=1}^{n} a_i d_i x_i \left(2 - W_i - W_{i-1}\right) + \sum_{i=1}^{n} \mu_i (1 - a_i) + \lambda \left(\sum_{i=1}^{n} a_i d_i x_i - F\right).$$

We are looking for a KKT-Point. Using Lagrange multiplier, we have to solve the following system of equations:

$$\begin{aligned}
&\text{(i)} && \frac{\partial L(a, \lambda, \nu)}{\partial a_i} = d_i x_i \left(2 - W_i - W_{i-1}\right) - \mu_i + \lambda d_i x_i = 0 \quad \text{for all } i \in \{1, \ldots, n\}, \\
&\text{(ii)} && \sum_{i=1}^{n} a_i d_i x_i - F = 0, \\
&\text{(iii)} && \mu_i \geq 0 \text{ for all } i \in \{1, \ldots, n\} \quad \text{(dual feasibility condition)}, \\
&\text{(iv)} && \mu_i (1 - a_i) = 0 \text{ for all } i \in \{1, \ldots, n\} \quad \text{(complementary slackness condition)}.
\end{aligned} \tag{66}$$

We differentiate between three different cases:

Case 1: $\lambda = -\Delta W_n = -(W_n - W_{n-1})$.
Then, by (66)(i), $\mu_n = 0$. Further, we can solve (66)(i) for μ_i, $i \in \{1, \dots, n-1\}$, and get

$$\begin{aligned}\mu_i &= d_i x_i \left(2 - W_i - W_{i-1}\right) + \lambda d_i x_i \\ &= d_i x_i \left(2 - W_i - W_{i-1}\right) - (W_n - W_{n-1}) d_i x_i \\ &= d_i x_i \left(W_n + W_{n-1} - W_i - W_{i-1}\right) > 0.\end{aligned} \tag{67}$$

So we know that in this case $\mu_n = 0$ and $\mu_i > 0$ for all $i \in \{1, \dots, n-1\}$. Due to the complementary slackness condition (66)(iv), $a_i = 1$ for all $i \in \{1, \dots, n-1\}$. The constraint $\sum_{i=1}^{n} a_i d_i x_i = F$ then yields $a_n = \frac{F - \sum_{i=1}^{n-1} x_i d_i}{d_n x_n}$.
In this case, all additional wealth is allocated to the already richest household.

Case 2: $\lambda > -\Delta W_n$.
Then, by (66)(i), $\mu_n > 0$. Solving (66)(i) for μ_i, $i \in \{1, \dots, n-1\}$ yields

$$\begin{aligned}\mu_i &= d_i x_i \left(2 - W_i - W_{i-1}\right) + \lambda d_i x_i \\ &> d_i x_i \left(W_n + W_{n-1} - W_i - W_{i-1}\right) > 0.\end{aligned} \tag{68}$$

So we know that in this case $\mu_i > 0$ for all $i \in \{1, \dots, n\}$. Due to the complementary slackness condition (66)(iv), $a_i = 1$ for all $i \in \{1, \dots, n\}$. Thus, $a_1 = \dots = a_n = 1$ which is not a feasible solution.

Case 3: $\lambda < -\Delta W_n$.
Then, by (66)(i), $\mu_n = \Delta W_n d_n x_n + \lambda d_n x_n < 0$, which contradicts (66)(iii). Thus, there exists no KKT-Point in this case.

Therefore, the solution to Problem (63) is given by $\boldsymbol{a} = \left(1, \dots, 1, \frac{F - \sum_{i=1}^{n-1} x_i d_i}{d_n x_n}\right)^T$.

Since we have assumed undercoverage, $\frac{F - \sum_{i=1}^{n-1} x_i d_i}{d_n x_n} > 1$. Thus, $\boldsymbol{a} = \left(1, \dots, 1, \frac{F - \sum_{i=1}^{n-1} x_i d_i}{d_n x_n}\right)^T$ also fulfills the constraint $a_i x_i - a_{i+1} x_{i+1} \leq 0$ for all $i \in \{1, \dots, n-1\}$ and we observe that this is also a feasible point for Problem (62). We conclude that this is, therefore, a solution to Problem (62). Hence, the upper bound for the Gini coefficient after adjustment is the case where we give all additional wealth to the wealthiest household, i.e. $\boldsymbol{a} = \left(1, \dots, 1, \frac{F - \sum_{i=1}^{n-1} x_i d_i}{d_n x_n}\right)^T$.

3.3.2 Lower bound

Now we want to establish the lower bound, i.e. we want to solve

$$\begin{cases} \min\limits_{\boldsymbol{a}\in\mathbb{R}^n} G(\boldsymbol{a},\boldsymbol{d},\boldsymbol{x}) \\ \text{subject to} \\ \quad -a_i \leq -1\,\forall i, \\ \quad a_i x_i - a_{i+1}x_{i+1} \leq 0\,\forall i \in \{1,\dots,n-1\}, \\ \quad \sum\limits_{i=1}^n d_i a_i x_i = F. \end{cases} \tag{69}$$

Again, we will first look at the more general problem, ignoring the constraint on households' ranking,

$$\begin{cases} \min\limits_{\boldsymbol{a}\in\mathbb{R}^n} G(\boldsymbol{a},\boldsymbol{d},\boldsymbol{x}) \\ \text{subject to} \\ \quad -a_i \leq 1\,\forall i, \\ \quad \sum\limits_{i=1}^n d_i a_i x_i = F. \end{cases} \tag{70}$$

Using (64), this is equivalent to

$$\begin{cases} \min\limits_{\boldsymbol{a}\in\mathbb{R}^n} -\sum\limits_{i=1}^n a_i d_i x_i \left(2 - W_i - W_{i-1}\right) \\ \text{subject to} \\ \quad -a_i \leq -1\,\forall i, \\ \quad \sum\limits_{i=1}^n d_i a_i x_i = F. \end{cases} \tag{71}$$

The Lagrange function is given by

$$L(a,\lambda,\nu) = -\sum_{i=1}^n a_i d_i x_i \left(2 - W_i - W_{i-1}\right) + \sum_{i=1}^n \mu_i(1-a_i) + \lambda\left(\sum_{i=1}^n a_i d_i x_i - F\right).$$

We are looking for a KKT-Point. Using Lagrange multiplier, we have to solve the following system of equations:

$$\begin{aligned} &\text{(i)} && \frac{\partial L(a,\lambda,\nu)}{\partial a_i} = -d_i x_i \left(2 - W_i - W_{i-1}\right) - \mu_i + \lambda d_i x_i = 0 \quad \text{for all } i \in \{1,\dots,n\}, \\ &\text{(ii)} && \sum_{i=1}^n a_i d_i x_i - F = 0, \\ &\text{(iii)} && \mu_i \geq 0 \text{ for all } i \in \{1,\dots,n\} \quad \text{(dual feasibility condition)}, \\ &\text{(iv)} && \mu_i(1-a_i) = 0 \text{ for all } i \in \{1,\dots,n\} \quad \text{(complementary slackness condition)}. \end{aligned} \tag{72}$$

We first observe from (72)(i) and (72)(iii) that a solution to (72) only exists if

$$\mu_i = -d_i x_i \left(2 - W_i - W_{i-1}\right) + \lambda d_i x_i \geq 0 \quad \text{for all } i \in \{1,\dots,n\}. \tag{73}$$

Thus, we get

$$\lambda \geq 2 - W_i - W_{i-1} \quad \text{for all } i \in \{1, \ldots, n\}, \tag{74}$$

which implies

$$\lambda \geq 2 - W_1 - W_0 = 2 - W_1. \tag{75}$$

From (72)(i) and (75), it follows that

$$\mu_n = -\Delta W_n d_n x_n + \lambda d_n x_n \geq -\Delta W_n d_n x_n + 2 d_n x_n - W_1 d_n x_n > 0. \tag{76}$$

Due to the complementary slackness condition (72)(iv), we must have $a_n = 1$. Similarly, for $i = n - 1$, we have

$$\begin{aligned} \mu_{n-1} &= -d_{n-1} x_{n-1} \left(2 - W_{n-1} - W_{n-2}\right) + \lambda d_{n-1} x_{n-1} \\ &\geq d_{n-1} x_{n-1} \left(W_{n-1} + W_{n-2} - W_1\right) > 0. \end{aligned} \tag{77}$$

Again due to the complementary slackness condition, we conclude $a_{n-1} = 1$. Iteratively, we get $a_i = 1$ for all $i \in \{2, \ldots, n\}$. In case of $i = 1$,

$$\mu_1 = -d_1 x_1 \left(2 - W_1\right) + \lambda d_1 x_1. \tag{78}$$

Thus, if $\lambda > 2 - \Delta W_1$, we would have $\mu_1 > 0$ and due to the complementary slackness condition, $a_1 = 1$. But this is only a solution to (70), if $\sum_{i=1}^{n} d_i x_i = F$. We conclude, that $\lambda = 2 - \Delta W_1$ must hold and we have $a_i = 1$ for all $i \in \{2, \ldots, n\}$.
But we can only increase a_1 until $a_1 x_1 = x_2$ because with a further increase, the order of the housholds' instrument holdings and therefore the formula for calculating the Gini coefficient would change[7]. Hence, for calculating the lower bound, we repeatedly have to increase the instrument holdings of the poorest household[8] until either the adjusted instrument holding equals the second poorest household or the full gap $F - \sum_{i=1}^{n} d_i x_i$ is allocated.

It is also worth pointing out that the optimization problems (62) and (69) could be solved with the simplex algorithm using generic solvers in R and matlab. Nonetheless, it is important and often useful to know these bounds analytically, since HFCS datasets are often large and thus resulting in long runtimes.

[7] Remember that the Gini coefficient is defined for $a_1 x_2 \leq a_2 x_2 \leq \ldots \leq a_n x_n$.

[8] Households with the same amount of instrument holdings are treated as one.

4 Case study covering household wealth of Germany

In this section, we investigate the data from the third wave covering household wealth in Germany. The survey was conducted between March 2017 and October 2017 ([7]). We compare the approach analyzed in Chapter 3, namely the proportional adjustment, with a second approach. As the Expert Group on Linking macro and micro data for the household sector (EG-LMM) ([19]) stated, a *"standard approach involves the calibration method"*. Therefore, we will use this multivariate calibration approach for comparison to the proportional adjustment method. Before we explain the multivariate calibration approach, note that the HFCS measures the wealth of households with several instruments. The assets are comprised of 'Deposits', 'Bonds', 'Shares', 'Funds', 'Voluntary Pension', 'Financial Business Wealth', 'Non-Financial Business Wealth' and 'Housing Wealth', whereas liabilities are divided into 'Mortgage Liabilities' and 'Other Liabilities'. The net wealth of a household is the difference between assets and liabilities. In Figure 4, we see the coverage ratio for each instrument, which is defined as the HFCS instrument total divided by the NtlA aggregate. Note that the HFCS data have already been adjusted for the missing wealthy. We see that in most of the cases there is undercoverage, i.e., the HFCS aggregate is lower than the corresponding NtlA instrument value. Nevertheless, in the case of 'Shares' and 'Bonds' the HFCS and the NtlA total is matched exactly via the added wealthy households whereas in case of 'Financial Business Wealth' we observe an overcoverage of approximately 4%, i.e. the HFCS aggregate is higher than the NtlA total. The undercoverage is most pronounced for 'Deposits' with a coverage ratio of roughly 55%, moderately pronounced for 'Funds, 'Voluntary Pension' and 'Other Liabilities' in the range of 77% to 82% and less pronounced for 'Non-Financial Business Wealth', 'Housing Wealth' and 'Mortgage Liabilities' with coverage ratios of roundabout 95%.

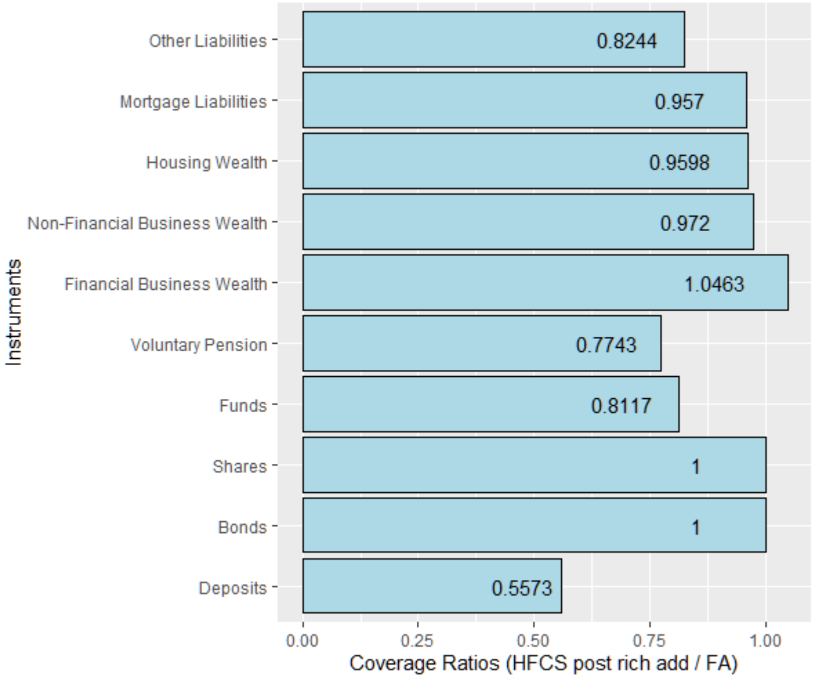

Figure 4: Coverage ratios of wealth instruments for German households.

4.1 Multivariate calibration

Now, we move on to introducing the multivariate calibration method of the EG-LMM. Calibration is typically used to adjust for unit non-response in a sample survey (cf. ([14])). The following summary is derived from internal R code and notes developed by ECB staff members. A similar approach can be found in ([3]).
This method is based on the following optimization problem,

$$\min_{\boldsymbol{a}\in\mathbb{R}^n} \chi^2(\boldsymbol{d},\boldsymbol{a}) = \min_{\boldsymbol{a}\in\mathbb{R}^n} \sum_{i=1}^{n} \frac{(d_i a_i - d_i)^2}{d_i}. \tag{79}$$

The objective function is chosen such that it minimizes the impact of the household specific correction factors $\boldsymbol{a}$ on the household weights $\boldsymbol{d}$. This impact is measured by the χ^2 distance of the weights before and after adjusting them by $\boldsymbol{a}$.
According to ([19]), *"if the calibration approach is selected, instruments with similar comparability should be calibrated together"*. Therefore, all instruments are divided into three disjoint groups, namely financial assets, business wealth assets, and housing wealth: {Deposits, Bonds, Shares, Funds, Voluntary Pension, Mortgage Liabilities, Other Liabilities}, {Financial Business Wealth, Non-Financial Business Wealth} and {Housing Wealth}.
As constraints, the adjusted HFCS instrument totals have to match the NtlA aggregates, for each instrument j,[9]

$$\sum_{i\in I_{\text{bottom}}} d_i a_i x_{ij} = F_{\text{calib.bot},j}, \tag{80}$$

$$\sum_{i\in I_{\text{top}}} d_i a_i x_{ij} = F_{\text{calib.top},j}, \tag{81}$$

where I_{bottom} denotes all households belonging to the bottom part of the wealth distribution and I_{top} all households belonging to the top part of the wealth distribution. This also means that $I_{\text{bottom}} \cap I_{\text{top}} = \emptyset$ and $I_{\text{bottom}} \cup I_{\text{top}} = \{1,\dots,n\}$ hold. To limit the change on a single household, some lower and upper bounds for all a_i are included, for our example[10] these bounds are

$$0.03 \le a_i \le 100. \tag{82}$$

Since we have three groups of instruments, we also have three different optimization problems. The first one matches the financial assets

$$\begin{cases} \min\limits_{\boldsymbol{a}\in\mathbb{R}^n} \chi^2(\boldsymbol{d},\boldsymbol{a}) = \min\limits_{\boldsymbol{a}\in\mathbb{R}^n} \sum\limits_{i=1}^{n} \dfrac{(d_i a_i - d_i)^2}{d_i} \\ \text{s.t.} \forall j \in \{\text{Deposits, Bonds, Shares, Funds,} \\ \qquad\qquad \text{Voluntary Pension, Mortgage Liabilities, Other Liabilities}\}, \\ \sum\limits_{i\in I_{\text{bottom}}} d_i a_i x_{ij} = F_{\text{calib.bot},j}, \\ \quad \sum\limits_{i\in I_{\text{top}}} d_i a_i x_{ij} = F_{\text{calib.top},j}, \\ \quad 0.03 \le a_i \le 100 \, \forall i. \end{cases} \tag{83}$$

[9] The staff members set up two constraints in the R code. They distinguish between the households from the original data set, I_{bottom}, and the households obtained from rich lists and sampling, I_{top}.

[10] These bounds are the default values in the internal R code provided by staff members of the EG-LMM.

Likewise, the second optimization problem matches the holdings in business wealth

$$\begin{cases} \min\limits_{\boldsymbol{a}^{\mathrm{biz}} \in \mathbb{R}^n} \chi^2(\boldsymbol{d}, \boldsymbol{a}^{\mathrm{biz}}) = \min\limits_{\boldsymbol{a}^{\mathrm{biz}} \in \mathbb{R}^n} \sum_{i=1}^{n} \frac{\left(d_i a_i^{\mathrm{biz}} - d_i\right)^2}{d_i} \\ \text{s.t.} \forall j \in \{\text{Financial Business Wealth, Non-Financial Business Wealth}\}, \\ \sum\limits_{i \in I_{\mathrm{bottom}}} d_i a_i^{\mathrm{biz}} x_{ij} = F_{\mathrm{calib.bot},j}, \\ \quad \sum\limits_{i \in I_{\mathrm{top}}} d_i a_i^{\mathrm{biz}} x_{ij} = F_{\mathrm{calib.top},j}, \\ \quad 0.03 \leq a_i^{\mathrm{biz}} \leq 100 \, \forall i, \end{cases} \tag{84}$$

and the third one matches the holdings in housing wealth

$$\begin{cases} \min\limits_{\boldsymbol{a}^{\mathrm{h}} \in \mathbb{R}^n} \chi^2(\boldsymbol{d}, \boldsymbol{a}^{\mathrm{h}}) = \min\limits_{\boldsymbol{a}^{\mathrm{h}} \in \mathbb{R}^n} \sum_{i=1}^{n} \frac{\left(d_i a_i^{\mathrm{h}} - d_i\right)^2}{d_i} \\ \text{s.t.} \forall j \in \{\text{Housing Wealth}\}, \\ \sum\limits_{i \in I_{\mathrm{bottom}}} d_i a_i^{\mathrm{h}} x_{ij} = F_{\mathrm{calib.bot},j}, \\ \quad \sum\limits_{i \in I_{\mathrm{top}}} d_i a_i^{\mathrm{h}} x_{ij} = F_{\mathrm{calib.top},j}, \\ \quad 0.03 \leq a_i^{\mathrm{h}} \leq 100 \, \forall i. \end{cases} \tag{85}$$

After solving those problems, the adjusted instrument holdings for each household i are given by

$$\begin{aligned} & a_i x_{ij} && \text{for all } j \in \{\text{Deposits, Bonds, Shares, Funds, Voluntary Pension,} \\ & && \qquad\qquad \text{Mortgage Liabilities, Other Liabilities}\}, \\ & a_i^{\mathrm{biz}} x_{ij} && \text{for all } j \in \{\text{Financial Business Wealth, Non-Financial Business Wealth}\}, \\ & a_i^{\mathrm{h}} x_{ij} && \text{for all } j \in \{\text{Housing Wealth}\}. \end{aligned} \tag{86}$$

4.2 Data analysis

By applying both the proportional allocation (PA) and the multivariate calibration (MC) to the adjusted HFCS data set, i.e. after adding the missing wealthy, the impact of both methods on the final households' instruments holdings can be analysed and compared. We start by providing some descriptive statistics, investigating in particular, whether the MC increases or decreases households' instrument holdings. While the PA increases (decrease) all households' instrument holdings in the case of undercoverage (overcoverage), the MC can lead to mixed effects. In Chapter 3, where we introduced the PA method, we wanted to ensure that each household does not have less instrument holdings after the adjustment in case of undercoverage and does not have higher instrument holdings in case of overcoverage. For that purpose we added the constraints $a_i \geq 1$ for all $i \in \{1, \ldots, n\}$ (undercoverage). This makes sense, because in the case of undercoverage (overcoverage) we have no further information that would justify to decrease (increase) holdings of certain

	lower instr. holdings	equal instr. holdings	higher instr. holdings
$\boldsymbol{a}$	12,675,437 (31.41%)	26,642,813 (66.03%)	1,034,157 (2.56%)
$\boldsymbol{a}^{\text{biz}}$	3,492,582 (8.65%)	36,469,765 (90.38%)	390,060 (0.97%)
$\boldsymbol{a}^{\text{h}}$	18,553,099 (45.98%)	1,558,559 (3.86%)	20,240,747 (50.16%)

Table 1: Number of households with $a_i < 1$, $a_i = 1$ and $a_i > 1$ divided into the three distinct groups.

	Multivariate Calibration						Proportional Adjustment	
	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Coefficient	Instrument
$\boldsymbol{a}$	0.030	0.990	1.015	1.102	1.128	16.613	1.794	Deposits
							1.000	Bonds
							1.000	Shares
							1.232	Funds
							1.291	Vol. Pension
							1.045	Mort. Liab.
							1.212	Other Liab.
$\boldsymbol{a}^{\text{biz}}$	0.211	1.000	1.000	0.995	1.000	2.311	0.956	F-Biz. W.
							1.029	NF-Biz. W.
$\boldsymbol{a}^{\text{h}}$	0.519	0.943	1.000	0.966	1.000	19.447	1.042	Hous. W.

Table 2: Summary statistics of the adjustment coefficients received by applying MC and values of PA coefficients for each instrument (rounded to 3 digits).

households in the respective instrument. Indeed, we have proved in Theorem 2 that for each instrument j the adjustment coefficient of the PA is given by $a_i = \frac{F_j}{\sum_{i=1}^{n} d_i x_{ij}}$ for all $i \in \{1, \ldots, n\}$. Therefore, in case of undercoverage, for each household with $x_{ij} > 0$ the holdings in an instrument j after adjustment increase and only those households with zero instrument holdings do not receive additional allocations.

When analyzing the adjustment coefficients resulting from the MC method, we immediately see one weakness: The adjustment coefficients have a higher range and whether or not there is under- or overcoverage, the MC causes simultaneously an increase, a decrease, and unchanged holdings for seemingly random groups of households (cf. Table 1). More precisely and both for the situation of an under- and overcoverage, before running the MC we cannot predict whether the instrument holdings of a particular household will be increased, decreased, or left unaffected. Going into detail, several points should be mentioned. With regards to the values of the adjustment coefficient, we see that on the one hand, the lower bound of 0.03 which we have introduced is even touched for the optimization problem of financial assets, e.g., let the household's deposits amount to 100,000 before the MC, then it will be worth 3,000 after the MC. In this case, the household's financial assets diminish almost completely. On the other, the highest adjustment coefficient is almost 20 for housing wealth, meaning that the household's housing wealth will increase by a factor of 20, e.g., let the household's housing wealth amount to 500,000 before the MC, then it will be worth 10,000,000 after the MC (cf. Table 2). Compared to the coverage ratios in Figure 4, these values appear rather extreme and unjustified.

Furthermore, we see that median and mean of the adjustment coefficients in the MC approach are close or equal to 1 (cf. Table 2). This is reasonable in case of business wealth where coverage ratios are also close to 1, but for example in the case of 'Deposits', where the coverage ratio is close to 50%, we would definitely expect higher values. These findings raise the question whether there are parts of the wealth distribution that have disproportionally more or less wealth after the adjustment. This aspect is also mentioned in ([19]): *"the gap, when positive, is allocated more than proportionally to rich households. A negative gap would be allocated more to poor households".* We want to give evidence for this observation by calculating the share of wealth in each instrument possessed by the 10% richest households (Figure 5) and the poorer half of households (Figure 6).
Indeed, we see that for each instrument, the share of the poorer households reduces significantly, whereas the share of the richest households increases by a noteworthy margin. The decrease for the poorest households is especially pronounced in 'Deposits' and 'Other liabilities' with a roughly 5% reduction, but in terms of relative decrease, this is also true for 'Funds', 'Shares' and 'Bonds' where the proportions have almost or more than halved. For the richest households, the aforementioned increase amounts to circa 10% for instruments belonging to the group of financial assets ('Deposits', 'Bonds', 'Shares', 'Funds', 'Voluntary Pension', 'Mortgage Liabilities', 'Other Liabilities'), roughly 7% for 'Housing Wealth' and 5% for 'Non-Financial Business Wealth'. Interestingly, there is also an increase for 'Financial Business Wealth' where we have already seen in Figure 4 that there is an overcoverage for this instrument. So in fact, these observations strongly suggest that the MC method treats poorer households less favorably in terms of allocating the wealth gaps.

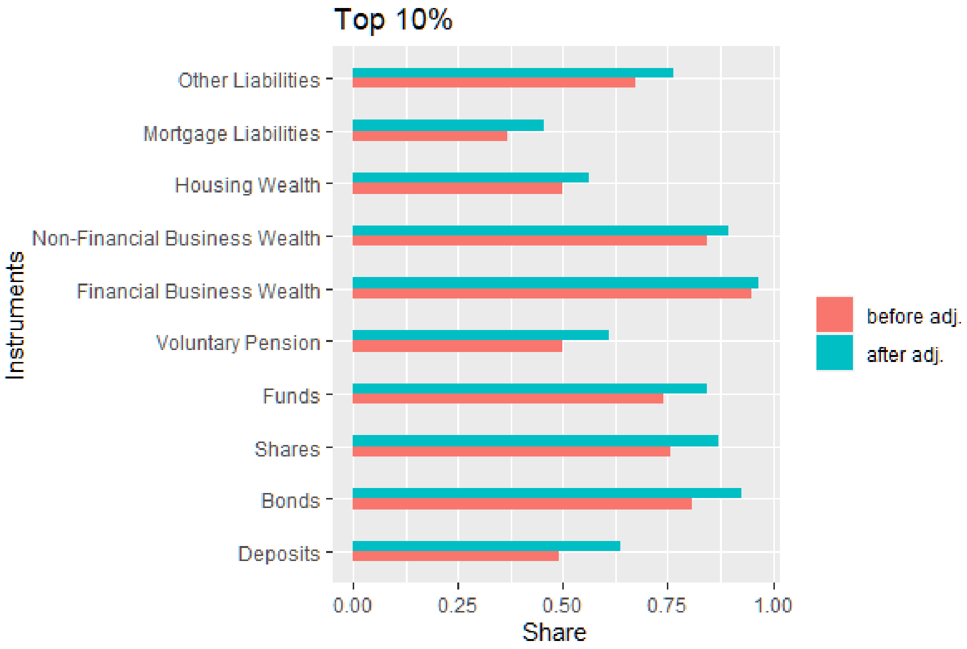

Figure 5: Share of instrument holdings possessed by the richest 10% of households before and after MC. Note that in the case of PA, the shares before and after adjustment do not change.

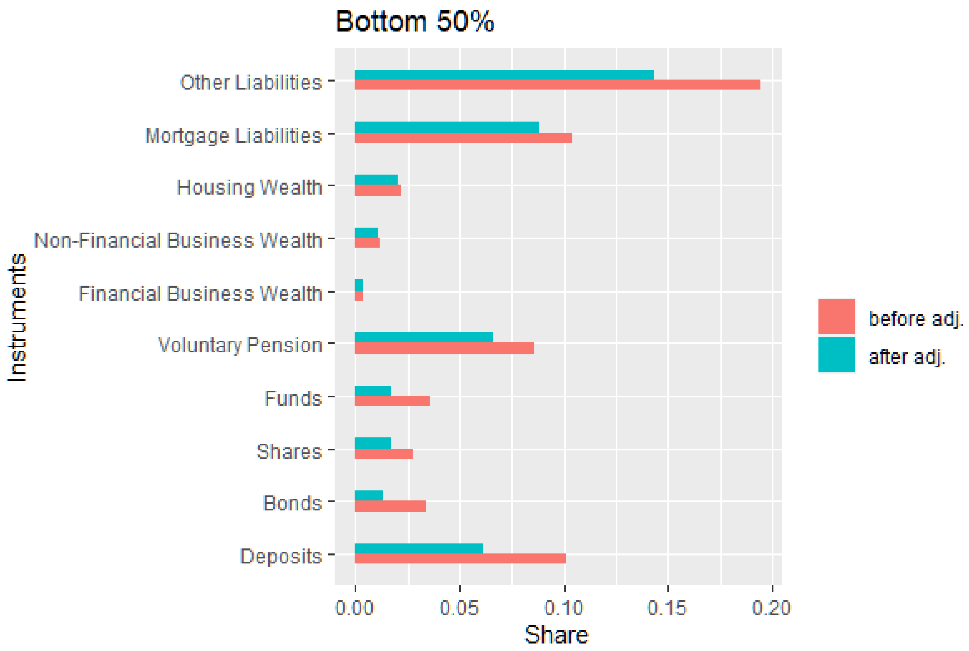

Figure 6: Share of instrument holdings possessed by the poorer half of households before and after MC. Note that in the case of PA, the shares before and after adjustment do not change.

This conclusion is also underpinned by the fact that the Gini coefficient for each instrument resulting from the MC is always higher than the one obtained from PA (cf. Figure 7). Obviously, the difference is very low for instruments that have a high coverage ratio, namely 'Bonds', 'Shares', 'Financial Business Wealth' and 'Non-Financial Business Wealth'.

We also calculated the minimal and maximum Gini coefficient for each instrument under the constraint that we do not want to change the order of wealth and not decreasing (increasing) any households' holdings in the case of undercoverage (overcoverage), see Figure 7. This was done by implementing the algorithms from Section 3.2 in R. The range is high for instruments that have a low coverage ratio, especially for 'Deposits' and 'Voluntary Pension'. We have already shown the reason for this relation in Section 3.2: The higher the wealth gap the more can be allocated to the poorer households or to the richest household. This makes the wealth distribution more equal in case of calculating the minimum or unequal when calculating the maximum. We can also see that the inequality of wealth is already skewed to the upper part, since the Gini coefficient is always closer to the maximum. Interestingly, for some instrument, the Gini coefficient after MC is even higher than the maximum Gini coefficient that preserves the wealth order. This is true for 'Funds', 'Non-Financial Business Wealth', 'Housing Wealth', 'Mortgage Liabilities' and 'Other Liabilities'.

For illustrative purposes, we have also plotted Lorenz curves for several instruments, namely 'Deposits', 'Housing Wealth' and 'Non-Financial Business Wealth', in Figures 8 to 10. Due to the high concentration of 'Non-Financial Business Wealth', Figure 10 only shows the share of population from 0.9 to 1. The plots clearly indicate that a higher under-

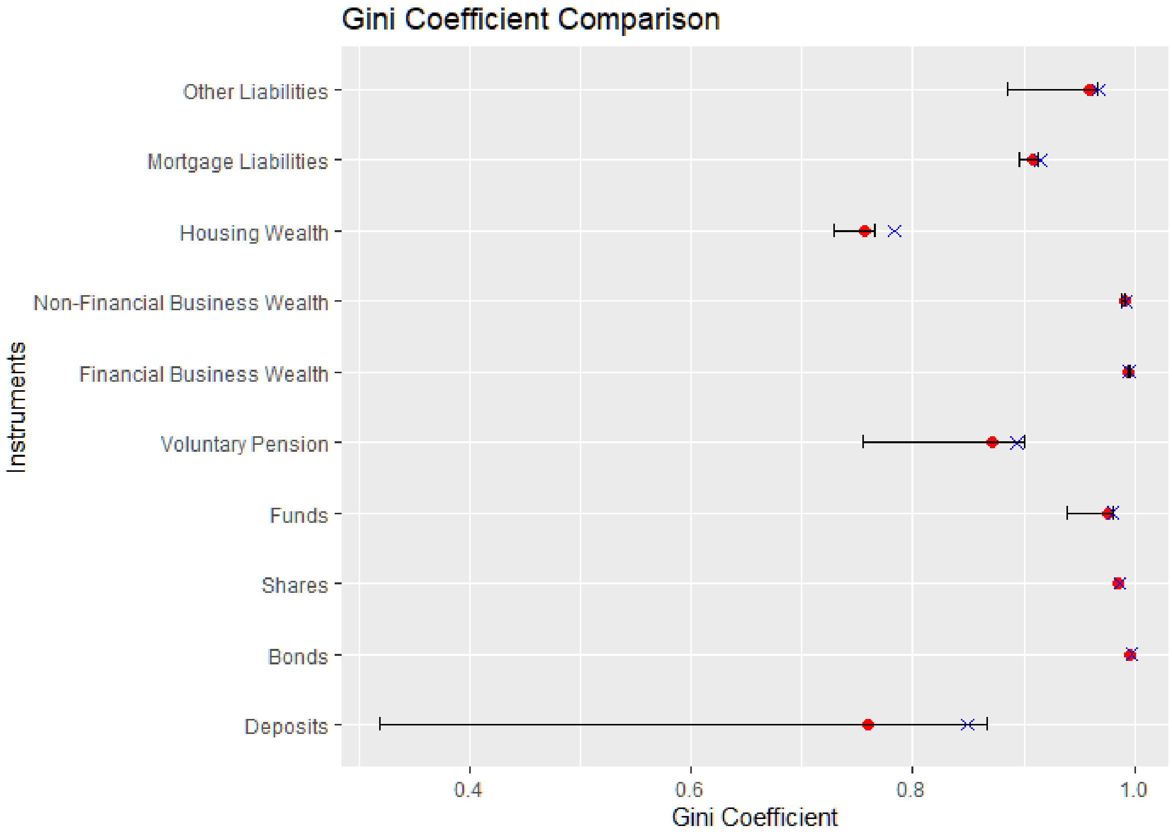

Figure 7: Gini Coefficients of each instrument resulting from proportional adjustment (red dot) and multivariate calibration (blue cross). The black lines denote the resulting range obtained from minimizing and maximizing according to the results in Section 3.2.

coverage implies more widespread Lorenz curves. The shape of the Lorenz curves is also influenced by the concentration of instrument holdings because wealth is only allocated to households with non-zero instrument holdings. For example, in the case of ‘Deposits’, where the undercoverage is high and more than 90% of the households actually possess ‘Deposits’, the Lorenz curve for the minimization method is similar to a straight line and much closer to the equality line than the other Lorenz curves. Even though the Lorenz curves for ‘Housing Wealth’ and ‘Non-Financial Business Wealth’ are closer together, we can observe that also for these instruments the Lorenz curve for the MC method is mostly below the Lorenz curve corresponding to the PA method. For ‘Housing Wealth’ and ‘Non-Financial Business Wealth’, this is even true when we compare the Lorenz curve of the MC approach with the Lorenz curve of the maximization method. This finding and the fact that the MC causes simultaneously an increase and a decrease of certain households immediately suggest that the MC method does not obey the assumption of wealth order perpetuation. Therefore, we would like to analyse to what extent the MC method preserves the rankings. In order to do so, we must introduce some rank correlation measures such as Kendall’s tau τ, which was first introduced by Maurice G. Kendall in ([11]).

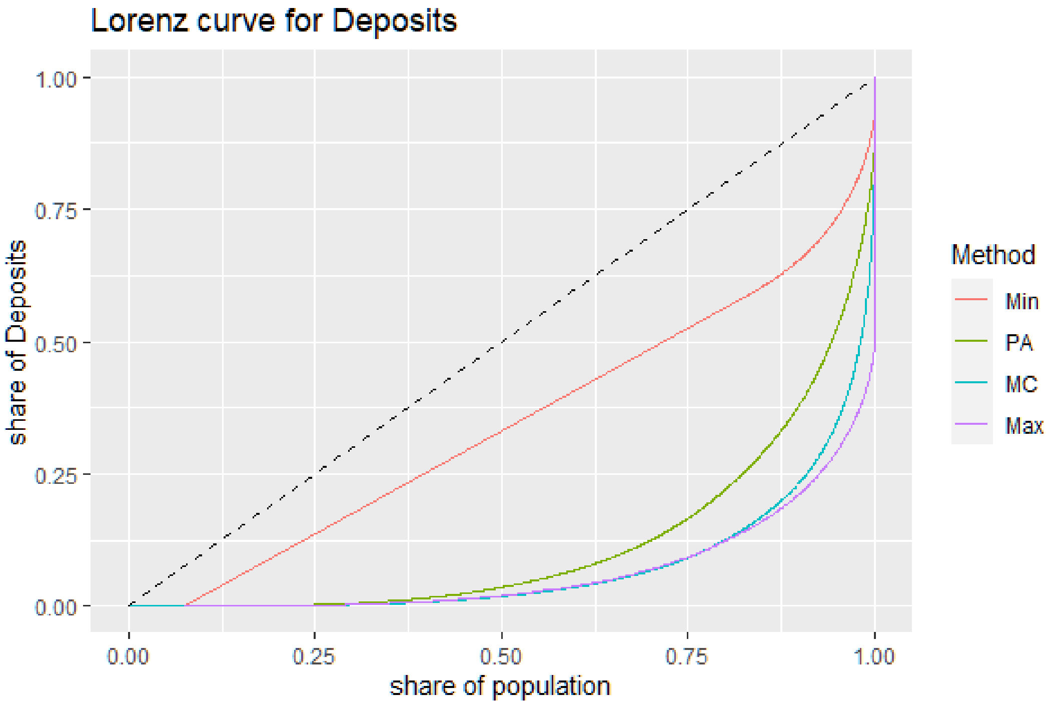

Figure 8: Lorenz curve for the different methods by the example of 'Deposits' (financial asset).

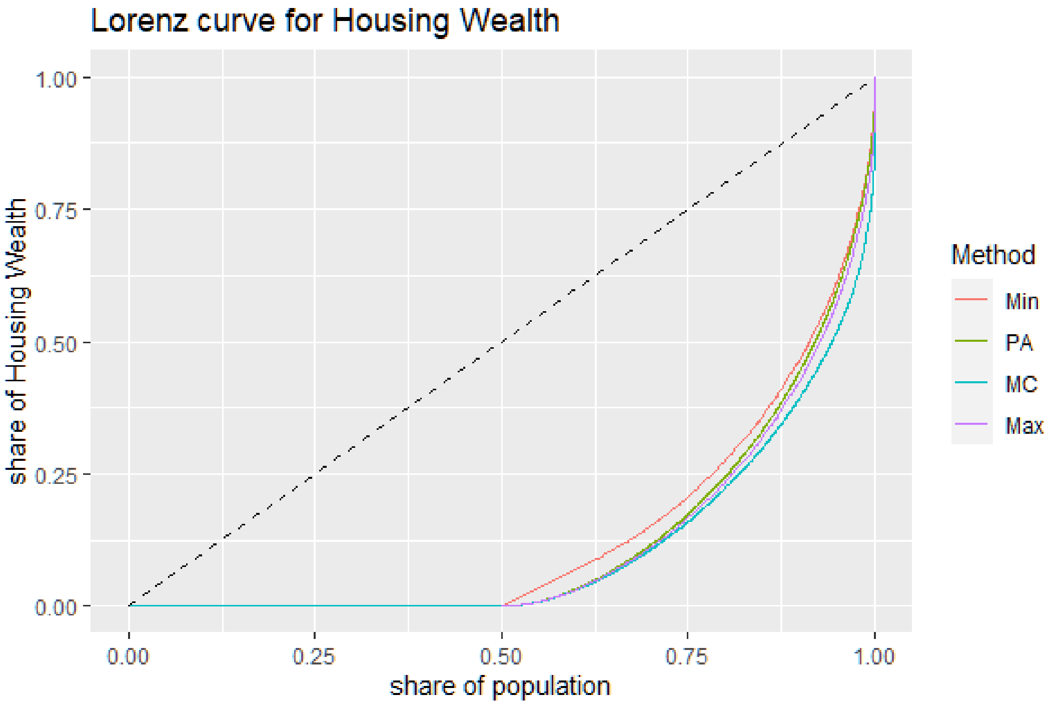

Figure 9: Lorenz curve for the different methods by the example of 'Housing Wealth'.

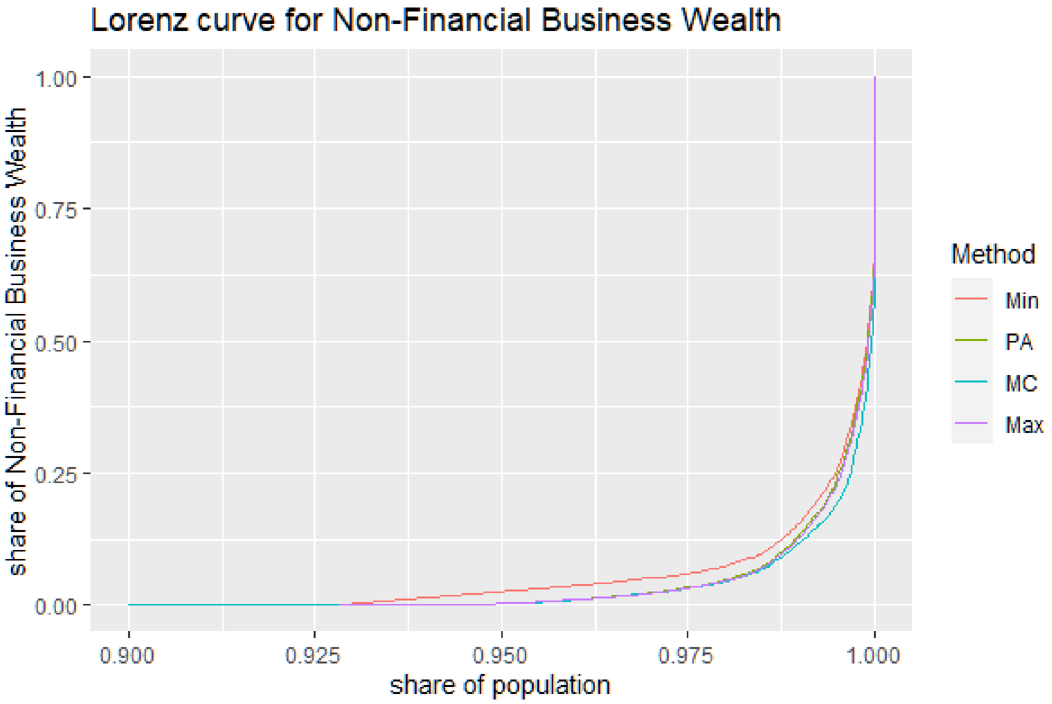

Figure 10: Lorenz curve for the different methods by the example of 'Non-Financial Business Wealth' (business wealth asset).

Multivariate Calibration	
Instrument	Kendall's tau
Other Liabilites	0.9538
Mortgage Liabilites	0.8747
Housing Wealth	0.9923
Non-Financial Business Wealth	0.9927
Financial Business Wealth	0.9955
Voluntary Pension	0.8865
Funds	0.8588
Shares	0.8656
Bonds	0.9325
Deposits	0.9170

Table 3: Kendall's tau for instrument holdings after applying multivariate calibration (rounded to 4 digits).

Definition 1. (cf. ([17], page 158))
Let $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ denote a set of observations from a vector (X, Y) of continuous random variables. A pair of observations (x_i, y_i) and (x_j, y_j) with $i < j$ is said to be concordant if $(x_i - x_j)(y_i - y_j) > 0$, i.e., if (x_i, x_j) and (y_i, y_j) have the same sort order. In any other case, the pair (x_i, y_i) and (x_j, y_j) is called discordant. Now, let c denote the number of concordant pairs and d the number of discordant pairs. Note that the number of distinct pairs (x_i, y_i) and (x_j, y_j) in the sample is given by $\frac{n(n-1)}{2} = \binom{n}{2}$. Then, the empirical version of Kendall's tau τ is given by

$$\tau = \frac{c-d}{\binom{n}{2}} \in [-1, 1]. \tag{87}$$

We have calculated Kendall's tau for each instrument after applying the MC in Table 3. Kendall's tau is 1 when no change in order takes place, which is true for each instrument after PA. When applying MC, we see that Kendall's tau ranges from approximately 0.85 for 'Funds' and Shares' to almost 1 for 'Housing Wealth', 'Financial Business Wealth' and 'Non-Financial Business Wealth'. We conclude that Kendall's tau suggests a rather high level of wealth order perpetuation. Nevertheless, there are cases where ranks are changed compared to proportional adjustment where Kendall's tau is by construction 1.

We also want to have a look at how the rankings change because of MC for a particular instrument. Therefore, for each group (financial assets, business wealth assets, and housing wealth), we picked an instrument and performed a scatter plot showing the rank of each observation on the x-axis and the change in the ranking after applying MC on the y-axis. As examples, we can see the scatterplots relating to 'Deposits', 'Financial Business Wealth' and 'Housing Wealth' in Figure 11 to Figure 13 (plots for other instruments are looking similar). In each of these plots, on the left-hand side, there is a flat line. These are the observations with zero instrument holdings and where the rank therefore does not change. The line is very short for instruments like 'Deposits' where there are only a few observations without any holdings but long for instruments like 'Financial Business Wealth' where instrument holdings are very concentrated. On the right-hand side of the plots, we can also see a rather flat section (approximately observations with a rank above 5000). These are the observations mainly belonging to I_{top}, i.e., the households obtained from rich lists and sampling. Thus, this section has almost the same length in all of the plots. Compared to Figure 11 and Figure 13, we can see some outliers in that part of the plot in Figure 12: Due to the fact that 'Financial Business Wealth' is very concentrated, we can find some observations belonging to the original HFCS data set in this range. In between these plot sections, there are the households belonging to I_{bottom}. For these observations, we clearly see changes in the rankings. The rank change for households obviously coincides with the value of the adjustment coefficient. For some of the households, the rank difference in the 'Deposits' and 'Housing Wealth' rankings is more than 3000. Compared to the absolute number of observation points (circa 6300), this is a high number meaning that the affected households lose almost all of their instrument holdings, e.g., a household that has a rank of 4000 suddenly drops to the rank 1000 after MC. Hence, even though we see that most samples are located within a certain acceptable range (e.g. -1000 to $+1000$ in the case of 'Deposits'), MC produces heavy outliers. One can doubt if these changes reflect the truth.

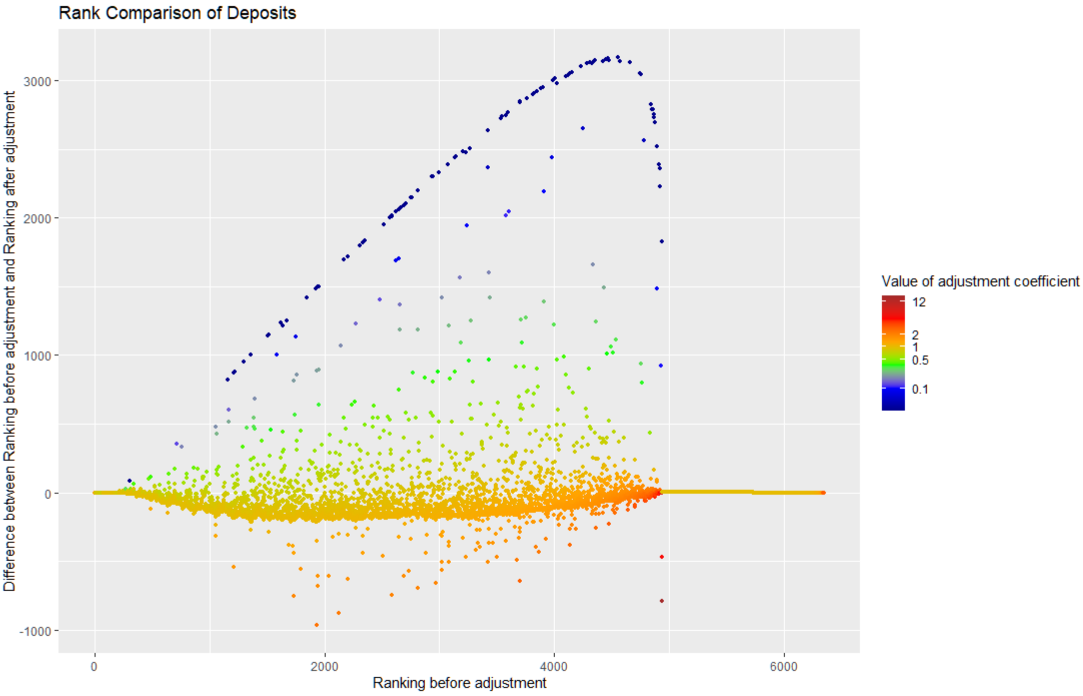

Figure 11: Rank comparison presented by the example of 'Deposits' (financial asset).

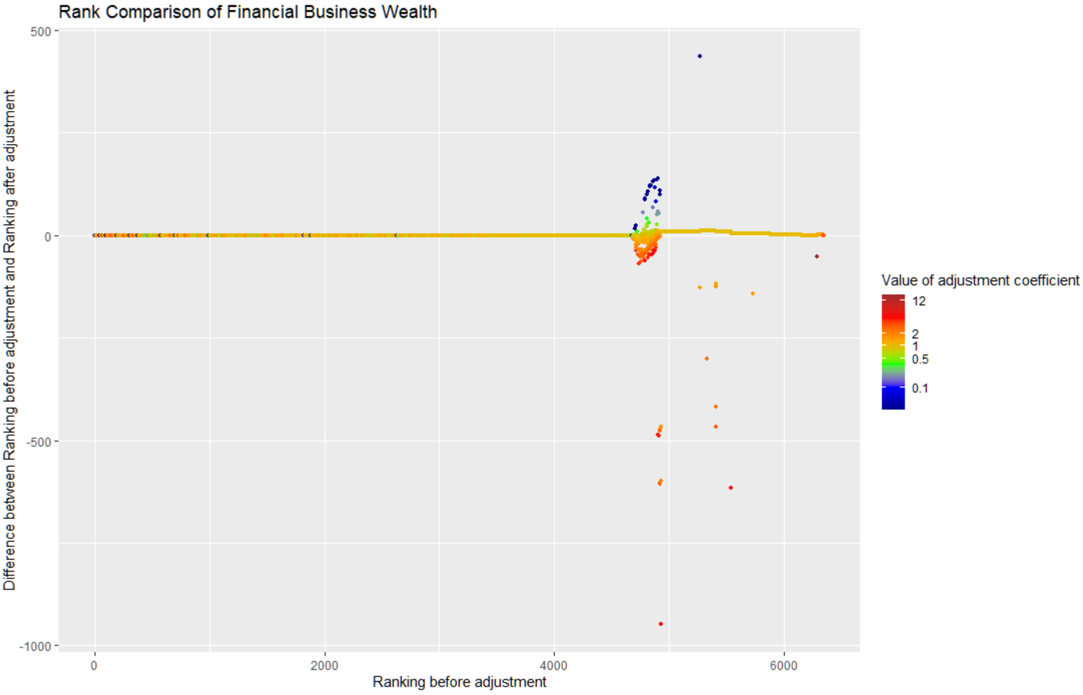

Figure 12: Rank comparison presented by the example of 'Financial Business Wealth' (business wealth asset).

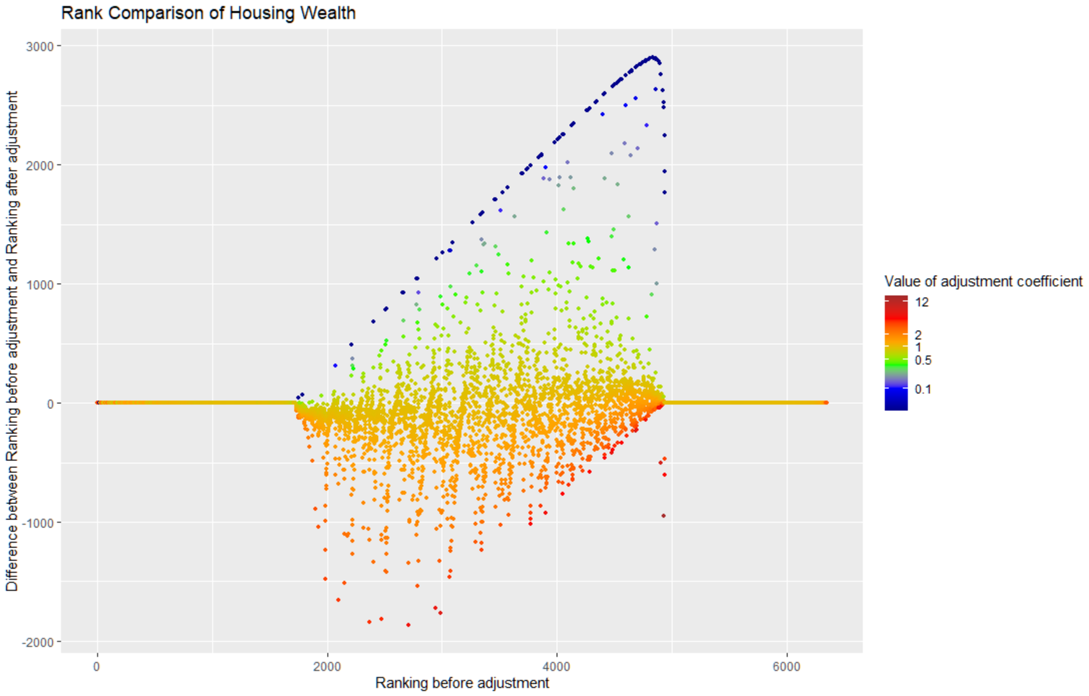

Figure 13: Rank comparison presented by the example of 'Housing Wealth'.

Concluding, we see that the MC leads to a more inhomogeneous allocation compared to the PA. This is a direct consequence of the fact that even in case of undercoverage there are always households with less wealth after adjustment. This can also be seen after calculating the Gini coefficient of net wealth. It is 0.7617 after the PA and 0.8006 after the MC. Furthermore, we have seen that the MC has indeed the tendency to preserve wealth rankings, but Kendall's tau is clearly below one. Thus, for each instrument, there are changes in the wealth order. One major concern using calibration techniques for this particular problem is that it is typically used in survey sampling to handle unit nonresponse (cf. ([14])). But as described in Section 4.1, here it is not used to adjust the weightings of the survey but to change the instrument holdings. Therefore, this approach is questionable. Due to the fact that we have no further information on how to allocate the wealth, preserving wealth inequality (measured by the Gini coefficient) and maintaining the order of instrument holdings seem to be appropriate assumptions that favor proportional adjustment.

5 Conclusion

This thesis provides mathematical insights for deriving sound distributional national accounts. Following the work of Vermeulen ([22], [23]), we have shown how to model the top of the wealth distribution with a Pareto distribution and how to sample additional households to overcome the underrepresentation of very rich households in the HFCS. As a theoretical basis, several wealth accumulation processes are presented that give an explanation of why the upper tail of wealth distribution follows a Pareto law. For example, one of those models is a wealth accumulation process with homogeneous investment talent, the others are an exponential wealth accumulation process with stable population as well as the generalized Lotka-Volterra model and the so-called Yard-Sale model.

Furthermore, it has been proven that $a_1 = \cdots = a_n = \frac{F}{\sum_{i=1}^{n} d_i x_i}$ is the unique solution to the optimization problem[11]

$$\begin{cases} \text{solve} \quad G^* - G(\boldsymbol{a}, \boldsymbol{d}, \boldsymbol{x}) = 0 \\ \text{subject to} \\ \qquad\qquad - a_i \leq -1 \, \forall i, \\ \quad a_i - a_{i+1} \leq 0 \quad \forall i \in \{1, \dots, n-1\}, \\ \quad \sum_{i=1}^{n} d_i a_i x_i = F. \end{cases}$$

This solution comes along with the proportional allocation method where the descrepancies between HFCS micro data and NtlA macro data are eliminated by adjusting household holdings of a certain instrument with the same adjustment coefficient.
The subsequently derived lower and upper bounds provide a clear picture of the leeway of the last step matching the micro with the macro data. While a lower range increases confidence a wide range can raise awareness of possibly wrong instrument allocations.
Additionally, the thesis provides for the first time a comparison of a multivariate calibration approach and the proportional adjustment by the example of German household data.
Altogether, the analysis and findings support academics and policymakers in deriving sound distributional national accounts.

Based on the insights of this thesis, similar investigations in household data of other countries are highly interesting, especially, whether such results would be similar to the findings of the case study covering household wealth in Germany.
Furthermore, with regards to the lower and upper bounds, a tighter range particularly for instruments with high under- or overcoverage is desirable. This raises the question of whether there are any further constraints that make these bounds more realistic. To answer this question, a more detailed study of the lower part of the wealth distribution of the HFCS household date could be useful. Chakrabarti et al. ([4], page 12) gathered wealth and income studies of the past and come to the conclusion that the *"lower part of the distribution follows one of the exponential (Gibbs) or gamma or log-normal (Gibrat)*

[11] Definitions of expressions and variables can be found in Chapter 3.

distributions". So it would be interesting to know whether the lower part of the distribution of the HFCS household wealth can be assigned to one of these distributions. If this is the case, confidence intervals could be helpful to make the lower and upper bound more realistic.

We see that this thesis provided a lot of insights into developing distributional accounts but there are still some other aspects to be examined.

A Appendix

The following Lemmas are used in the proof of Theorem 2. Note that we define $\Delta W_k = W_k - W_{k-1}$ and $H = \sum_{i=1}^{n} d_i x_i$.

Lemma 1. The solution space of the intersection of the hyperplanes $G^* - G(\boldsymbol{a}, \boldsymbol{d}, \boldsymbol{x}) = 0$ and $\sum_{i=1}^{n} d_i a_i x_i = F$ for $a_1, \dots, a_{n-2} \in \mathbb{R}$ is determined by

$$a_{n-1} = \frac{\sum_{k=1}^{n} \Delta W_k F \left(2 \sum_{\ell=1}^{k-1} d_\ell x_\ell + d_k x_k\right) - H \Delta W_n F - \sum_{i=1}^{n-2} a_i \left(H(\Delta W_i - \Delta W_n) d_i x_i + \sum_{k=i+1}^{n} 2H \Delta W_k d_i x_i\right)}{H(\Delta W_{n-1} + \Delta W_n) d_{n-1} x_{n-1}}, \tag{88}$$

and

$$a_n = \frac{FH(\Delta W_{n-1} + \Delta W_n) - L}{H d_n x_n \left(\Delta W_{n-1} + \Delta W_n\right)} + \frac{\sum_{i=1}^{n-2} a_i \left(H(\Delta W_i + \Delta W_{n-1}) d_i x_i + \sum_{k=i+1}^{n-2} 2H \Delta W_k d_i x_i\right)}{H d_n x_n \left(\Delta W_{n-1} + \Delta W_n\right)}. \tag{89}$$

Proof. Note that $G(\boldsymbol{a}, \boldsymbol{d}, \boldsymbol{x})$ can be written as

$$G(\mathbf{a}, \mathbf{d}, \mathbf{x}) = 1 - \left[\sum_{k=1}^{n} (W_k - W_{k-1}) \frac{2 \sum_{\ell=1}^{k-1} a_\ell d_\ell x_\ell + a_k d_k x_k}{F}\right], \tag{90}$$

where we inserted the constraint $\sum_{i=1}^{n} d_i a_i x_i = F$.
Doing some manipulations yields

$$\begin{aligned}
G^* - G(\boldsymbol{a}, \boldsymbol{d}, \boldsymbol{x}) &= G^* - 1 + \left[\sum_{k=1}^{n} (W_k - W_{k-1}) \frac{2 \sum_{\ell=1}^{k-1} a_\ell d_\ell x_\ell + a_k d_k x_k}{F}\right] \\
&= - \left[\sum_{k=1}^{n} \Delta W_k \frac{2 \sum_{\ell=1}^{k-1} d_\ell x_\ell + d_k x_k}{\sum_{i=1}^{n} d_i x_i}\right] + \left[\sum_{k=1}^{n} \Delta W_k \frac{2 \sum_{\ell=1}^{k-1} a_\ell d_\ell x_\ell + a_k d_k x_k}{F}\right] \\
&= - \left[\sum_{k=1}^{n} \Delta W_k \frac{2 \sum_{\ell=1}^{k-1} d_\ell x_\ell + d_k x_k}{H}\right] + \left[\sum_{k=1}^{n} \Delta W_k \frac{2 \sum_{\ell=1}^{k-1} a_\ell d_\ell x_\ell + a_k d_k x_k}{F}\right] \\
&= \frac{1}{HF} \left[- \sum_{k=1}^{n} \Delta W_k F \left(2 \sum_{\ell=1}^{k-1} d_\ell x_\ell + d_k x_k\right) + \sum_{k=1}^{n} \Delta W_k H \left(2 \sum_{\ell=1}^{k-1} a_\ell d_\ell x_\ell + a_k d_k x_k\right)\right] \\
&= \frac{1}{HF} \left[- \sum_{k=1}^{n} \Delta W_k F \left(2 \sum_{\ell=1}^{k-1} d_\ell x_\ell + d_k x_k\right) + \sum_{i=1}^{n} a_i \left(H \Delta W_i d_i x_i + \sum_{k=i+1}^{n} 2H \Delta W_k d_i x_i\right)\right].
\end{aligned}$$

Due to linearity, the solution space of $G^* - G(\boldsymbol{a}, \boldsymbol{d}, \boldsymbol{x}) = 0$ is a hyperplane. The constraint $\sum_{i=1}^{n} d_i a_i x_i = F$ also yields a hyperplane. We now want to calculate the $n-2$ dimensional solution space of the intersection of these two hyperplanes:

$$\begin{aligned}
&\text{(E1)} \quad \sum_{i=1}^{n} a_i \left(H\Delta W_i d_i x_i + \sum_{k=i+1}^{n} 2H\Delta W_k d_i x_i \right) = \sum_{k=1}^{n} \Delta W_k F \left(2 \sum_{\ell=1}^{k-1} d_\ell x_\ell + d_k x_k \right), \\
&\text{(E2)} \quad \sum_{i=1}^{n} d_i a_i x_i = F.
\end{aligned}$$

We calculate (E1') = (E1) $- H\Delta W_n$(E2) and set (E2') = (E2) to get

$$\begin{aligned}
&\text{(E1')} \quad \sum_{i=1}^{n-1} a_i \left(H(\Delta W_i - \Delta W_n) d_i x_i + \sum_{k=i+1}^{n} 2H\Delta W_k d_i x_i \right) = \sum_{k=1}^{n} \Delta W_k F \left(2 \sum_{\ell=1}^{k-1} d_\ell x_\ell + d_k x_k \right) - H\Delta W_n F, \\
&\text{(E2')} \quad \sum_{i=1}^{n} d_i a_i x_i = F.
\end{aligned}$$

From (E1') it follows that

$$\begin{aligned}
a_{n-1} &= \frac{\sum_{k=1}^{n} \Delta W_k F \left(2 \sum_{\ell=1}^{k-1} d_\ell x_\ell + d_k x_k \right) - H\Delta W_n F - \sum_{i=1}^{n-2} a_i \left(H(\Delta W_i - \Delta W_n) d_i x_i + \sum_{k=i+1}^{n} 2H\Delta W_k d_i x_i \right)}{H(\Delta W_{n-1} - \Delta W_n) d_{n-1} x_{n-1} + 2H\Delta W_n d_{n-1} x_{n-1}} \\
&= \frac{\sum_{k=1}^{n} \Delta W_k F \left(2 \sum_{\ell=1}^{k-1} d_\ell x_\ell + d_k x_k \right) - H\Delta W_n F - \sum_{i=1}^{n-2} a_i \left(H(\Delta W_i - \Delta W_n) d_i x_i + \sum_{k=i+1}^{n} 2H\Delta W_k d_i x_i \right)}{H(\Delta W_{n-1} + \Delta W_n) d_{n-1} x_{n-1}}.
\end{aligned} \tag{91}$$

Now we define

$$\begin{aligned}
L &:= \sum_{k=1}^{n} \Delta W_k F \left(2 \sum_{\ell=1}^{k-1} d_\ell x_\ell + d_k x_k \right) - H\Delta W_n F = F \left[\sum_{i=1}^{n} \left(\Delta W_i d_i x_i + \sum_{k=i+1}^{n} 2\Delta W_k d_i x_i \right) - \sum_{i=1}^{n} \Delta W_n d_i x_i \right] \\
&= F \left[\sum_{i=1}^{n} \left((\Delta W_i - \Delta W_n) d_i x_i + \sum_{k=i+1}^{n} 2\Delta W_k d_i x_i \right) \right] = F \left[\sum_{i=1}^{n-1} \left((\Delta W_i - \Delta W_n) d_i x_i + \sum_{k=i+1}^{n} 2\Delta W_k d_i x_i \right) \right] \\
&= F \left[\sum_{i=1}^{n-1} \left((\Delta W_i + \Delta W_n) d_i x_i + \sum_{k=i+1}^{n-1} 2\Delta W_k d_i x_i \right) \right].
\end{aligned} \tag{92}$$

Inserting the expression we got for a_{n-1} in Equation (91) into (E2'), we get

$$
\begin{aligned}
a_n &= \frac{F - \sum_{i=1}^{n-1} a_i d_i x_i}{d_n x_n} \\
&= \frac{F - \sum_{i=1}^{n-2} a_i d_i x_i}{d_n x_n} - \frac{L - \sum_{i=1}^{n-2} a_i \left(H(\Delta W_i - \Delta W_n) d_i x_i + \sum_{k=i+1}^{n} 2H\Delta W_k d_i x_i\right)}{H d_n x_n \left(\Delta W_{n-1} + \Delta W_n\right)} \\
&= \frac{FH(\Delta W_{n-1} + \Delta W_n) - L}{H d_n x_n \left(\Delta W_{n-1} + \Delta W_n\right)} + \frac{\sum_{i=1}^{n-2} a_i \left(H(\Delta W_i - \Delta W_n) d_i x_i + \sum_{k=i+1}^{n} 2H\Delta W_k d_i x_i\right) - \sum_{i=1}^{n-2} a_i d_i x_i H \left(\Delta W_{n-1} + \Delta W_n\right)}{H d_n x_n \left(\Delta W_{n-1} + \Delta W_n\right)} \\
&= \frac{FH(\Delta W_{n-1} + \Delta W_n) - L}{H d_n x_n \left(\Delta W_{n-1} + \Delta W_n\right)} + \frac{\sum_{i=1}^{n-2} a_i \left(H(\Delta W_i - \Delta W_{n-1} - 2\Delta W_n) d_i x_i + \sum_{k=i+1}^{n} 2H\Delta W_k d_i x_i\right)}{H d_n x_n \left(\Delta W_{n-1} + \Delta W_n\right)} \\
&= \frac{FH(\Delta W_{n-1} + \Delta W_n) - L}{H d_n x_n \left(\Delta W_{n-1} + \Delta W_n\right)} + \frac{\sum_{i=1}^{n-2} a_i \Big(H(\Delta W_i - \Delta W_{n-1} - 2\Delta W_n) d_i x_i + \sum_{k=i+1}^{n-2} 2H\Delta W_k d_i x_i + 2H\Delta W_{n-1} d_i x_i + 2H\Delta W_n d_i x_i\Big)}{H d_n x_n \left(\Delta W_{n-1} + \Delta W_n\right)} \\
&= \frac{FH(\Delta W_{n-1} + \Delta W_n) - L}{H d_n x_n \left(\Delta W_{n-1} + \Delta W_n\right)} + \frac{\sum_{i=1}^{n-2} a_i \Big(H(\Delta W_i + \Delta W_{n-1}) d_i x_i + \sum_{k=i+1}^{n-2} 2H\Delta W_k d_i x_i\Big)}{H d_n x_n \left(\Delta W_{n-1} + \Delta W_n\right)}.
\end{aligned}
\tag{93}
$$

□

Lemma 2. Assume that the constraints of (60) hold. Then, $a_1 \leq \frac{F}{H}$.

Proof. Due to $a_1 \leq a_{n-1}$, we have

$$F = \sum_{i=1}^{n} a_i d_i x_i \geq \sum_{i=1}^{n} a_1 d_i x_i = a_1 H.$$

Thus, we get

$$a_1 \leq \frac{F}{H}. \tag{94}$$

□

Lemma 3. Assume that the constraints of (60) hold. Then, $a_k \leq \frac{F}{H}$ for all $k \in \{1, \ldots, n-2\}$.

Proof. In the following, we will prove by induction that $a_k \leq \frac{F}{H}$ for all $k \in \{1, \ldots, n-2\}$. The induction basis $a_1 \leq \frac{F}{H}$ is already given by Lemma 2.
We assume that $a_1, \ldots a_{m-1} \leq \frac{F}{H}$ for fixed $m \leq n-2$. We will now show that $a_m \leq \frac{F}{H}$.

Due to our required constraint, $a_m \leq a_{n-1}$, we have by (91),

$$\begin{aligned} a_m &\leq \frac{\sum_{k=1}^{n} \Delta W_k F \left(2 \sum_{\ell=1}^{k-1} d_\ell x_\ell + d_k x_k\right) - H \Delta W_n F - \sum_{i=1}^{n-2} a_i \left(H\left(\Delta W_i - \Delta W_n\right) d_i x_i + \sum_{k=i+1}^{n} 2 H \Delta W_k d_i x_i\right)}{H\left(\Delta W_{n-1} + \Delta W_n\right) d_{n-1} x_{n-1}} \\ &\leq \frac{L - \sum_{i=1}^{n-2} a_i \left(H\left(\Delta W_i - \Delta W_n\right) d_i x_i + \sum_{k=i+1}^{n-1} 2 H \Delta W_k d_i x_i\right)}{H\left(\Delta W_{n-1} + \Delta W_n\right) d_{n-1} x_{n-1}}, \end{aligned} \tag{95}$$

since $H\left(\Delta W_i - \Delta W_n\right) d_i x_i + \sum_{k=i+1}^{n} 2 H \Delta W_k d_i x_i \geq 0$ for all $i \in \{1, \ldots, n-2\}$.

It follows,

$$\begin{aligned} & a_m \leq \frac{L - \sum_{i=1}^{n-2} a_i \left(H(\Delta W_i - \Delta W_n) d_i x_i + \sum_{k=i+1}^{n} 2 H \Delta W_k d_i x_i\right)}{H(\Delta W_{n-1} + \Delta W_n) d_{n-1} x_{n-1}} \\ & \Leftrightarrow a_m \left(H(\Delta W_{n-1} + \Delta W_n) d_{n-1} x_{n-1}\right) \leq L - \sum_{i=1}^{n-2} a_i \left(H(\Delta W_i - \Delta W_n) d_i x_i + \sum_{k=i+1}^{n} 2 H \Delta W_k d_i x_i\right) \\ & \Leftrightarrow a_{m-1} \left(H(\Delta W_{m-1} - \Delta W_n) d_{m-1} x_{m-1} + \sum_{k=m}^{n} 2 H \Delta W_k d_{m-1} x_{m-1}\right) \\ & \leq L - \sum_{i=1, i \neq m-1}^{n-2} a_i \left(H(\Delta W_i - \Delta W_n) d_i x_i + \sum_{k=i+1}^{n} 2 H \Delta W_k d_i x_i\right) - a_m \left(H(\Delta W_{n-1} + \Delta W_n) d_{n-1} x_{n-1}\right) \\ & \Leftrightarrow a_{m-1} \leq \frac{L - \sum_{i=1, i \neq m-1}^{n-2} a_i \left(H(\Delta W_i + \Delta W_n) d_i x_i + \sum_{k=i+1}^{n-1} 2 H \Delta W_k d_i x_i\right) - a_m \left(H(\Delta W_{n-1} + \Delta W_n) d_{n-1} x_{n-1}\right)}{H(\Delta W_{m-1} + \Delta W_n) d_{m-1} x_{m-1} + \sum_{k=m}^{n-1} 2 H \Delta W_k d_{m-1} x_{m-1}}. \end{aligned} \tag{96}$$

We define $M := FH(\Delta W_{n-1} + \Delta W_n) - L$. One can verify that $d_n x_n L - d_{n-1} x_{n-1} M \geq 0$.
Then, due to $a_{n-1} \leq a_n$ and using the expressions (91) and (93)

$$\begin{aligned}
&\frac{L - \sum_{i=1}^{n-2} a_i \left(H(\Delta W_i + \Delta W_n) d_i x_i + \sum_{k=i+1}^{n-1} 2H\Delta W_k d_i x_i\right)}{H(\Delta W_{n-1} + \Delta W_n) d_{n-1} x_{n-1}} \leq \frac{M + \sum_{i=1}^{n-2} a_i \left(H(\Delta W_i + \Delta W_{n-1}) d_i x_i + \sum_{k=i+1}^{n-2} 2H\Delta W_k d_i x_i\right)}{H d_n x_n \left(\Delta W_{n-1} + \Delta W_n\right)} \\
&\Leftrightarrow d_n x_n \left(L - \sum_{i=1}^{n-2} a_i d_i x_i \left(H(\Delta W_i + \Delta W_n) + \sum_{k=i+1}^{n-1} 2H\Delta W_k\right)\right) \leq d_{n-1} x_{n-1} \left(M + \sum_{i=1}^{n-2} a_i d_i x_i \left(H(\Delta W_i + \Delta W_{n-1}) + \sum_{k=i+1}^{n-2} 2H\Delta W_k\right)\right) \\
&\Leftrightarrow d_n x_n \left(L - \sum_{i=1, i\neq m-1}^{n-2} a_i d_i x_i \left(H(\Delta W_i + \Delta W_n) + \sum_{k=i+1}^{n-1} 2H\Delta W_k\right)\right) - d_{n-1} x_{n-1} \left(M + \sum_{i=1, i\neq m-1}^{n-2} a_i d_i x_i \left(H(\Delta W_i + \Delta W_{n-1}) + \sum_{k=i+1}^{n-2} 2H\Delta W_k\right)\right) \\
&\leq d_{n-1} x_{n-1} a_{m-1} d_{m-1} x_{m-1} \left(H(\Delta W_{m-1} + \Delta W_{n-1}) + \sum_{k=m}^{n-2} 2H\Delta W_k\right) + d_n x_n a_{m-1} d_{m-1} x_{m-1} \left(H(\Delta W_{m-1} + \Delta W_n) + \sum_{k=m}^{n-1} 2H\Delta W_k\right) \\
&\Leftrightarrow \left\{ d_n x_n L - d_{n-1} x_{n-1} M - \sum_{i=1, i\neq m-1}^{n-2} a_i \left[d_{n-1} x_{n-1} d_i x_i \left(H(\Delta W_i + \Delta W_{n-1}) + \sum_{k=i+1}^{n-2} 2H\Delta W_k\right) + d_n x_n d_i x_i \left(H(\Delta W_i + \Delta W_n) + \sum_{k=i+1}^{n-1} 2H\Delta W_k\right)\right]\right\} \\
&\left\{ d_{n-1} x_{n-1} d_{m-1} x_{m-1} \left(H(\Delta W_{m-1} + \Delta W_{n-1}) + \sum_{k=m}^{n-2} 2H\Delta W_k\right) + d_n x_n d_{m-1} x_{m-1} \left(H(\Delta W_{m-1} + \Delta W_n) + \sum_{k=m}^{n-1} 2H\Delta W_k\right)\right\}^{-1} \leq a_{m-1}.
\end{aligned} \tag{97}$$

Define $D := d_{n-1}x_{n-1}d_{m-1}x_{m-1}\left(H(\Delta W_{m-1}+\Delta W_{n-1})+\sum_{k=m}^{n-2}2H\Delta W_k\right)+d_nx_nd_{m-1}x_{m-1}\left(H(\Delta W_{m-1}+\Delta W_n)+\sum_{k=m}^{n-1}2H\Delta W_k\right)$ and $E := H(\Delta W_{m-1}+\Delta W_n)d_{m-1}x_{m-1}+\sum_{k=m}^{n-1}2H\Delta W_k d_{m-1}x_{m-1}$.

Combining the upper and lower bound for a_{m-1} (see (96) and (97)), we get

$$\begin{aligned}
&\Bigg\{d_nx_nL-d_{n-1}x_{n-1}M-\sum_{i=1,i\neq m-1}^{n-2}a_i\Bigg[d_{n-1}x_{n-1}d_ix_i\left(H(\Delta W_i+\Delta W_{n-1})+\sum_{k=i+1}^{n-2}2H\Delta W_k\right)+d_nx_nd_ix_i\left(H(\Delta W_i+\Delta W_n)+\sum_{k=i+1}^{n-1}2H\Delta W_k\right)\Bigg]\Bigg\}\\
&\Bigg\{d_{n-1}x_{n-1}d_{m-1}x_{m-1}\left(H(\Delta W_{m-1}+\Delta W_{n-1})+\sum_{k=m}^{n-2}2H\Delta W_k\right)+d_nx_nd_{m-1}x_{m-1}\left(H(\Delta W_{m-1}+\Delta W_n)+\sum_{k=m}^{n-1}2H\Delta W_k\right)\Bigg\}^{-1}\leq a_{m-1}\\
&\leq\frac{L-\sum_{i=1,i\neq m-1}^{n-2}a_i\left(H(\Delta W_i+\Delta W_n)d_ix_i+\sum_{k=i+1}^{n-1}2H\Delta W_kd_ix_i\right)-a_m\left(H(\Delta W_{n-1}+\Delta W_n)d_{n-1}x_{n-1}\right)}{H(\Delta W_{m-1}+\Delta W_n)d_{m-1}x_{m-1}+\sum_{k=m}^{n-1}2H\Delta W_kd_{m-1}x_{m-1}}\\
&\Leftrightarrow E\Bigg\{d_nx_nL-d_{n-1}x_{n-1}M-\sum_{i=1,i\neq m-1}^{n-2}a_i\Bigg[d_{n-1}x_{n-1}d_ix_i\left(H(\Delta W_i+\Delta W_{n-1})+\sum_{k=i+1}^{n-2}2H\Delta W_k\right)+d_nx_nd_ix_i\left(H(\Delta W_i+\Delta W_n)+\sum_{k=i+1}^{n-1}2H\Delta W_k\right)\Bigg]\Bigg\}\\
&\leq D\left(L-\sum_{i=1,i\neq m-1}^{n-2}a_i\left(H(\Delta W_i+\Delta W_n)d_ix_i+\sum_{k=i+1}^{n-1}2H\Delta W_kd_ix_i\right)-a_m\left(H(\Delta W_{n-1}+\Delta W_n)d_{n-1}x_{n-1}\right)\right)\\
&\Leftrightarrow a_mD\left(H(\Delta W_m+\Delta W_n)d_mx_m+\sum_{k=m+1}^{n-1}2H\Delta W_kd_mx_m\right)+a_mD\left(H(\Delta W_{n-1}+\Delta W_n)d_{n-1}x_{n-1}\right)\\
&-a_mE\Bigg[d_{n-1}x_{n-1}d_mx_m\left(H(\Delta W_m+\Delta W_{n-1})+\sum_{k=m+1}^{n-2}2H\Delta W_k\right)+d_nx_nd_mx_m\left(H(\Delta W_m+\Delta W_n)+\sum_{k=m+1}^{n-1}2H\Delta W_k\right)\Bigg]\\
&\leq DL-E(d_nx_nL-d_{n-1}x_{n-1}M)-D\left(\sum_{i=1,i\neq m-1,m}^{n-2}a_i\left(H(\Delta W_i+\Delta W_n)d_ix_i+\sum_{k=i+1}^{n-1}2H\Delta W_kd_ix_i\right)\right)\\
&+E\left(\sum_{i=1,i\neq m-1,m}^{n-2}a_i\Bigg[d_{n-1}x_{n-1}d_ix_i\left(H(\Delta W_i+\Delta W_{n-1})+\sum_{k=i+1}^{n-2}2H\Delta W_k\right)+d_nx_nd_ix_i\left(H(\Delta W_i+\Delta W_n)+\sum_{k=i+1}^{n-1}2H\Delta W_k\right)\Bigg]\right)
\end{aligned}\tag{98}$$

Continuing with (98), brings us to the following inequality.

$$\begin{aligned}
&a_m\Bigg\{(D - d_n x_n E)\left(H(\Delta W_m + \Delta W_n)d_m x_m + \sum_{k=m+1}^{n-1} 2H\Delta W_k d_m x_m\right) + D\left(H(\Delta W_{n-1} + \Delta W_n)d_{n-1}x_{n-1}\right)\\
&- E\left(d_{n-1}x_{n-1}\left(H(\Delta W_m + \Delta W_{n-1})d_m x_m + \sum_{k=m+1}^{n-2} 2H\Delta W_k d_m x_m\right)\right)\Bigg\}\\
&\leq DL - E(d_n x_n L - d_{n-1}x_{n-1}M) + \Bigg[\sum_{i=1,i\neq m-1,m}^{n-2} a_i\Bigg((d_n x_n E - D)\left(H(\Delta W_i + \Delta W_n)d_i x_i + \sum_{k=i+1}^{n-1} 2H\Delta W_k d_i x_i\right)\\
&+ d_{n-1}x_{n-1}E\left(H(\Delta W_i + \Delta W_{n-1})d_i x_i + \sum_{k=i+1}^{n-2} 2H\Delta W_k d_i x_i\right)\Bigg)\Bigg].
\end{aligned} \tag{99}$$

Shortly, we will also have to distinguish between the following two cases:
Case 1: $i \leq m-2$:

$$\left(\Delta W_i + \sum_{k=i+1}^{n-2} 2\Delta W_k\right) - \left(\Delta W_{m-1} + \sum_{k=m}^{n-2} 2\Delta W_k\right) = \Delta W_i + \sum_{k=i+1}^{m-1} 2\Delta W_k - \Delta W_{m-1} = \Delta W_i + \sum_{k=i+1}^{m-2} 2\Delta W_k + \Delta W_{m-1} \geq 0. \tag{100}$$

Case 2: $i \geq m+1$:

$$\left(\Delta W_i + \sum_{k=i+1}^{n-2} 2\Delta W_k\right) - \left(\Delta W_{m-1} + \sum_{k=m}^{n-2} 2\Delta W_k\right) = \Delta W_i - \sum_{k=m}^{i} 2\Delta W_k - \Delta W_{m-1} = -\Delta W_{m-1} - \sum_{k=m}^{i-1} 2\Delta W_k - \Delta W_i \leq 0. \tag{101}$$

Let us have a closer look at the last part of (99),

namely$\Bigg((d_nx_nE-D)\left(H(\Delta W_i+\Delta W_n)d_ix_i+\sum_{k=i+1}^{n-1}2H\Delta W_kd_ix_i\right)+d_{n-1}x_{n-1}E\left(H(\Delta W_i+\Delta W_{n-1})d_ix_i+\sum_{k=i+1}^{n-2}2H\Delta W_kd_ix_i\right)\Bigg)$:

$$
\begin{aligned}
&d_{n-1}x_{n-1}E\left(H(\Delta W_i+\Delta W_{n-1})d_ix_i+\sum_{k=i+1}^{n-2}2H\Delta W_kd_ix_i\right)+(d_nx_nE-D)\left(H(\Delta W_i+\Delta W_n)d_ix_i+\sum_{k=i+1}^{n-1}2H\Delta W_kd_ix_i\right)\\
&=d_{n-1}x_{n-1}\left(H(\Delta W_{m-1}+\Delta W_n)d_{m-1}x_{m-1}+\sum_{k=m}^{n-1}2H\Delta W_kd_{m-1}x_{m-1}\right)\left(H(\Delta W_i+\Delta W_{n-1})d_ix_i+\sum_{k=i+1}^{n-2}2H\Delta W_kd_ix_i\right)\\
&-d_{n-1}x_{n-1}\left(H(\Delta W_{m-1}+\Delta W_{n-1})d_{m-1}x_{m-1}+\sum_{k=m}^{n-2}2H\Delta W_kd_{m-1}x_{m-1}\right)\left(H(\Delta W_i+\Delta W_n)d_ix_i+\sum_{k=i+1}^{n-1}2H\Delta W_kd_ix_i\right)\\
&=d_{n-1}x_{n-1}H^2d_{m-1}x_{m-1}d_ix_i\Bigg[\left((\Delta W_{m-1}+\Delta W_n)+\sum_{k=m}^{n-1}2\Delta W_k\right)\left((\Delta W_i+\Delta W_{n-1})+\sum_{k=i+1}^{n-2}2\Delta W_k\right)\\
&-\left((\Delta W_{m-1}+\Delta W_{n-1})+\sum_{k=m}^{n-2}2\Delta W_k\right)\left((\Delta W_i+\Delta W_n)+\sum_{k=i+1}^{n-1}2\Delta W_k\right)\Bigg]\\
&=d_{n-1}x_{n-1}H^2d_{m-1}x_{m-1}d_ix_i\Bigg[\left((\Delta W_{m-1}+2\Delta W_{n-1}+\Delta W_n)+\sum_{k=m}^{n-2}2\Delta W_k\right)\left((\Delta W_i+\Delta W_{n-1})+\sum_{k=i+1}^{n-2}2\Delta W_k\right)\\
&-\left((\Delta W_{m-1}+\Delta W_{n-1})+\sum_{k=m}^{n-2}2\Delta W_k\right)\left((\Delta W_i+2\Delta W_{n-1}+\Delta W_n)+\sum_{k=i+1}^{n-2}2\Delta W_k\right)\Bigg]\\
&=d_{n-1}x_{n-1}H^2d_{m-1}x_{m-1}d_ix_i\Bigg[(\Delta W_{n-1}+\Delta W_n)\left((\Delta W_i+\Delta W_{n-1})+\sum_{k=i+1}^{n-2}2\Delta W_k\right)-\left((\Delta W_{m-1}+\Delta W_{n-1})+\sum_{k=m}^{n-2}2\Delta W_k\right)(\Delta W_{n-1}+\Delta W_n)\Bigg]\\
&=d_{n-1}x_{n-1}H^2d_{m-1}x_{m-1}d_ix_i\,(\Delta W_{n-1}+\Delta W_n)\Bigg[\left(\Delta W_i+\sum_{k=i+1}^{n-2}2\Delta W_k\right)-\left(\Delta W_{m-1}+\sum_{k=m}^{n-2}2\Delta W_k\right)\Bigg].
\end{aligned}
\tag{102}
$$

Let us come back to inequality (99). We insert (102) and be aware of (100) and (101) to get:

$$\begin{aligned}
&a_m\Bigg\{(D-d_nx_nE)\Bigg(H(\Delta W_m+\Delta W_n)d_mx_m+\sum_{k=m+1}^{n-1}2H\Delta W_kd_mx_m\Bigg)+D\,(H(\Delta W_{n-1}+\Delta W_n)d_{n-1}x_{n-1})\\
&-E\Bigg(d_{n-1}x_{n-1}\Bigg(H(\Delta W_m+\Delta W_{n-1})d_mx_m+\sum_{k=m+1}^{n-2}2H\Delta W_kd_mx_m\Bigg)\Bigg)\Bigg\}\\
&\le DL-E(d_nx_nL-d_{n-1}x_{n-1}M)+\Bigg[\sum_{i=1,i\neq m-1,m}^{n-2}a_i\Bigg((d_nx_nE-D)\Bigg(H(\Delta W_i+\Delta W_n)d_ix_i+\sum_{k=i+1}^{n-1}2H\Delta W_kd_ix_i\Bigg)+\\
&d_{n-1}x_{n-1}E\Bigg(H(\Delta W_i+\Delta W_{n-1})d_ix_i+\sum_{k=i+1}^{n-2}2H\Delta W_kd_ix_i\Bigg)\Bigg)\Bigg]\\
&\Leftrightarrow a_m\Bigg\{D\,(H(\Delta W_{n-1}+\Delta W_n)d_{n-1}x_{n-1})-d_{n-1}x_{n-1}H^2d_{m-1}x_{m-1}d_mx_m\,(\Delta W_{n-1}+\Delta W_n)\Bigg[\Bigg(\Delta W_m+\sum_{k=m+1}^{n-2}2\Delta W_k\Bigg)-\Bigg(\Delta W_{m-1}+\sum_{k=m}^{n-2}2\Delta W_k\Bigg)\Bigg]\Bigg\}\\
&\le DL-E(d_nx_nL-d_{n-1}x_{n-1}M)+\Bigg\{\sum_{i=1,i\neq m-1,m}^{n-2}a_id_{n-1}x_{n-1}H^2d_{m-1}x_{m-1}d_ix_i\,(\Delta W_{n-1}+\Delta W_n)\Bigg[\Bigg(\Delta W_i+\sum_{k=i+1}^{n-2}2\Delta W_k\Bigg)-\Bigg(\Delta W_{m-1}+\sum_{k=m}^{n-2}2\Delta W_k\Bigg)\Bigg]\Bigg\}\\
&\le DL-E(d_nx_nL-d_{n-1}x_{n-1}M)+\Bigg\{\sum_{i=1}^{m-2}a_id_{n-1}x_{n-1}H^2d_{m-1}x_{m-1}d_ix_i\,(\Delta W_{n-1}+\Delta W_n)\Bigg(\Delta W_i+\sum_{k=i+1}^{m-2}2\Delta W_k+\Delta W_{m-1}\Bigg)\Bigg\}\\
&+\Bigg\{\sum_{i=m+1}^{n-2}a_md_{n-1}x_{n-1}H^2d_{m-1}x_{m-1}d_ix_i\,(\Delta W_{n-1}+\Delta W_n)\Bigg(-\Delta W_{m-1}-\sum_{k=m}^{i-1}2\Delta W_k-\Delta W_i\Bigg)\Bigg\}.
\end{aligned}\tag{103}$$

By assumption, $a_i \leq F/H$ for all $i \leq m-1$. Thus, inequality (103) leads to

$$\begin{aligned} &a_m \left\{ D\left(H(\Delta W_{n-1}+\Delta W_n)d_{n-1}x_{n-1}\right) + d_{n-1}x_{n-1}H^2 d_{m-1}x_{m-1}\left(\Delta W_{n-1}+\Delta W_n\right)\left\{\sum_{i=m}^{n-2} d_i x_i \left(\Delta W_{m-1} + \sum_{k=m}^{i-1} 2\Delta W_k + \Delta W_i\right)\right\}\right\} \\ &\leq DL - E(d_n x_n L - d_{n-1}x_{n-1}M) + \left\{\sum_{i=1}^{m-2} F d_{n-1}x_{n-1} H d_{m-1}x_{m-1} d_i x_i \left(\Delta W_{n-1}+\Delta W_n\right)\left(\Delta W_i + \sum_{k=i+1}^{m-2} 2\Delta W_k + \Delta W_{m-1}\right)\right\}. \end{aligned} \tag{104}$$

We observe,

$$\begin{aligned} DL - E(d_n x_n L - d_{n-1}x_{n-1}M) &= DL - E\big((d_n x_n + d_{n-1}x_{n-1})\, L - d_{n-1}x_{n-1}FH(\Delta W_{n-1}+\Delta W_n)\big) \\ &= L(D - (d_n x_n + d_{n-1}x_{n-1})E) + EFHd_{n-1}x_{n-1}(\Delta W_{n-1}+\Delta W_n) \\ &= L\left(d_{n-1}x_{n-1}\left(H(\Delta W_{m-1}+\Delta W_{n-1})d_{m-1}x_{m-1} + \sum_{k=m}^{n-2} 2H\Delta W_k d_{m-1}x_{m-1} - E\right)\right) + d_{n-1}x_{n-1}EFH(\Delta W_{n-1}+\Delta W_n) \\ &= L\left(d_{n-1}x_{n-1}\left(H(\Delta W_{n-1}-\Delta W_n)d_{m-1}x_{m-1} - 2H\Delta W_{n-1}d_{m-1}x_{m-1}\right)\right) + d_{n-1}x_{n-1}EFH(\Delta W_{n-1}+\Delta W_n) \\ &= FHd_{n-1}x_{n-1}(\Delta W_{n-1}+\Delta W_n)\left(E - d_{m-1}x_{m-1}\frac{L}{F}\right). \end{aligned} \tag{105}$$

Inserting (105) into (104) yields

$$
\begin{aligned}
a_m &\leq \frac{DL - E(d_n x_n L - d_{n-1}x_{n-1}M) + \left\{\sum_{i=1}^{m-2} F d_{n-1}x_{n-1}Hd_{m-1}x_{m-1}d_i x_i \left(\Delta W_{n-1} + \Delta W_n\right)\left(\Delta W_i + \sum_{k=i+1}^{m-2} 2\Delta W_k + \Delta W_{m-1}\right)\right\}}{D\left(H(\Delta W_{n-1} + \Delta W_n)d_{n-1}x_{n-1}\right) + d_{n-1}x_{n-1}H^2 d_{m-1}x_{m-1}\left(\Delta W_{n-1} + \Delta W_n\right)\left\{\sum_{i=m}^{n-2} d_i x_i \left(\Delta W_{m-1} + \sum_{k=m}^{i-1} 2\Delta W_k + \Delta W_i\right)\right\}} \\
&= \frac{FHd_{n-1}x_{n-1}(\Delta W_{n-1} + \Delta W_n)\left(E - d_{m-1}x_{m-1}\frac{L}{F}\right) + \left\{\sum_{i=1}^{m-2} F d_{n-1}x_{n-1}Hd_{m-1}x_{m-1}d_i x_i \left(\Delta W_{n-1} + \Delta W_n\right)\left(\Delta W_i + \sum_{k=i+1}^{m-2} 2\Delta W_k + \Delta W_{m-1}\right)\right\}}{D\left(H(\Delta W_{n-1} + \Delta W_n)d_{n-1}x_{n-1}\right) + d_{n-1}x_{n-1}H^2 d_{m-1}x_{m-1}\left(\Delta W_{n-1} + \Delta W_n\right)\left\{\sum_{i=m}^{n-2} d_i x_i \left(\Delta W_{m-1} + \sum_{k=m}^{i-1} 2\Delta W_k + \Delta W_i\right)\right\}} \\
&= \frac{FHd_{n-1}x_{n-1}(\Delta W_{n-1} + \Delta W_n)\left[\left(E - d_{m-1}x_{m-1}\frac{L}{F}\right) + \left\{\sum_{i=1}^{m-2} d_i x_i\left(\Delta W_i + \sum_{k=i+1}^{m-2} 2\Delta W_k + \Delta W_{m-1}\right)\right\}\right]}{\left(H(\Delta W_{n-1} + \Delta W_n)d_{n-1}x_{n-1}\right)\left[D + Hd_{m-1}x_{m-1}\left\{\sum_{i=m}^{n-2} d_i x_i \left(\Delta W_{m-1} + \sum_{k=m}^{i-1} 2\Delta W_k + \Delta W_i\right)\right\}\right]} \\
&= \frac{F\left[\left(E - d_{m-1}x_{m-1}\frac{L}{F}\right) + \left\{\sum_{i=1}^{m-2} d_i x_i d_{m-1}x_{m-1}\left(\Delta W_i + \sum_{k=i+1}^{m-2} 2\Delta W_k + \Delta W_{m-1}\right)\right\}\right]}{\left[D + Hd_{m-1}x_{m-1}\left\{\sum_{i=m}^{n-2} d_i x_i \left(\Delta W_{m-1} + \sum_{k=m}^{i-1} 2\Delta W_k + \Delta W_i\right)\right\}\right]} \\
&= \frac{F\left[\sum_{i=1}^{n} d_i x_i \left(\Delta W_{m-1} + \Delta W_n + \sum_{k=m}^{n-1} 2\Delta W_k\right) - \sum_{i=1}^{n-1} d_i x_i \left(\Delta W_i + \Delta W_n + \sum_{k=i+1}^{n-1} 2\Delta W_k\right) + \left\{\sum_{i=1}^{m-2} d_i x_i\left(\Delta W_i + \sum_{k=i+1}^{m-2} 2\Delta W_k + \Delta W_{m-1}\right)\right\}\right]}{H\left[d_{n-1}x_{n-1}\left(\Delta W_{m-1} + \Delta W_{n-1} + \sum_{k=m}^{n-2} 2H\Delta W_k\right) + d_n x_n \left(\Delta W_{m-1} + \Delta W_n + \sum_{k=m}^{n-1} 2\Delta W_k\right) + \left\{\sum_{i=m}^{n-2} d_i x_i \left(\Delta W_{m-1} + \sum_{k=m}^{i-1} 2\Delta W_k + \Delta W_i\right)\right\}\right]}.
\end{aligned}
\tag{106}
$$

Continuing with (106),

$$a_m \le \frac{F\left[\left(\sum_{i=1}^{n} d_i x_i \left(\Delta W_{m-1}+\Delta W_n+\sum_{k=m}^{n-1} 2\Delta W_k\right) - \sum_{i=1}^{n-1} d_i x_i \left(\left(\Delta W_i+\Delta W_n\right)+\sum_{k=i+1}^{n-1} 2\Delta W_k\right)\right)+\sum_{i=1}^{m-2} d_i x_i\left(\Delta W_i+\sum_{k=i+1}^{m-2} 2\Delta W_k+\Delta W_{m-1}\right)\right]}{H\left[\sum_{i=m}^{n} d_i x_i\left(\Delta W_{m-1}+\sum_{k=m}^{i-1} 2\Delta W_k+\Delta W_i\right)\right]}. \tag{107}$$

We will now show that the last quotient equals $\frac{F}{H}$. Note that the following holds,

$$\begin{aligned}
&\left[\left(\sum_{i=1}^{n} d_i x_i \left(\Delta W_{m-1}+\Delta W_n+\sum_{k=m}^{n-1} 2\Delta W_k\right) - \sum_{i=1}^{n-1} d_i x_i \left(\left(\Delta W_i+\Delta W_n\right)+\sum_{k=i+1}^{n-1} 2\Delta W_k\right)\right)+\left\{\sum_{i=1}^{m-2} d_i x_i\left(\Delta W_i+\sum_{k=i+1}^{m-2} 2\Delta W_k+\Delta W_{m-1}\right)\right\}\right]\\
&=\left[d_n x_n\left(\Delta W_{m-1}+\Delta W_n+\sum_{k=m}^{n-1} 2\Delta W_k\right)+\left(\sum_{i=1}^{n-1} d_i x_i\left(\Delta W_{m-1}+\Delta W_n+\sum_{k=m}^{n-1} 2\Delta W_k-\Delta W_i-\Delta W_n-\sum_{k=i+1}^{n-1} 2\Delta W_k\right)\right)\right.\\
&\left.+\left\{\sum_{i=1}^{m-2} d_i x_i\left(\Delta W_i+\sum_{k=i+1}^{m-2} 2\Delta W_k+\Delta W_{m-1}\right)\right\}\right]\\
&=\left[d_n x_n\left(\Delta W_{m-1}+\Delta W_n+\sum_{k=m}^{n-1} 2\Delta W_k\right)+\left(\sum_{i=1}^{n-1} d_i x_i\left(\Delta W_{m-1}+\sum_{k=m}^{n-1} 2\Delta W_k-\Delta W_i-\sum_{k=i+1}^{n-1} 2\Delta W_k\right)\right)+\left\{\sum_{i=1}^{m-2} d_i x_i\left(\Delta W_i+\sum_{k=i+1}^{m-2} 2\Delta W_k+\Delta W_{m-1}\right)\right\}\right]\\
&=d_n x_n\left(\Delta W_{m-1}+\Delta W_n+\sum_{k=m}^{n-1} 2\Delta W_k\right)+\left(\sum_{i=1}^{m-2} d_i x_i\left(\Delta W_{m-1}-\Delta W_i-\sum_{k=i+1}^{m-1} 2\Delta W_k\right)\right)+\\
&+\left(\sum_{i=m}^{n-1} d_i x_i\left(\Delta W_{m-1}-\Delta W_i+\sum_{k=m}^{i} 2\Delta W_k\right)\right)+\left\{\sum_{i=1}^{m-2} d_i x_i\left(\Delta W_i+\sum_{k=i+1}^{m-2} 2\Delta W_k+\Delta W_{m-1}\right)\right\}\\
&=\left(\sum_{i=m}^{n} d_i x_i\left(\Delta W_{m-1}-\Delta W_i+\sum_{k=m}^{i} 2\Delta W_k\right)\right)\\
&=\left(\sum_{i=m}^{n} d_i x_i\left(\Delta W_{m-1}+\Delta W_i+\sum_{k=m}^{i-1} 2\Delta W_k\right)\right).
\end{aligned} \tag{108}$$

Thus, we have shown that

$$a_m \leq \frac{F}{H}. \tag{109}$$

By induction, $a_k \leq \frac{F}{H}$ for all $k \in \{1, \ldots, n-2\}$. □

Lemma 4. Assume that the constraints of (60) hold. Then, $a_k = \frac{F}{H}$ for all $k \in \{1, \ldots, n-2\}$.

Proof. We have already shown in Lemma 3 that $a_k \leq \frac{F}{H}$ for all $k \in \{1, \ldots, n-2\}$. Now, we show that $a_k \geq F/H$ for all $k \in \{1, \ldots, n-2\}$.

Due to $a_{n-1} \leq a_n$ and using the expressions (91) and (93), we get

$$\begin{aligned}
&\frac{L - \sum_{i=1}^{n-2} a_i \left(H(\Delta W_i + \Delta W_n) d_i x_i + \sum_{k=i+1}^{n-1} 2H\Delta W_k d_i x_i \right)}{H(\Delta W_{n-1} + \Delta W_n) d_{n-1} x_{n-1}} \leq \frac{M + \sum_{i=1}^{n-2} a_i \left(H(\Delta W_i + \Delta W_{n-1}) d_i x_i + \sum_{k=i+1}^{n-2} 2H\Delta W_k d_i x_i \right)}{H d_n x_n \left(\Delta W_{n-1} + \Delta W_n \right)} \\
&\Leftrightarrow d_n x_n L - d_{n-1} x_{n-1} M \leq d_{n-1} x_{n-1} \sum_{i=1}^{n-2} a_i d_i x_i \left(H(\Delta W_i + \Delta W_{n-1}) + \sum_{k=i+1}^{n-2} 2H\Delta W_k \right) + d_n x_n \sum_{i=1}^{n-2} a_i d_i x_i \left(H(\Delta W_i + \Delta W_n) + \sum_{k=i+1}^{n-1} 2H\Delta W_k \right) \\
&\Leftrightarrow d_n x_n L - d_{n-1} x_{n-1} M \leq a_{n-2} \left(d_{n-1} x_{n-1} \sum_{i=1}^{n-2} d_i x_i \left(H(\Delta W_i + \Delta W_{n-1}) + \sum_{k=i+1}^{n-2} 2H\Delta W_k \right) + d_n x_n \sum_{i=1}^{n-2} d_i x_i \left(H(\Delta W_i + \Delta W_n) + \sum_{k=i+1}^{n-1} 2H\Delta W_k \right) \right).
\end{aligned} \tag{110}$$

Observe that,

$$\begin{aligned}
d_nx_nL - d_{n-1}x_{n-1}M &= L(d_nx_n + d_{n-1}x_{n-1}) - d_{n-1}x_{n-1}FH(\Delta W_{n-1} + \Delta W_n) \\
&= F\left((d_nx_n + d_{n-1}x_{n-1})\sum_{i=1}^{n-1}\left((\Delta W_i + \Delta W_n)d_ix_i + \sum_{k=i+1}^{n-1} 2\Delta W_k d_ix_i\right) - d_{n-1}x_{n-1}\sum_{i=1}^{n} x_id_i(\Delta W_{n-1} + \Delta W_n)\right) \\
&= F\Bigg[d_{n-1}x_{n-1}\left(d_{n-1}x_{n-1}(\Delta W_{n-1} + \Delta W_n) + \sum_{i=1}^{n-2} x_id_i\left(\Delta W_i + 2\Delta W_{n-1} + \Delta W_n + \sum_{k=i+1}^{n-2} 2\Delta W_k\right)\right) \\
&+ d_nx_n\left(d_{n-1}x_{n-1}(\Delta W_{n-1} + \Delta W_n) + \sum_{i=1}^{n-2} x_id_i\left(\Delta W_i + \Delta W_n + \sum_{k=i+1}^{n-1} 2\Delta W_k\right)\right) \\
&- d_{n-1}x_{n-1}\sum_{i=1}^{n} x_id_i(\Delta W_{n-1} + \Delta W_n)\Bigg] \\
&= F\Bigg[d_{n-1}x_{n-1}\sum_{i=1}^{n-2} x_id_i\left(\Delta W_i + \Delta W_{n-1} + \sum_{k=i+1}^{n-2} 2\Delta W_k\right) + d_nx_n\sum_{i=1}^{n-2} x_id_i\left(\Delta W_i + \Delta W_n + \sum_{k=i+1}^{n-1} 2\Delta W_k\right)\Bigg].
\end{aligned} \tag{111}$$

Inserting (111) into (110) yields

$$a_{n-2} \geq \frac{F}{H}.$$

Together with Lemma 3, it follows that

$$a_{n-2} = \frac{F}{H}.$$

In a smiliar way, one can show that $a_{n-k} \geq \frac{F}{H}$ if $a_{n-\ell} = \frac{F}{H}$ for all $2 \leq \ell \leq k-1$.
Again, in combination with Lemma 3, this gives $a_k = \frac{F}{H}$ for all $k \in \{1, \ldots, n-2\}$. □

List of Figures

References

[1] B. M. Boghosian. Fokker-Planck description of wealth dynamics and the origin of Pareto's law. *International Journal of Modern Physics C*, Vol. 25(No. 11), 2014.

[2] B. M. Boghosian. Kinetics of wealth and the Pareto law. *Physical review. E, Statistical, nonlinear, and soft matter physics*, Vol. 89(Iss. 4), 2014.

[3] M. Cantarella, A. Neri, and M. G. Ranalli. Mind the wealth gap: a new allocation method to match micro and macro statistics for household wealth. *Working Paper Series*, 2021.

[4] B. K. Chakrabarti, A. Chakraborti, S. R. Chakravarty, and A. Chatterjee. *Econophysics of income and wealth distributions*. Cambridge University Press, 2013.

[5] R. Chakraborty and S. R. Waltl. Missing the wealthy in the HFCS: micro problems with macro implications. *Working Paper Series*, (No 2163), June 2018.

[6] R. N. Costa and S. Pérez-Duarte. Not all inequality measures were created equal. *Statistics Paper Series*, No 31, December 2019.

[7] Household Finance and Consumption Network. The Household Finance and Consumption Survey: Results from the 2017 wave. *Statistics Paper Series*, (No 36), March 2020.

[8] B. Hayes. Follow the money. *American Scientist*, Vol. 90(No. 5):pp. 400–405, 2002.

[9] Ch. I. Jones. Simple models of Pareto income and wealth inequality, 2014. https://web.stanford.edu/ chadj/SimpleParetoJEP.pdf.

[10] Ch. I. Jones. Pareto and Piketty: The macroeconomics of top income and wealth inequality. *Journal of Economic Perspectives*, Vol. 29(1):pp. 29–46, 2015.

[11] M. G. Kendall. A new measure of rank correlation. *Biometrika*, Vol. 30(No. 1/2):pp. 81–93, June 1938.

[12] O. S. Klass, O. Biham, M. Levy, O. Malcai, and S. Solomon. The Forbes 400 and the Pareto wealth distribution. *Economics Letters*, 90(2):pp. 290–295, 2006.

[13] O. S. Klass, O. Biham, M. Levy, O. Malcai, and S. Solomon. The Forbes 400, the Pareto power-law and efficient markets. *The European Physical Journal B*, Vol. 55(2):pp. 143–147, 2007.

[14] P. Kott and T. Chang. Using calibration weighting to adjust for nonignorable unit nonresponse. *Journal of the American Statistical Association*, Vol. 105(491):pp. 1265–1275, September 2010.

[15] M. Levy and H. Levy. Investment talent and the Pareto wealth distribution: theoretical and experimental analysis. *The Review of Economics and Statistics*, Vol. 85(No. 3):pp. 709–725, 2003.

[16] A. J. Lotka. Relation between birth rates and death rates. *Science*, Vol. 26(653):pp. 21–22, 1907.

[17] R. B. Nelsen. *An introduction to copulas*. Springer, 2nd edition edition, 2006.

[18] T. Ogwang. Power laws in top wealth distributions: evidence from Canada. *Empirical Economics*, Vol.41(2):pp. 473–486, 2011.

[19] Expert Group on Linking macro and micro data for the household sector. Understanding household wealth: linking macro and micro data to produce distributional financial accounts. *Working Paper Series*, (NO 37), July 2020.

[20] S. Solomon and P. Richmond. Power laws of wealth, market order volumes and market returns. *Physica A*, Vol. 299(1):pp. 188–197, 2001.

[21] S. Solomon and P. Richmond. Stable power laws in variable economies; Lotka-Volterra implies Pareto-Zipf. *The European Physical Journal B - Condensed Matter and Complex Systems*, Vol. 27(2):pp. 257–261, 2002.

[22] Philip Vermeulen. How fat is the top tail of the wealth distribution? *Working Paper Series*, (NO 1692), July 2014.

[23] Philip Vermeulen. Online appendix for: How fat is the top tail of the wealth distribution?, January 2016.

[24] H. O. A. Wold and P. Whittle. A model explaining the Pareto distribution of wealth. *Econometrica*, Vol.25(4):pp. 591–595, 1957.