Inequality, Risk, and Wealth

Inaugural-Dissertation zur Erlangung des akademischen Grades
eines Doktors der Wirtschaftswissenschaft

Fachbereich Wirtschaftswissenschaft
Freie Universität Berlin

vorgelegt von

Maximilian Longmuir, M.Sc.
geb. Wenzel,

Berlin, November 22, 2021

Erstgutachter
Prof. Dr. Carsten Schröder, DIW Berlin & Freie Universität Berlin

Zweitgutachterin
Prof. Dr. Katharina Jenderny, Umeå University

Datum der Disputation: 05.11.2021

Collaboration with Coauthors and Publications

Chapter 1: Investment Losses and Inequality

- Chapter 1 is based on an unpublished article that was written in equal parts with Johannes König.
- Longmuir and König (2019)

Chapter 2: De-routinization of Jobs and Polarization of Earnings – Evidence from 35 Countries

- Chapter 2 is based on an unpublished article that was written in equal parts with Carsten Schröder and Matteo Targa.
- Longmuir, Schröder and Targa (2021)

Chapter 3: The Distribution of Augmented Wealth in Australia from 2002 to 2018

- Chapter 3 is based on an unpublished article that was made in sole authorship.
- Longmuir (2021)

Chapter 4: Wage Risk and Portfolio Choice: The (Ir)relevance of Correlated Returns

- Chapter 4 is based on an unpublished article that was written in equal parts with Johannes König.
- König and Longmuir (2021)

Acknowledgements

Throughout the writing of this dissertation, I have received immense support and assistance.

Foremost, I want to thank my supervisors Prof. Carsten Schröder and Prof. Katharina Jenderny. I consider myself as very lucky that I had such supportive and approachable supervisors. Carsten initiated my first steps into the research world by employing me as a student assistant in the GSOEP department at the DIW Berlin, supervising my master thesis, helping me to apply for my scholarship, and guiding me throughout the last years. His supervision and feedback have increased and spurred my understanding of theoretical and empirical economics. Katharina has always provided excellent feedback and discussions on my work, which greatly improved this thesis and from which I learned tremendously. She employed me as a teaching assistant in Public Economics at the beginning of my PhD, and she has supported me ever since. I will never forget, while I was still a nervous first year student, how she organized my first seminar presentation outside of Berlin at Umeå University. I especially want to thank both for their great flexibility and ongoing support throughout the difficult time of the Covid19 pandemic during the last year of this thesis.

I also want to thank my co-authors Johannes König and Matteo Targa. Johannes and I have co-authored the first and the fourth chapter of this thesis. Johannes was always great to work with and he somehow managed to find a balance between being a most helpful mentor and a very good friend. My experience from our co-operation also spilled over to my other projects and it is not overstated that this thesis would be very different without him. Matteo and I co-authored the second chapter. Besides Matteo's amazing expertise, his positive spirit and winning attitude made this cooperation very fruitful and enjoyable. Since our cooperation I am sure: good food is the basis of good work.

I want to thank my PhD companions from the PhD program "Public economics and inequality", from the FU Berlin, and the GSOEP department of the DIW Berlin, namely Patrick Burauel, Peihua Deng, Benjamin Fischer, Daniel Graeber, Hannes Fauser, Christoph Halbmeier, Luisa Hammer, Viola Hilbert, Dominik Hügle, Edison Jakurti, Markus Krecik, Theresa Neef, Felicitas Schikora, Johannes Seebauer, and Puja Singhal. I especially thank my PhD colleagues and co-organizers of our conference "But Some are More Equal than Others. Interdisciplinary Perspectives on Growing Inequality and the Need for Solidarity", i.e., Mattis Beckmannshagen, Rick Glaubitz, Christine Körner, Astrid Pape, and Marie Wachinger. It was, and still is, amazing to work with you.

I also want to thank the scientific advisory board of my PhD program: Prof. Corneo, Prof. Schöb, and Prof. Steiner. I thank Holger Lüthen for his work as the program's coordinator. I especially want to thank Prof. Timm Bönke for his discussions and feedback over the last years. I thank Prof. Natalia Danzer for agreeing to join my doctoral committee. Thanks also to Nelli Anan for being the backbone and soul of the department.

Many thanks to the people at the Melbourne Institute, who have engaged with my work and made my stay abroad such a rewarding experience. I had a great time in Melbourne and I hope I see you all again soon. In particular, I want to thank Prof. Buly Cardak, Prof. Michael Coelli, Sarah Dahmann, Alexandra de Gendre, Prof. Bruce Headey, Nicolas Hérault, Johannes Kunz, Jordy Meekes, Prof. John P De New, Nicolás Salamanca, Felix Schenuit, Antonia Settle, Prof. Mark Wooden, Ou Yang, and Federico Zilio. I also want to thank Jo Steinle and the whole administrative staff. I especially thank Prof. Roger Wilkins for being an excellent host and a great partner for discussions on my research. The third chapter of this thesis would not have been possible without your help.

I gratefully acknowledge the financial and administrative support from the Hans-Böckler-Foundation.

A great thank you to my dear friends outside my PhD world, namely Felix, Jan, Nils, Patrick M., Patrick N., René, Simon, Sören, and, of course, the whole Viktoria Mitte Football team.

I want to thank my family, in Germany and in Australia, for your lovely encouragement and your help. I could not have done it without you!

And finally, I want to express my unending love and gratitude to my wife and best friend Sophie. Thank you for supporting and helping me unconditionally throughout all the highs and lows along the way.

Contents

1 Investment Losses and Inequality

1.1 Introduction

Household investment behavior is a central issue in economics and finance.[1] Households use their savings to finance future consumption and to pass wealth to their offspring – both important channels for individual welfare. Thus, losing out on potential returns from investments is a substantial source of welfare loss. Economic policy-makers should aim to improve households' investment behavior if it is found to be sub-optimal. Furthermore, inequality in investment losses affects wealth inequality, as found in Lusardi and Mitchell (2011a). If high-return investment opportunities are only available to high-wealth households (Calvet et al., 2007a; Goetzmann and Kumar, 2008), wealth inequality will grow over time. The macro-economic literature is in the process of identifying the main drivers of wealth inequality. Heterogeneous returns have emerged as a focal point (Benhabib and Bisin, 2018; Benhabib et al., 2011). In turn, as wealth inequality affects both the macroeconomy (Kaplan et al., 2018b) and political institutions (Piketty, 2018), it must be considered in economic policy-making.

In this paper, we analyze whether European households invest efficiently and the implications for financial wealth inequality. Using the Global Capital Asset Pricing Model (GCAPM) framework, we calculate return losses, i.e., the return lost by the inefficient investment choices of a household compared to an efficient benchmark. We then turn to potential factors driving return losses, analyze the inequality of return losses, and we show the distributional consequences of making household investment behavior more efficient.

Researchers analyze whether investors hold efficient financial portfolios by considering the fundamental trade-off between the expected return of a portfolio (mean return) and its risk (return variance). According to the GCAPM, investors behave efficiently if they take on the least amount of risk for any given return. This principle defines a set of efficient portfolios: those giving the highest return at any given level of risk. An efficient benchmark portfolio is typically defined by the mean return and the variance of a fully diversifying investor. Convex combinations of the risk-free asset with the benchmark portfolio trace out the mean-variance-efficient frontier. Households should hold a portfolio on this frontier to avoid unrewarded, unsystematic financial market risk.[2] If the household does not hold a portfolio on

[1]Major contributions to this literature include Markowitz (1952), Sharpe (1964), Lintner (1965), Benartzi and Thaler (2001). Calvet et al. (2007a), Cochrane (2009), Grinblatt et al. (2011), Von Gaudecker (2015), Fagereng et al. (2017b), and Bianchi (2018).

[2]This description of efficient behavior comes with a caveat. In the consumption-CAPM, assets are held based on their covariance with the marginal utility of consumption, which is risky because

the efficient frontier, it suffers a return loss, which is the difference in return between household and a defined benchmark portfolio measured at the household's chosen level of risk.

Empirical studies of return losses are limited to a few, select countries, namely France (Bianchi, 2018), the Netherlands (Von Gaudecker, 2015), Norway (Fagereng et al., 2016, 2017b,a), and Sweden (Calvet et al., 2007a).[3] The limiting factor for studies of this type is data availability. Constructing the relevant measures of inefficient investment requires a dataset with complete information on each asset in a household's portfolio. For most countries, these data requirements are not met.

On this front, our paper makes a methodological contribution. We use data on asset classes rather than every single asset in the portfolio to derive households' return losses. This lifts the data availability constraint and allows us to use the Household Finances and Consumption Survey (HFCS), which covers many central and peripheral European countries. We evaluate the reliability of this method by replicating the analysis of Von Gaudecker (2015), who used the CentERpanel alongside the Dutch Household Survey (DHS). The Dutch data contains every single asset, so that we can apply a traditional, unclassed GCAPM. To evaluate the class-based GCAPM, we artificially aggregate the data into classes, comparable to the HFCS, and compute return losses. Finally, we compare the distribution of return losses from both unclassed and classed data. Our results show that we obtain very similar distributions across both methods and that the errors we introduce are roughly symmetric and concentrated on zero. Building on this result, we apply the class-based GCAPM to the portfolios of households in Austria, Belgium, Germany, Ireland, and Spain. We choose these countries for reasons of data availability and because they are a balanced mix of central and peripheral countries.

Using the rich socio-demographic information from the HFCS dataset, we show that class-based return losses vary over several household characteristics. Households with higher levels of financial wealth incur smaller return losses. When we examine the components of the return losses, we see that this result is due both to a smaller risky portfolio share, i.e., the share of total financial wealth in risky assets, and a more diversified investment strategy. A more cautious risk attitude is associated with a smaller return loss due to a smaller risky portfolio share, but also less diversification.[4] Households with highly educated household heads have greater return losses because they have larger risky portfolio shares, although they hold more diversified portfolios compared to household with less educated heads. The pattern that more financially adept households, as shown by more diversified portfolios, still tend to have higher return losses because they also hold a larger risky

of, for example, labor income risk. The consumption-CAPM is a dynamic model that we cannot implement with our current data.

[3]The French data does not cover complete financial portfolios because it stems from the client records of one large financial institution.

[4]Risk attitudes are elicited on a 5-item-scale and self-reported.

portfolio share also carries through to country fixed effects. The clearest examples are Austria and Spain. Austrians have small return losses compared to Spaniards, but only because their risky portfolio share is small and not because they hold well-diversified portfolios.

In our distributional analysis, we show that the Gini coefficient of return losses is quite high, ranging from 0.48 to 0.61 across the countries in our sample. Inequality is high in countries with lower risk-taking, like Austria, and low in high risk-taking countries, like Spain. Further, along the financial wealth distribution, return losses accrue disproportionately to the less wealthy in several countries, while returns are higher for those high up in the financial wealth distribution. Finally, we investigate whether making household investment behavior more efficient has a sizable impact on financial wealth inequality. This is not the case. Although raising investment efficiency causes both an increase in total financial wealth and a progressive effect on the distribution, the latter effect turns out to be small. Further, this distributional change only affects the upper tail of the distribution of financial wealth. Across all countries, most households in the bottom 40% of the financial wealth distribution neither benefit from the increase in efficiency nor does their relative position in the distribution change. The implication from this experiment is that government programs aiming to increase investment efficiency will likely have a desirable effect on welfare but will not have strong effects on wealth inequality.

With this paper, we contribute to the literature on investment behavior and its effects on the private wealth distribution. Calvet et al. (2007a) and Von Gaudecker (2015) find in cross-sectional analyses that Swedish and Dutch households, respectively, are reasonably diversified. Both studies find return losses across all chosen levels of risk. Large absolute losses are associated with higher wealth, however, not due to under-diversification. Rather, wealthier households invest a larger fraction into risky assets. This implies, that less wealthy households incur larger diversification losses, but due to a smaller risky share of their overall portfolio absolute losses are smaller.

Our analysis focuses on individuals participating in risky financial markets. Therefore, we do not incorporate inefficiencies that potentially arise from non-participation. However, this should not affect our major findings. Welfare losses from non-participation are found to be relatively small, as non-participating households would be unlikely to invest efficiently (Calvet et al., 2007a). Fagereng et al. (2017b) estimate a life-cycle model using Norwegian data. Their model indicates that non-participation can be an efficient choice, if households face per-period participation costs and the risk of a financial crisis. Bianchi (2018) finds higher participation rates and higher expected returns for more financially literate households in France.

Other researchers analyze portfolio choices empirically without estimating an asset-pricing-model. They find a positive correlation between wealth, education, and risky asset-market participation in the U.S. (Vissing-Jorgensen, 2002; Goetzmann and Kumar, 2008; Kimball and Shumway, 2010) and Europe (Guiso et al., 2003).

Guiso et al. (2008) show that trust in equity markets also influences participation decisions. These differences in investment decisions and, hence, the heterogeneity of investment returns affects the wealth distribution.[5] Bilias et al. (2005) find for the U.S. that equity holding is important to explain the distribution of net wealth. However, a spread of asset market participation does not necessarily reduce wealth inequality. Bach et al. (2018) show with Swedish administrative data that heterogeneity in returns explains most of the historical increase in top wealth shares. Kuhn et al. (2017) find similar results in a long-term analysis in the U.S. using the Survey of Consumer Finances. Fagereng et al. (2016), using Norwegian tax-register data, find that returns are heterogeneous and positively correlated with wealth. Benhabib and Bisin (2018) fit an OLG-model with heterogeneous returns to wealth to U.S. data. They find that heterogeneous returns, also when correlated with wealth, are necessary to fit the tails of the wealth distribution in the cross-section.

The paper is structured as follows: in Section 1.2 we give a short review of the GCAPM, introduce the class-based GCAPM, and describe how we will apply it to the data. In section 1.3, we introduce our three data sources. Section 1.4 provides the validation of the class-based GCAPM and its application to the HFCS, while Section 1.5 investigates the inequality of return losses and their impact on the financial wealth distribution. Section 1.7 concludes.

1.2 Methodology

This section provides a summary of the theoretical underpinnings of the Global CAPM model and the route we take to empirically implementing it using the DHS and HFCS data.

The CAPM The CAPM is a popular model for analyzing portfolio choice. It was developed in the 1950s and 1960s, with the foundational contribution of the mean-

[5]The literature on return losses is closely connected to the literature on financial literacy, since financial literacy seems to mitigate return losses. Bernheim et al. (1998) and Hilgert et al. (2003) are pioneers in this area, providing empirical evidence on the positive effect of financial education on households' finances, especially in terms of insuring against income shocks and efficient investment strategies. Lusardi and Mitchell (2007), Lusardi and Mitchell (2011a), Lusardi and Mitchell (2011b), and Lusardi and Mitchell (2014) extend this analysis, contributing greatly to the literature by establishing standardized financial literacy measurements and applying them in countries around the world. The importance of financial literacy to explain wealth inequality is shown in Lusardi et al. (2017), where the authors establish a life-cycle model and show that 30 to 40 percent of retirement wealth inequality is accounted for by financial knowledge in the U.S. With our research, we show how wealth inequality in Europe is affected by inefficient investment strategies.

variance-efficiency analysis by Markowitz (1952) and the extension with the separation theorem due to Tobin (1958).[6]

The CAPM incorporates two main dimensions of interest for the investor: the expected return of an asset and its volatility. The expected return is the percentage change in value of an asset anticipated by the investor. The volatility, commonly measured by the standard deviation, is the measure of risk. An asset with zero volatility is risk-free and its expected return equals the realized return. Empirically, both the expected return and the standard deviation are constructed from the historic time series of the assets. Return losses in the CAPM can be computed as a portfolio's shortfall in expected return compared to a portfolio on the mean-variance-efficient frontier. Convex combinations between the risk-free asset and an efficient benchmark portfolio will trace the mean-variance-efficient frontier (Markowitz, 1952). Portfolio choice theory states that rational investors will hold a portfolio that locates on this frontier and that the location of that portfolio on the frontier is determined by risk aversion (Merton, 1969).

The Global CAPM The GCAPM is an extension of the CAPM, developed by Solnik (1974), Sercu (1980), and Adler and Dumas (1983). It assumes that investors can access a fully integrated, international financial market. Accordingly, the model takes currency fluctuation risk into account. We choose the GCAPM, because the countries in our analysis are open economies and expected returns as well as risk on local markets are driven to a large extent by global market conditions.

In the GCAPM, asset prices are set on world markets in an international currency, namely the U.S. Dollar. Our dollar-denominated benchmark is the Morgan Stanley Capital International Europe Index (MSCI Europe), which is motivated both by choosing a benchmark natural to the European context and the replication of Von Gaudecker (2015).[7] Currency-hedging of the benchmark portfolio, which converts the dollar-denomination into Euros, is implemented in the following set of steps according to MSCI methodology: Let the price of the unhedged, dollar-denominated benchmark index at time t be denoted by $p^{\$}_{m,t}$ and let the exchange rate from Dollars to Euros be $ex^{€}_{\$,t}$. Then the hedged benchmark price $p^{€}_{m,t}$ is given by

$$\begin{aligned} p^{€}_{m,t} &= p^{\$}_{m,t} \times (1 + CF_t) \qquad (1.1) \\ CF_t &= \frac{p^{\$}_{m,t}}{p^{\$}_{m,t-1}} \times \frac{ex^{€}_{\$,t} - ex^{€}_{\$,t-1}}{ex^{€}_{\$,t-1}}. \end{aligned}$$

[6] See Treynor (1961), Sharpe (1964), Lintner (1965) and Mossin (1966).

[7] Further information on the chosen index is provided in Morningstar (2015).

To construct the benchmark excess return series, we need a risk-free rate of return $r_{f,t}$. We choose the EURIBOR 1 month rate. The choice is motivated by the same line of arguments as our choice of the benchmark. Then, excess returns of the benchmark are,

$$r^e_{m,t} = \frac{p^{€}_{m,t} - p^{€}_{m,t-1}}{p^{€}_{m,t-1}} - r_{f,t}. \tag{1.2}$$

The volatility of the hedged benchmark is the same as the volatility of the return of the dollar-denominated one ($p^{\$}_{m,t}$). The equation for the factor model to implement the GCAPM takes the form:

$$r^e_{i,t} = \beta_i r^e_{m,t} + \epsilon_{i,t}, \tag{1.3}$$

where the excess return of an asset i in period t is $r^e_{i,t}$ and $\epsilon_{i,t}$ is white noise. The coefficient β_i is a measure of comovement between the benchmark portfolio and the asset i. For example, when β_i is exactly one, excess returns of the asset and the benchmark have perfect comovement. In this case, holding the asset does not carry any idiosyncratic risk. When β_i equals zero, the asset is driven by idiosyncratic risk only. As a result, the expected returns on this asset are at the same level of return as the risk-free asset.[8] Using the mean excess return of the benchmark μ_m and the estimate of β_i, we can construct the mean excess return of the asset μ_i and the vector of all excess returns $\boldsymbol{\mu}$.[9] Using the variance-covariance matrix of the $\epsilon_{i,t}$, the vector of beta-factors $\boldsymbol{\beta}$ and the variance of the benchmark returns σ_m, we can derive the variance-covariance matrix of these returns Σ with the variances of each asset σ_i^2 on the main diagonal.[10]

Unclassed Asset Information When we have the vector of household portfolio weights $\boldsymbol{\omega}_h$, i.e., the share of each asset in the household's financial wealth portfolio, we can construct $\mu_h = \boldsymbol{\omega}_h' \boldsymbol{\mu}$, the household portfolio's expected excess return, and $\sigma_h = \left(\boldsymbol{\omega}_h' \Sigma \boldsymbol{\omega}_h\right)^{1/2}$, the household portfolio's standard deviation. The weights $\boldsymbol{\omega}_h$ are available in the DHS data we use in Section 1.4.1.

Classed Asset Information In the HFCS, the dataset we use for our analysis of return losses, information on assets is grouped into classes. Our method to proceed is to calculate an excess return series for each of the classes and treat them like individual assets. This is equivalent to imputing asset-specific weights for each household from the class-specific weights. We use two procedures to construct class-specific excess return series. Procedure (I) aggregates all historic prices of the

[8] β_i can become negative, such that the excess return of the asset moves in the opposite direction of the benchmark. According to the GCAPM, the asset provides a lower expected return than the risk free asset.

[9] Bold symbols denote vectors or matrices.

[10] The full method is outlined in Calvet et al. (2007b).

assets we have available in a class to a class-specific mean, which is then used to calculate the class-specific return series. Procedure (II) proxies this class-specific return series by the returns of an index, e.g., the German DAX. Further details on the two procedures are provided in Section 1.3. In Section 1.4.1, where we evaluate our class-based approach, we follow procedure (I) to artificially aggregate returns within a class.[11]

Let ω_h^c be the weight of a given asset class in the household's financial portfolio. Then, for procedure (I), imputed weights for each asset within a class $\omega_{h,i}^c$ are

$$\omega_{h,i}^c = \omega_h^c \times \frac{1}{N_c}, \tag{1.4}$$

where N_c is the number of assets within an asset class. Thus, we give equal weight to each asset within a class. When we follow procedure (II), $\omega_{h,i}^c$ is calculated in an analogous fashion, except that $\frac{1}{N_c}$ is swapped for the weights specific to the index we use. The class-based measures of households' expected excess returns and portfolio standard deviations are,

$$\begin{aligned} \mu_h^c &= \boldsymbol{\omega_h^c}' \boldsymbol{\mu} \qquad (1.5) \\ \sigma_h^c &= \left(\boldsymbol{\omega}_h^{c\prime} \Sigma \boldsymbol{\omega}_h^c \right)^{1/2}. \end{aligned}$$

In Section 1.4.1, we can compute the difference between classed and unclassed values of expected returns and portfolio standard deviations.

$$\begin{aligned} \Delta\mu_h &= \sum_{i=1}^{N} (\omega_{h,i} - \omega_{h,i}^c)\beta_i \mu_m \qquad (1.6) \\ \Delta\sigma_h &= \sqrt{\sum_{i=1}^{N} \omega_{h,i}^2 \sigma_i^2 + \left(\sum_{i=1}^{N} \omega_{h,i}\beta_i \right)^2 \sigma_m^2} - \sqrt{\sum_{i=1}^{N} (\omega_{h,i}^c)^2 \sigma_i^2 + \left(\sum_{i=1}^{N} \omega_{h,i}^c \beta_i \right)^2 \sigma_m^2}. \end{aligned}$$

Return Loss Ultimately, we want to quantify the return loss. The return loss is the vertical distance between the expected return of the household's portfolio and the mean-variance-efficient frontier. Thus, return losses are measured in percentage points. Given a certain level of risk, a higher return loss represents a less efficient

[11] We provide an example on calculating excess returns in Appendix 1.8.1.

investment strategy compared to the benchmark. With unclassed asset information we can calculate the return loss as stated in Calvet et al. (2007a).

$$\begin{aligned} RL_h &= \mu_m \times \beta_h \times w_h \times DL_h, \\ \beta_h &= \boldsymbol{\omega}_h' \boldsymbol{\beta}, \\ DL_h &= \frac{S_m - S_h}{S_h}. \end{aligned} \tag{1.7}$$

RL_h is the household's return loss. β_h is the household portfolio's beta factor. w_h is the proportion of financial wealth invested in risky assets, the risky portfolio share.[12] DL_h is the diversification loss, which is determined by the Sharpe ratio of the benchmark $S_m = \mu_m/\sigma_m$ and the Sharpe ratio of the household $S_h = \mu_h/\sigma_h$.

The return losses using class-based portfolio weights are calculated analogously:

$$\begin{aligned} RL_h^c &= \mu_m \times \beta_h^c \times w_h \times DL_h^c, \\ \beta_h^c &= \boldsymbol{\omega}_h^{c\prime} \boldsymbol{\beta}, \\ DL_h^c &= \frac{S_m - S_h^c}{S_h^c}. \end{aligned} \tag{1.8}$$

In Section 1.4.1, we replicate the analysis of Von Gaudecker (2015), comparing the unclassed and classed return losses RL_h and RL_h^c. Moreover, we separately describe each of their components, namely the beta factor, the risky portfolio share, and the diversification loss. In Section 1.4.2, we conduct our class-based analysis of return losses using the HFCS data. We perform a regression analysis of RL_h^c and, by applying the natural logarithm, transform equation 1.8 into:

$$ln\ RL_h^c = ln\ \mu_m + ln\left|\beta_h^c\right| + ln\ w_h + ln\ \left|DL_h^c\right|. \tag{1.9}$$

We run OLS regressions on the log return loss and the three components:

$$\begin{aligned} &ln\ RL_h^c = \mathbf{X}_h \boldsymbol{\gamma}_1 + \varepsilon_{h,1}, \\ &ln\ w_h = \mathbf{X}_h \boldsymbol{\gamma}_2 + \varepsilon_{h,2}, \\ &ln\ \left|\beta_h^c\right| = \mathbf{X}_h \boldsymbol{\gamma}_3 + \varepsilon_{h,3}, \\ &ln\ \left|DL_{h,c}\right| = \mathbf{X}_h \boldsymbol{\gamma}_4 + \varepsilon_{h,4}. \end{aligned} \tag{1.10}$$

The vector $\mathbf{X}_i$ is comprised of socio-economic characteristics of the household, country fixed effects, and a constant. The $\varepsilon_{h,\bullet}$ denote residuals. These regressions enable us to determine which variables have relevant associations with the return loss and

[12] Risky assets are mutual funds, stocks, managed accounts, and bonds following Von Gaudecker (2015).

which of the components are responsible for the overall effect. Further, we can directly compare these results to Calvet et al. (2007a).

Our methodology implies that the more asset classes we can define, the better the class-based approach will perform. Both quantities in Eq. (1.6) would be zero if $\omega_{h,i} = \omega^c_{h,i}$. This could arise if households invest the wealth they allocate to a class c equally between all assets in that class. Similarly, if the number of asset classes approaches the number of individual assets, $\omega^c_{h,i}$ would tend to $\omega_{h,i}$.

1.3 Data

We use three data sources: 1) the CentERpanel linked with the Dutch Household Survey (DHS); 2) the Household Finances and Consumption Survey (HFCS); and 3) the Datastream and Eikon database by Thomson Reuters.

1.3.1 CentERpanel and DHS

The CentERpanel is a panel survey in the Netherlands comprising about 2,000 households. It is representative of the Dutch population and conducted regularly via the internet. The CentERpanel can be linked to the annually conducted DHS, which provides information on numerous financial characteristics, e.g., income, pensions, loans, real assets, financial assets, and personal characteristics, e.g., age, education, etc. For our replication analysis, the information on financial assets is especially important. The DHS includes data with the quantity, names, and the amount invested in mutual funds and shares. Due to this detailed information on financial portfolios, it is suited for an application of the GCAPM. A more comprehensive description of the dataset is found in Von Gaudecker (2015).

1.3.2 HFCS

We use the second wave of the Household Finance and Consumption Survey to calculate return losses in Section 1.4.2. The HFCS is a representative survey provided by the European Central Bank. The second wave was surveyed between 2011 and 2014 in 20 European countries.[13]

As stated at the outset, the HFCS dataset does not provide information on single assets, but a wealth module that contains total financial wealth and the share of several asset classes in financial wealth. We include twelve asset classes that are provided in the HFCS in our class-based GCAPM: mutual funds investing in equity, bonds, money market instruments, real estate, hedge funds, and others, as well as bonds issued by governments, banks, and non-financial institutions, publicly traded stocks, managed accounts, and savings accounts. We restrict our dataset to

[13]For most countries the survey was conducted in 2014.

households from Austria, Belgium, Germany, Ireland, and Spain. The HFCS dataset is multiply imputed to correct for item non-response.

Table 1.1 shows descriptive statistics of several important household characteristics for the countries in our working sample. The "HFCS working sample" column shows descriptives for all households in the working sample pooling all countries. For an observation to be included in the working sample, it must have a positive risky portfolio share, otherwise the household is *out* of the financial market. Our working sample includes 5,521 households. Looking at the country-specific statistics, sample sizes vary greatly between countries.

Table 1.1. Descriptive Statics

Variable	HFCS working sample	HFCS AUT	HFCS BE	HFCS ESP	HFCS GER	HFCS IR
Average Members in HH	2.25	2.38	2.36	2.60	2.05	2.89
	(1.13)	(1.23)	(1.20)	(1.13)	(1.04)	(1.47)
Households with Children in %	19.05	19.22	22.19	23.69	16.12	33.34
	(39.27)	(39.40)	(41.55)	(42.52)	(36.77)	(47.14)
Average Age Women (HH Heads)	57.48	58.67	57.80	56.51	58.11	49.50
	(15.27)	(15.30)	(16.28)	(14.22)	(15.29)	(16.20)
Average Age Men (HH Heads)	55.37	53.84	56.57	56.39	55.17	48.75
	(15.17)	(15.78)	(15.72)	(14.20)	(15.33)	(14.75)
Percentage of HH married	59.43	61.38	55.41	65.86	57.26	67.64
	(49.10)	(48.69)	(49.71)	(47.42)	(49.47)	(46.79)
Migration background in %	10.32	10.98	9.030	n/a	10.49	n/a
	(30.42)	(31.27)	(28.66)	(n/a)	(30.65)	(n/a)
Average monthly hh income	3,724	3,488	3,349	2,465	4,245	4,579
	(4,229)	(2,260)	(2,126)	(3,171)	(4,798)	(4,947)
Average net wealth in €	550,856	654,105	597,032	646,892	502,173	456,975
	(1,330,461)	(2,855,263)	(673,374)	(1,601,360)	(1,114,172)	(690,605)
Average financial wealth in €	155,511	106,905	217,237	160,470	147,600	127,739
	(508,375)	(141,997)	(332,734)	(896,216)	(322,823)	(375,454)
Average real asset wealth in €	415,301	604,518	425,277	460,191	378,139	491,177
	(1,077,963)	(2,930,779)	(498,720)	(906,336)	(929,094)	(894,918)
Observations	5,521	440	686	2,011	1,328	1,056

Note. Authors' calculations using the HFCS. The table shows weighted, descriptive statistics for selected variables. The statistics are given for the working sample in our analysis. Migration information is not provided by HFCS for Spain and Ireland. All financial variables are PPP adjusted, using the OECD PPP statistics from 2014, as recommended by Brandolini (2007) and Davies et al. (2010). The income variable incorporates the household's equivalent income using the OECD-modified scale (Hagenaars et al., 1994). Standard deviations are in parentheses. We calculated point estimates and standard deviations taking multiple imputation into account.

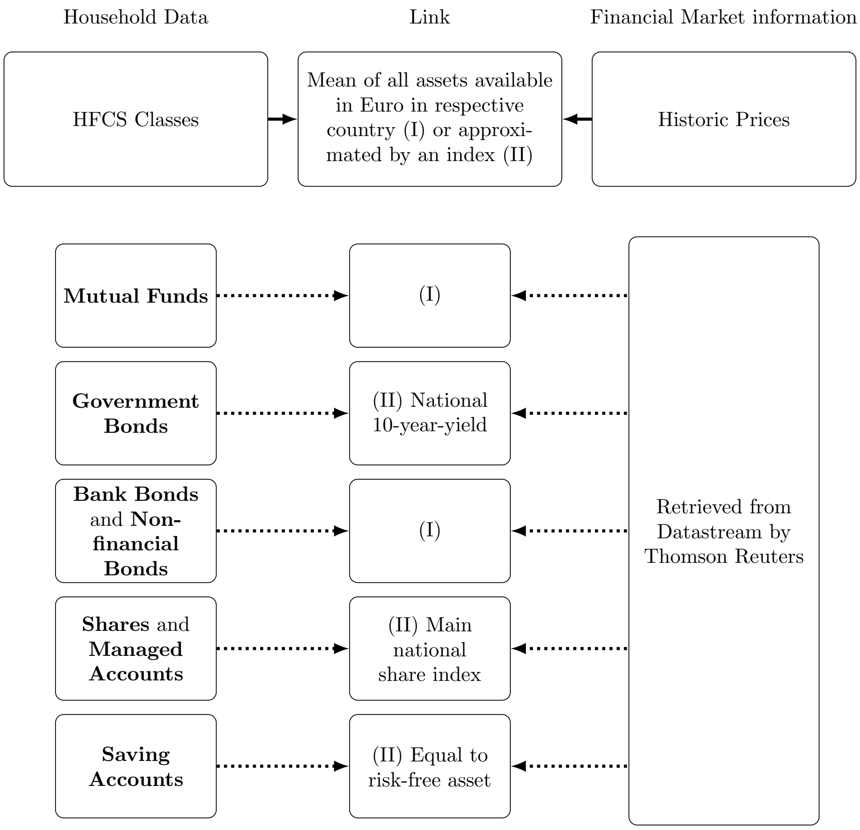
The descriptives on the financial variables show the selective nature of our sample: average monthly household income is 60% higher compared to the full sample (see Table 1.6 in Appendix 1.8.2). At the same time, the higher standard deviation shows that income and wealth is more unevenly distributed in our sample. This is unsurprising, since we drop those without risky assets, which includes many households with zero financial wealth. Furthermore, the intentional oversampling of the rich had varying degrees of success across countries. Austria, for instance, does not apply any oversampling strategy at all, whereas the Spanish survey used wealth tax data in their oversampling procedure to much success. However, the increase of net and financial wealth is relatively similar across all countries.[14]

To proceed we need to link the historic, excess return time series with the HFCS asset classes. We use the two procedures outlined in Section 1.2: The first procedure (I) uses all prices of the assets we have available and estimates the class-specific mean return. These class-specific returns differ for each country. The second procedure (II) proxies this class-specific mean return by an index. Figure 1.1 provides an overview of the linking procedure and to which asset classes we apply either procedure. On the left side, we have the HFCS asset classes, the middle shows the procedure (I) or (II), the right side represents the origin of the historic returns, i.e., the Datastream database by Thomson Reuters. If we follow procedure (II), it means that procedure (I) did not appear reasonable or feasible for this particular asset class. Government Bonds may have different maturities, which we do not observe. This lack of data is remedied by assuming a home-biased investment strategy[15] and approximating the historic returns with the ten-year-treasury bond of the respective country. We implement procedure (II) for stocks and managed accounts. Thus, we assume that stockholders and consultants (managed accounts) can perform as well as the national stock index. Saving accounts are set equal to the risk-free asset, which is in line with Calvet et al. (2007a) and Von Gaudecker (2015). A schematic example of the construction of the return series is provided in Figure 1.14.

We only include assets traded in Euros; a restriction that is only relaxed in the case of the hedge fund category; for which we assume a currency-hedge. Even if we could use the information on single assets, we would not be able to observe whether households hedge currency fluctuations or bear the currency risk. Relying on the home bias argument from before, this should approximate the real investment environment as close as possible.

[14]More information about the full dataset and the working sample is provided in Appendix 1.8.2.

[15]The home-biased strategy, found in Calvet et al. (2007a) and Von Gaudecker (2015), is analyzed in detail by Graham et al. (2009) and Kimball and Shumway (2010).

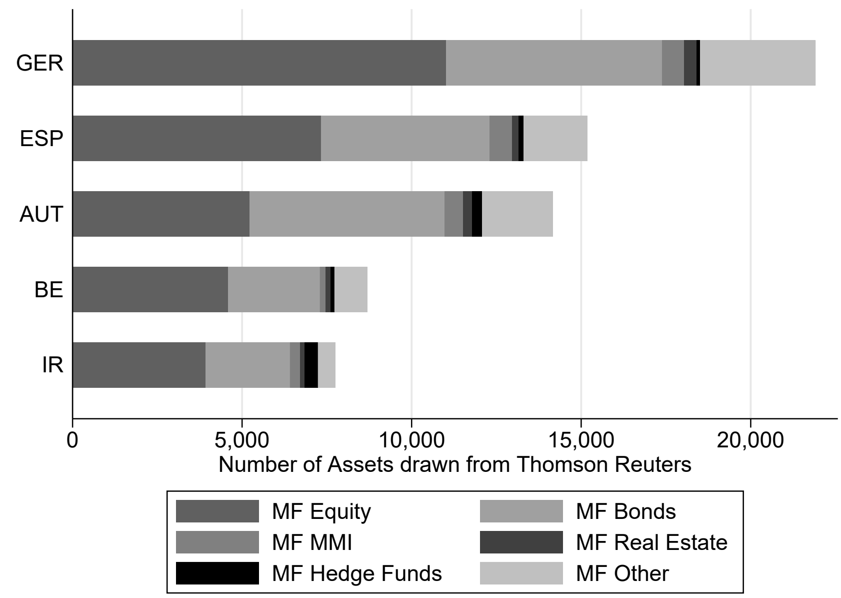


Note. Compiled by authors. The first column shows the financial portfolio categories. The third column represents the information from Eikon and Thomson Reuters Datastream, while the second column shows how they are linked.

Figure 1.1. Matching Financial Classes with Historic Returns

1.3.3 Datastream and Eikon

We construct the class-based return series using information on the assets in each class. For implementation, we rely on the Datastream and Eikon Database by Thomson Reuters. Thomson Reuters Eikon is a financial data screening platform providing data on equities, bonds, stock market indices, currencies, and many other financial products and related information. It is one of the most comprehensive financial databases available. Given that it tracks several million asset types and it is used by professional financial investors all over the world, it is reasonable to assume that it provides the most relevant assets. We use the Eikon database to derive the International Securities Identification Number (ISIN) of all assets available in 2014 in the countries represented in the HFCS working sample. With the ISIN, we can look for the asset's historic returns in the Datastream database. Other authors in the portfolio analysis literature also use the Datastream database (Calvet et al., 2007a; Von Gaudecker, 2015).



Note. Author's calculation based on Eikon and Thomson Reuters Datastream. Figure shows the number of assets used for mutual funds asset classes by country.

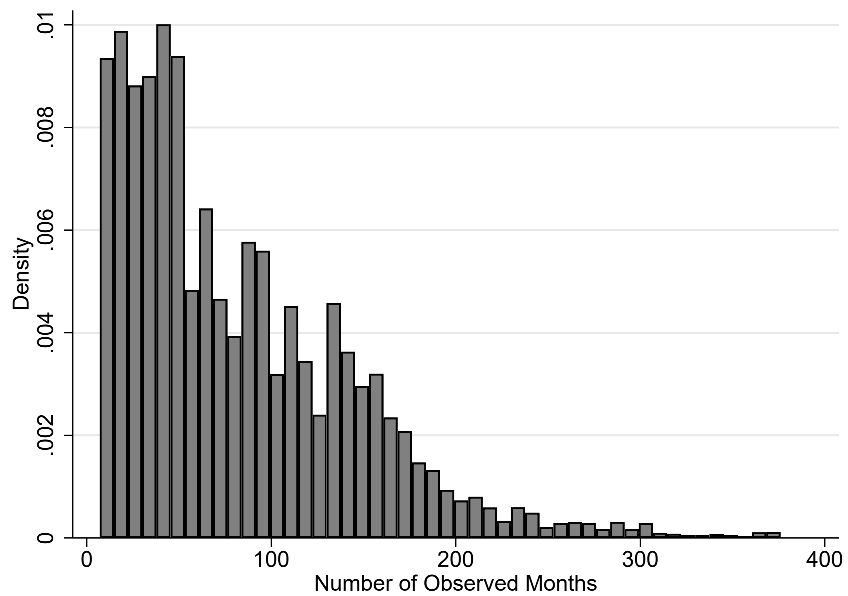
Figure 1.2. Number of Assets per Country

We collect the information on 69,395 financial assets. Figure 1.2 shows the number of assets used to construct the time series for the mutual funds classes[16] by country. The most exhaustive asset information is available for Germany, followed

[16]The same procedure is applied for the Bank-Bonds and Non-financial Bonds. These are not reported in Figure 1.2.

by Spain and Austria. Datastream has a relatively low number of assets available for Belgium and Ireland. The data shows that the Irish financial market includes fewer small mutual funds than, for example, Germany. This fact goes toward explaining the smaller number of assets available for Belgium and Ireland. Based on our home-bias assumption, we only include assets available in Euros, except for the hedge fund category. Many funds traded in Ireland are available in pounds or dollars only, which we do not include.

Figure 1.3 shows the distribution of the time-series length for the historic return series we have available. We observe assets on average for 82 months with a standard deviation of 63 month. Note that we drop those assets where we have less than 7 months of observations. Therefore, our return series have a minimum observation period of 7 month and a maximum observation period of 375 months. Calvet et al. (2007a) cut off their time series after 120 months. We follow Von Gaudecker (2015) by using the full length of the time series. We provide an overview about the differences in the return loss using different benchmarks and time lags in Appendix 1.8.2.



Note. Author's calculation based on Eikon and Thomson Reuters Datastream. Distribution of the time-series length of historic returns in month using all 69,395 financial assets.

Figure 1.3. Observation Time of Assets

1.4 GCAPM Estimation

We divide the presentation of the GCAPM estimation into two subsections. First, we show how our class–based GCAPM approach performs in comparison to the unclassed GCAPM using DHS data. Second, we provide descriptive results of our class–based GCAPM application for several countries in Europe working with the HFCS and the historic prices from Datastream. We perform a regression analysis to understand which socio-economic characteristics are associated with return losses and how they differ across countries.

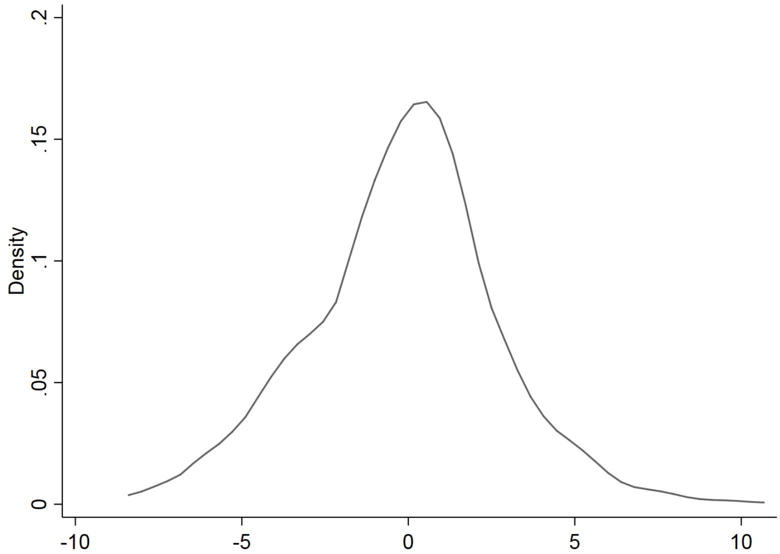
1.4.1 Comparing the Unclassed and the Classed GCAPM

We start with the comparison between the expected returns, portfolio standard deviations, and return losses calculated from the unclassed and the artificially classed DHS data. As described in Section 1.3, we use the hedged MSCI Europe index as the benchmark and the EURIBOR 1 Month rate for the safe return.

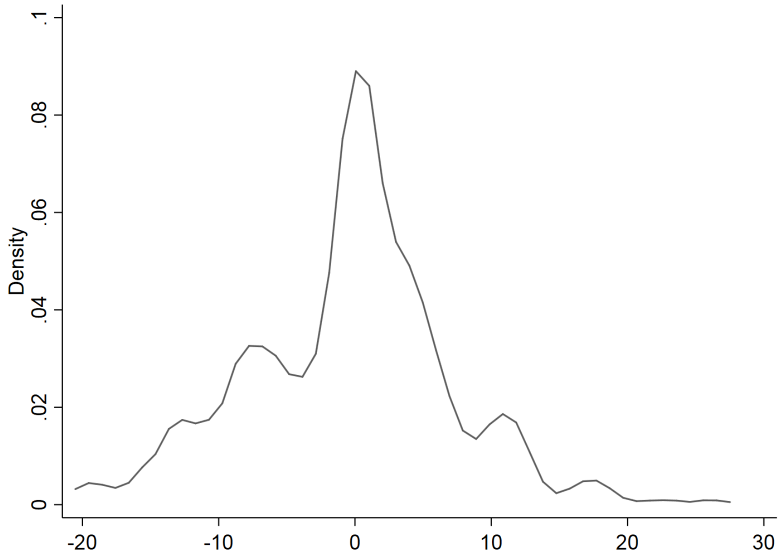
Excess Returns and Standard Deviations First, we evaluate the distributions of the central characteristics of the household portfolios according to the GCAPM: the expected excess return and the portfolio standard deviation. The quantities $\Delta\mu_h$ and $\Delta\sigma_h$ in equation (1.6) give the difference introduced by using classed weights.

In Figure 1.4 we show the density plots for both the difference in the expected return, $\Delta\mu_h$, and the standard deviation, $\Delta\sigma_h$, based on the DHS. Both distributions are centered around zero with the mode of the distribution very close to zero. While the distributions are certainly not normal, both exhibit tails that are fairly thin; less so for the distribution of $\Delta\sigma_h$. The distribution of the $\Delta\sigma_h$ exhibits a particular lumpiness: several values, around 10 and -10, show distinct humps in the density plot. It is $\Delta\mu_h$ and $\Delta\sigma_h$ that will carry through to the difference in the classed and the unclassed return losses. Specifically, as one can see from equation (1.8), the beta factor β_h and the diversification loss DL_h will be affected.

Return Losses We proceed by showing the distribution of the two return loss variables. Table 1.2 shows the distribution of return losses and the distribution of their components (Equations (1.7) and (1.8)) in terms of quintile-specific means. The most noticeable feature is that we match the moments of the unclassed return loss distribution well using the class-based approach. This is a quite remarkable result, since the implication is that we can learn much about the distribution of the unclassed from the classed return losses. Further, the means of the β_hs in each quintiles also track each other well. The only component that does not track between classed and dis-aggregated data is the diversification loss. Within each quantile, it is about half as large in the classed data. This is due to the fact that the classed measure



(a) Difference in the Expected Return $\Delta\mu_h$



(b) Bias in the Standard Deviation $\Delta\sigma_h$

Note. Authors' calculation based on the DHS. We plot the kernel density of (a) the difference in the expected return, $\Delta\mu_h$, and (b) the bias in the standard deviation of the expected return, $\Delta\sigma_h$, for each observation in the DHS sample using an Epanechnikov kernel with a bandwidth parameter of 1.

Figure 1.4. Density Plot of $\Delta\mu_h$ and $\Delta\sigma_h$

refers to the diversification across classes instead of individual assets, which incurs considerably less variation.

Table 1.2. Distribution of Return Losses according to Unclassed and Classed Household Portfolios

	unclassed				classed			
	RL_h	w_h	β_h	DL_h	RL_h^c	w_h	β_h^c	DL_h^c
1.Q	0.04	0.08	0.77	0.30	0.03	0.09	0.80	0.18
	(0.00)	(0.01)	(0.03)	(0.06)	(0.00)	(0.01)	(0.03)	(0.01)
2.Q	0.14	0.23	0.83	0.47	0.13	0.18	0.87	0.21
	(0.00)	(0.02)	(0.04)	(0.07)	(0.00)	(0.01)	(0.03)	(0.01)
3.Q	0.29	0.43	0.89	0.27	0.27	0.37	0.82	0.23
	(0.01)	(0.02)	(0.03)	(0.04)	(0.00)	(0.02)	(0.03)	(0.01)
4.Q	0.52	0.54	0.91	0.50	0.48	0.59	0.83	0.23
	(0.01)	(0.03)	(0.03)	(0.07)	(0.01)	(0.02)	(0.03)	(0.01)
5.Q	1.76	0.60	1.01	0.89	1.62	0.65	1.10	0.39
	(0.16)	(0.02)	(0.05)	(0.08)	(0.09)	(0.02)	(0.02)	(0.01)

Note. Authors' calculation based on the DHS. Replication of Table 1 in Von Gaudecker (2015). Values are means in the quintiles of the distribution of return losses. Standard errors in parentheses. As in Von Gaudecker (2015) we winsorize the diversification loss from above at the 95th percentile.

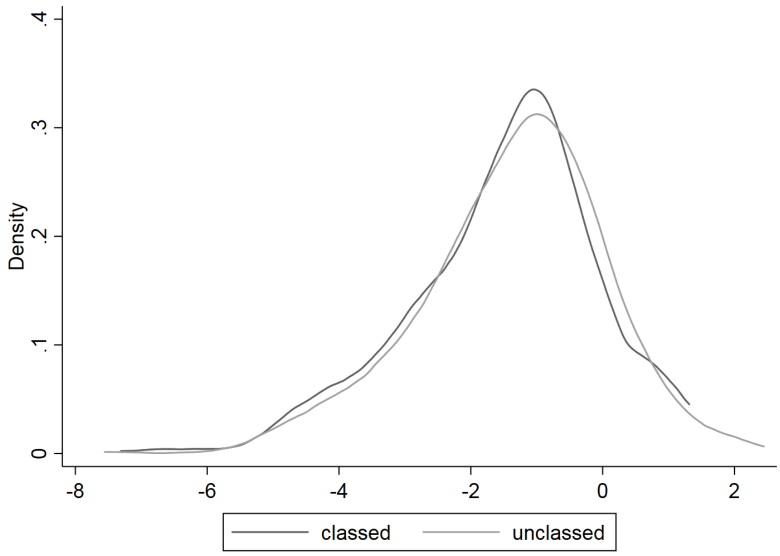
In Section 1.5, we consider various statistics relating to the inequality of return losses, starting with the Gini coefficient and the coefficient of variation. In Table 1.3, we show the Gini coefficient and coefficient of variation computed from the classed and unclassed data both to compare the distributions and to give an indication about the reliability of our results in Section 1.5. Remarkably, the point estimates of the Gini coefficients for both the classed and the unclassed data are very close to each other and the confidence bands of both estimates overlap. The same does not hold for the point estimates of the coefficient of variation; however, the confidence bands of the two estimates still overlap.

Table 1.3. Inequality of Return Losses according to Unclassed and Classed Household Portfolios

	unclassed			classed		
	CI(-)	PE	CI(+)	CI(-)	PE	CI(+)
Gini	0.578	0.615	0.653	0.581	0.606	0.631
CoV	1.376	1.695	2.014	1.323	1.415	1.507

Note. Authors' calculation based on the DHS. Point estimate (PE) and bootstrapped confidence intervals (CI) of Gini coefficient (Gini) and coefficient of variation (CoV) for unclassed and classed data. For the bootstrap, we use 1000 replicate weights and compute the confidence intervals according to the normal distribution.

To give a more immediate impression where the two return loss distributions diverge, we graph the kernel density of the log return losses in Figure 1.5.[17]

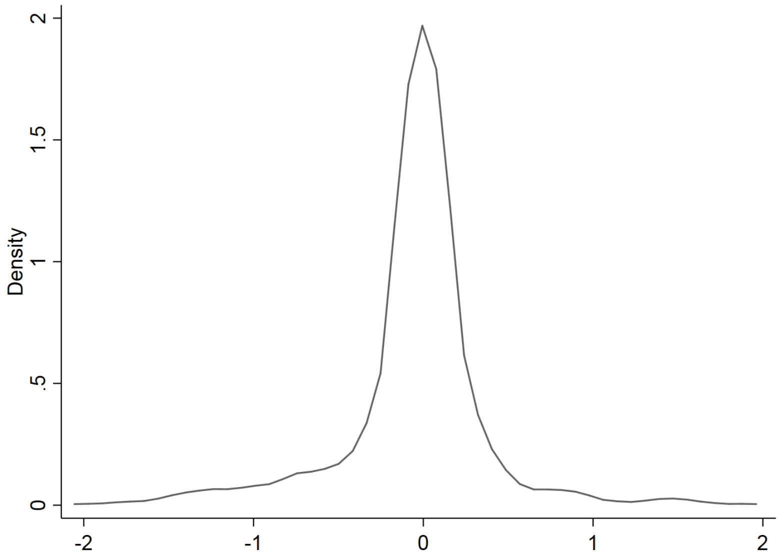


Note. Authors' calculations based on the DHS. We plot the kernel density of the classed and unclassed log return loss variables using an Epanechnikov kernel with a bandwidth parameter of 0.4.

Figure 1.5. Density Plot of the two Return Loss Variables

The figure shows that by artificially aggregating the asset data to asset classes, we lose some, but not much, of the dispersion of the unclassed return losses. The fit is fairly good. The classed variable is certainly more concentrated around the mean and it lacks very large values in the right tail of the distribution, but other than that – and especially in the left tail – the densities line up very well. However, there are still differences in the return losses within individuals. It could be the case that we match the distributional features of the unclassed return loss distribution but sort individuals into incorrect ranks. For this reason, we show the kernel density of the within-individual difference between the classed and unclassed return losses in Figure 1.6. The differences within each individual are very close to zero for most of the sample. Furthermore, the difference is evenly distributed around zero. This allays concerns that we can reproduce the distribution, but are sorting individuals into incorrect ranks. We provide some summary statistics of the interpersonal difference that go along with Figure 1.6 in Table 1.4. The summary statistics show that both the mean and the median error are very small as well as that the 25th and 75th quantile fall fairly symmetrically around the median. All this affirms

[17]We use the log-transformation because the variables appear approximately log-normal and the differences in the distributions become more apparent in logs.



Note. Authors' calculations based on the DHS. We plot the kernel density of the difference between the classed and unclassed return loss variables using an Epanechnikov kernel with a bandwidth parameter of 0.1. We restrict the sample to values between -2 and 2. This eliminates 20 observations.

Figure 1.6. Density Plot of $RL_{h,c} - RL_{h,i}$

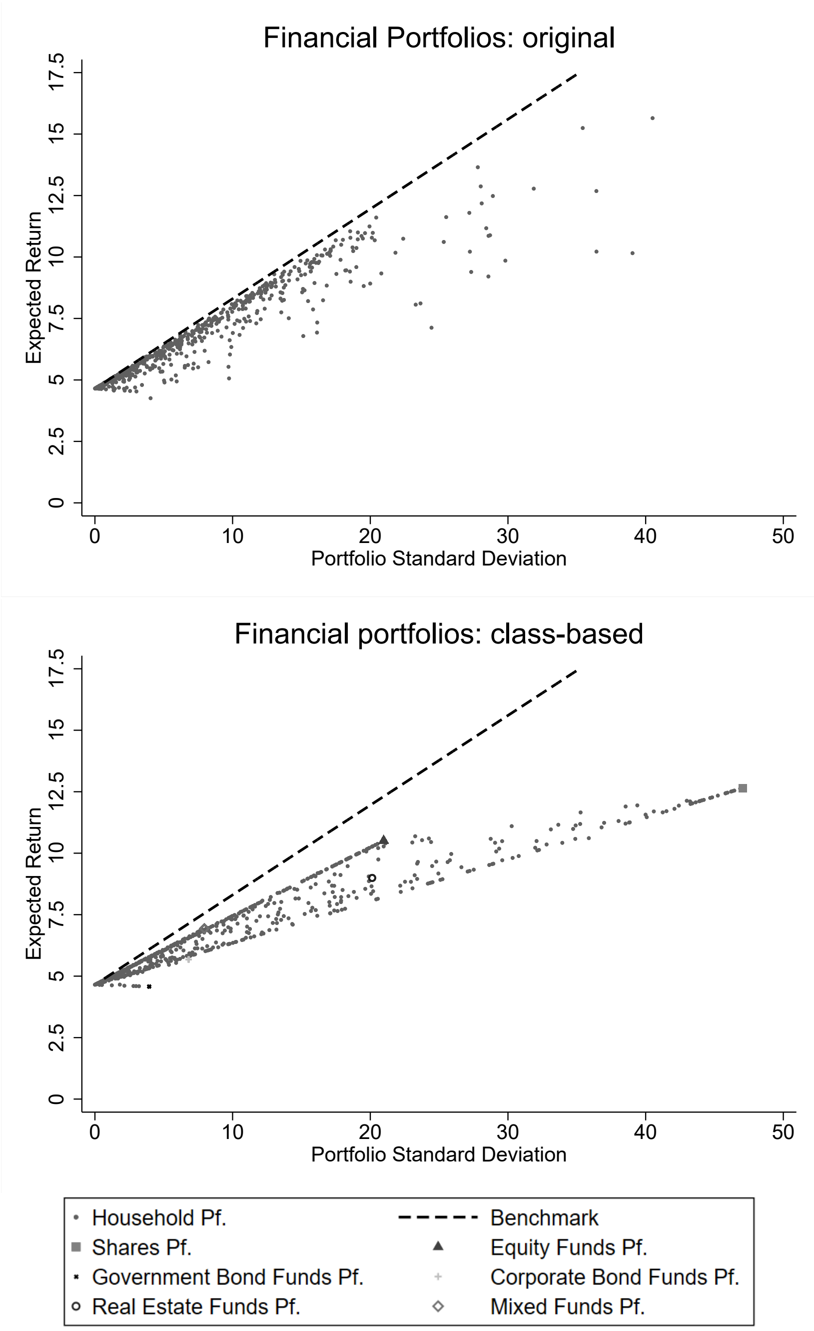
Table 1.4. Summary Statistics of Return Loss Differences

	mean	std. dev.	25th Q.	median	75th Q.	Obs.
$RL_h^c - RL_h$	-0.049	0.414	-0.110	-0.007	0.088	579

Note. Authors' calculations based on the DHS. We restrict the sample to values between -2 and 2. This eliminates 20 observations.

that the error we introduce is small and not biased. In Appendix 1.8.3, we show that the fraction of observations with large relative differences in return losses is small. In sum, we find that the errors we introduce to the return losses by using the class-based approach are small and symmetrical, preserving the shape of the distribution of return losses quite well. The drawback remains that we cannot match the distributions of all components of return losses, especially the diversification loss. However, these results are based on seven asset classes, whereas our main analysis with HFCS is conducted with twelve asset classes. Hence, these results should be seen as a lower-bound in terms of the reliability of our class-based approach.

Overall, the class-based approach delivers a desirable fit to the unclassed return losses both on the aggregate and individual level. Thus, we have constructed a meaningful measure of return losses based on classed household portfolio data.



Note. Authors' calculations based on Von Gaudecker (2015) analysis of the DHS. The first (second) graph provides the unclassed (classed) expected returns and the portfolio standard deviation of the Dutch financial asset portfolios (gray dots). The benchmark represent the currency-hedged MSCI Europe.

Figure 1.7. Expected Return and Portfolio Standard Deviation with the DHS

Graphing the Return Losses The above findings are also reflected in Figure 1.7, which shows the position of the household portfolios in relation to the mean-variance-efficient frontier. The upper panel shows the unclassed portfolios, while the lower panel shows the classed portfolios. Dots represent the return-risk combination of a household's portfolio. The dotted black line represents the mean-variance-efficient, currency-hedged benchmark, i.e., the hedged MSCI Europe index. The vertical distance to the dotted line defines the return loss. The lower panel reveals that households rarely diversify their portfolios between classes. Shares and equity funds dominate households' portfolios, indicated by the grey square and the black triangle. Further, households seem to invest into two asset classes, which results mostly in a "savings account plus one"-strategy: they split their financial wealth between a savings account and one other asset class. Approximately sixty percent of the Dutch households in the analysis by Von Gaudecker (2015) follow this strategy.[18] Households do not hold a large risky portfolio share, as we can determine from the bulk of observations clustering close to the safe return with very low standard deviations.

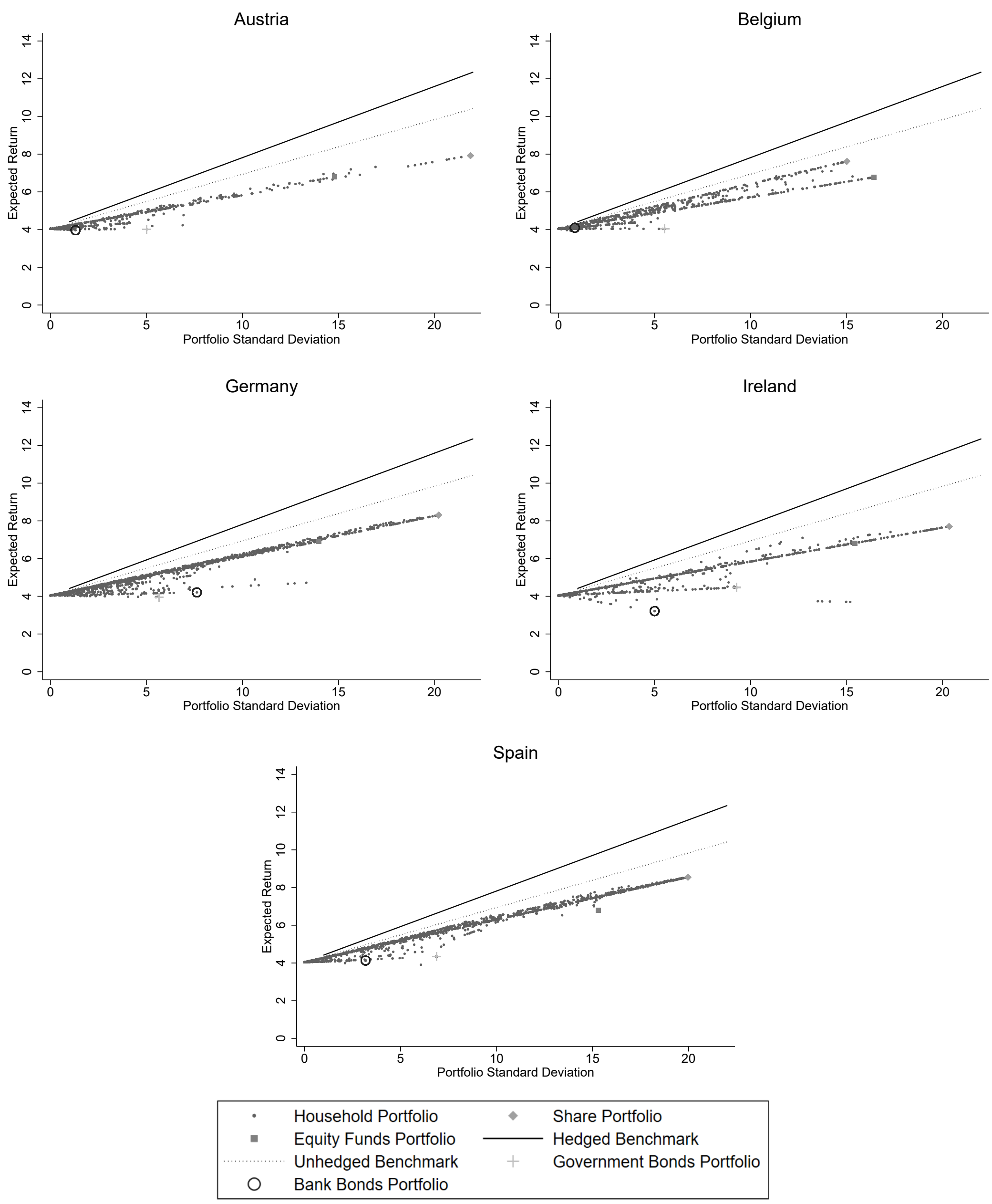
1.4.2 Classed GCAPM in Europe

We apply the classed GCAPM to the second wave of the HFCS (reference year 2014). Figure 1.8 shows the position of private household portfolios based on the classed approach in Austria, Belgium, Germany, Ireland, and Spain, analogous to Figure 1.7.[19] The black line represents the currency-hedged benchmark (MSCI Europe) and the dotted line is the unhedged equivalent. For clarity, we show just four out of the twelve class-specific portfolios in the figure.

Variation between portfolios is driven by the different return-risk combinations of the classed portfolios and the household-specific weights corresponding to the wealth share in a class. As a result, most portfolios are located on one of four lines: The first two are directly below the unhedged benchmark frontier, spanning between the origin and the gray diamond or the gray square, respectively, which mark the stock and the equity funds portfolios. These lines are relatively similar in all observed countries. Only in Belgium is the share portfolio visibly above the equity funds portfolio, meaning that the stocks bare less idiosyncratic risk than equity funds. The third and fourth lie horizontal to the origin and, in some countries, show a negative slope. They are indicated by the gray cross, which marks the government bonds portfolio and the black circle, which marks the bank bonds portfolio.

[18] Further, 29 percent invested in a savings account plus two other asset classes; while 11 percent invested in a savings account and more than two other asset classes.

[19] For a direct comparison of these graphs, one has to keep in mind that the HFCS dataset provides twelve different asset classes, whereas the Dutch assets could only be aggregated to seven different asset classes. Moreover, there are eight years between the DHS survey and the second wave of the HFCS.



Note. Authors' calculations based on Thomson-Reuters Datastream and the working sample constructed from the HFCS. The graphs provide the expected return and the portfolio standard deviation of the class-based portfolio of private households. Selected classes are presented in the graphs. The hedged and unhedged benchmarks are calculated using the MSCI Europe Index.

Figure 1.8. Expected Return and Portfolio Standard Deviation with the HFCS

Most household portfolios are located on one of these lines because the households follow a "savings account plus one"-strategy. This is less predominant compared to the Dutch analysis above as we now include twelve asset classes instead of seven. However, about 73% of the working sample invest in savings account, of which 52% invest in only one more asset class. This indicates a low degree of between-class diversification. Furthermore, many households bear unsystematic risk, even compared to the unhedged benchmark, and many hold a small risky portfolio share. Hence, we confirm the finding of households being *down* and close to *out*, as in Calvet et al. (2007a). However, the results should be interpreted with caution. Considering stocks, for example, households certainly hold a variety of stocks in their portfolio. Thus, they would provide more varied risk-return combinations than indicated in our figure. So it is not clear that all of the households shown as such are really *down*, but on average they are.

These figures also show the differences of investment opportunities in the respective countries. None of the stocks indices outperform the unhedged benchmark. Following a home-biased investment strategy, therefore, leads to inefficient household portfolios.

1.4.3 Regression Analysis

In this subsection, we provide a regression analysis of European households' return losses. We regress log return losses and their components on several socio–economic characteristics as outlined in Equation 1.10.

Many characteristics refer to the household head, who is defined as the household member with the largest gross income. $\mathbf{X}_h$ includes information on employment status, gender, the number of household members above the age of 16, and the age of the household head. Moreover, we include a dummy-set for the household head's level of education. A medium level of education is coded when the household head holds an upper secondary or post-secondary degree. A high level of education follows when the first stage of tertiary education has been achieved. The financial characteristics consist of the log of gross financial and real assets. Real asset wealth captures the gross wealth held in private business or real estate. As an addition to the literature, we include information on self-reported investment attitudes. The HFCS asks the household head to define their level of risk-taking. They can choose between 'no risk,' 'average,' 'above average,' and 'high.' We also include a set of country fixed effects.

Table 1.5 contains the results. Starting with the financial characteristics, we find that European households that hold 1% more financial wealth have 0.205% lower return losses. The total effect is induced by the decreasing risky portfolio share and a lower household beta coefficient. The diversification loss also decreases in financial wealth. More wealth held in real assets is associated with a higher return loss; yet, the coefficient is a little more than half the size of the financial wealth coefficient.

Table 1.5. Regression of Return Losses

	(1) RL_h^c	(2) w_h	(3) $\left\|\beta_h^c\right\|$	(4) $\left\|DL_h^c\right\|$
Financial Characteristics				
Gross Wealth Financial Assets (LOG)	-0.205***	-0.101***	-0.083***	-0.021**
	(0.026)	(0.018)	(0.013)	(0.001)
Gross Wealth Real Assets (LOG)	0.109***	0.085***	0.048***	-0.020*
	(0.029)	(0.021)	(0.015)	(0.011)
Montly Gross Equivalent Income (LOG)	-0.022	0.014	-0.016	-0.020
	(0.053)	(0.034)	(0.027)	(0.017)
No Risk Attitude	-0.584***	-0.512***	-0.292***	0.204***
	(0.076)	(0.052)	(0.039)	(0.030)
Above Average Risk Attitude	0.703***	0.578***	0.333***	-0.194***
	(0.104)	(0.078)	(0.053)	(0.040)
High Risk Attitude	0.761***	0.659***	0.332***	-0.193***
	(0.248)	(0.150)	(0.123)	(0.072)
Dummy Countries				
Austria	-0.279**	-0.938***	-0.143**	0.653***
	(0.119)	(0.103)	(0.064)	(0.066)
Belgium	0.001	-0.495***	0.119*	0.306***
	(0.120)	(0.103)	(0.061)	(0.066)
Spain	1.029***	0.874***	0.564***	-0.392***
	(0.094)	(0.063)	(0.048)	(0.034)
Ireland	-0.353**	-0.335***	-0.370***	0.329***
	(0.137)	(0.080)	(0.070)	(0.037)
Demographic Characteristics				
Age HH Head under 30	0.275	0.215	0.147	-0.036
	(0.261)	(0.183)	(0.133)	(0.103)
Age HH Head 30-39	0.008	-0.020	0.016	0.029
	(0.160)	(0.101)	(0.080)	(0.049)
Age HH Head 50-59	0.315**	0.062	0.171***	0.084**
	(0.122)	(0.078)	(0.062)	(0.042)

Continued on next page

Table 1.5 – continued from previous page

	(1) RL_h^c	(2) w_h	(3) $\left\|\beta_h^c\right\|$	(4) $\left\|DL_h^c\right\|$
Age HH Head 60-69	0.321**	0.113	0.145**	0.045
	(0.144)	(0.089)	(0.073)	(0.046)
Age HH Head 70-79	0.650***	0.245**	0.305***	0.087
	(0.168)	(0.107)	(0.085)	(0.059)
Age HH Head above80	0.761***	0.262**	0.325***	0.144**
	(0.179)	(0.117)	(0.091)	(0.066)
Female HH Head	-0.050	-0.131**	-0.002	0.081***
	(0.086)	(0.056)	(0.044)	(0.031)
Numb. HH Members Over 16	-0.037	0.006	-0.021	-0.018
	(0.045)	(0.030)	(0.023)	(0.016)
HH Head Employment Dummy	-0.256*	-0.176**	-0.146**	0.053
	(0.136)	(0.083)	(0.069)	(0.045)
HH Head Self-Employment Dummy	0.070	-0.001	0.020	0.042
	(0.122)	(0.080)	(0.063)	(0.043)
HH Head Retired	-0.108	-0.122	-0.051	0.071
	(0.131)	(0.081)	(0.065)	(0.045)
Medium Educated HH Head	0.369***	0.235**	0.188***	-0.046
	(0.131)	(0.092)	(0.067)	(0.049)
Highly Educated HH Head	0.598***	0.432***	0.305***	-0.123***
	(0.123)	(0.083)	(0.062)	(0.044)
Constant	-0.471	-2.481***	-0.726***	0.763***
	(0.519)	(0.345)	(0.264)	(0.191)
Observations	5,521	5,521	5,521	5,521
Adjusted R^2	0.084	0.103	0.155	0.154

Note. Authors' calculations based on the HFCS working sample. Four OLS regressions as defined in eq. 1.10. The final three coefficients in a given row add up to the coefficient on return loss, except those of the constant. All financial variables are PPP adjusted, using the OECD PPP statistics from 2014, as recommended by Brandolini (2007) and Davies et al. (2010). The income variable is the household's equivalent income constructed using the OECD-modified scale (Hagenaars et al., 1994). All results are estimated using all implicates of the multiply imputed dataset. Robust standard errors are shown in parentheses. Significance levels: *p < 0.10, ** p < 0.05, *** p < 0.01.

The variables on individual investment attitudes have to be interpreted in relation to the 'average risk-taker' category. Our estimates reveal significant results at the 1%-level for all categories. Investors declaring a 'no-risk' strategy have 0.584% less return losses than households that follow an average risk strategy. Investors reporting to invest above average risk have a return loss that 0.703% higher compared to the average risk-taker. Households that declare a high risk attitude incur the highest return losses with 0.761% above the average risk category.

The country-specific variables show some heterogeneity. The base category are German households. Austrian households have 0.279% lower return losses compared to German households, which is due to far less risky investments and smaller household betas. However, the coefficient on diversification loss is 0.653, which indicates lower diversification between asset classes for Austrians. The return losses of Belgian households differ neither economically nor statistically from German households. However, they are less engaged in risky markets but also, they incur a significantly higher diversification loss. In the end, the two effects offset each other. The household beta coefficient is slightly positive and significant. In Appendix 1.8.5, we estimate a fractional probit model to analyze the differences in the portfolio choice using the same explanatory variables as above. This provides further insights regarding the results above.[20] Spanish households realize significantly higher return losses. This is, Table 1.5 shows, due to a higher risky portfolio share and a higher household beta. The lower diversification loss reduces the overall return loss by 0.392%. This is likely to be driven by the successful oversampling of the Spanish wealthy households, as our samples selects households depending on their risky financial investments. Under the assumption that the Spanish dataset includes a relatively high number of wealthy households, this indicates that wealthier households invest in a more sophisticated manner. Irish households' portfolios show a lower household beta, which accounts, to a large extent, for a lower return loss. Their risky portfolio share is also lower, yet their diversification loss higher than German households.

With respect to demographic characteristics, we find significant associations for the age of the household head, gender, the employment status, and the level of education. We include six age dummies, with 40 to 49 year-olds being the reference category. The return loss increases with age, as do household betas and the risky portfolio share. The higher risky portfolio share appears to be contradictory to the precepts of portfolio-choice theory, which predicts high-risk portfolio shares at young ages and the opposite close to retirement (Merton, 1975).[21] However, we cannot separate age from cohort effects. Further, Guiso and Paiella (2008) argue that with concave risk tolerance the age-risk portfolio profile is upward sloping. For households with a female head, the effect on the return loss is not significant and close to zero. Employed household heads seem to choose less risky asset classes

[20] In this case, the higher beta can be explained by the fact that Belgian households are more likely to invest into equity funds, which performed relatively well in our observation period, as seen in Figure 1.8.

[21] This is due to the decreasing relative importance of human capital relative to total assets. Calvet et al. (2007a) discuss the issue of portfolio choice and age in their online appendix. See Fagereng et al. (2017b) for an additional discussion.

compared to their unemployed counterparts. The level of education has a significant positive effect on the return loss due to larger risky portfolio shares and higher betas. Households with a high level of education, however, incur lower diversification losses, which is evidence that the highly educated diversify more.

Discussion Compared to the results above, Calvet et al. (2007a) find that more financially sophisticated households, i.e., households with higher education and more financial wealth, hold larger fractions of risky portfolio share, thus incurring higher return losses from their risky portfolios. Based on their results, they theorize that less sophisticated Swedish households are aware of their limitations and invest rather conservatively producing lower return losses. Our results show that European households with higher financial wealth incur lower return losses induced by a lower risky portfolio share and a lower household beta. It appears that the switch in sign of the coefficient of financial wealth in the two studies is a result of the differences in risky portfolio shares. This is an interesting outcome as it is not driven by our assumptions.[22] Further, much like in Calvet et al. (2007a), our estimates show that a higher level of education is associated with more sophisticated investment strategies, as indicated by a higher risky portfolio share, higher betas, and a lower diversification loss. Using the level of education as a proxy for financial literacy, we provide further evidence of its importance for efficient investment choice. The real asset coefficient is positive but relatively small. This is also in line with the results in Calvet et al. (2007a). The investment attitude variables also play a role. Households seem to understand the concept of risk regarding financial investments and they seem to have an idea of their risk preference compared to others that corresponds with their portfolio choice. A unique feature of our analysis is a direct comparison of European households. It shows that they follow different investment strategies and face different financial market conditions. Nearly all measures show significant differences compared to those of German households. In the following section, we closely examine the link between the heterogeneity in return losses and financial wealth inequality.

[22] The results are based on a portfolio shift from the stocks class toward the non-risky savings account class and the less risky fund class. See the fractional probit model in appendix 1.8.5 for a more detailed discussion.

1.5 Inequality of Return Losses

In this section, we quantify the level of return loss inequality, shed some light on how it relates to the financial wealth distribution, and calculate how efficient investment would influence the financial wealth distribution. We should point out one fact at the outset: if households were efficient in their investment strategies from a neoclassical perspective, inequality in return losses would be zero, as would the return losses themselves.

Describing the Distribution of Return Losses We start by quantifying the inequality in return losses across countries. We choose the Gini coefficient and the coefficient of variation to summarize inequality.[23] Heterogeneity in the inequality of return losses across countries is present but not large. Figure 1.9 shows a Gini coefficient in the range of 0.47 to 0.61. Spain has the smallest Gini (0.47) and the smallest CoV (0.84).[24] Belgium, Germany, and Ireland have very similar Gini coefficients and CoVs. Austria has the highest inequality across both measures.

The coefficient of variation and the Gini coefficient rank the countries similarly. Overall, inequality in return losses is rather large.[25] Reassuringly, the Gini coefficients are in a similar range as the Gini coefficient found in our replication exercise (see Table 1.3 above).

To gain more insight into the overall distribution of return losses, we calculate generalized Lorenz curves. The generalized Lorenz curves reveal to which parts of the distribution accumulated return losses accrue.

Figure 1.10 displays the generalized Lorenz curve for each country. Except for Spain and Ireland, the curves look fairly similar. All curves show a flat region until about the second decile. The reason is certainly the large fraction of the population that chooses to hold particularly safe assets, i.e., savings accounts. Most of the mass of the distribution of return losses lies past the eighth decile. Return losses become

---

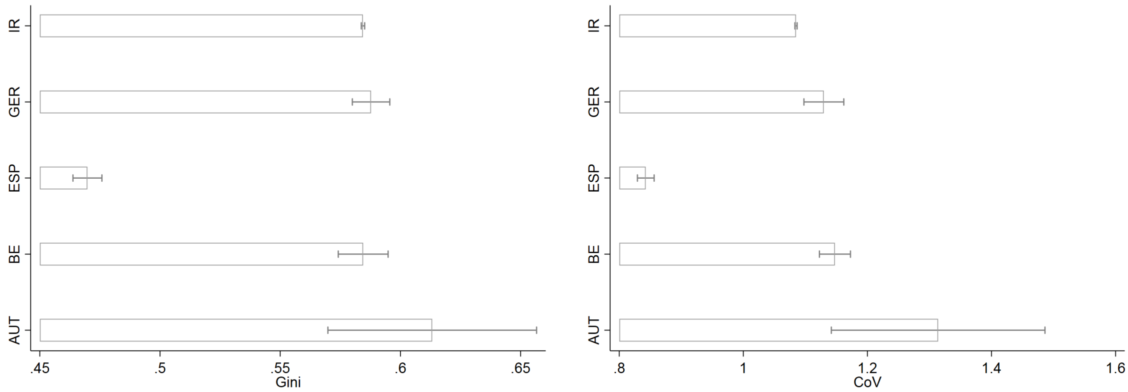
[23] The Gini coefficient is defined as

$$G = \frac{1}{2H\sum_{k=1}^{H} RL_k} \sum_{k=1}^{H}\sum_{j=1}^{H} \left|RL_k - RL_j\right|,$$

where RL is the return loss of a given household and H is the total number of households in the distribution. The Gini coefficient is 1 if the distribution is maximally unequal and 0 at complete equality. The coefficient of variation is the standard deviation of return losses divided by the mean.

[24] As before, the result for Spain must be cautiously interpreted: the Spanish data are oversampled for rich individuals more effectively than in other countries.

[25] In appendix 1.8.6, we decompose the Gini coefficient of return losses by two sources. We can define one as stemming from differences between the international and the national asset market and another one as stemming from the lack of investment efficiency within the country. We find that the Gini coefficient is largely driven by the first component.



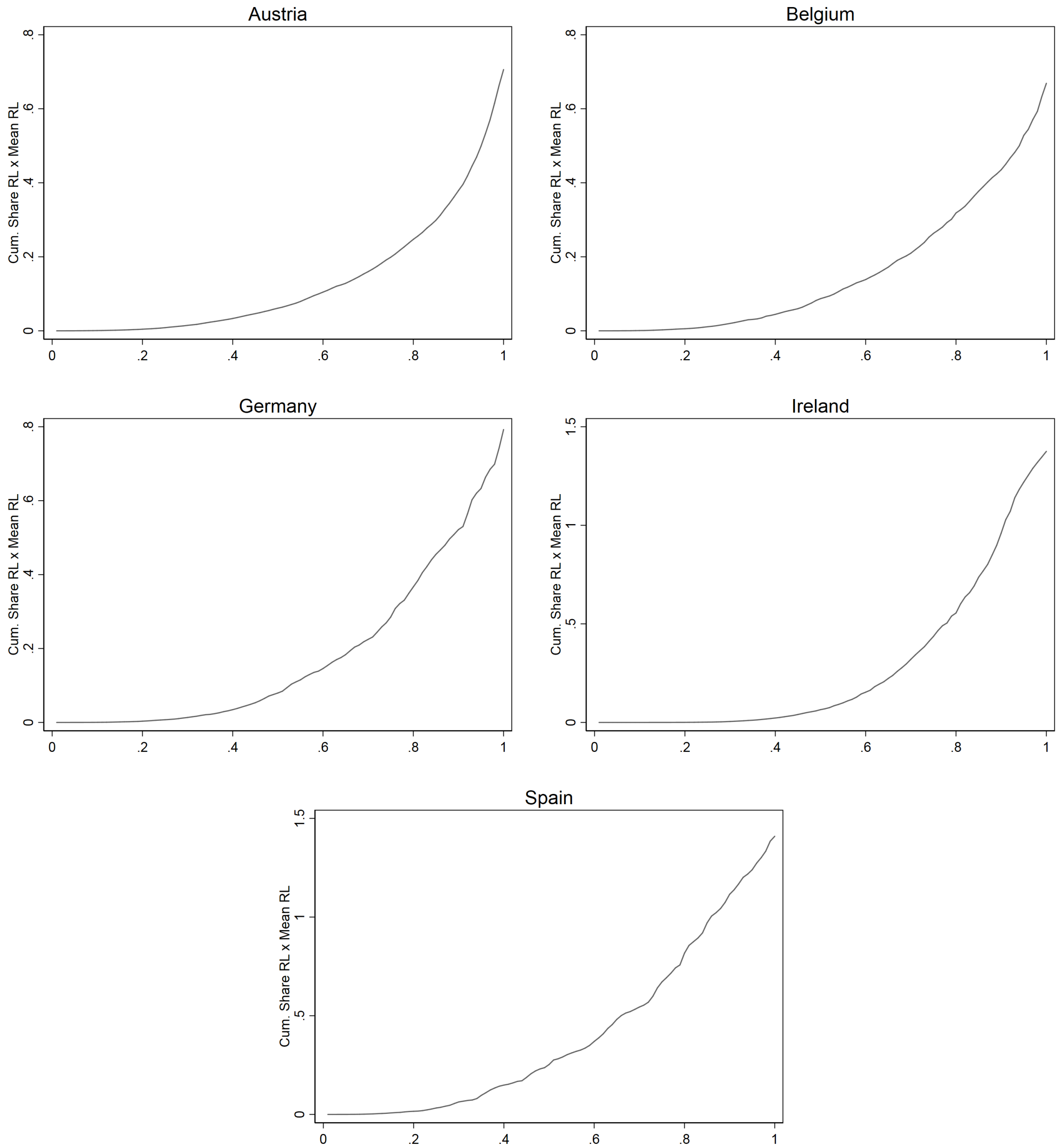
Note. Authors' calculations using the HFCS working sample. The figure plots the point estimates and the 95% confidence intervals of the Gini coefficients and the coefficients of variation for all countries. The confidence intervals are bootstrapped using the replicate weights of the HFCS and are adjusted for multiple imputation of the data.

Figure 1.9. Inequality Measures Across Countries

considerable after that point: mean values at the eighth decile range between 0.2 to about 0.7 percentage points of return lost compared to the benchmark. Although the Spanish curve indicates a distribution in which return losses are more equally distributed, we can state the following: return losses in all countries are fairly unequally distributed with a large portion of the population having close to zero return loss. This is either due to the very low-risk portfolios they choose or due to very efficient diversification; most likely the former. Conversely, a few households accrue most of the return losses in the distribution.

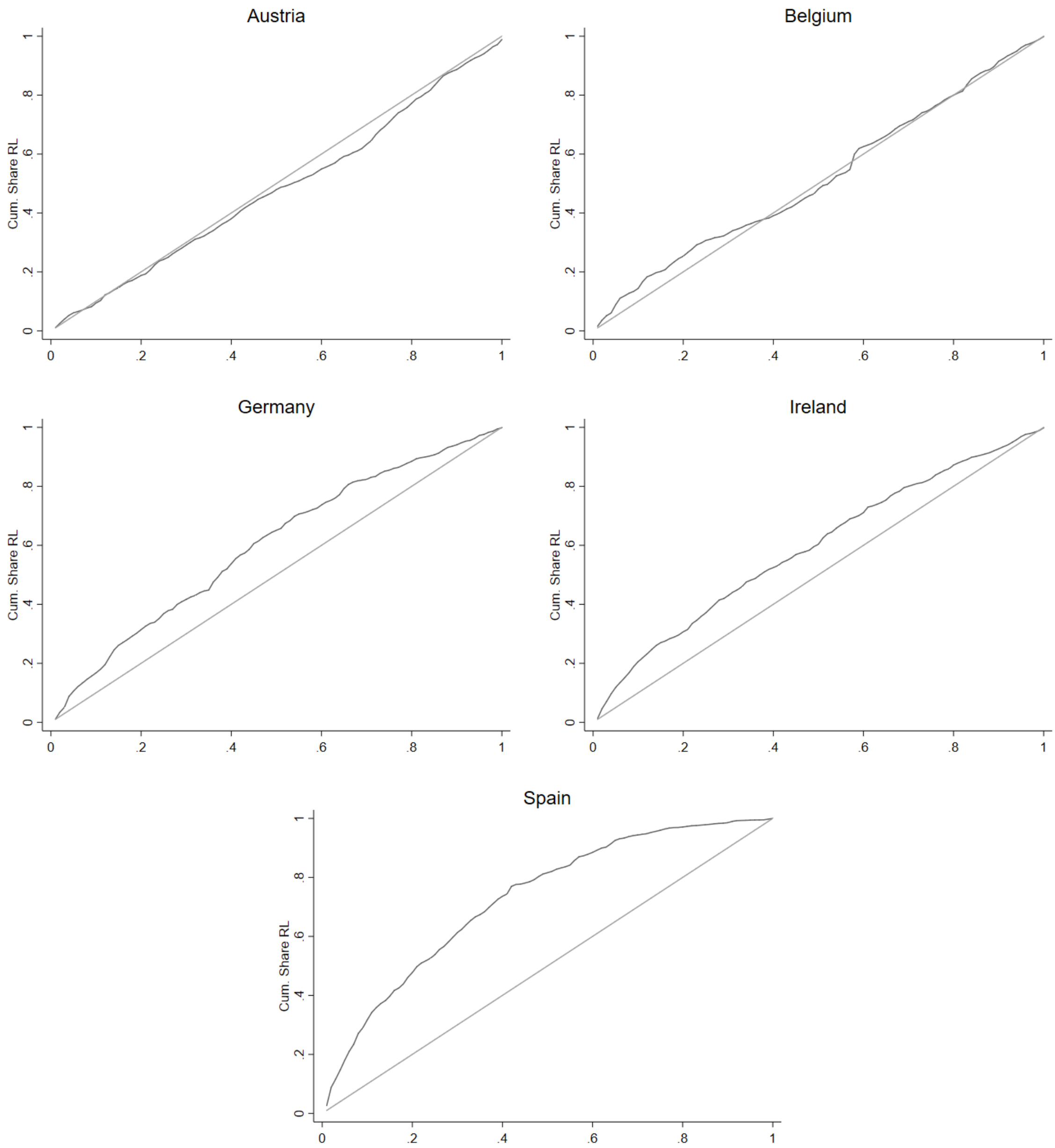
Inequality along the Financial Wealth Distribution Who are the individuals exposed to these grave return losses? Are those at the bottom of the financial wealth distribution hardest hit or is it those at the top? To address these questions. we compute the concentration curve of return losses ordered by financial wealth. Figure 1.11 displays these concentration curves. There is considerable heterogeneity across countries in the distribution of return losses along financial wealth. Spain, again, is an outlier, having the most unequal distribution with roughly 50% of the return losses accruing to the poorest 20% of households. Germany and Ireland are not that different: in both, the poorest 40% of households hold roughly 50% of all return losses. In Austria and Belgium the concentration curves indicate very little inequality. We can conclude that wealthier households fare better than their poorer counterparts across several countries and that this is very likely due to the smaller degree of risk that richer households are willing to bear.

The above analysis can only be performed in our working sample, i.e., those households with a positive risky portfolio share. Further, it does not provide us



Note. Authors' calculations using the HFCS working sample. The figure plots the generalized Lorenz curves of the return losses. We collapsed the curves over the separate imputations by calculating 100 quantiles of the ranking variable in each imputation and then calculating the mean of the generalized Lorenz curve in that quantile. Then we mean the values across all five imputations of the data.

Figure 1.10. Generalized Lorenz Curves of Return Losses Curves Across Countries



Note. Authors' calculations using the HFCS working sample. The figure plots the concentration curves of the return losses along the distribution of financial wealth. We collapsed the curves over the separate imputations by calculating 100 quantiles of the ranking variable in each imputation and then calculating the mean of the concentration curve in that quantile. Then we mean the values across all five imputations of the data.

Figure 1.11. Concentration Curves of Return Losses Sorted by Financial Wealth Across Countries

with information about the returns themselves. In the end, expected returns will determine how wealth inequality changes over time. Therefore, we use all observations with at least some financial wealth to compute the gross return factor $r_h^c = \mu_h^c w_h + r_f(1 - w_h)$ and plot its mean along twenty quantiles of financial wealth. Most of the sample holds no risky assets, so that they are assigned the risk-free rate as the gross return factor on their portfolio. Figure 1.12 displays the results.

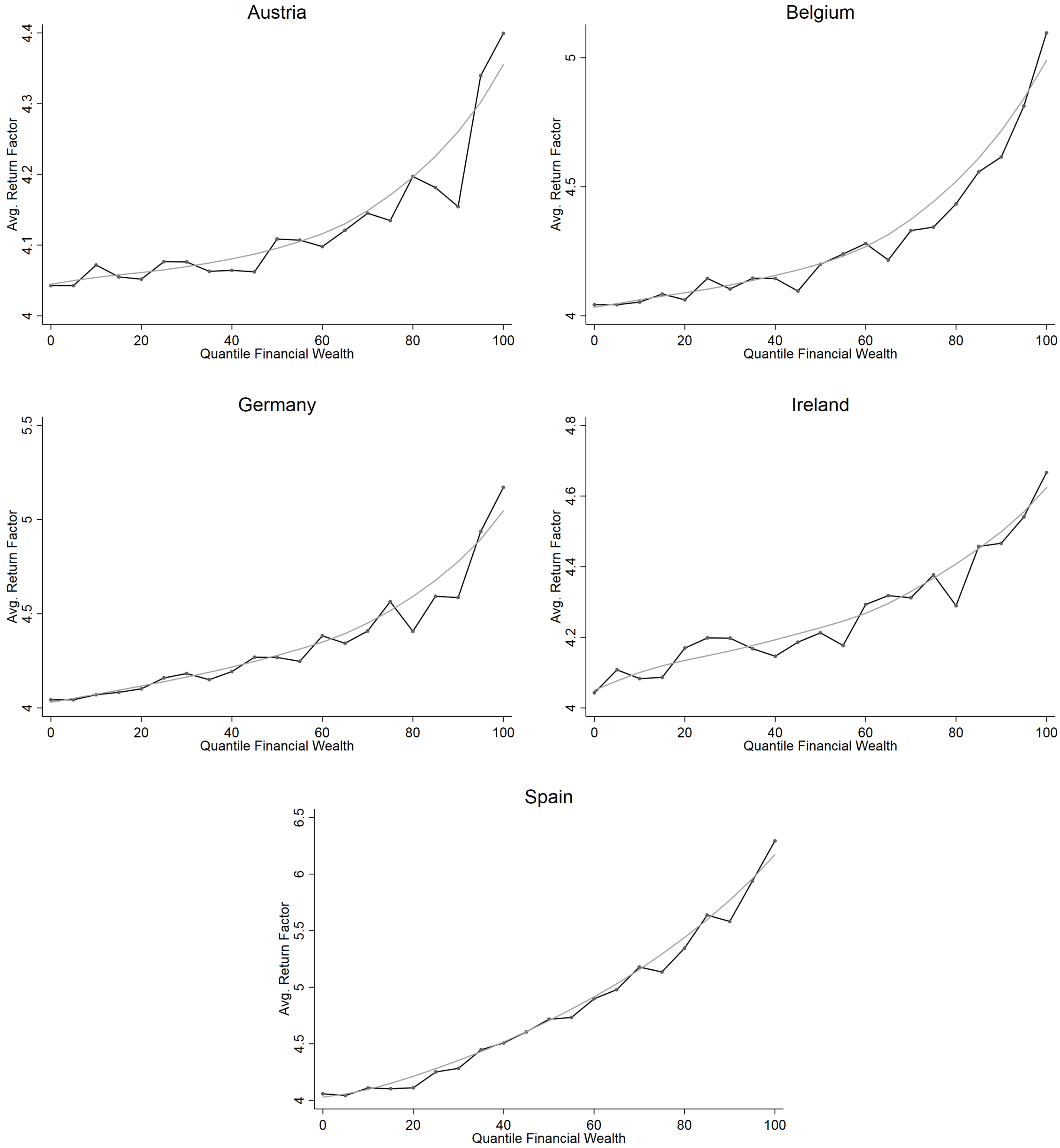
In all countries, the return factor rises with the quantiles of financial wealth; however, the slopes are quite different. Spain has the strongest rise over the quantiles with an average return at the bottom of about 4%, essentially the risk-free rate, to about 6.2% at the top. Belgium and Germany follow right behind Spain. Their profiles are flat until the middle of the distribution, after which most of the increase in return factors takes place. The average return factor in the top quantile in these countries is slightly higher than 5%. Although Austria and Ireland have flatter profiles, even in these countries return factors rise. These results are qualitatively in line with Fagereng et al. (2016).[26]

The Effect of Efficient Investment In some countries, relevant sorting of return losses along the wealth distribution exists. In all countries, return factors rise along the wealth distribution. How important is this sorting? How much would the wealth position of the inefficient investors change if they suddenly made efficient investment choices? This is important to understand the dynamic effects of return loss inequality. As we pointed out at the outset: all else being equal, if high return investment opportunities are available only to the wealthy few, wealth inequality will grow over time. To gauge the effect of efficient investment on the distribution of wealth returns, we calculate the following two measures:

$$\begin{aligned} WR_h &= FW_h \mu_h^c w_h + FW_h r_f(1 - w_h), \qquad (1.11) \\ WR_e &= FW_h \mu_m w^* + FW_h r_f(1 - w^*), \end{aligned}$$

where WR_h is the inefficient return on the household's financial wealth FW_h we calculate from the GCAPM and WR_e is the efficient counterfactual. We calibrate the optimal risky portfolio share w^* to 0.62, which is based on Merton (1975) and his model of portfolio choice with CRRA preferences. We set the parameter of relative risk aversion to 2. We use the Sharpe ratio and expected return of the benchmark portfolio to calibrate w^*. Just like when we calculated average return factors, we broaden our sample to include every household with positive financial wealth. Using WR_h and WR_e, we can calculate generalized concentration curves sorted by financial wealth to investigate whether 1) a change in the relative distribution of the wealth returns and 2) a change in the average wealth return will result. In Figure 1.13, we

[26]Fagereng et al. (2016) construct realized returns instead of using an asset pricing model. Hence, they do not need to model what rate of return assets earn.



Note. Authors' calculations using the HFCS. The figure plots the mean of the gross return along the distribution of financial wealth. We collapsed the gross return along 20 quantiles of the ranking variable in each imputation. Then we mean the values across all five imputations of the data. Finally, we fit a LOWESS-regression through the data given by the gray curve.

Figure 1.12. Association of Gross Return and Financial Wealth Across Countries

display the results. In each panel we show three generalized concentration curves: the black line at the bottom shows the distribution of WR_h, the dashed line shows the distribution of $WR_h \times \frac{\bar{WR}_e}{\bar{WR}_h}$,[27] and the gray line at the top shows the distribution of WR_e, our counterfactual. The difference between the dashed line and the gray line allows us to assess whether relative inequality has changed.

Except for Spain, the figures show a uniform pattern: First, all countries make efficiency gains, i.e., the mean of the wealth return grows considerably. Second, the change in relative inequality, shown by the change from the dashed to the gray line, is small. The change is slightly progressive, but mostly affects the upper tail of the distribution. Across all countries, most households in the bottom 40% of the financial wealth distribution neither benefit from the increase in efficiency nor does their relative position in the distribution change. That the WR_e-distribution Lorenz-dominates the WR_h-distribution is unsurprising, since the counterfactual describes a situation where everyone improves their investment behavior at no cost to anyone else.[28] Our counterfactual equalizes the relative return to wealth but not the initial distribution. Therefore, any residual inequality in the distribution of wealth returns of the counterfactual results from the inequality in the financial wealth distribution. As we have seen, the distribution hardly moves in terms of the relative ordering of households, so that we can conclude that the inequality in wealth returns mostly mirrors the inequality in financial wealth. In appendix 1.8.7 we compare the progressive effect of letting the efficient return factor compound for ten years instead of just one. It turns out that the effect is still small. From this, we can draw the conclusion that policies aiming to improve the investment strategies of European households will likely raise efficiency, however, the distributional gains are predominant in the upper tail of the financial wealth distribution.

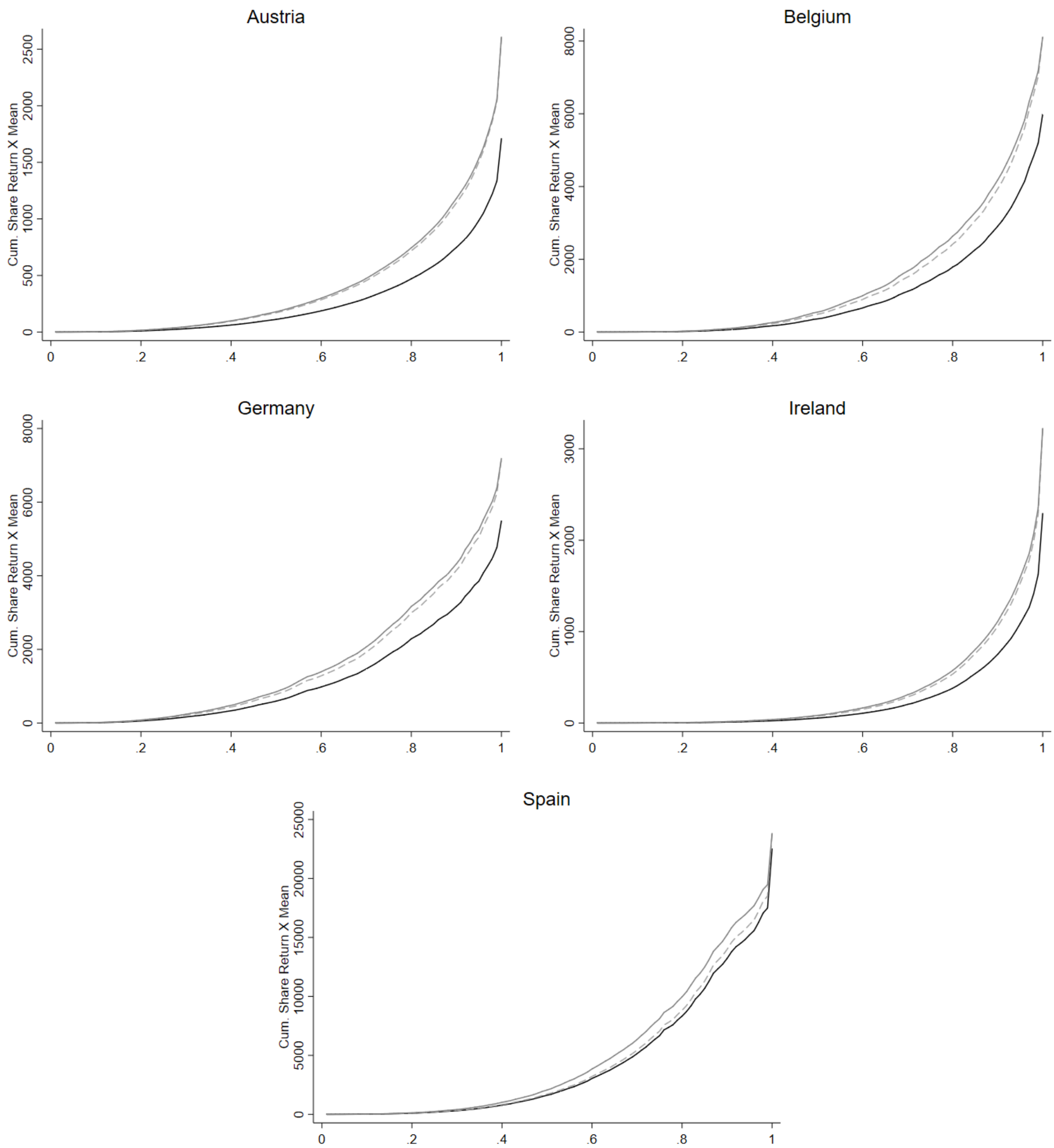
1.6 Qualification and Extensions

This section qualifies the results of this chapter by discussing assumptions, data restrictions, and their implications for our results.

We require several theoretical strong assumptions for our analysis. Aside from the neo-classical framework, i.e., rational, risk-averse individuals maximize their utility from consumption given a budget constraint, the CAPM, and the GCAPM, assume a broad set of additional assumptions: individuals are price-takers, investment products are generally available and are discretionary divisible, there are neither transaction costs nor is there incomplete information (Markowitz, 1952; Sharpe,

[27] $\bar{WR}_e$ is the sample mean of the efficient wealth return and $\bar{WR}_h$ is the analogue for the inefficient wealth return.

[28] Real-world social policies trying to raise investment efficiency will have at least some fiscal costs that need to be covered by public funds. Therefore, whether these policies can cause a true Pareto improvement is unclear.



Note. Authors' calculations using the HFCS working sample. The figure plots the generalized concentration curve of the wealth return along the distribution of financial wealth. The black line shows the distribution of returns, WR_h, the dashed line shows the distribution of $WR_h \times \frac{\bar{WR}_e}{\bar{WR}_h}$, where $\bar{WR}_e$ ($\bar{WR}_h$) is the sample mean of the efficient (inefficient) wealth return, and the gray line depicts the counterfactual distribution of WR_e. We collapsed the curves over the separate imputations by calculating 100 quantiles of the ranking variable in each imputation and then calculating the mean of the concentration curve in that quantile. Then we mean the values across all five imputations of the data.

Figure 1.13. Generalized Concentration Curves of Inefficient and Efficient Wealth Returns Across Countries

1964; Lintner, 1965; Mossin, 1966; Cochrane, 2009). Moreover, the GCAPM assumes that investors can access fully integrated, international financial markets (Solnik, 1974; Sercu, 1980; Adler and Dumas, 1983).

The question arises if the efficient benchmark based on the GCAPM is feasible for an investor and, consequently, if the return losses can be reduced to zero. The empirical GCAPM, as applied in Calvet et al. (2007a), and Von Gaudecker (2015), evaluates investment performances in the long run in relation to a currency-hedged, internationally diversified benchmark portfolio, typically the MSCI World or Europe. While investment in these benchmark portfolios is possible, a perfectly currency-hedged strategy would require complete information and foresight about changes of exchange rates. Naturally, this cannot be observed in this long run perspective. A "more feasible" benchmark, the unhedged benchmark portfolio, would decrease the return loss for all households but, in absolute terms, more return loss for those choosing higher levels of risk.[29] We follow Calvet et al. (2007a), and Von Gaudecker (2015), by applying the currency-hedged benchmark. However, re-running the analysis with an unhedged benchmark would ultimately reduce the effect of making investments efficient in our counterfactual analysis and, therefore, contribute to our conclusion, that increasing investment efficiencies mitigates wealth inequality.

In comparison to Calvet et al. (2007a) and Von Gaudecker (2015), our empirical analysis imposes further assumptions to match the financial classes in the HFCS with historical returns. We either calculate the mean of all assets, available in Euro in the respective country, or include a corresponding index. A caveat in this approach is the comparability of the two procedures: due to incomplete data information in the Thomson Reuters dataset, we can only calculate unweighted mean returns - while an index typically takes market volumes into account. This means that returns of smaller funds and assets are over-represented in these calculations. Most of the returns correlate with each other, as we observe relatively dis-aggregated classes, but we cannot rule out some imprecision in these estimates. Another concern might be the 'home-bias' assumption regarding the shares and managed accounts categories. This assumption excludes the possibility in our empirical model to reach the hedged benchmark portfolio, as every national share index is outperformed in the long run. As discussed above, our method declares: first, better investment performances for individuals with under-diversified share portfolios, and second, worse investment performances for those with internationally diversified share portfolios. This inaccuracy cannot be avoided, however, our analysis of the Dutch DHS survey[30] shows that our analysis is applicable for large parts of the return, risk, and return loss distribution.

[29]This can be illustrated by Figure 1.8, as a shift in the benchmark portfolio proportionally changes the vertical difference to the household portfolios, i.e., the return loss.

[30]See Figure 1.4, and 1.5.

Our cross-country comparison is affected by different oversampling procedures in the national surveys of the HFCS. As we select our working sample over individual participation in risky financial markets, we potentially draw more individuals out of the population in countries with larger efforts in top wealth adjustments. The clearest example for this is Spain, which includes the largest number of observations in the dataset. We, hence, refer to this in the data description and in the discussion of our results. The main implications of our distributional findings are not necessarily affected, as this caveat, as we find similar patterns in all countries. The comparability of the absolute values is, however, potentially a concern. The problem remains, that we conceivably compare very different groups of investors across countries.

Finally, one may be concerned whether our measure of risk, i.e., return volatility, or our definition of efficiency is accurately represented in our approach. Generally, volatility of returns seems to capture financial market risk reasonably well, as risky investments as mutual funds, shares or even riskier derivatives can be ranked by their return volatility. Moreover, the measure is established in the literature. Nevertheless, it only provides one type of risk and leaves other risks, e.g., health or labor market risk, aside. In terms of the evaluation of the investment efficiency of private households in our analysis, background risk is neglected. Consequently, inefficiencies in our analysis can overall be seen as efficient from a household's perspective. For example, if a household only holds some corporate bonds, the investment would be considered as under-diversified with significant return losses. However, the household may have re-balanced the portfolio to react to a potential lay-off or a health shock. Our analysis does not account for these types of risks, as it focuses solely on investment risk in the cross-section. For this reason, we analyze background risk, especially labor market risk, and portfolio choice in more detail in the chapter "Wage Risk and Portfolio Choice: The (Ir)relevance of Correlated Returns".

1.7 Conclusion

In this paper, we implement a novel method to derive the return losses of European households' financial portfolios: the class-based GCAPM. Instead of relying on detailed asset information about the household portfolio, the class-based GCAPM only requires information about the number of asset classes in the household portfolio. To validate the method, we use the Dutch Household Survey to replicate the unclassed return losses computed in Von Gaudecker (2015) and compare them to return losses we compute after artificially aggregating the data to classes. The results are reassuring: the distribution of classed matches the distribution of unclassed return losses well and the errors we introduce are small and symmetric. Further, the measures of inequality we compute for classed and unclassed return losses are also very similar.

Building on this result, we apply the class-based GCAPM to the Household Finances and Consumption Survey, which has data on household portfolio classes for many European countries. We examine return losses in Austria, Belgium, Germany, Ireland, and Spain. We find that household portfolio returns generally fall short of the efficient benchmark, although many households in all countries follow a fairly low-risk investment strategy, leading to smaller return losses. We conduct a regression analysis of return losses and their components in the style of Calvet et al. (2007a). We find that return losses are smaller for those with higher financial wealth and those with a less risky investment attitude. These effects are mostly driven by a smaller risky portfolio share. We find higher return losses among the elderly and the highly educated, which is driven by a higher risky portfolio share. These results, except for the finding on financial wealth, are in line with Calvet et al. (2007a). That the results on financial wealth diverge is explained by a directly observed difference in risky portfolio shares. Return losses differ substantially between countries. Austrians and the Irish receive far smaller return losses than does the baseline, Germany. The Spanish sub-sample receives far greater return losses. These results are not just driven by the risky portfolio share, but also by the beta factor and the diversification loss. In Spain, the risky portfolio share and the beta factor are much larger, but the diversification loss is a lot smaller. Nevertheless, we cannot rule out that some of the differences stem from differences in oversampling procedures.

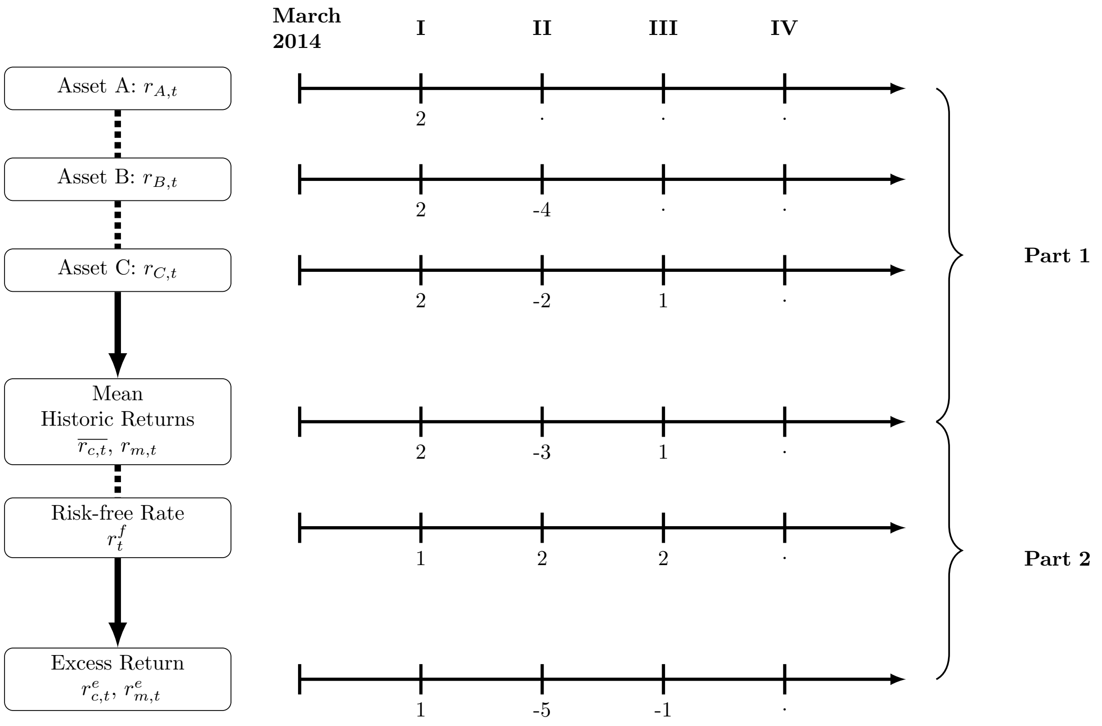
Finally, we investigate the inequality of return losses and how, as a consequence, heterogeneous rates of return influence the distribution of financial wealth. First, we see that the distribution of return losses is quite unequal for all countries, which is primarily due to the fact that only a couple of households have relevant risky portfolio shares. However, these return losses disproportionately accrue to less wealthy households. This is the distribution pattern for Germany, Ireland, and Spain, but it is not for Austria and Belgium. Thus, it seems that the status quo of household investment behavior is skewed toward more wealth inequality, since households with larger financial wealth holdings will experience smaller return losses. However, return losses are measures of portfolio performance relative to a benchmark at a given level of risk. Therefore, we investigate the distribution of return factors along the financial wealth distribution. In all countries, return factors rise over the financial wealth distribution. In the case of Spain and Germany even quite strongly. We find that inefficient investment behavior does not markedly influence the relative distribution of expected wealth return, i.e., the increase in value of a household's financial wealth. We calculate the wealth return for the status quo, where households invest as derived from the GCAPM and, for an efficient scenario, where households invest an optimal share of their wealth in the benchmark portfolio. The status-quo distribution of wealth returns is quite unequal, but switching to the efficient scenario only has a slight progressive effect. This is especially true for countries with low participation in financial markets. For countries with more participation, like Spain, the progressive effect is larger. In a dynamic perspective, this small progressive effect would become more relevant.

We show that the structure of household portfolios in Europe implies not only return losses, but also substantial differences in returns along the financial wealth distribution. The implication is that the mechanisms driving wealth inequality, as described in Benhabib et al. (2011) and Benhabib and Bisin (2018), are bound to be relevant for central and peripheral European countries as well. Our results suggests that the findings by Bach et al. (2018) also apply for other European states. In the counterfactual experiment we pursue, differences in returns are eliminated. We show a progressive effect of eliminating this heterogeneity, however, it is small, which points to the fact that wealth inequality is persistent. Thus, our paper has some implications for the effects of programs that are supposed to raise household investment efficiency. While such programs would increase total financial wealth returns, most gains are made in the upper tail of the wealth distribution. Therefore, such programs are unlikely to have an immediate effect on wealth inequality, since the distribution of wealth returns mostly mirrors the distribution of wealth.

Benhabib et al. (2011) pursue tax experiments (capital income and estate taxes) in their model of the wealth distribution with heterogeneous returns. They find that capital income taxes, which lower incentives to save by lowering the net return, reduce inequality at the top of the wealth distribution. However, they yield only moderate effects on inequality for the entire distribution. Our counterfactual experiment delivers some complementary evidence: inequality slightly drops and this drop is fairly local in the upper half of the distribution of wealth. This comparison should not be overstretched, since we do not look into the dynamic or inter-generational effects they investigate. However, both papers indicate that wealth inequality is not easily malleable.

1.8 Appendix

1.8.1 Example of Deriving Excess Returns



Note. Compiled by authors. The figure provides a schematic example how to derive excess returns for asset classes and the benchmark portfolio.

Figure 1.14. Example: Excess Returns

Figure 1.14 shows a simple example how we derive historic excess returns. We consider three types of assets in one asset class and four scenarios (I – IV) that can occur. In Part 1, we receive the historic returns $r_{A,t}$, $r_{B,t}$, and $r_{C,t}$ from Thomson Reuters and calculate the mean of the historic returns $r_{c,t}$ of the asset class c. The historic returns of the benchmark, $r_{m,t}$, can be directly drawn from the Thomson Reuters database. In Part 2, we subtract the risk-free rate from the historic returns to conduct the excess return $r_{c,t}^e$ and $r_{m,t}^e$, respectively. These are used in the fundamental CAPM regression 1.3 in the main article.

1.8.2 Further Descriptive Statistics

Table 1.6 compares several means of the full HFCS dataset with those of the working sample (WS) used in our analysis. We provide a comparison for the complete datasets and for each country respectively. The non-financial variables remain relatively stable. Household heads in our sample are a little older and more household heads are married. The share of households with a migration background is smaller; the HFCS does not provide information for Ireland or Spain on this matter. Selecting our sample for those who possess risky financial assets, the means of financial variables increase heavily as we drop those households with zero values out of the sample. Nonetheless, the increase of the average net wealth and the financial wealth is relatively stable across countries. Differences occur, e.g., in Austria, for the aforementioned reasons, namely missing oversampling of the wealthy. On the contrary, the differences between the full and the working sample in the Spanish data is in line with those of the other countries. A selective dataset is a normal when analyzing household's financial portfolios. Given the descriptive results, we argue that the survey data provides a close approximation to those households holding risky assets, but cross-country comparisons should not be overstretched.

Table 1.7 provides information on the distribution of Return Losses. The first configuration is used in the main paper with the MSCI Europe and a 375 month observation period of asset returns. Using the MSCI World instead of the MSCI Europe as benchmark does not seem to change much from the distribution. Reducing the observation period of the asset returns reduces the return loss along the distribution, which can be seen in the last configuration. A main reason for this is relatively high importance of the financial crisis in 2008, where returns fell in all categories. This reduces the RL, as this is a time of high co-movement between the returns of asset classes and the benchmark index. We use the fist configuration, because it is less affected by tail-events.

1.8.3 Relative Differences in Return Losses

Figure 1.15 displays the cumulative density function of the absolute value of the relative difference between the classed and unclassed return loss variables. The figure reveals that more than half of the sample show relative errors that are below 0.5. Further, only a couple of observations have large relative errors greater than 1, roughly 10% of the sample.

Table 1.6. Descriptive Statics

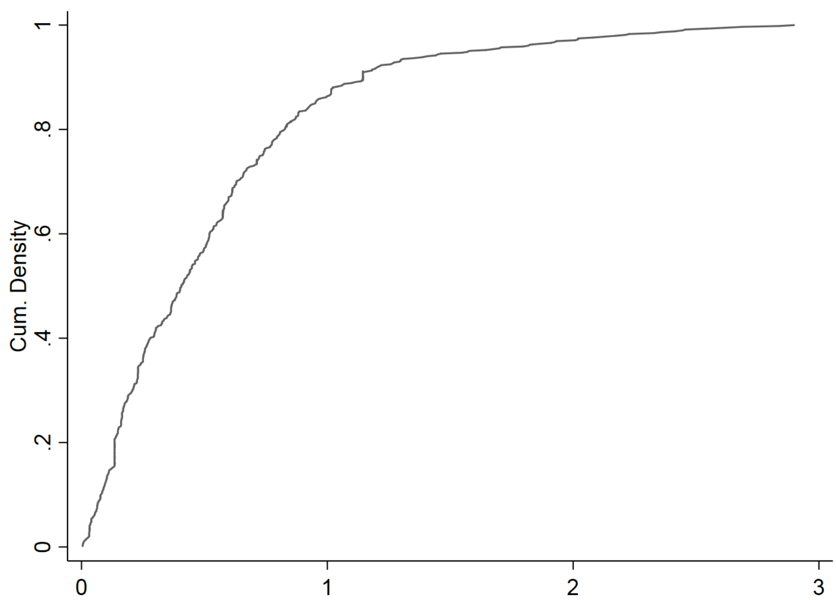
Variable	Full Complete	WS Complete	Full AUT	WS AUT	Full BE	WS BE	Full ESP	WS ESP	Full GER	WS GER	Full IR	WS IR
Average Members in HH	2.220	2.250	2.140	2.380	2.320	2.360	2.630	2.600	2.020	2.050	2.720	2.890
	(1.22)	(1.13)	(1.24)	(1.23)	(1.39)	(1.20)	(1.22)	(1.13)	(1.13)	(1.04)	(1.45)	(1.47)
Percentage of Households with Children	22.24	19.05	18.96	19.22	24.42	22.19	29.38	23.69	18.67	16.12	33.72	33.34
	(41.58)	(39.27)	(39.20)	(39.40)	(42.96)	(41.55)	(45.55)	(42.52)	(38.97)	(36.77)	(47.28)	(47.14)
Average Age Women (HH Heads)	54.83	57.48	56.05	58.67	57.37	57.80	56.02	56.51	54.33	58.11	46.88	49.50
	(18.68)	(15.27)	(18.14)	(15.30)	(18.21)	(16.28)	(17.92)	(14.22)	(18.99)	(15.29)	(17.75)	(16.20)
Average Age Men(HH Heads)	51.77	55.37	52.45	53.84	53.39	56.57	51.74	56.39	51.72	55.17	47.43	48.75
	(16.57)	(15.17)	(16.82)	(15.78)	(16.61)	(15.72)	(15.60)	(14.20)	(16.98)	(15.33)	(15.86)	(14.75)
Percentage of HH married	49.41	59.43	49.46	61.38	43.89	55.41	56.43	65.86	46.85	57.26	52.47	67.64
	(50.00)	(49.10)	(50.00)	(48.69)	(49.63)	(49.71)	(49.58)	(47.42)	(49.90)	(49.47)	(49.94)	(46.79)
Percentage with migration background	17.92	10.32	12.98	10.98	16.16	9.030	0	0	18.61	10.49	0	0
	(38.35)	(30.42)	(33.60)	(31.27)	(36.81)	(28.66)	(0.00)	(0.00)	(38.92)	(30.65)	(0.00)	(0.00)
Everage monthly hh income	2,301	3,724	2,446	3,488	2,680	3,349	1,383	2,465	2,624	4,245	2,745	4,579
	(2,595)	(4,229)	(1,466)	(2,260)	(1,953)	(2,126)	(1,686)	(3,171)	(2,948)	(4,798)	(2,787)	(4,947)
Average net wealth in Euros	240,286	550,856	258,414	654,105	330,410	597,032	273,107	646,892	214,259	502,173	216,349	456,975
	(744,304)	(1,330,461)	(1,241,976)	(2,855,263)	(525,713)	(673,374)	(752,232)	(1,601,360)	(704,626)	(1,114,172)	(469,955)	(690,605)
Average financial wealth in Euros	51,059	155,511	40,182	106,905	88,243	217,237	41,799	160,470	52,187	147,600	40,717	127,739
	(228,516)	(508,375)	(90,223)	(141,997)	(201,514)	(332,734)	(361,300)	(896,216)	(155,288)	(322,823)	(175,040)	(375,454)
Average real assets wealth in Euros	241,248	415,301	293,304	604,518	320,621	425,277	245,184	460,191	221,495	378,139	280,139	491,177
	(688,986)	(1,077,963)	(1,380,796)	(2,930,779)	(467,316)	(498,720)	(510,825)	(906,336)	(691,844)	(929,094)	(569,438)	(894,918)
Observations	21,211	5,521	2,997	440	2,235	686	6,099	2,011	4,461	1,328	5,419	1,056

Note. Authors' calculations. The table shows weighted, descriptive statistics for selected variables drawn from the 2nd wave of the HFCS dataset. The "complete" dataset includes Austria, Belgium, Spain, Germany and Ireland. "full" represents the statistics with all observations. "WS" refers to the dataset used in our analysis. All financial variables are PPP adjusted, using the OECD PPP statistics from 2014, as recommended by Brandolini (2007) and Davies et al. (2010). The income variable incorporates the household's equivalent income using the OECD-modified scale (Hagenaars et al., 1994). Standard deviation is provided in parentheses.

Table 1.7. Descriptive Statics: Return Loss

Benchmark	length	Mean	Variance	p25	p50	p75	p90
MSCI Europe	375 month	0.934	1.047	0.096	0.533	1.451	2.889
MSCI World	375 month	0.998	1.263	0.099	0.541	1.526	3.239
MSCI Europe	120 month	0.816	0.864	0.081	0.461	1.307	2.256

Note. Author's calculation using the working sample. All estimates are weighted using the HFCS household weights. We calculate each statistic for each implicate and then we estimate the mean of the values across all five imputations of the data.



Note. Authors' calculation based on the DHS. We plot the cumulative density of the absolute value of the relative return loss difference. We restrict the sample to values between 0 and 2. This eliminates 13 observations.

Figure 1.15. Cumulative Density Plot of $\left|\frac{RL_h^c - RL_h}{RL_h}\right|$

1.8.4 Further Regression Analysis

In table 1.8, we provide the same OLS-Regressions on the return loss measure and its components as described in 1.10. We use the same characteristics as before only leaving out the country dummies. It is obvious, that we lose significance for many coefficients. Only the risk attitude variables and the estimates for the older age brackets are still clearly significant at the 1 percent level. The signs of the financial asset estimates point in the same direction, but they become smaller. The education variables seem to be in line with the overall countries estimation, but they are not significant. It is difficult to draw further conclusions from this regression compared to our main analysis but it raises the question as to which countries contribute more to the results of our main analysis.

Table 1.8. Regression of Return Losses in Germany (OLS)

	(1) RL_h^c	(2) w_h	(3) $\lvert\beta_h^c\rvert$	(4) $\lvert DL_h^c\rvert$
Financial Characteristics				
Gross Wealth Financial Assets (LOG)	-0.063	-0.022	-0.010	-0.032
	(0.063)	(0.043)	(0.032)	(0.025)
Gross Wealth Real Assets (LOG)	0.033	0.044	0.009	-0.015
	(0.040)	(0.030)	(0.021)	(0.020)
Montly Gross Equivalent Income (LOG)	-0.024	0.023	-0.022	-0.022
	(0.106)	(0.070)	(0.054)	(0.037)
No Risk Attitude	-0.704***	-0.560***	-0.376***	0.216***
	(0.153)	(0.095)	(0.078)	(0.057)
Above Average Risk Attitude	1.078***	0.739***	0.496***	-0.151**
	(0.194)	(0.125)	(0.098)	(0.073)
High Risk Attitude	2.052***	0.988***	0.907***	0.156
	(0.230)	(0.326)	(0.141)	(0.290)
Demographic Characteristics				
Age HH Head under 30	0.001	-0.191	-0.013	0.196
	(0.557)	(0.350)	(0.283)	(0.206)
Age HH Head 30-39	-0.032	-0.194	-0.008	0.178*
	(0.324)	(0.224)	(0.159)	(0.107)

Continued on next page

Table 1.8 – continued from previous page

	(1) RL_h^c	(2) w_h	(3) $\lvert\beta_h^c\rvert$	(4) $\lvert DL_h^c\rvert$
Age HH Head 50-59	0.516**	-0.041	0.285**	0.256***
	(0.220)	(0.136)	(0.111)	(0.068)
Age HH Head 60-69	0.327	0.051	0.166	0.083
	(0.278)	(0.174)	(0.142)	(0.096)
Age HH Head 70-79	0.836***	0.271	0.427***	0.123
	(0.315)	(0.202)	(0.161)	(0.124)
Age HH Head above80	1.405***	0.229	0.643***	0.426**
	(0.341)	(0.245)	(0.178)	(0.171)
Female HH Head	-0.087	-0.259**	-0.044	0.208***
	(0.188)	(0.116)	(0.095)	(0.068)
Numb. HH Members Over 16	-0.019	0.014	-0.024	-0.009
	(0.087)	(0.063)	(0.044)	(0.041)
HH Head Employment Dummy	0.094	0.014	0.040	0.044
	(0.203)	(0.129)	(0.103)	(0.080)
HH Head Self-Employment Dummy	0.468**	0.219*	0.240***	0.014
	(0.182)	(0.119)	(0.092)	(0.073)
HH Head Retired	0.435*	0.0601	0.232*	0.165
	(0.261)	(0.176)	(0.133)	(0.103)
Medium Educated HH Head	0.328	0.044	0.106	0.092
	(0.434)	(0.308)	(0.223)	(0.232)
Highly Educated HH Head	0.400	0.190	0.140	-0.008
	(0.424)	(0.294)	(0.218)	(0.226)
Constant	-1.657	-2.923***	-1.207**	0.548
	(1.143)	(0.724)	(0.588)	(0.443)
Observations	1,328	1,328	1,328	1,328
Adjusted R^2	0.084	0.103	0.155	0.154

Note. Authors' calculations based on the German dataset of the working sample constructed from the HFCS. Table shows four OLS Regressions, with the dependent weight variables in the top row. All gamma coefficients add um to the return loss in column, except those of the constant. All results are estimated using all implicates provided by the HFCS dataset. Robust standard errors are shown in parentheses. Significance levels: *p < 0.10, ** p < 0.05, *** p < 0.01.

In table 1.9, we run the same regression as in 1.8 for each country, focusing on the overall effects on the return loss. We see that the most significant results are provided by the Spanish dataset. This is not a surprise, as we have previously discussed the successful oversampling in this dataset. The association with higher financial wealth and lower return losses is clearly significant in Ireland and Spain. Austria and Germany do not show a significant relation, and the Belgium analysis coefficient is positive and significant the 10 percent level. The latter might be affect by the income variable, as its gamma coefficient becomes negative and significant at the one percent level. The effect of real assets on the return loss is only highly significant in Spain. The investment attitude gamma coefficients differ in terms of size and significance, but they show in the same direction. The age coefficients are mainly driven by Spain and Germany. Hence, the other countries may do not show a comparably strong cohort effect, as described in the main analysis. The education coefficients do all show in the same direction but only the Spanish dataset provides significant results.

Table 1.9. Return Loss Regression for Each Country

	(1) RL AUT	(2) RL BE	(3) RL ESP	(4) RL GER	(5) RL IR
Financial Characteristics					
Gross Wealth Financial Assets (LOG)	-0.088	0.159*	-0.214***	-0.063	-0.369***
	(0.108)	(0.085)	(0.032)	(0.063)	(0.059)
Gross Wealth Real Assets (LOG)	0.074	0.084	0.177***	0.033	0.193*
	(0.071)	(0.071)	(0.043)	(0.040)	(0.107)
Montly Gross Equivalent Income (LOG)	0.277	-0.325**	-0.084	-0.024	0.035
	(0.243)	(0.143)	(0.062)	(0.106)	(0.171)
No Risk Attitude	-0.284	-1.009***	-0.512***	-0.704***	-0.408
	(0.256)	(0.186)	(0.100)	(0.153)	(0.249)
Above Average Risk Attitude	0.526*	0.323	0.572***	1.078***	0.895***
	(0.289)	(0.242)	(0.126)	(0.194)	(0.345)
High Risk Attitude	0.855	0.960***	0.366	2.052***	0.759
	(0.637)	(0.331)	(0.327)	(0.230)	(0.997)
Demographic Characteristics					
Age HH Head under 30	0.367	0.550	0.280	0.001	0.279
	(0.461)	(0.995)	(0.362)	(0.557)	(0.563)
Age HH Head 30-39	0.148	-0.605	0.151	-0.032	-0.084
	(0.397)	(0.467)	(0.275)	(0.324)	(0.321)

Continued on next page

Table 1.9 – continued from previous page

	(1) RL AUT	(2) RL BE	(3) RL ESP	(4) RL GER	(5) RL IR
Age HH Head 50-59	-0.046	0.025	0.625***	0.516**	-0.103
	(0.400)	(0.325)	(0.179)	(0.220)	(0.344)
Age HH Head 60-69	-0.128	0.330	0.518***	0.327	-0.272
	(0.496)	(0.476)	(0.190)	(0.278)	(0.519)
Age HH Head 70-79	0.342	0.527	0.764***	0.836***	-0.071
	(0.608)	(0.584)	(0.213)	(0.315)	(0.713)
Age HH Head above80	0.661	0.791	0.830***	1.405***	-0.456
	(0.559)	(0.553)	(0.239)	(0.341)	(0.689)
Female HH Head	-0.084	0.238	0.077	-0.087	-0.307
	(0.252)	(0.195)	(0.106)	(0.188)	(0.252)
Numb. HH Members Over 16	-0.208	-0.300**	-0.010	-0.019	0.039
	(0.140)	(0.125)	(0.047)	(0.087)	(0.136)
HH Head Employment Dummy	-0.031	0.132	-0.400**	0.094	-0.640
	(0.447)	(0.405)	(0.182)	(0.203)	(0.447)
HH Head Self-Employment Dummy	-0.186	0.120	-0.144	0.468**	-0.351
	(0.439)	(0.388)	(0.177)	(0.182)	(0.381)
HH Head Retired	0.406	0.0366	-0.360*	0.435*	-0.438
	(0.523)	(0.314)	(0.184)	(0.261)	(0.469)
Medium Educated HH Head	0.226	0.326	0.308*	0.328	0.007
	(0.461)	(0.286)	(0.175)	(0.434)	(0.454)
Highly Educated HH Head	0.549	0.311	0.652***	0.400	0.191
	(0.476)	(0.254)	(0.144)	(0.424)	(0.452)
Constant	-4.432**	-0.589	0.289	-1.657	0.165
	(2.130)	(1.679)	(0.626)	(1.143)	(1.886)
Observations	433	685	2,011	1,328	1,056
Adjusted R^2	0.355	0.295	0.432	0.304	0.251

Note. Authors' calculations based on the working sample constructed from the HFCS. Table shows four OLS Regressions, with the dependent Return loss variable. All financial variables are PPP adjusted, using the OECD PPP statistics from 2014, as recommended by Brandolini (2007) and Davies et al. (2010). The income variable incorporates the household's equivalent income using the OECD-modified scale (Hagenaars et al., 1994). All results are based on the five implicates of the HFCS dataset. Robust standard errors are shown in parentheses. Significance levels: *p < 0.10, ** p < 0.05, *** p < 0.01.

1.8.5 Fractional Probit Model

So far, we have only run regressions on return losses and its components. In table 1.10, we provide the results of the fractional probit model using the household h's portfolio share $\omega_{g,h}$, where the four different groups g are (1) funds, (2) bonds, (3) shares including managed accounts), and (4) saving accounts as dependent variables. We choose the fractional probit model as it is suited to analyze these fractional response variables, which are bounded between zero and one (Papke and Wooldridge, 1996). Hence, the regressions are purely statistical and do not rely here on any asset pricing model. The fractional probit model is given by

$$w_h = \mathbf{X}_h \boldsymbol{\gamma}_1 + \varepsilon_h \tag{1.12}$$

X_h includes the same characteristics as used in the main OLS regression 1.9 in the main paper, plus a constant, while ϵ_h is a mean zero error term. The portfolio weights w_h add up to one for each household. This means that an increase (decrease) of the estimated gamma coefficients necessarily corresponds to a decrease (increase) of the gamma coefficients of the other regression results. Table 1.10 provides the marginal effect at the mean.

Table 1.10. Regression of Portfolio Weights (Fractional Probit)

	(1) Funds	(2) Bonds	(3) Shares	(4) Savings Ac.
Financial Characteristics				
Gross Wealth Financial Assets (LOG)	0.067***	-0.013	-0.120***	0.073***
	(0.014)	(0.017)	(0.013)	(0.011)
Gross Wealth Real Assets (LOG)	-0.054***	-0.001	0.088***	-0.032***
	(0.012)	(0.016)	(0.013)	(0.011)
Montly Gross Equivalent Income (LOG)	-0.025	-0.017	0.016	0.002
	(0.027)	(0.031)	(0.024)	(0.021)
No Risk Attitude	-0.115***	0.137***	-0.221***	0.221***
	(0.033)	(0.044)	(0.033)	(0.028)
Above Average Risk Attitude	0.044	-0.018	0.292***	-0.286***
	(0.057)	(0.081)	(0.057)	(0.048)
High Risk Attitude	0.032	-0.003	0.349**	-0.386***
	(0.139)	(0.237)	(0.136)	(0.125)

Continued on next page

Table 1.10 – continued from previous page

	(1) Funds	(2) Bonds	(3) Shares	(4) Savings Ac.
Country Dummies				
Austria	0.113**	0.021	-0.507***	0.227***
	(0.056)	(0.082)	(0.073)	(0.049)
Belgium	0.275***	0.009	-0.246***	-0.089**
	(0.059)	(0.069)	(0.055)	(0.044)
Spain	-0.061	-0.168***	0.837***	-0.707***
	(0.043)	(0.058)	(0.043)	(0.039)
Ireland	-0.583***	0.194***	0.320***	-0.011
	(0.055)	(0.058)	(0.046)	(0.039)
Demographic Characteristics				
Age HH Head under 30	-0.054	-0.080	0.141	-0.019
	(0.111)	(0.169)	(0.116)	(0.087)
Age HH Head 30-39	-0.158**	0.097	0.074	0.019
	(0.072)	(0.090)	(0.065)	(0.053)
Age HH Head 50-59	-0.025	0.141**	0.109**	-0.135***
	(0.052)	(0.071)	(0.052)	(0.042)
Age HH Head 60-69	-0.066	0.028	0.207***	-0.159***
	(0.060)	(0.085)	(0.058)	(0.052)
Age HH Head 70-79	-0.036	0.042	0.336***	-0.299***
	(0.074)	(0.106)	(0.071)	(0.063)
Age HH Head above80	-0.063	0.090	0.365***	-0.330***
	(0.077)	(0.108)	(0.075)	(0.066)
Female HH Head	0.063*	0.105**	-0.090**	-0.008
	(0.037)	(0.047)	(0.036)	(0.031)
Numb. HH Members Over 16	-0.021	-0.023	0.010	0.014
	(0.020)	(0.024)	(0.018)	(0.017)
HH Head Employment Dummy	-0.091	0.063	-0.069	0.115**
	(0.060)	(0.076)	(0.053)	(0.049)
HH Head Self-Employment Dummy	0.027	0.009	0.054	-0.078*
	(0.057)	(0.072)	(0.051)	(0.046)

Continued on next page

Table 1.10 – continued from previous page

	(1) Funds	(2) Bonds	(3) Shares	(4) Savings Ac.
HH Head Retired	-0.080	0.248***	-0.065	0.034
	(0.056)	(0.072)	(0.052)	(0.046)
Medium Level of Educ	0.073	-0.035	0.128**	-0.187***
	(0.060)	(0.076)	(0.055)	(0.048)
High Level of Educ	0.116**	-0.102	0.189***	-0.242***
	(0.055)	(0.070)	(0.051)	(0.046)
Constant	-0.471*	-1.286***	-0.944***	-0.284
	(0.259)	(0.340)	(0.240)	(0.207)
Observations	5,521	5,521	5,521	5,521
Pseudo R^2	0.043	0.023	0.109	0.064

Note. Authors' calculations based on the working sample constructed from the HFCS. Table shows the marginal effect at the mean in four fractional probit regressions, with the dependent weight variables in the top row. All gamma coefficients add um to zero, except those of the constant. All financial variables are PPP adjusted, using the OECD PPP statistics from 2014, as recommended by Brandolini (2007) and Davies et al. (2010). The income variable incorporates the household's equivalent income using the OECD-modified scale (Hagenaars et al., 1994). All results are based on the five implicates of the HFCS dataset. Robust standard errors are shown in parentheses. Significance levels: *p < 0.10, ** p < 0.05, *** p < 0.01.

We show the regression results in Table 1.10. The first part of the table provides the coefficients of financial characteristics, the second part provides country–specific dummies, and the third part includes demographic characteristics. The financial characteristics reveal several significant effects. An increase of the financial gross wealth variable of one percent is associated with an significant increase of the probability to invest in mutual funds 6.65 percentage points at the mean, a decrease of the probability to invest in shares by 12.0 percentage points, and an increase of the probability to invest in saving accounts 7.31 percentage points. As discussed in the main paper, this leads to lower return losses for those households with higher levels of financial wealth, as they invest more into more diversified funds. They are also less engaged in risky markets.

The results of Table 1.10 also show that a higher share of gross wealth in real assets seems to be associated with relatively lower fund and savings account investments as well as a higher probability of investing into shares. The coefficients for income are small and not significant. The self declared risk attitudes correspond to the observations in the main analysis. The average–risk–attitude category is not included in the regressions; thus, all gamma coefficients are expressed in relation to this category. Households that declare investing with 'no risk' hold with significant probability of 22.1 percentage points more wealth in in secure savings account and they are 13.7 percentage points more likely to invest in relatively save bonds. In contrast, the probability investing into mutual funds and shares decreases nearly at the same extent, respectively. The estimated coefficients of the other risk categories vary into the expected direction with less wealth in saving accounts and more in the share asset class.

The country–specific dummy variables show significant differences across countries. As German households are the reference unit, Austrian households are 50.7 percentage points less likely to invest in shares but 11.3 and 22.7 percentage points more likely to invest in mutual funds and in savings accounts respectively. Belgian households seem to invest more in mutual funds and less in shares. Irish and Spanish households hold less wealth in funds than their German counterparts, but they both hold significantly more wealth in shares. Spanish household portfolios are especially remarkable, as they are 70.7 percentage points less likely to invest in savings accounts and about to 83.07 percentage points more likely to invest in shares than German households. As discussed in the main paper, this is to some extent driven by the differences of the oversampling techniques in the HFCS.

The demographic variables show that the older cohorts in our sample have a higher probability to invest in shares and a lower probability to hold wealth in saving accounts. This translates to the higher return loss in the main analyses and seems to be a cohort effect. The retirement dummy reveals that retired household heads are 24.8 percentage points more likely to invest in a bond category, compared to unemployed counterparts. This is in line with optimal portfolio choice over the lifecycle, which predicts rebalancing towards relatively safe assets, i.e., bonds, when human capital decreases and individuals face retirement (Merton, 1975). This does not translate into the regression in the main paper, as bonds perform relatively badly and incur a high return loss. The education variables show higher probability of engagement in the share and fund classes when the level of education increases. Households with a high level of education are also more likely to invest in funds showing a higher degree of between-asset-classes diversification. This is completely in line with the implication of the main analysis. All in all, these results support the results we observe in regression 1.5 and they provide some additional information about the portfolio choice in our working sample.

1.8.6 Decomposing Return Loss Inequality

In addition to the main analysis, we also attempted to understand whether the heterogeneity in the return loss distributions are driven by country specific circumstances; for example features of local asset markets or whether the population within these countries has specific features that determine the distribution of return losses.

To implement such a comparison, we will additive decompose the Gini coefficient into (1) a component that measures the return loss between a country-specific and the overall benchmark; and (2) a component that gives the return loss between the country-specific benchmark and the household. The country-specific benchmark gives the best possible performance in the country, measured by the highest Sharpe ratio in the country under consideration. Let the country-specific benchmark be defined by its Sharpe ratio S_l. Then the decomposition of the household's return loss is written as:

$$RL_h^c = \underbrace{w_h \sigma_h^c (S_m - S_l)}_{RL_{h,l}^c} + \underbrace{w_h \sigma_h^c \left(S_l - S_h^c\right)}_{RL_{h,p}^c}, \tag{1.13}$$

where the first term is the country-specific component ($RL_{h,l}$) and the second the population specific component ($RL_{h,p}$). The decomposition of the Gini coefficient in a country following Lerman and Yitzhaki (1985) takes the following form:

$$G_{RL_h^c} = R_l G_l S_l + R_p G_p S_p, \tag{1.14}$$

where R_k is the Gini correlation of component $k \in \{l, p\}$, G_k is the Gini coefficient in component k, and S_k is the ratio of component k's mean to the overall mean.[31] Table 1.11 contains the fractions of the total Gini coefficient attributable to either the

Table 1.11. Fractions of Total Gini Coefficient due to Between and Within Components

Country	Between	Within
AT	0.81	0.19
BE	0.59	0.41
ES	0.80	0.20
DE	0.82	0.18
IR	0.63	0.37

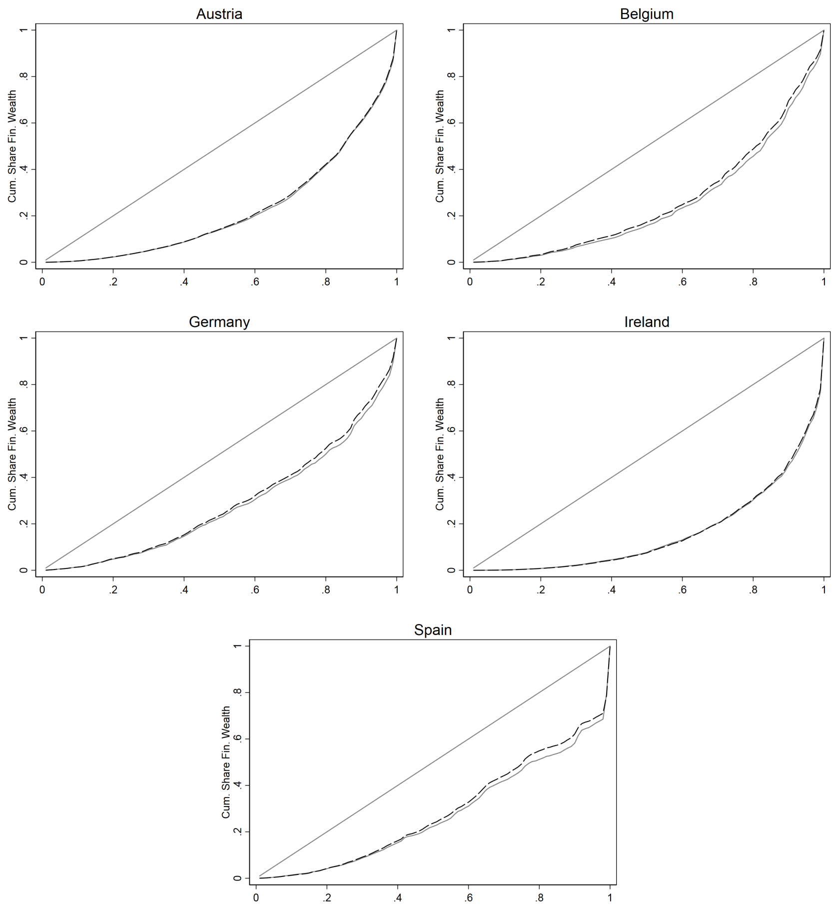
Note. Authors' calculations based on the HFCS. The table gives the fractions of the total Gini coefficient attributable to either the between and the within component as in equation (1.14).

between or within component. Across all countries most of the contribution to the Gini is made by the between component, although there is significant heterogeneity. For Austria, Germany, and Spain, about 80% of return loss inequality is attributable to the between country component. For Belgium and Ireland, it ranges around 60%. We can conclude that the majority of the inequality in each country is driven by factors determining the difference between the local and the international financial market.

[31] Let y be a variable, decomposable into some components y_k and let F denote the cumulative distribution function (CDF) of y and F_k be the CDF of component y_k. Then the Gini correlation of component k is $R_k = \frac{cov(y_k,F)}{cov(y_k,F_k)}$. The Gini coefficient in component k is $G_k = \frac{2cov(y_k,F_k)}{m_k)}$.

1.8.7 Longer Run Counterfactual Effect

Figure 1.16 displays the concentration curve of the efficient return factor in equation 1.11 compounded over 1 and over 10 years along the distribution of financial wealth. Like in Figure 1.13, the total population holding any financial wealth is included. The figure shows that the 10-year return factor (dashed line) is more equally distributed than the 1-year return factor (solid line), however, the difference is small.



Note. Authors' calculations using the HFCS working sample. The figure plots the concentration curve of the 1-year (solid) and 10-year (dashed) wealth return along the distribution of financial wealth. We collapsed the curves over the separate imputations by calculating 100 quantiles of the ranking variable in each imputation and then calculating the mean of the concentration curve in that quantile. Then we mean the values across all five imputations of the data.

Figure 1.16. Concentration Curves of 1-Year and 10-Year Efficient Wealth Returns Across Countries

2 De-routinization of Jobs and the Distribution of Earnings – Evidence from 35 Countries

2.1 Introduction

Dynamics of occupational changes in the labor force is a central topic of economic research. In particular, technological change is historically identified as a key explanation for major shifts in the workforce, through the creation and disruption of jobs.[1] Autor et al. (2003) proposed the Routine-Biased Technological Change (henceforth, RBTC) hypothesis, which relates improvements in information and communications technologies (henceforth, ICT) with de-routinization of the workforce. According to the RBTC hypothesis, the decreasing prices of technology over the last decades have exogenously driven the substitution of workers operating routine tasks by computer algorithms or machines.[2] Simultaneously, the relative demand for workers who perform complementary non-routine tasks has increased. Typical non-routine tasks include problem-solving, creativity, situational adaptability, and in-person interactions. Recent empirical literature supports the RBTC hypothesis (Acemoglu and Autor, 2011; Goos et al., 2014; De La Rica and Gortazar, 2016), finding that the increasing adaption of ICT as labor input has contributed to the de-routinization of jobs globally over the last decades.

Acemoglu and Autor (2011) empirically investigate how job de-routinization alters the distribution of skills. Because routine jobs are typically middle-skilled jobs while non-routine jobs mostly concentrate at the tails of the skill distribution, de-routinization results in job polarization: increasing employment shares of high and low skilled jobs relative to middle skilled.

The link between job de-routinization and job polarization opened the field to empirical investigation of its consequences for the wage distribution. Acemoglu and Autor (2011) and Autor and Dorn (2013) provide evidence that the RBTC framework explains overall wage polarization experienced in the US since the 1960s. The authors define wage polarization as u-shaped earnings growth along the wage distribution, which results in a reduction of bottom-half -, and an increase in top-half inequality. Following their definition, overall distributional consequences depend

[1] See (Vivarelli, 2014) for a detailed survey of the literature.

[2] Routine intense occupations include, for example, clerical work, repetitive production, and monitoring jobs.

on which of the two margins dominate.[3] Moreover, Autor and Dorn (2013) conclude form their empirical analyses that "labor specialization [...] play[s] a critical role as a driver of rising employment and wage polarization in the US and, potentially, in other countries" (p. 1591). However, this generalization is contested (Dustmann et al., 2009; Massari et al., 2014; Green and Sand, 2015; De La Rica and Gortazar, 2016; Hunt and Nunn, 2019; Taber and Roys, 2019; Böhm, 2020). This is because occupations are not systematically sorted along the wage distribution.

We recognize three major reasons for the debated nexus between job polarization and wage inequality. First, the global phenomena of de-routinization of jobs potentially has diverse distributional consequences as the number of routine and non-routine workers differs across countries. Hence, an extensive cross-country comparison can shed more light on the link between job polarization and inequality. Second, several studies focus on comparisons of average wages by occupations (Acemoglu and Autor, 2011; Autor and Dorn, 2013). Focusing on averages disregards that job polarization may also alter occupational class specific wage inequalities (Hunt and Nunn, 2019; Taber and Roys, 2019). As this paper shows, class-specific wage and earnings inequalities respond to the changing demands for different occupational classes of workers. Hence, the quantification of the nexus between job polarization and wage inequality requires a comprehensive assessment of both wage inequalities within and between occupations. Third, embedding variation within-occupations acknowledges that workers in routine and non-routine occupational classes can overlap along the wage distribution (Böhm et al., 2019; Böhm, 2020). In this sense, de-routinization of jobs does not just displace workers in middle-, but also at the bottom and at the top of the wage distribution. Consequently, one needs to account for different occupational composition and return effects along the quantiles of the wage distribution over time to understand the overall distributional effects of job de-routinization.

This paper contributes to the literature by providing a comprehensive international assessment of job de-routinization processes and their relevance for changes in hourly wages and annual labor earnings[4] inequalities within and between occupations. A novel and harmonized dataset for 35 countries, provided by the Luxembourg Income Study (LIS) and the Economic Research Forum (ERF), the so-called LIS-ERF dataset, provides the empirical base for our analysis. The LIS is the largest available income database of harmonized micro data from countries

[3]In RBTC literature, polarization does not rely on the traditionally applied concepts of identification and alienation (Esteban and Ray, 1994), rather it simply refers to differentiated u-shaped growth patterns along the wage distribution. In this sense, the wage polarization notion used in RBTC literature is strictly bi-polar, looking at the dispersion of the distribution from the middle position, and does not contemplate the possibility of multi-polar polarization, defined as the bunching of the population into any number of income subgroups clustered around local means of the income distribution (Chakravarty et al., 2015).

[4]Henceforth referred to as earnings.

around the world. Technically, we estimate the Re-centered Influence Functions (RIF) decomposition method (Firpo et al., 2009, 2011, 2018) to measure ceteris paribus effects of job de-routinization for percentiles of the country specific earnings distributions, accounting for both within and between occupational variation. Further, we characterize the RIF decomposition results in the light of changes in occupational composition and returns. Finally, we are the first to quantify the relative importance of changes in earnings inequalities within and between occupations induced by job polarization around the globe. The distinction of inequalities within and between occupations is motivated by the findings for the US by Hunt and Nunn (2019), Taber and Roys (2019), and Böhm (2020), who show that the overall distributional effect is unclear if occupational groups are scattered, and job displacements effects are not homogeneous along the wage distribution.

We show that job polarization occurs in 30 out of the 35 countries under investigation with different time frames ranging from the 1990s to the 2010s. Our results support the RBTC hypothesis as suited for explaining the observed shifts of employment shares in the workforce. In a cross-country perspective, we show that de-routinization is ambiguously linked with inequality within and between occupational groups. Moreover, variation in overall inequality mostly stems from variation within occupational groups. Applying the RIF decomposition method, country-specific earnings distributions have developed heterogeneously. In 14 (eleven) countries earnings growth rates are monotonically increasing (decreasing) over the quantile distribution, resulting in increasing (decreasing) overall inequality. Only five countries over 35 show u-shaped growth patterns along the earnings distribution following the definition of polarization adopted in Acemoglu and Autor (2011), and Autor and Dorn (2013). In five countries we find no substantial changes in inequality.

We show that the weak link between job polarization and earnings inequality is for the following reason: Overall earnings inequality is determined by inequalities between and within occupational classes. Changing the average pay of a particular occupational class will unambiguously change the between-class inequality component – as determined by differences in class-specific average earnings. However, because employees from a certain occupational class are not perfectly stratified but scattered along the earnings distribution, the implication for within-class inequalities and – ultimately - for overall inequality are ambiguous. Contrary to the RBTC framework (Acemoglu and Autor, 2011; Autor and Dorn, 2013), our results also do not support that job polarization contributed to reduce inequality at the bottom of the earnings distribution.

The paper is organized as follows: Section 2.2 provides a literature review. Section 2.3 discusses data sources and harmonization processes. Section 2.4 describes the methodology and the wave selection. Section 2.5 provides the results. Section 2.6 presents the results using hourly wages instead of yearly gross-income. Section 2.7 discusses the assumptions of our analysis. Section 2.8 concludes.

2.2 Literature Review

This section reviews the empirical literature on job polarization and its debated implications for earnings inequality.

Job polarization and its direct link to ICT adoption is extensively studied in both advanced and emerging economies. In their widely recognized work, Autor et al. (2003) find evidence of job de-routinization between the 1960s and 2000s in the US. Goos and Manning (2007), analyzing different models of labor market changes for the UK between 1975 and 1999, conclude that the RBTC hypothesis by Autor et al. (2003) works best for explaining shifts in occupational classes. Autor (2019) updates these findings, also describing an increasing divide in wages between non-college and college workers in the US. Goos et al. (2014) show de-routinization in the workforce due to ICT adaption in 16 Western European countries between 1993 and 2010. Green and Sand (2015) find similar patterns between the 1980s and 2005 in Canada and Coelli and Borland (2016) between the 1980s and 1990s in Australia. Aedo et al. (2013), analyzing eight developing countries over time, find a strong correlation between economic development and the skill intensity of non-routine cognitive, analytical, and interpersonal skills, as well as strong negative correlations with routine and non-routine manual skills. De La Rica and Gortazar (2016) focus on a set of OECD developed countries around the world and find evidence for job polarization due to ICT adaption; Hardy et al. (2018) do so for Central and Eastern Europe. Mahutga et al. (2018) describe de-routinization of jobs primarily as a phenomenon of the global north. Their analysis bases on 38 aggregated LIS countries. Even though they use the same data source, Mahutga et al. (2018) do not explore country-specific effects, a fundamental difference to our approach.

In sum, most previous research finds empirical evidence for job polarization due to ICT adaption in many countries around the world. We contribute to this strand of literature by using a harmonized dataset up to the year 2016 for 35 countries.

Several empirical studies investigate the nexus between job polarization and its distributional consequences. The evidence is mixed.

One stream of the literature finds that de-routinization due to ICT adaption implies wage polarization defined as u-shaped earnings growth along the wage distribution. In the US, Autor and Dorn (2013) show that the hourly wage of non-college workers employed in service occupations, with relatively high routine-task intensity, rose significantly between 1980 and 2005. They also find positive wage growth for all the others occupational categories characterized by low routine task intensity. Highly routinized employment experienced wage losses. The authors conclude that job de-routinization polarizes the returns to skills *between* occupational classes and can explain a substantial share of aggregated polarization. In Europe, evidence for wage polarization is provided for Germany (Dustmann et al., 2009) and the UK (Machin, 2010). Mahutga et al. (2018) state that de-routinization contributes to earnings polarization in rich democracies.

Apart from the country-specific results, the findings also depend on the time span under analysis. Focusing on the US, Firpo et al. (2011) find that technological change was skill-biased[5] in the 1980s, while it was routine-biased[6] in the 1990s. In the 2000s, they only find a modest effect. Our results extent their analysis by adding an additional decade. As this paper shows, we do not find that job de-routinization is associated with wages and earnings polarization in the 2010s. Although our results do not exclude temporary influences of ICT adaption on the earnings distribution in line with RBTC, we cannot observe a close nexus in the long run.

Several studies contest the link between de-routinization and earnings polarization. Goos and Manning (2007) do not find evidence for a relationship between de-routinization and wage inequality in the UK and raise doubts as the literature typically does not account for heterogeneous wages distributions within occupations. Green and Sand (2015) find similar results for Canada. Böhm et al. (2019), Hunt and Nunn (2019), and Taber and Roys (2019) suggest that the RBTC hypothesis is generally not suitable for studying the evolution of wages and earnings inequality, raising similar concerns as Goos and Manning (2007). Böhm et al. (2019) find skill selection effects between occupation entrants and leavers, as they earn lower wages than stayers, suggesting that wage effects are negative for growing occupations and positive for shrinking ones. This selection cannot be captured by focusing on between-occupational changes alone. According to Hunt and Nunn (2019), 86% of the increase in wage inequality in US between 1973 and 2018 stems from variation within occupations. Taber and Roys (2019) argue that labor-demand changes between occupations explain only a small part of changes of the wage distribution between 1979 and 2017 in the US, concluding that skill price changes within occupation are far more important. Massari et al. (2014) do not find wage polarization in Europe and find only weak polarizing effects of technological change, suggesting that deterioration of labor institutions, e.g., increasing part-time and temporary jobs, may play a more important role by hindering wage growth at the bottom. According to De La Rica and Gortazar (2016), differences in ICT adoption explain an important and significant part of wage differentials but have little explanatory power for wage inequalities in OECD countries. In a theoretical analysis, (Böhm, 2020) shows that job polarization leads to a polarization of task prices, which does not translate into wage polarization. He suggests that the overall distributional effect is unclear if occupational groups are scattered and job displacements effects are not homogeneous along the wage distribution.

[5] Wage growth strictly increases with skills.

[6] Wage growth was lower in the middle than at the tails of the skill distribution

Our analysis of a large set of countries captures these heterogeneous findings and sets them analytically into perspective compared to the results of Goos and Manning (2007), Böhm et al. (2019), Hunt and Nunn (2019), and Böhm (2020).

2.3 Data

Our empirical analyses rely on the LIS-ERF joint dataset, the largest available international harmonized income micro-database based on repeated cross-sections from over fifty countries. Compared to the standard LIS dataset, LIS-ERF includes additional data for seven countries: Egypt, Iraq, Jordan, Palestine, Somalia, Sudan, and Tunisia. The LIS cross-national data center acquires, harmonizes, and documents microdata from different national statistical institutions.[7] In addition to detailed income information, it includes a broad set of individual and household characteristics – including occupational and socio-demographic information of household members. Our final working sample includes 35 countries, which are selected based on two criteria:

1. Availability of repeated cross-sections: the minimum data requirement for a country to be included in the working sample is availability of at least two waves, since the empirical testing of our hypotheses requires measures of differences in earnings and employment shares over time.

2. Availability of focal variables: labor income and job information are necessary to define quantiles and occupational classes used in the analysis.

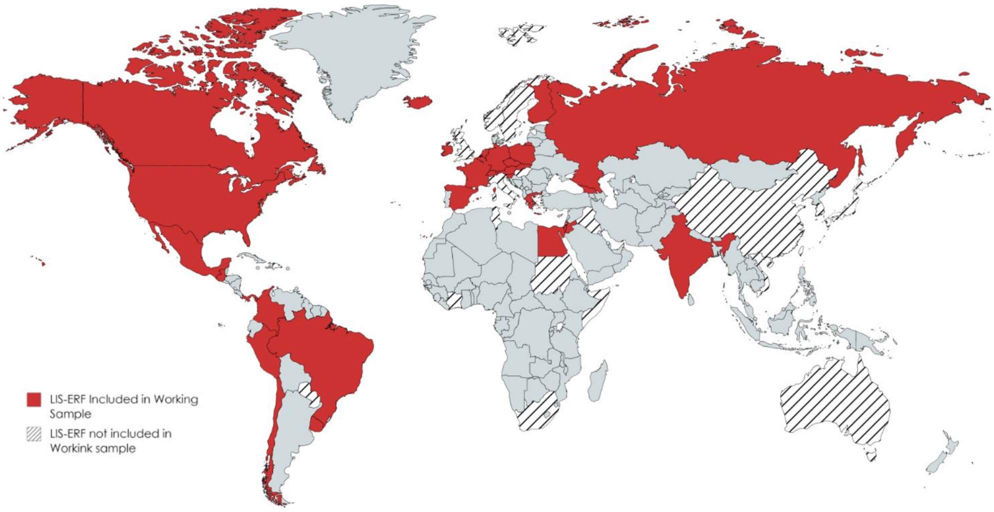
Our working sample focuses on prime-age employed individuals aged 25-55. Missing values are imputed in all LIS and ERF countries. The imputation is conducted by the individual survey institute in each country. Most countries follow a simple random sampling or a two-stage area sampling procedure. Although the imputation procedures are not completely standardized, we acknowledge a high comparability across waves and countries, as guaranteed by LIS and ERF. Top- or bottom-coding procedures do not apply.

Figure 2.1 depicts a map of the countries included in LIS-ERF and our working sample. A detailed overview of the country-specific waves compatible with our selection criteria are reported in Table 2.2.

For most of the countries, the LIS-ERF database provides various cross-sectional waves. To avoid an arbitrary selection of the base ($t = 0$) and ending period ($t = 1$) in the decomposition exercises, we opt for the longest available time span, which fulfills our availability criteria of the focal variables.[8]

[7] Access to the harmonized dataset is available to registered users and a detailed description of the variables included can be found online: https://www.lisdatacenter.org/frontend#/home.

[8] We also run our analysis for shorter time spans if they are available. In this paper, we provide the results for the US. The estimates for the other countries are available in supplementary materials.



Note. Countries included in the working sample are in red: Austria, Belgium, Brazil, Canada, Chile, Colombia, Czech Republic, Denmark, Egypt, Estonia, Finland, France, Georgia, Germany, Greece, Guatemala, Iceland, India, Ireland, Israel, Jordan, Luxembourg, Mexico, Netherlands, Panama, Peru, Poland, Russia, Serbia, Slovakia, Slovenia, Spain, Switzerland, US, and Uruguay.

Figure 2.1. Countries in Working Sample

2.3.1 Focal Variable - Earnings

We rely on individual yearly gross and net labor incomes, which are defined for all LIS countries as the total income from the main job. This includes cash payments as well as the values of goods and services received from dependent employment, plus the profits/losses and values of goods from self-employment. ERF countries provide information on labor income at the household level. Therefore, for these countries, we proxy the individual income by dividing the household income[9] by the number of members in the household who receive a salary. LIS waves that do not provide individual labor income information[10] are excluded from the analysis.

Although most of the literature on distributional analysis of the RBTC hypothesis focuses on hourly wages, our main variable of interest in the later analysis is yearly earnings. The reason we opt for this is twofold: first, LIS provide wages and hours information for a more restricted number of countries. Since one of the aims of the analysis is to test RBTC theory internationally, we choose the largest harmonized sample of countries possible. Second, the earnings information in LIS is more reliable than wages that suffer of higher item non-response rates. Nevertheless, in Section

[9]ERF provides net household income for Egypt, gross for Jordan.
[10]Estonia in 2000, Ireland in 1987, and Poland in 1999.

2.6, we replicate the analysis using hourly wages as dependent variable in order to provide closer comparability with the previous literature. Compatible hourly wage information is available for 21 countries. Our hourly wage variable is calculated dividing the personal labor income by the number of actual working hours usually worked during the week multiplied by 4.33.

As the earnings information is not harmonized across countries, we include:

- Net earnings countries: Belgium, Chile, Egypt, Georgia, India, Mexico, Russia, Slovenia, and Uruguay.
- Gross earnings countries: Austria, Brazil, Colombia, Czech Republic, Denmark, Finland, Germany, Guatemala, Iceland, Israel, Jordan, Panama, Peru, Slovakia, Switzerland, and the US.
- "Mixed income information": France and Poland have a "mixed" income information.[11]
- Greece, Spain, Estonia, Ireland, and Luxembourg do not have harmonized earnings information across the available time span. Thus, we separate gross from net earnings waves.[12]

We adjust the income variables for inflation using yearly Consumer Price Index data provided by the LIS and trimmed the distribution at 1st and 99th percentiles.[13]

2.3.2 Focal Variable – Occupation

The literature on job polarization proposes two main approaches to characterize job de-routinization and occupation definition according to task requirements. The most frequently used approach relies on the so-called Routine-Task-Index (RTI). Developed for the US by Autor et al. (2003) and later refined in Autor and Dorn (2013), the index "merges job tasks requirements from the fourth edition of the US Department of Labor's Dictionary of Occupational Titles (DOT 1977) to their corresponding (US) Census occupation classification to measure routine, abstract, and manual task content by occupation" (Autor and Dorn (2013), p. 1570). The index is typically normalized around 0: high positive RTI values indicate jobs that are highly routinized and, consequently, more prone to the risk of being displaced according to RBTC hypothesis. Negative RTI values characterize non-routine occupations. Goos et al. (2014) mapped the RTI index from US-specific occupational

[11] According to the codebook: "total income does not account for full taxes and contributions.".

[12] Greece and Spain have gross earnings information available only from 2007 onward. Estonia, Ireland, and Luxembourg switched from net to gross earnings starting in 2000.

[13] `https://www.lisdatacenter.org/data-access/web-tabulator/methods/ppp/`. CPI series for the Czech Republic and Slovakia are not complete, so we use World Bank data available at `https://data.worldbank.org/indicator/FP.CPI.TOTL`.

classification to ISCO-88 (2-digitis)[14] in order to allow for international cross county comparison. According to their metrics, RTI is highest for office clerks and lowest for managers of small enterprises. Mahutga et al. (2018) generalized the RTI index metrics adopted in Goos et al. (2014) for 38 LIS countries, providing correspondence tables to harmonize national occupational schemes to the two-digits ISCO-88 scheme.

In our view, the use of RTI-based classifications has several drawbacks. First, RTI lacks a unique metric. Since numerous potential task scales exist, there is no obvious measure that represents a given group of tasks efficiently (Acemoglu and Autor, 2011). This also makes it difficult to interpret the regression coefficient for the RTI in econometric assessments. Second, in a cross-country perspective, RTI values rely on the assumption that tasks content and exposure to automation is the same for all jobs in all countries of interest. While this assumption might hold for a homogeneous group of highly developed countries, it is difficult to justify it for a set of heterogeneous countries.

For these reasons, we cluster specific occupations into three main job classes, i.e., service, routine, and abstract job classes. With this classification, we follow Acemoglu and Autor (2011). Table 2.1 provides a detailed overview about the definition of the occupational classes in our analysis, the original formulation by Acemoglu and Autor (2011). Moreover, we provide the corresponding ISCO-88 (2-digits) codes and their respective RTI value, as applied by Mahutga et al. (2018).

The Acemoglu and Autor (2011) classification is particularly convenient since it is more flexible for cross-countries comparison: it does not rely on US-centered metrics and it is easily implementable in those countries where ISCO classification is not available and harmonization processes must be applied.[15]

Our classification deviates in two ways from Acemoglu and Autor (2011). We merge the "routine abstract" and the "routine manual" into one "routine" occupational class as done by Massari et al. (2014) and Böhm (2020). Furthermore, we do not drop agricultural occupations entirely from our working sample. Even though we focus on service, routine and abstract occupations, we still control for agricultural occupations in the decomposition analysis. We argue that several countries in our working sample rely considerably on the agriculture sector, hence, it would be inappropriate to exclude them.

[14] The International Standard Classification of Occupations (ISCO) is an International Labor Organization (ILO) classification structure for organizing information on labor and jobs. The current version, known as ISCO-08, was published in 2008 and is the fourth iteration, following ISCO-58, ISCO-68 and ISCO-88.

[15] In some cases, complete harmonization from national to ISCO scheme is not possible. Un-matched occupations from the national occupational scheme can, however, still be assigned to the appropriate routine/non-routine, manual/abstract class based on Acemoglu and Autor (2011) classification. Such manual imputations typically involve around 1-5% of the employed workforce in the wave-specific country and are available upon request.

Table 2.1. Occupational Classes Based on 2-digts ISCO

Occupational Class: Longmiur, Schroeder, Targa	Occupational Class: Acemoglu and Autor	ISCO-88 Label	ISCO-88 Code	RTI
Abstract Occupations	**Non Routine Abstract**	Legislators and senior officials	11	-0.57
		Corporate managers	12	-0.65
		Managers of small enterprises	13	-1.45
		Physical, mathematical and engineering professionals	21	-0.73
		Life science and health professionals	22	-0.91
		Teaching professionals	23	-1.47
		Other professionals	24	-0.64
		Physical and engineering science associate professionals	31	-0.29
		Life science and health associate professionals	32	-0.23
		Teaching associate professionals	33	-1.37
		Other associate professionals	34	-0.34
Routine Occupations	**Routine Abstract**	Office clerks	41	2.41
		Customer services clerks	42	1.56
		Models, salespersons and demonstrators	52	0.17
	Routine Manual	Extraction and building trades workers	71	-0.08
		Metal, machinery and related trades workers	72	0.58
		Precision, handicraft, craft, printing and related trades workers	73	1.74
		Other craft and related trades workers	74	1.38
		Stationary plant and related operators	81	0.45
		Machine operators and assemblers	82	0.62
		Drivers and mobile plant operators	83	-1.42
		Labourers in mining, construction, manufacturing and transport	93	0.57
Service Occupations	**Non Routine**	Personal and protective services workers	51	-0.50
		Sales and services elementary occupations	91	0.14
Agricultural	—	Skilled agricultural and fishery workers	61	0.14

Note. The table shows the correspondence between ISCO-88 2 digits codes and the main occupational classes as proposed in Acemoglu and Restrepo (2017). Last column on the right provides RTI vales before weighting provided in Mahutga et al. (2018). Drivers and mobile plant operators (83) and Extraction and building trades workers (71), in the decomposition analysis have been separated with a specific class dummy. The two categories have negative RTI indexes in Goos et al. (2014), pointing non-routine characteristics, and both categories have wage and hours profile is typically different from the average non routine manual worker.

The main limitation of the 4-classes classification adopted in Acemoglu and Autor (2011) is that it neglects the routine-intensity gradient between different occupations: RTI scores in Table 2.1 ranges from 0.17 for models, salespersons, and demonstrators, to 2.41 for office clerks within the routine abstract occupational class. This heterogeneity in the routine-intensity scale suggests important difference in the nature of the tasks performed by workers and, therefore, potential heterogeneity in the exposure to technological change and to the risk of being subject to automation processes. In this sense, RTI scores can be interpreted as a measure of risk and, therefore, they are particularly suitable in sensitivity analysis seeking to detect the differences in the degree of exposure to the risk of displacements effects between regional and local labor markets. Since we are interested in the distributional effects of *realized* job de-routinization and not on the *potential* risk of layoffs. Thus, we argue that, for our analysis, the aggregated occupational classes adequately characterize the composition of the workforce.

For the assignment of employees to the aforementioned occupational classes, LIS-ESR's harmonized 1-digit occupational variable (9 clusters), *occb1*, is not appropriate since routine and non-routine occupations are mixed together within the same class.[16] For this reason, we classify workers using the country-specific, non-harmonized occupational variable, *occ1_c*. In many countries this variable is directly available and coded in the ISCO-88 two or more digits format. For those countries that rely on national occupational coding schemes, we use the conversion tables provided by Mahutga et al. (2018). This is necessary for Brazil, Canada, Colombia, Finland, France, India, Ireland (87), Israel, Mexico, Panama, and the US. Once the harmonization process is completed, we assign each ISCO-88 occupation to the respective class according to Table 2.1.

Several major changes in the ISCO coding schemes occurred following the year 2010 (ISCO 08). Since a solid harmonization of ISCO 88 and ISCO 08 occupational schemes is not possible at the 2-digit level, we do not include these survey years in our working sample.

Table 2.2 provides a full overview over all countries and waves used in our working sample given the criteria described in this section. The full set of country-specific waves is included in the investigation of job de-routinization over time. The waves used for our decomposition analysis are bold.

2.4 Methodology

In the following section, we present our main methodological framework. In Section 2.4.1, we introduce the descriptive approach for the analysis of job de-routinization. In the following Section 2.4.2, we present the methods to investigate correlations between job polarization and overall inequality patterns across countries. Section 2.4.3 presents the unconditional RIF decomposition technique proposed by Firpo et al. (2009) and Firpo et al. (2018), then applied in Firpo et al. (2011), which constitutes

[16]This is the case for ISCO category 5 "services and sales workers," comprising both "personal and protective services workers" (ISCO 51) and "models, salespersons and demonstrators" (ISCO 52). According to the existent literature, the former should be classified as manual non-routine (RTI index=-.60) and the latter as abstract routine (RTI=+.05). Similar problems exist for ISCO class 8. We need to distinguish between "machine operators and assemblers" (82), who are highly routinized (RTI=0.49), from "drivers and mobile plant operators" (83). who are highly non-routinized (RTI=-1.50). Then in class 9, we need to distinguish between "sales and services elementary occupations" (91), which are non-routinized (RTI=0.03), from agricultural jobs (92 and RTI=n/a) and routinized "laborers in mining, construction, manufacturing and transport" (93) with RTI=+0.53.

Table 2.2. Countries and Waves in the Working Sample

Austria	2004	2007	2010	2013								
Belgium	1995	2000										
Brazil	2006	2009	2013									
Canada	1994	1997	1998	2004	2007	2010						
Chile	1992	1994	1996	1998	2000	2003	2006	2009	2011	2013	2015	
Colombia	2004	2007	2010	2013								
Czech Republic	1992	1996	2002	2004	2007	2010	2013					
Denmark	2004	2007	2010	2013								
Estonia	2000	2007	2010	2013								
Egypt	1999	2008	2010									
Finland	1987	1991	1995	2000	2004	2007	2010	2013				
France	1984	1989	1994	2000	2005	2010						
Georgia	2010	2013	2016									
Germany	1984	1987	1989	1991	1994	1995	1998	2000	2001	2002	2003	2004
	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	
Greece	2004	2007	2010	2013								
Guatemala	2006	2011	2014									
Iceland	2004	2007	2010									
India	2004	2011										
Ireland*	1994*	1995	1996	2000*	2004	2007	2010					
Israel	2007	2010	2012									
Jordan	2002	2006	2008	2010	2013							
Luxembourg*	1997	2000	2004	2007	2010	2013						
Mexico	1984	1989	1992	1994	1996	1998	2000	2004	2008	2010	2012	
Netherlands	1990	1993	2004	2007	2010	2013						
Panama	2007	2010	2013									
Peru	2004	2007	2010	2013								
Poland	2004	2007	2010	2013	2016							
Russia	2000	2004	2007	2010								
Serbia	2006	2010	2013	2016								
Slovakia	1992	2004	2007	2010	2013							
Slovenia	1997	1999	2004	2007	2010	2012						
Spain*	1980	1990	2000*	2004	2007	2010	2013	2016				
Switzerland	1992	2007	2010	2013								
US	1974	1979	1986	1991	1994	1997	2000	2004	2007	2010	2013	2016
Uruguay	2004	2007	2010	2013	2016							

Note. The table shows the countries used in our analysis and provides the waves available in the LIS-ERF data. The waves used for the decomposition analysis are bold. LIS-ERF waves in which the occupational coding scheme is updated to ISCO 08 are marked in blue and have been excluded in the decomposition exercise. We use the remaining set of waves for the analysis of the evolution of employment shares over time. Countries marked with an asterisk changed gross/net classification of earnings as explained in Section 2.3.1. Estonia's and Greece's first waves have been dropped because not consistent with earnings information in later waves.

our empirical framework of the distributional consequences of job de-routinization within each country under analysis. Section 2.4.4 provides the procedure to analyze effects from occupational class-specific composition and returns.[17]

2.4.1 Assessing De-routinization of Jobs

We start our analysis by scrutinizing country specific changes in the composition of the workforce over time. Falling employment shares characterize job de-routinization. Accordingly, we define employment shares as

$$ES_t^{Occ} = \frac{N_t^{Occ}}{N_t^{Serivce} + N_t^{Routine} + N_t^{Abstract}}, \tag{2.1}$$

where *Occ* refers to service, routine, and abstract occupations, N_t^{Occ} is the total number of workers in each occupational class in each period t, as defined in Table 2.2.

RBTC hypothesis suggests that occupational classes follow a strict hierarchy in earnings, with abstract workers earning, on average, more than routine workers, who earn on average more than service. To provide descriptive evidence of it, we provide the mean earnings of occupational classes over time.

2.4.2 Analysis of De-routinization and Inequality

We then describe how changes in the employment structure correlate with overall inequality within and between occupational groups across countries. The aim is to provide suggestive evidence of the importance of both within- and between occupational class dynamics for distributional analysis. Specifically, since there exists a hierarchy in the average returns of the service, routine, and abstract classes, job de-routinization should decrease (increase) inequality *between* service (abstract) and routine occupations by composition effects. As a consequence of job polarization, inequality *between* service (abstract) and routine occupations should decrease (increase). Hence, the evolution of inequality of the *overall* population is ambiguous as it depends on which of these two effects dominate. For this reason, we study correlations between de-routinization and inequality for the lower (service + routine) and upper (abstract + routine) pole separately. We focus on workers employed in routine and service occupations. Complementary analysis for the routine and abstract sub-population is provided in Figure 2.12 in Appendix.

We consider the relative country-specific drop of the employment shares in routine occupations as the measure of job de-routinization, formally:

[17] Formulas provided in this section are all country-specific. For the sake of clarity, we do not include a country index.

$$\Delta ES^{Routine} = \frac{ES_0^{Routine} - ES_1^{Routine}}{ES_0^{Routine}}. \tag{2.2}$$

The higher $\Delta ES^{Routine}$, the stronger is the de-routinization process in that country between period $t = 0$ and $t = 1$.[18] Countries that did (not) experience job de-routinization exhibit negative (positive) $\Delta ES^{Routine}$ growth rates.

We use the variation of the Theil index in the Routine-Service population as measure of earnings inequality since it complies with the decomposition principle (Bourguignon, 1979) and to distinguish inequality within and between occupational classes:

$$\Delta T = \frac{(T_1 - T_0)}{T_0} = \frac{T_1^b + T_1^w}{T_0} - \frac{T_0^b + T_0^w}{T_0} = \frac{T_1^b - T_0^b}{T_0} + \frac{T_1^w - T_0^w}{T_0} = \Delta T^b + \Delta T^w, \tag{2.3}$$

where T is the overall Theil in the routine-service population, T^b is the between component, and T^w the one within.

Exploiting the heterogeneity across countries in our sample, we study correlations between job de-routinization ($\Delta ES^{Routine}$) and changes in between (ΔT^b), within (ΔT^b), and overall (ΔT) inequality for the Service and Routine sub-population. These components enable us to unravel the nexus of de-routinization on inequality by focusing on occupational classes. We see this cross-country evidence as a contribution to the literature, as this link, to the best of our knowledge, has not yet been analyzed in this way.

2.4.3 RIF-Regression Methods

Firpo et al. (2009, 2018) introduced RIF regressions as a generalization of the traditional Oaxaca-Blinder decomposition method. This technique allows for the estimation of a broad set of distributional parameters (e.g., quantiles, Gini index, or variance) and, following Firpo et al. (2011) and Massari et al. (2014), builds a central element in our empirical analysis. We provide a detailed explanation of the methodology in Appendix 2.9.1.

We apply two different decompositions, i.e., the unconditional quantile decomposition for estimating changes along the entire distribution and the P-shares decomposition for four main earnings bins. The unconditional quantile decomposition allows us to present the results intuitively in graphs, while the P-shares decomposition provides a formal proof of our findings providing comprehensive numeric estimates of the distributional effects.

[18] Time periods are defined using the first and the last available harmonized waves.

RIF-unconditional quantile decomposition allows the comparison of observed quantile growth with the counterfactual growth that each quantile of the earnings distribution would have experienced driven by *ceteris paribus* de-routinization effects. We interpret u-shaped patterns in the growth curves of quantiles as evidence of overall earnings polarization.

P-shares are points on the Lorenz curve that represent the share of total earnings going to a pre-defined segment of the earnings distribution. In our analysis, we focus on four main segments: the lower (below the 10th percentile), the lower-middle (between the 10th and 25th), the middle (between 25th and 75th), the upper earnings segment (above the and 75th). More specifically, P-shares are calculated as differences of Lorenz ordinates, such that the middle segment earnings share is the difference between the Lorenz ordinate at the 25th and the 75th percentiles of the cumulative population distribution. A decreasing middle segment share and simultaneously rising shares of upper- and lower-earnings segments indicate earnings polarization (u-shaped pattern).

The decomposition for quantiles takes the following form:

$$\begin{aligned}\Delta^p = q_1^p - q_0^p &= E[RIF(y, q_t^p, F)|T=1] - E[RIF(y, q_t^p, F)|T=0] \\ &= \textstyle\sum_i [\overline{Occ_{i1}}(\widehat{\gamma}_{1,i}^p - \widehat{\gamma}_{0,i}^p) + (\overline{Occ_{i1}} - \overline{Occ_{i0}})\widehat{\gamma}_{0,i}^p] \\ &+ \bar{X}_1(\widehat{\beta}_1^p - \widehat{\beta}_0^p) + (\bar{X}_1 - \bar{X}_0)\widehat{\beta}_0^p,\end{aligned} \tag{2.4}$$

where q_t^p represents the p-quantile at time t, Occ_i is a set of occupational class dummies[19] and X indicates the list of further controls included in the model. We opt for a list of covariates that are fully comparable across time and countries. Specifically, we control for gender, age (six 5-years classes), education (3 classes), and industry affiliation (9 industry classes).[20] Time indexes $t = 1$ and $t = 0$ are defined over the longest time span available as explained in Section 2.3.

In the case of P-shares, $\Delta^v = L(q_t^p)^1 - L(q_r)^0$, where $L(q_t^p)^t$ is the Lorenz curve ordinate at the population p-quantile in time t. The same controls and time spans definition apply for both quantiles and P-shares decomposition.

There are several reasons why we apply the RIF decomposition methodology. First, as in the Oaxaca-Blinder, the RIF decomposition allows for disentangling two distinct channels through which job polarization may affect earnings: first, the

[19] In the model, we include a dummy variable for each category where i: service, routine, abstract, agriculture.

[20] For Canada and Mexico we include a three classes industry categorization (variable $inda1$) since more detailed classifications (variable $indb1$) suffer from considerable missing observations. Russia, Serbia, and Switzerland are the exceptions since early waves do not provide any industry information.

coefficient effect accounts for the change in covariates returns on Δ^p;[21] the *composition effect* shows how much changes in Δ^p can be explained by over-time differences in the level of covariates.[22] Second, the methodology is designed for regression analysis on distributional statistics over the detailed list of covariates X. This means that, for each LIS-ERF country, it is possible to estimate how much of the variation in the statistic of interest can be explained by de-routinization, which is captured by composition and coefficient effects of the class dummies. Simultaneously, we are able to control for other characteristics, X, that might have distributional effects, such as female participation, education, aging, etc. Third, these decomposition methods are robust to non-linearity in the wage setting equation once re-weighted as the counterfactual (Firpo et al., 2018).

It is important to stress two main limitations of the RIF decomposition exercises. First, decomposition methods are accounting exercises that lack of a formal identification strategy so that the estimates should not be interpreted in a strict causal sense (Fortin et al., 2011). Nevertheless, decomposition methods represent a well-established estimation tool to deliver elaborated, descriptive investigation of aggregated phenomena based on counterfactuals. Second, as is well known for the standard Oaxaca-Blinder decomposition, decomposition results depend on the choice of the base group. As highlighted by Fortin et al. (2011), there exists no final remedy to this problem and some arbitrariness is unavoidable, even if normalization strategies are applied (Yun, 2008).[23]

For the sake of clarity, we do not provide confidence intervals for our RIF estimates in our main results section. Nevertheless, we provide the confidence intervals for the estimates of the composition and the coefficient effects in the US in Figure 2.13 in Appendix 2.9.2. We provide robust, instead of bootstrapped standard errors, which should be interpreted as a lower-bound.[24]

In the following sections and in the results tables, we use the term *Total Change* for defining the overall difference in the dependent variables, Δ^p. For RIF-quantiles, it is calculated as the difference in (log)-quantiles between two reference years. Moreover, we refer to *Occupational Effect* for indicating the *sum* of the composition and coefficient effect due to changes in occupational classes. Such effects jointly

[21] In our framework, a reason for this may be that returns of non-routine occupations grow at a faster pace than routine ones inflicted by changes in relative labor demand.

[22] Here, composition effects account for over time differences in the employment shares between routine and non-routine occupations. Specifically, we can estimate the effect on Δ^p of the pure re-allocation of jobs away from routine toward non-routine abstract and service occupations.

[23] In our model the baseline group is represented by male workers between 35 and 39 years old, working in routine occupations, in manufacturing, mining and quarrying industries. Results proved to be robust to different base group specifications and are available upon request.

[24] The confidence intervals are compiled using the Stata command oaxaca_rif provided by Rios-Avila (2020). Bootstrapped standard errors are typically larger than robust standard errors (Firpo et al. (2018) and Rios-Avila (2020)). Therefore, if confidence intervals based on robust standard errors include zero values, those based on bootstrapped standard errors would as well.

account for within- and between-occupation determinants on earnings (Firpo et al., 2009).

2.4.4 Analysis of Occupational Composition and Return Effects

RIF decomposition measures the joint effect of occupational changes on earnings growth. As our interest is also a description of how each of the three main occupational classes (service, routine, and abstract occupations) contribute to shape the overall *Occupational Effects*. Therefore, we first study how the quartile-specific earnings share of each occupational *class e*volved over the time span considered:

$$s_{t,Q}^{Occ} = \frac{\sum_{i=1}^{N_Q^{Occ}} y_{i,t}^{Occ}}{\sum_{i=1}^{N_Q} y_{i,t}} \qquad if\ F(y_{i,t}) \leq Q. \tag{2.5}$$

$s_{t,Q}^{Occ}$ is the quartile-specific earnings share of each occupational class, i.e., service, routine, and abstract. Q indicates the quartile of the earnings distribution. N_Q is the total number of workers in each quartile, while N_Q^{Occ} is the number of those in one of the three occupational classes. We calculate changes in the quartile-specific earnings share for each occupational class as:

$$\begin{aligned} &\Delta\, s_Q^{Occ} = s_{1,Q}^{Occ} - s_{0,Q}^{Occ}, \\ &with\ \Delta\, s_Q^{service} + \Delta\, s_Q^{routine} + \Delta\, s_Q^{abstract} = 1. \end{aligned} \tag{2.6}$$

$\Delta\, s_Q^{Occ} > 0$ indicates that that class increased their earnings share in quartile Q over the time period considered.

Additionally, we explore the dynamics in composition and returns of the three different occupational classes. To describe the changes of the composition of the workforce over time, we estimate the population share of each occupational class below each ventile V of the (log) monthly earnings distribution y in period t=1 and t=0:

$$ES_{t,V}^{Occ} = \frac{N_t^{Occ}}{N_t^{Serivce} + N_t^{Routine} + N_t^{Abstract}} \quad if\ y \leq v. \tag{2.7}$$

We describe the changes of the composition below each ventile of the distribution as

$$\Delta ES_V^{Occ} = ES_{1,V}^{Occ} - ES_{0,V}^{Occ}. \tag{2.8}$$

Positive (negative) values of ΔES_V^{Occ} would imply, that the concentration of workers employed in the occupational class has increased (decreased) below ventile V over time. Aside from composition effects, differences in occupational returns shape the overall *Occupational Effect*. To estimate how the returns of each occupational classes evolved along the ventiles of the earnings distribution, we run the following unconditional quantile regressions $Q_{i,t}$:

$$V_{i,t} = X_{i,t}\beta_{t,V} + \gamma_{t,V}^{Service} * Service_{i,t} + \gamma_{t,V}^{Abstract} * Abstract_{i,t} + \varepsilon_{t,V}. \tag{2.9}$$

As $Service_{i,t}$ ($Abstract_{i,t}$) is equal to one if individual i belongs to the service (abstract) class, $\gamma_{t,V}^{Occ}$ represent the return of the occupation in comparison to the routine class in period t, at the venitle V. We run the regression above for the first and the last period in our dataset. Since routine occupations are generally more clustered at the middle of the distribution, we expect negative values for $\gamma_{t,V}^{Service}$ and positive values for $\gamma_{t,V}^{Abstract}$.

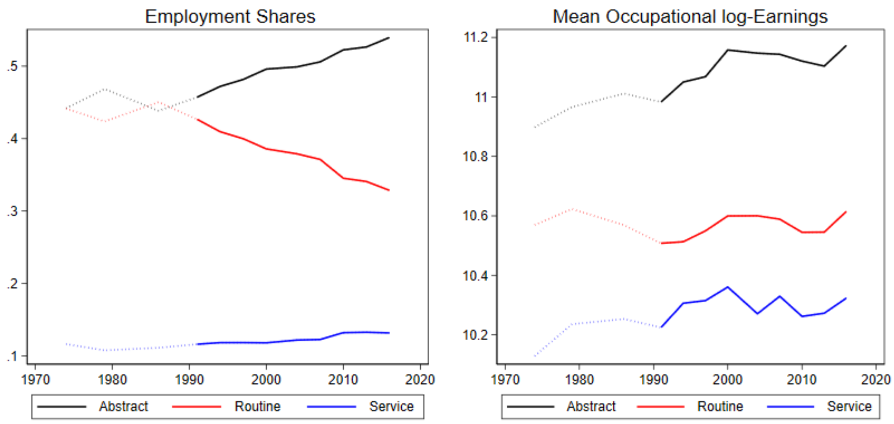
2.5 Results

This section provides the results for de-routinization and its distributional consequences. First, we investigate if de-routinization of jobs is a common feature in our working sample by describing how occupational classes evolved over time in all countries under analysis. Specifically, job polarization is defined as decreasing employment and earnings shares in routine occupations over time. Second, we provide cross-country correlations between de-routinization and inequality between and within occupational groups. Third, we analyze how de-routinization affects the country-specific earnings distributions, based on decomposition methods described in Section 2.4.3. Fourth, we expand RIF results, scrutinizing composition and return effects of each occupational class.

We first present the country-specific results explanatory for the US before we discuss the other countries in our sample. The reason is that the RBTC hypotheses are typically studied for the US and there is not a general consensus regarding the distributional effects of de-routinization of jobs. Moreover, focusing on one country facilitates the interpretation of our results. Detailed country-specific estimations are provided in Appendix 2.9.4.

2.5.1 De-routinization of Jobs

This section provides descriptive evidence for de-routinization of jobs. Figure 2.2 depicts class-specific inter-temporal changes in the employment (left panel) and class-specific average log-earnings (right panel) in the US. Dotted lines indicate waves incurring methodological changes in the main variable, e.g., major changes in the occupational coding scheme, that may decrease their degree of comparability over time. Solid lines, however, are fully harmonized over the entire period.



Note. Compiled by authors based on LIS data for the prime-aged, employed population. This table summarizes the results of our analysis of job-polarization for the US. The left panel shows the change of the employment share for each occupational class over time. The right panel depicts average log earnings over time. Dotted lines indicate waves that incur methodological changes in the main variables. Results of the other countries are provided in the Appendix.

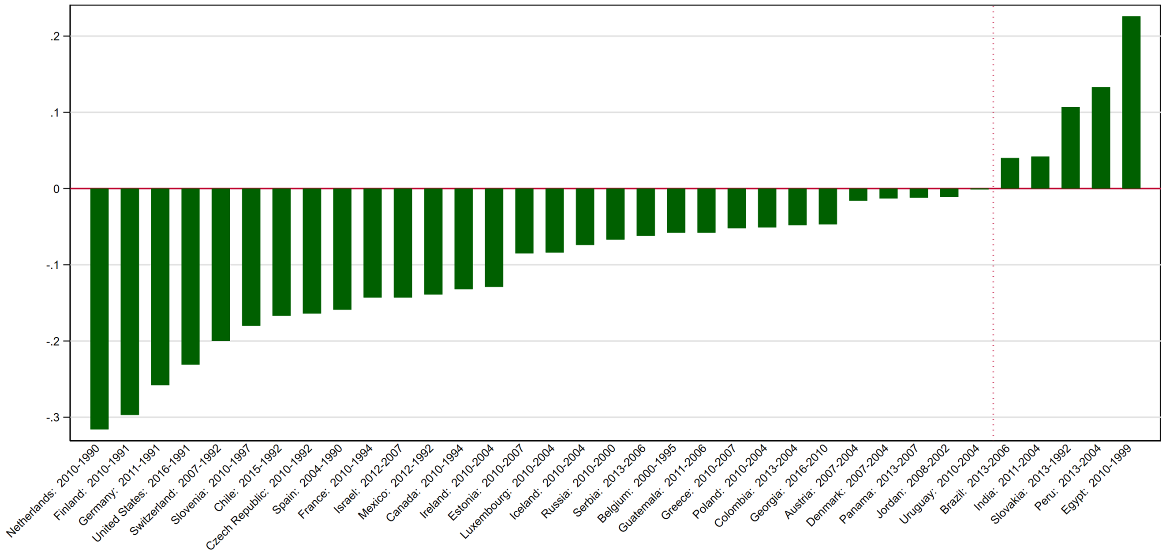
Figure 2.2. Employment and Earnings Shares over Time in the US

The left panel of Figure 2.2 suggests that routine jobs make up a decreasing share of the work force since the 1990s, decreasing from 43% in 1991 to 33% in 2016. Service occupations, marginally increase their employment shares, from 12.2% to 13.6%, while abstract employment share grew from 45% to 53%. These findings support the results of Acemoglu and Autor (2011) regarding the secular decline of routine and abstract occupations between 1959 and 2007.

Average earnings curves in the right panel confirm a hierarchy between occupational classes consistent with the RBTC framework, where abstract occupations are, on average, located at the top, routine in the middle, and service occupations at the bottom of the earnings distribution.

Figure 2.3 summarizes the relative change in the share of workers employed in routine occupations in all countries under analysis. Job polarization, as reflected by a decreasing share of employees in routine task, is present in 30 of 35 countries. These findings are in line with the aggregated analysis by Mahutga et al. (2018). The results for countries where harmonized waves are available for long periods (e.g., Chile, Finland, Germany, and the US) suggest that de-routinization is a long-lasting phenomenon. Only five countries exhibit increasing employment shares in routine tasks, i.e., Brazil, Egypt, India, Peru, and Slovakia. These countries are economies where recent industrialization may explain increases in the production sector and, therefore, higher demand for operative jobs.

The figures with class-specific average log-earnings for each country are provided in Appendix 2.9.4. The hierarchy found in the US is also confirmed for all remaining



Note. Compiled by authors based on LIS data for prime-age, employed population. This table summarizes the results of our analysis of job-polarization. Y-axis is the percentage change of the employment share in routine occupations over time. The X-axis specifies for each country the time span considered.

Figure 2.3. Changes in the Employment Shares of Routine Classes.

countries in our sample. Nevertheless, average earnings do not provide information on the dispersion of earnings levels within occupational classes. Consequently, they show between-class differences, but they are not informative about within-class inequalities or about the overall inequality trend.

2.5.2 De-routinization and Inequality: A Cross-Country Perspective

This section provides correlations between job de-routinization and earnings inequality in a cross-country perspective. As explained in Section 2.4.2, we focus on employees in routine and service occupations. In Figure 2.12 in the Appendix, we provide results for the complementary routine and abstract sub-population.

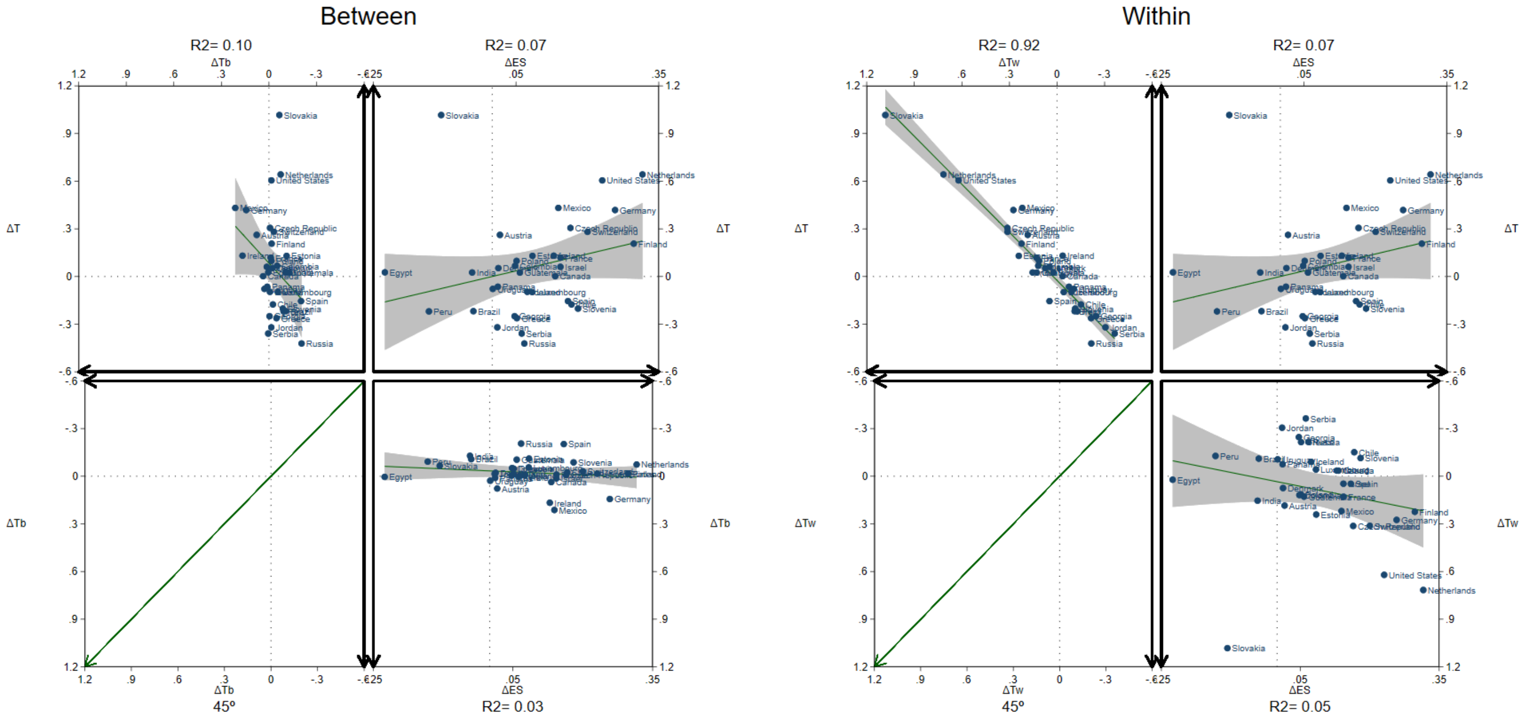
Figure 2.4 summarizes the results for the between- and within-class inequalities by means of two 4-quadrant diagrams. Each diagram includes three dimensions: the measure of de-routinization $\Delta ES^{Routine}$, the overall Theil of the sub-population ΔT, and the Theil variation between (within) the two subgroups ΔT^b (ΔT^w).

Let us first turn to the results for the between component of the Theil index (left 4-quadrant diagram). Here, the upper-right quadrant shows, for all countries in our sample, the relationship between de-routinization ($\Delta ES^{Routine}$) and changes in the Theil index for the Service-Routine sub-sample (ΔT). The relationship is positive indicating that job de-routinization does not coincide with a systematic reduction in inequality at the lower end of the earnings distribution. However, the correlation is weak. R^2 from the binary regression including the 35 countries are low and confidence intervals bands are wide. The lower-right quadrant reenforces this result: the correlation between de-routinization ($\Delta ES^{Routine}$) and changes of the *between*-occupations margin of the Theil index (ΔT^b) is close to zero. So, although the vast majority of the countries under analysis experienced job de-routinization, the *between*-occupations margin of the Theil index exhibits very small variation, which is contrary to RBTC predictions. Eventually, the upper-left quadrant shows the relationship between the Theil index for the Service-Routine sub-sample, ΔT, and the *between*-occupations margin of the Theil index, ΔT^b. The correlation is positive, but strongly driven by few observations. Most of the analyzed countries exhibit no, or only little, variation in the *between*-occupations Theil component. We see this as suggestive evidence that inequality *between* occupations does not sufficiently approximate changes in inequality at the lower end of the earnings distribution.

The right 4-quadrant diagram on the right in Figure 2.4 provides analogous estimates for the Theil variation between *within* service and routine occupations (ΔT^w). The upper-right quadrant is the same as above, showing the correlation between the de-routinization measure, $\Delta ES^{Routine}$, and changes of the Theil index for the Service-Routine sub-sample, ΔT. The lower-right quadrant shows the relationship between the de-routinization measure, $\Delta ES^{Routine}$, and the variation of the Theil between *within* service and routine occupation ΔT^w. Differently from the *between* perspective, the lower-right quadrant shows slightly positive gradient, which mirrors the relationship between job de-routinization and the change in the overall Theil in the upper-right quadrant. Eventually, the upper-left quadrant confirms that the changes in earnings inequality at the lower end of the earnings distribution (ΔT) correlates strongly with the changes over time of *within*-occupations earnings inequality (ΔT^w).

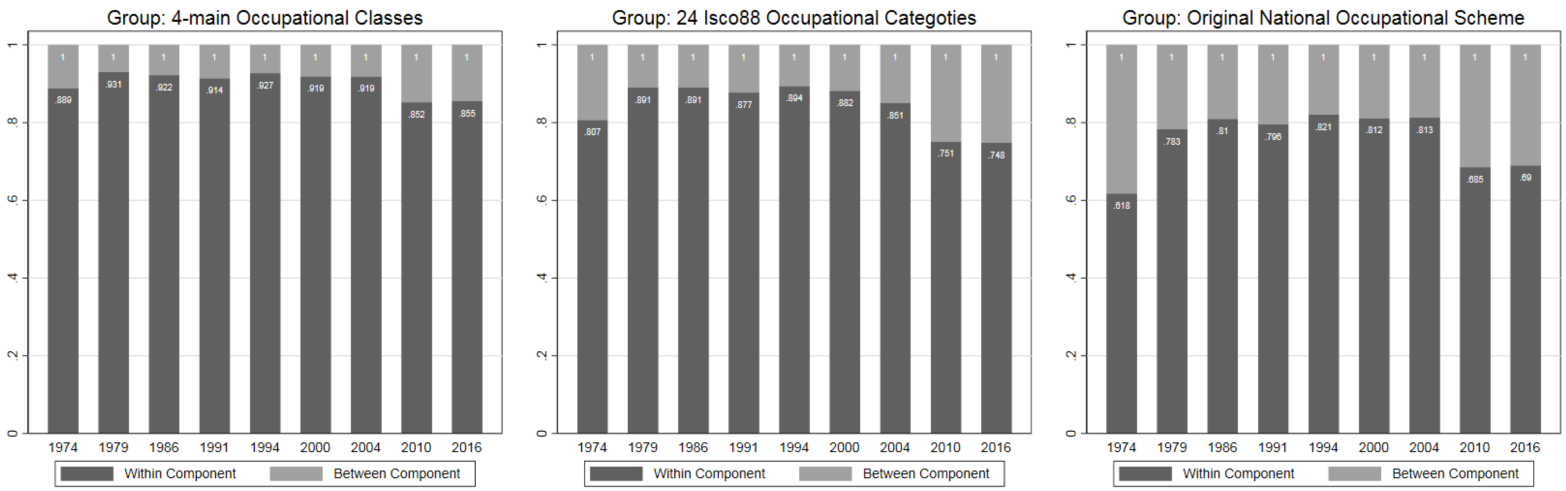
Figure 2.12 in the Appendix provides the 4-quadrant graphs for employees in routine-and abstract occupations. In sum, they confirm the previous findings. Job de-routinization and changes in the inequality at the upper end of the earnings distribution are slightly positively correlated. Again, changes of inequality at the upper end of the distribution emerge from variation *within*, rather than *between* routine and abstract occupations.

Disentangling the effect of de-routinization of jobs on both *between*- and *within*-class inequality on the aggregated country level, we arrive at two major findings: first, there is only little evidence for a quantitatively important link between job and earnings polarization. Second, within occupations dynamics seem to play a major role for the evolution of the earnings distribution over time.



Note. Compiled by authors based on LIS data for prime-aged employed population. The construction of the figure is described in detail in Section 2.4.2 and relates changes of the employment share of workers employed in routine occupations with changes in the overall Theil index and in its between-(within-)occupations component. Occupational classes are defined using the LIS variable $occ1_c$, which is the most detailed occupation information available in LIS. Confidence intervals are reported at the 95% confidence level. R^2 are calculated regressing the y-variable on the x-variable in each graph.

Figure 2.4. Linking H-JP and H-EP: Service and Routine Sub-Population



Note. Compiled by authors based on LIS data for prime-aged employed population. The figure shows the relative composition of Theil index once decomposed in its between (light gray bar) and within (dark grey bar) occupations components. Different clusters of occupations are considered. The panel on the left considers 4 main occupational classes (non-routine service, routine manual, routine abstract and non-routine abstract). The panel in the middle decomposed the Theil index in the 24 ISCO-88 occupation categories. The panel on the right uses 4-digits occupational codes. Results of the remaining countries are provided in the Appendix.

Figure 2.5. Theil Decomposition Within and Between Occupational Classes in the US

The importance of the within-group component for overall inequality is valid under different definitions of occupational groups. Figure 2.5 shows the Theil decomposition *within* and *between* occupational classes for the US. The three panels consider different classifications of occupational classes, from the most aggregated (4 main clusters of workers) on the left, to the least aggregated (4-digits classification) on the right. Even with dis-aggregated occupational information (right panel), overall inequality is mostly determined by inequalities *within* rather than *between* occupations. The same result holds for all countries in our working sample.[25]

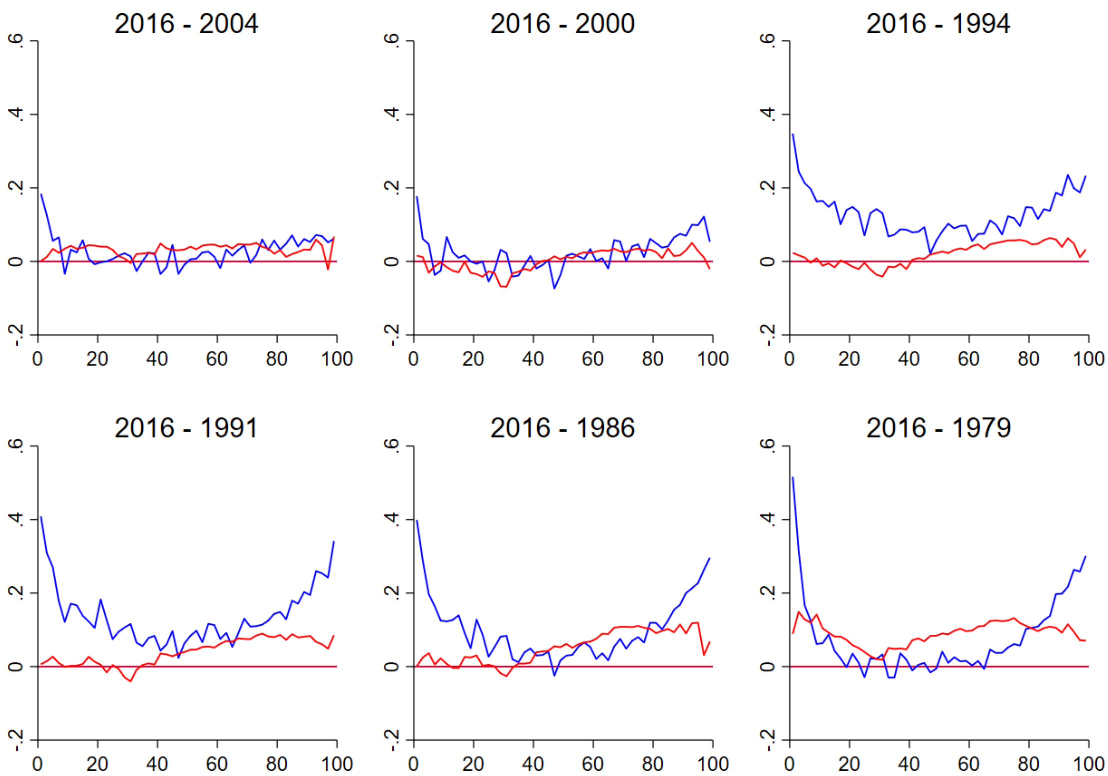
2.5.3 Country-specific Distributional Consequences of De-routinization

The previous section provides static descriptions of earnings dispersion. To investigate the role of de-routinization on earnings distributions in more detail, we turn to our estimates from unconditional quantile decompositions. Figure 2.6 is a comprehensive summary of the results for the US. Figures 2.7 to 2.9b summarize the results for the remaining countries. The blue lines, the *Total Change*, show the unconditional quantile specific earnings growth over the respective time span. The red lines, *Occupational Effect*, indicate growth rates in earnings quantiles that we would observe if only de-routinization of jobs had occurred and all other control variables were fixed at their levels in the baseline reference period. Parallel movements between the *Occupational Effect* and the *Total Change* provide evidence for the determinant role played by de-routinization shaping the earnings distribution. We choose this graphical representation because it enables us to analyze two important dimensions: the (dis)connection of the *Occupational Effect* and the *Total Change*, as well as the evolution of overall inequality over time. The results from the P-share decomposition are reported in Table 2.3. Countries are grouped according to the trend of the *Total Change*, i.e., increasing or decreasing earnings inequality, a polarizing earnings distribution, or no change over time.

Figure 2.6 includes six panels, showing quantile decomposition for US over different time spans. Specifically, the x-axis reports earnings quantiles while the y-axis reports the size of the *Total Change* in blue and the *Occupational Effect* in red. The panels show that the longer the time span, the more distinct are the u-shaped polarization trends exhibited by the *Total Change*. Simultaneously, the *Occupational Effect* growth along the earnings quantiles, implying increasing inequality, and does not exhibit any polarizing pattern. This means that employment de-routinization *per se* cannot explain the observed overall polarization trend in the US.[26]

[25]Country-specific results are presented in the Appendix.

[26]We provide the figures with confidence intervals based on robust standard errors in Appendix 2.9.2. Here, the *Occupational Effect* is divided into the composition and coefficient effects. The



Note. Compiled by authors based on LIS data for prime-age, employed population. The figure shows the total percentile earnings growth (*Total Change*, blue line) and the counterfactual earnings growth (*Occupational Effect*, red line) for the US based on RIF quantiles decomposition explained in Section 2.4.3. The base group is represented by male workers, with a HS diploma, working in routine occupations in manufacturing, mining, or quarrying industry, aged between 35 and 39 years.

Figure 2.6. Quantile Decompositions Results for the US

Our results are in line with those of Hunt and Nunn (2019) and Böhm et al. (2019), showing the ambiguous distributional consequences of the RBTC framework: by including within-group variation, the *Occupational Effect* does not correlate with the *Total Change*. Our estimates also suggest that increased labor demand for non-routine occupations did not necessarily lead to higher returns for service workers at the bottom of the distribution. Moreover, *Occupational Effects* are positive in the middle of the distribution, meaning that workers in middle quantiles experienced earnings growth driven by changes in the occupational composition. As *Occupational Effects* do not explain the *Total Change*, labor market institutions, like unions (Firpo et al., 2018) and minimum wages (DiNardo et al., 1996), might have played an important role for shaping earnings distribution over time, especially at the bottom.

Detailed P-shares decomposition in Table 2.3 corroborate the results: positive coefficients in the lower and upper P-shares indicate shift from the middle toward the end of the distribution. Since the 1980s, an increasing share of middle class labor income is redistributed toward the tails of the distribution, resulting in simultaneous reductions of inequality in the bottom-half and increases in the upper half.

For the remaining countries, we find various overall distributional outcomes, but no close link between job de-routinization and changes in the earnings distribution over time. We discuss the country-specific trends with regards to changes in the overall earnings distribution, i.e., increased and decreased inequality, polarization, and no change in inequality. Specifically, we observe increasing (decreasing) inequality if the *Total Change* is monotonically increasing (decreasing) over the earnings quantiles. Moreover, we refer to polarization if the *Total Change* is u-shaped over the earnings quantiles. Finally, we characterize no change in inequality, if the *Total Change* is constantly close to zero. Our working sample consists of countries that are differently embedded in the world economy, which are observed over various time spans. Interpreting the magnitude and the sources for heterogeneous earnings percentiles growth for every single country, however, is not within the scope of this paper.

Like Figure 2.6, Figures 2.7 to 2.9 report RIF-quantile decomposition results for all the countries in our sample. We provide results for the longest time span available.

Figure 2.7 includes estimates for those countries in our working sample that experienced increased inequality: Austria, Czech Republic, Denmark, Estonia, Finland, France, Germany, India, Mexico, Netherlands, Poland, Slovakia, Slovenia, and Spain. With the unique exception of India, we find evidence for overall job de-routinization in all these countries; however, our RIF decomposition results show that the *Occupational Effect* does not explain the *Total Change* along the earnings quantiles. In some countries, like Finland, Germany, Mexico, and Spain, *Occupational Effects* are

confidence intervals are narrow and do not affect the interpretation of our results. We provide the confidence intervals for other countries upon request.

positive at the bottom of the distribution. This is consistent with the RBTC framework, as bottom-tail earnings would have increased if only occupational changes had occurred. However, other mechanisms offset the impact of job de-routinization on the overall *Total Change*. In several countries, i.e., Austria, Czech Republic, Denmark, Estonia, France, Poland, Slovakia, and Slovenia, de-routinization effects are close to zero along the entire distribution. The Netherlands is the only country where we observe that the *Occupational Effects* and the *Total Change* are very similar. Nevertheless, they are both monotonically increasing along the earnings distribution, which is not in line with the RBTC framework.

Figure 2.8 reports RIF decomposition results for those countries that experienced decreasing inequality. It includes Brazil, Chile, Colombia, Georgia, Guatemala, Jordan, Panama, Peru, Russia, Serbia, and Uruguay. The *Total Change* show that lower quantiles are growing at a faster rate compared to upper quantiles. Although we find evidence of job de-routinization in all these countries, except for Brazil and Peru, *Occupational Effects* are generally weak and, again, they do not explain the decreasing *Total Change*.

Figure 2.9a shows the results for countries that exhibit overall earnings polarization: Belgium, Canada, Ireland, Switzerland, and the United States. We find evidence of employment *and* earnings polarization in all these countries. The u-shaped *Total Change* are less extreme in comparison to the United States, suggesting that strong earnings polarization is a phenomenon limited to the latter. Ireland and Switzerland, however, seem to be the only countries in our sample where the *Total Change* at the bottom of the earnings distribution is fully explained by *Occupational Effects*, which is in line with the RBTC framework.

Figure 2.9b plots the results for Egypt, Greece, Iceland, Israel, and Luxembourg. These countries show rather stable inequality over the considered time horizons.

Table 2.3 provides a summary of the P-shares decompositions for all 35 countries. The *Total Change*, TC, reports the estimates of four main earnings bins: lower segment (between the 1st and 10th percentiles), lower-middle segment (between the 10th and 25th percentiles) middle segment (between the 25th and 75th percentiles), and the upper segment (between the 75th and 99th percentiles). The *Occupational Effect*, OE, provides the estimates of the counterfactual. The coefficients are multiplied by 100.[27] Table 2.3 confirms our graphical results by reporting heterogeneous pattern in inequality growth between the different countries under analysis and the weak distributional impact of job de-routinization. Moreover, the *Total Changes*, as well as the *Occupational Effects* vary considerably across countries, implying that a generalization of the nexus between de-routinization of jobs and the earnings distribution is not achievable. We confirm these findings for hourly wages; discussion is provided in Section 2.6.

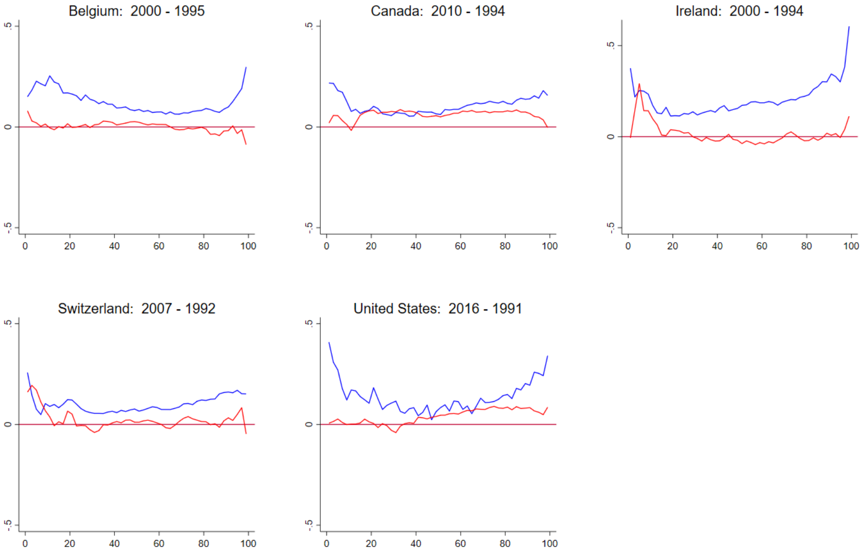
[27] The complete decomposition results for each country are provided in the Appendix 2.9.4.

Note. Compiled by authors based on LIS data for prime-age, employed population. The figure shows the total percentile earnings growth (*Total Change*, blue line) and the counterfactual earnings growth (*Occupational Effect*, red line) for countries with increasing inequality based on RIF quantiles decomposition explained in Section 2.4.3. The base group is represented by male workers, with a HS diploma, working in routine occupations in manufacturing, mining, or quarrying industry, aged between 35 and 39 years.

Figure 2.7. Increased Inequality - *Total Change* and *Occupational Effect* from RIF Quantiles Decomposition.

Note. Compiled by authors based on LIS data for prime-age, employed population. The figure shows the total percentile earnings growth (*Total Change*, blue line) and the counterfactual earnings growth (*Occupational Effect*, red line) for countries with decreasing inequality based on RIF quantiles decomposition explained in Section 2.4.3. The base group is represented by male workers, with a HS diploma, working in routine occupations in manufacturing, mining, or quarrying industry, aged between 35 and 39 years.

Figure 2.8. Decreased Inequality - *Total Change* and *Occupational Effect* from RIF Quantiles Decomposition.



(a) Polarization

(b) No change in inequality

Note. Compiled by authors based on LIS data for prime-age, employed population. The figure shows the total percentile earnings growth (*Total Change*, blue line) and the counterfactual earnings growth (*Occupational Effect*, red line) for countries exhibiting (a) polarization or (b) no change in inequality, based on RIF quantiles decomposition explained in Section 2.4.3. The base group is represented by male workers, with a HS diploma, working in routine occupations in manufacturing, mining, or quarrying industry, aged between 35 and 39 years.

Figure 2.9. *Total Change* and *Occupational Effect* from RIF Quantiles Decomposition.

Table 2.3. P-shares Decomposition

Country	Time Span	1-10		10-25		25-75		75-99	
		TC	OE	TC	OE	TC	OE	TC	OE
				Increasing Inequality					
Austria	2007 - 2004	-0.56	0.00	-1.29	-0.33	-1.95	-1.14	3.80	1.47
Czech Rep.	2010 - 1996	-0.58	-0.09	-0.65	-0.07	-0.71	-0.74	1.93	0.90
Denmark	2007 - 2004	-0.25	-0.02	-0.13	0.03	-0.11	-0.10	0.49	0.09
Estonia	2010 - 2007	-0.47	-0.03	-0.49	-0.11	-0.31	-0.63	1.27	0.76
Finland	2010 - 2000	-0.76	0.43	-0.56	0.07	0.52	-0.06	0.79	-0.44
France	2010 - 1989	-1.58	-0.21	-1.41	-0.47	-0.40	-0.89	3.39	1.57
Germany	2011 - 1995	-0.73	0.01	-1.45	0.05	-1.08	-1.58	3.25	1.52
India	2011 - 2004	-0.34	0.09	-0.18	0.25	1.92	2.00	-1.40	-2.33
Mexico	2012 - 1996	-0.58	0.12	-1.05	-0.12	0.18	-1.55	1.44	1.55
Netherlands	2010 - 1990	-0.84	-0.41	-2.12	-0.36	-1.37	-0.42	4.33	1.19
Poland	2010 - 2004	-0.57	-0.18	-0.22	-0.17	0.09	0.22	0.70	0.13
Slovakia	2013 - 1992	-0.72	0.11	-0.81	-0.09	-1.40	-0.38	2.93	0.36
Slovenia	2010 - 1997	-0.63	-0.33	0.21	-0.67	0.89	0.07	-0.47	0.93
Spain	2004 - 1990	-0.20	0.22	-0.46	0.08	-1.47	0.38	2.13	-0.67
				Decreasing Inequality					
Brazil	2013 - 2006	0.35	-0.16	1.02	-0.01	2.81	0.11	-4.19	0.05
Chile	2015 - 2000	0.44	-0.08	1.32	-0.03	3.21	0.80	-4.97	-0.70
Colombia	2013 - 2004	0.19	-0.26	0.74	-0.07	0.76	0.60	-1.70	-0.27
Georgia	2016 - 2010	0.13	-0.05	0.66	0.21	1.34	3.37	-2.13	-3.52
Guatemala	2011 - 2006	0.15	0.22	0.80	0.42	1.98	-0.84	-2.93	0.20
Jordan	2008 - 2002	2.50	-0.17	1.48	-0.34	-1.54	0.78	-2.44	-0.27
Panama	2013 - 2007	-0.02	-0.17	0.89	-0.32	0.03	-1.35	-0.90	1.84
Peru	2013 - 2004	0.11	0.10	0.85	0.63	3.20	1.21	-4.16	-1.94
Russia	2010 - 2000	1.11	0.02	1.93	-0.11	5.99	0.33	-9.03	-0.25
Serbia	2013 - 2006	0.80	-0.43	1.40	-0.47	-0.46	0.04	-1.74	0.86
Uruguay	2010 - 2004	0.06	-0.05	0.38	-0.05	2.48	1.03	-2.92	-0.93
				Polarization					
Belgium	2000 - 1995	0.36	0.02	0.67	0.01	-1.43	0.80	0.40	-0.84
Canada	2010 - 1994	0.04	-0.11	-0.17	-0.02	-1.33	0.49	1.47	-0.35
Ireland	2000 - 1994	0.01	0.23	-0.84	0.25	-3.62	-0.99	4.45	0.51
Switzerland	2007 - 1992	0.00	0.11	-0.04	-0.01	-1.57	-0.15	1.61	0.05
US	2016 - 1991	0.11	-0.13	-0.18	-0.30	-3.15	-0.15	3.22	0.58
				No Change					
Greece	2010 - 2007	0.27	0.00	1.09	-0.00	1.63	0.79	-2.98	-0.79
Iceland	2010 - 2004	-0.07	-0.33	0.37	-0.49	-0.50	-0.17	0.19	0.99
Egypt	2010 - 1999	0.12	0.03	0.55	-0.23	1.66	-1.35	-2.32	1.55
Israel	2012 - 2007	-0.14	0.08	-0.17	-0.05	0.12	0.29	0.18	-0.31
Luxembourg	2010 - 2004	0.10	0.17	0.23	-0.24	0.61	-2.60	-0.94	2.67

Note. Compiled by authors based on LIS data for prime-age, employed population. The table presents detailed result for the P-share decomposition, as explained in Section 2.4.3, with the estimates of four main earnings bins: lower segment (between the 1st and 10th percentiles), lower-middle segment (between the 10th and 25th percentiles) middle segment (between the 25th and 75th percentiles), and the upper segment (between the 75th and 99th percentiles). TC columns in black report estimates of *Total Change* in four wage bins considered. OE columns in light gray report estimates for *Class Effect*. Coefficents are multiplied by 100.

2.5.4 Occupational Composition and Return Effects

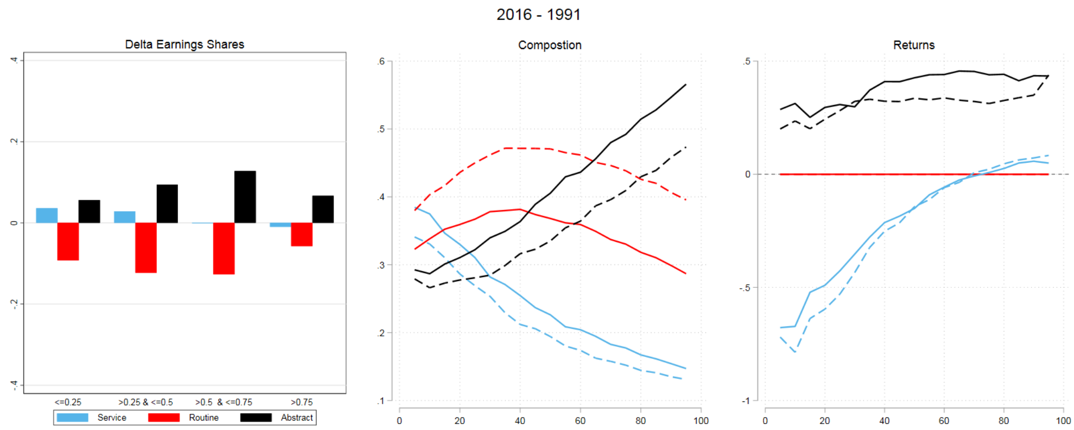
Our results so far show that job de-routinization does not imply generalized distributional consequences. We argue that employees from a certain occupational class are not perfectly stratified but scattered along the earnings distribution. Consequently, de-routinization does not just shift jobs from the middle toward the tails, but it replaces routine occupations along the entire earnings distribution. The weak link, therefore, arises form simultaneous movements of different occupational classes within same quantiles that can counteract and enforce each other resulting in ambiguous distributional effects. We present and discuss these arguments with the aid of three case studies - i.e., the US, Ireland, and Switzerland. We chose to focus on these countries for two reasons: first, the US case is highly debated in the literature and, with the following analysis, we contribute a novel perspective. Second, Ireland and Switzerland represent interesting study cases since *Occupational Effect* predicts well the *Total Change* at the bottom of the distribution, although, as we show it this section, the changes in occupational composition and returns are distinctively different. We provide the results for the other countries in the Appendix 2.9.4.

For each country in our case study, Figures 2.10 and 2.11 provide the results in three panels. The left panels provide the quartile-specific earnings share of the three occupational classes. The middle panels describe the change of the composition of employment shares along the earnings distribution. The right panels depict the returns along the earnings distribution with the routine class as base category.[28] In all panels, blue represents the service class, red the routine class, and black the abstract class. Dotted lines indicate the estimates for the initial period.

Figure 2.10 provides the results for the US. The left and middle panels show that the share of employees in routine occupations has reduced evenly along the earnings distribution. Hence, workers in routine jobs have been replaced equally by both workers in service occupations, with lower returns, and workers in abstract occupations, with higher returns, along the entire distribution. From the right panel, we observe that the hierarchy of returns between occupational classes has not changed over time. Thus, within each quantile, there are service (abstract) workers who replace routine workers and, therefore, reduce (increase) earnings growth. These shifts seem to neutralize each other, especially at the bottom, explaining why we find an *Occupational Effect* close to zero in lower quantiles for the US, as shown in the section above.

Although Switzerland and Ireland exhibit both similar positive *Total Change* and *Occupational Effect* at the lower end of the earnings distribution, the underlying mechanisms differ considerably. Figure 2.11 depicts the results for Ireland in the upper three panels and for Switzerland in the lower three panels. In Ireland, the left panel shows that routine jobs lost their earning shares to the service and abstract

[28] Formulas are discussed in detail in Section 2.4.4.



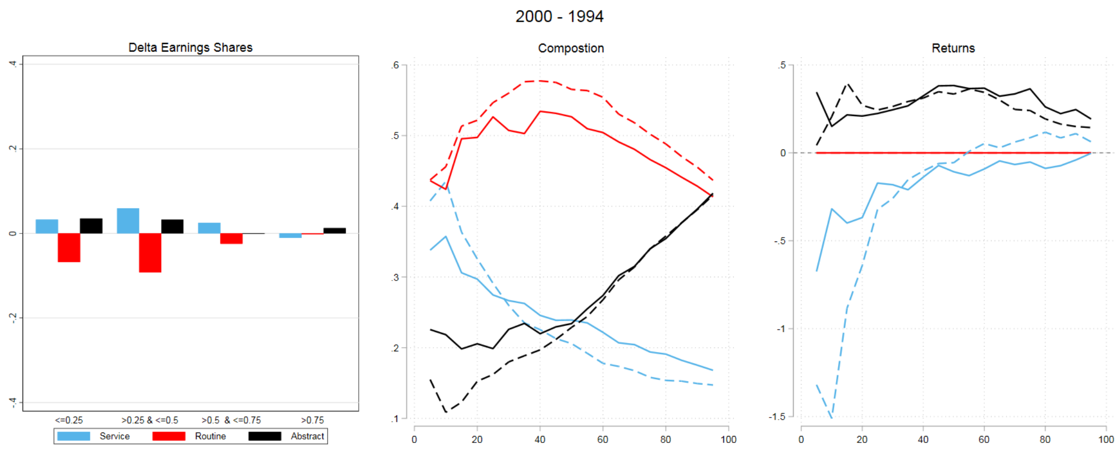
Note. Compiled by authors based on LIS data for prime-age, employed population. The figure provides the results of the US. The left panel provides the change of earnings shares by occupational class for the quartiles of the earnings distribution over time. The central panel depicts the changes in occupational composition along the ventiles of the earnings distribution. The right panel shows the changes of occupational returns using the routine occupation as baseline category. Dashed lines indicate the estimates in the base year.

Figure 2.10. Occupational Composition and Return Effects in the US

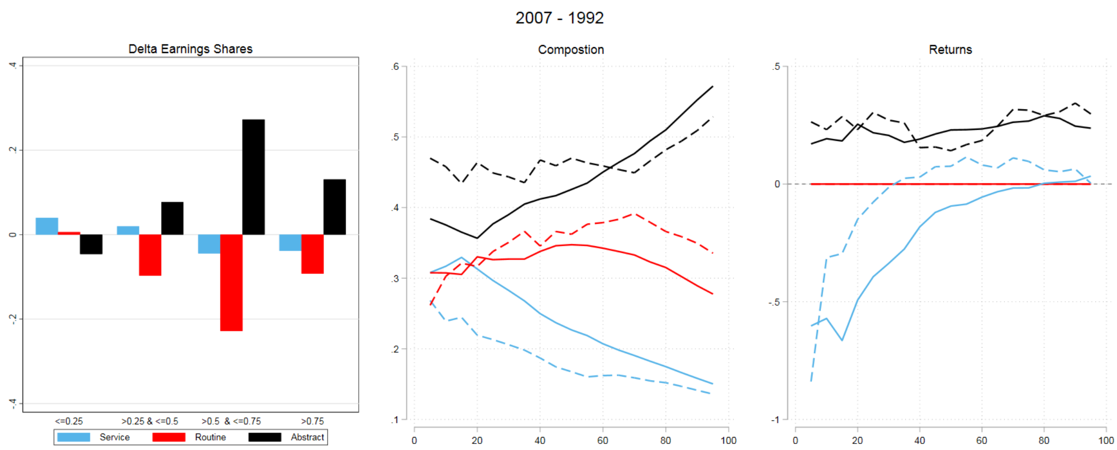
classes. This is due to both a relative reduction in the composition (middle panel) and the returns (right panel). The *Total Change* at the lower end of the distribution seems to be driven by both a large increase of the composition of abstract workers and increased returns for service workers.

In Switzerland, the earning shares of routine jobs decrease especially at the top half of the distribution. For the lower 25 percent, the earnings share of abstract workers decreases, those of routine occupations remain constant, while the earnings shares of service jobs increase. Abstract workers with higher returns were clustered at the upper end of the lower part of the distribution in 1992 and left it over time while service jobs increased their share. As the *Total Change* at the lower end of the distribution is positive, the returns of all classes have increased, despite the fact that abstract workers are moving up the earnings distribution.

These results suggest large differences in the composition of the workforce between and within countries. Ireland's composition comes close to patterns described by RBTC hypothesis, with service jobs at the bottom and more abstract occupations at the top. However, routine jobs still dominate in all parts of the distribution. In Switzerland, abstract occupations dominate along the whole earnings distribution. Considerable shares of abstract workers at the lower end of the distribution can be also found in Austria, Belgium, Czech Republic, Denmark, Egypt, Estonia, Finland, France, Georgia, Germany, Greece, Israel, Netherlands, Russia, Slovakia, Slovenia.



(a) Ireland



(b) Switzerland

Note. Compiled by authors based on LIS data for prime-age, employed population. The figure provides the results of Ireland and Switzerland. The left panel provides the change of earnings shares by occupational class for the quartiles of the earnings distribution over time. The central panel depicts the changes in occupational composition along the ventiles of the earnings distribution. The right panel shows the changes of occupational returns using the routine occupation as baseline category. Dashed lines indicate the estimates in the base year.

Figure 2.11. Occupational Composition and Return Effects

These results suggest several insights on the link of job de-routinization and the overall earnings distribution. We find evidence for a persistent hierarchy of returns, i.e., abstract workers gaining the highest returns, routine workers in the middle, and service workers at the bottom, which is consistent with the RBTC framework. Nevertheless, occupational classes are scattered along the whole distribution and, therefore, job de-routinization is not necessarily displacing workers only in the middle of the earnings distribution. Additionally, we find that job de-routinization does not displace routine workers evenly along the earnings quantiles. As shown for Ireland and Switzerland, routine occupations have been displaced only in middle and higher quantiles, keeping their employment shares relatively unchanged at the bottom of the distribution. Similarly, increasing demand of abstract and service occupations is not necessarily concentrated only at the top and bottom of the earnings distribution, respectively. Dynamics within the abstract workers' share is potentially critical for understanding the evolution of earnings at the bottom of the distribution. A point that is commonly disregarded in the literature.

2.6 Robustness Checks - Wages instead of Yearly Gross-Income

In this section, we replicate the analysis explained in Section 2.5.3 using hourly wages as the dependent variable in order to provide closer comparability with the existing literature. Due to data constraints explained in Section 2.3, we can reproduce the analysis on hourly wages for only 21 countries.

We plot tables and figures of the wage analysis in Appendix 2.9.3. Figure 2.14 and Figure 2.15 provide detailed unconditional quantiles decomposition results for the United States and for eight selected countries. Table 2.5 reports P-shares decomposition results using wages as dependent variable.

The results for wages confirm our main findings for earnings and we do not observe critical differences. In Figure 2.14, the wage decomposition for the US shows very similar patterns as in Figure 2.6 for earnings: u-shaped *Total Change* curves indicating overall polarization of wages, which are not driven by *Occupational Effects*. Similar parallelism can be observed in Figure 2.15 for wage and in Figures 2.7, 2.8, and 2.9 for earnings. This suggests that working hours did not affect the estimation results and they contributed only marginally to the evolution of inequality in our working sample. Studying the long-run relationship of de-routinization and working hours is outside the scope of this paper, but we invite future research to provide more evidence on this matter.

There are two interesting exceptions that are important to discuss. In Ireland, our results show that *Occupational Effects* on the hourly wage distribution are negative at the bottom of the distribution, despite u-shaped patterns in *Total Change*. Once hourly wages are taken as dependent variable, *Occupational Effects* at the bottom of the earnings distribution are close to zero. Such results might be explained by the strong replacement of service with abstract jobs at the bottom of the distribution experienced in Ireland, as seen in Section 2.5.4 and in Figure 2.11. Workers in abstract occupations work, on average, more hours than individuals in the service sector, achieving higher earnings for similar hourly wage levels. Once working hours are ruled out, the *Occupational Effect* turns to zero. Overall, these findings suggest that the *earnings* growth experienced in Ireland at the bottom of the distribution results from compositions effects, i.e., the substitution of service jobs with abstract jobs, characterized by more working hours, rather than relevant wage increases.[29]

The second notable exception between earnings and wage analysis is Greece: Table 2.15 suggest strong wage increases at the bottom and strong wage drops at the top of the wage distribution, which should result in decreased inequality. Such results are, however, compensated by changes in the structure of working hours, so that in Table 2.3, we observe limited changes in overall earnings inequality.

2.7 Qualification and Extensions

The analysis applied in this paper requires several assumptions for its methodological approach and the comparability of the data for some countries needs to be treated cautiously.

A central assumption of the RIF decomposition is the in-variance of the conditional distribution. It means that there are no equilibrium effects between the two different periods (Firpo et al. (2009), Firpo et al. (2011) and Fortin et al. (2011)). This assumption is relatively strong and potentially violated in our analysis, as we include several time periods, where structural changes took place. This is a major caveat of this approach, especially as job de-routinization is typically a long-lasting effect. This can potentially result in biased estimates and we cannot exclude that some long-run time frames are affected by this.

The results based on the RIF decomposition are sensitive to the choice of the base group. This arbitrary choice could potentially mean that the elements of the decomposition are viewed as arbitrary as well.[30] In our main analysis, the base group is defined as "male workers, with a HS diploma, working in routine occupations in manufacturing, mining, or quarrying industry, aged between 35 and 39 years". We re-run our analysis for different base groups, but the implications remained the

[29]Specific figures on the composition of occupational classes along the wage distribution, in spirit of Figure 2.11, can be provided upon request.

[30](Oaxaca and Ransom, 1999) see this as an identification problem, Fortin et al. (2011) and Gelbach (2016) refer to it as a conceptual problem.

same.[31] Nevertheless, there is no standardized method for choosing the base group and we cannot avoid an arbitrary selection.

Although the literature refers to the counterfactual estimates as *effects*, the RIF decomposition is not suitable for causal interpretation (Firpo et al. (2009), Firpo et al. (2011), Fortin et al. (2011), and Firpo et al. (2018)). The estimated counterfactual is a local approximation for the effect of changes in occupational classes on the quantiles growth over time. Fortin et al. (2011) argue that its accuracy depends on the application at hand. As we run the decomposition for each quantile of the earnings distribution, we see the interpretation of our results in the light of a local approximation as reasonable.

We do not provide confidence intervals of our estimates in our main analysis. The reason for this is the computational limits of bootstrapping the RIF results. Therefore, we only provide robust standard errors for the US. in Appendix 2.9.2. As discussed above, we choose this simplification for computational reasons, as we would have to run about 500 times 20 RIF regressions per country otherwise. The literature, especially in Firpo et al. (2011) and Firpo et al. (2018), typically does not provide confidence bands for quantile regressions, since they are mostly interested in the "shape of the effect". We do not think that is a strong argument and, therefore, we included intervals based on robust standard errors, which are typically smaller than bootstrapped standard errors and should be seen as a lower bound.

Apart from the RIF methodology, we also do not include a formal test of polarization as provided in the polarization literature by (Esteban and Ray, 1994) and (Chakravarty et al., 2015). We do not apply it for two reasons: first, the analysis of polarization is not the main concern of this analysis, as we find various distributional consequences in our analysis. Second, the RBTC literature typically refers to "u-shaped patterns" of wage growth along the wage distribution. We follow this very generalized definition to be consistent to the literature. As we typically focus on the lower or upper part of the distribution, this should not affect our results and interpretation.

Another concern is the comparability of the datasets. The LIS-ERF dataset is an internationally recognized and well received collection of surveys, but they are not perfectly harmonized, which effects the comparability of our vocal variables, i.e., earnings and occupation, between countries and over time. We apply a restrictive selection scheme, excluding several countries and waves from our analysis. The remaining 35 countries, however, still rely to some extent on net-, or mixed earnings information, as pointed out in Section 2.3.1. Furthermore, we rely on the crosswalks of national occupational coding schemes to the generalized isco88 scheme. As occupation is a central variable in our analysis, we do not transform isco88 into the isco08 scheme, as definition of service, routine, and abstract jobs cannot not easily be translated without strong assumptions, which ultimately would affect our results.

[31] Results are available upon request.

This comes with a cost, as the exclusion of isco08 reduces the possibilities to analyze the most recent years. However, the two periods chosen in each country for our analysis should be highly comparable.

From a general perspective, one could also criticize that we argue against a theory without proposing an alternative explanation for the drivers of earnings inequality. Furthermore, the analysis focuses on occupational changes and neglects other labor market transitions over time. These would be interesting extensions for follow-up research in the future.

2.8 Conclusion

In this paper, we analyze whether de-routinization of the workforce can be observed internationally and if this is explains changes in earnings inequalities within and between occupations. Our analysis focuses on 35 LIS-ERF countries characterized by different economic and political systems. We confirm shifts from routine-intense jobs toward non-routine occupations in 30 countries, but we do not find a close link between de-routinization of jobs and changes in the earnings distribution.

We provide two major reasons for our findings: first, we find that, on an aggregated country level, the intensity of de-routinization does not correlate with changes in inequality between and within occupational classes. Factors *within* - rather than *between* - occupational groups determine overall inequality trends, indicating that differences in returns between occupational classes do not changes to the earnings distribution. Second, our case studies show that, although we confirm a hierarchy in their average returns, service, routine, and abstract jobs are jointly distributed along the earnings distribution. Therefore, de-routinization not only affects jobs at the middle, but it also displaces workers in all earnings quantiles. We argue that such shifts in occupational shares within each quantile ultimately defines the *Occupational Effect* on overall earnings.

Our results highlight that de-routinization induced by ICT adoption is a process most countries face. Given the heterogeneous composition and returns of occupational classes within and between countries, policy makers need to take these multifaceted patterns into account. We see a further investigation of the channels through which within-occupational variation affect the earnings distribution, as a relevant field for further research to understand the effect of job de-routinization on inequality of labor market outcomes.

2.9 Appendix

2.9.1 RIF-Regression Methods

Assume a generic wage structure function, that depends on some observed components X_i, some unobserved components ϵ_i and time $t = 0,1$:

$$Y_{it} = g_t(X_i, \epsilon_i) \tag{2.10}$$

From observed data on (Y, T, X) we can identify the distributions of $Y_t|T = t \overset{d}{\sim} F_t$ for $t = 0,1$. The framework proposed by Firpo et al. (2009, 2018) is a generalization of Oaxaca-Blinder that allows the estimation of a broad set of distributional parameters $v_t = v(F_t)$ including quantiles, the variance, or the Gini index under very general assumptions on the earnings setting equation 2.10. The central innovation is the use of Recentered Influence Functions (RIF). RIFs give the influence that each observation has on the calculation of $v(F_t)$ and have the property of integrating up to the parameter of interest $v(F_t)$. Therefore, it is possible to express group/time specific functions, v_1 and v_0, as conditional expectations:

$$v(F_t) = E[RIF(y_t, v_t, F_t)|X, T = t] \tag{2.11}$$

Firpo et al. (2009, 2018) prove that using the estimated $\widehat{RIF_{it}}$ as a dependent variable in a linear model, it is possible to estimate coefficients via standard OLS:

$$E[RIF(y_t, v_t, F_t)|X, T = t] = X_t \widehat{\gamma}_t^v \tag{2.12}$$

$$\widehat{\gamma}_t^v = E[XX'|T = t]^{-1} E[RIF(y_t, v_t, F_t)|X, T = t] \tag{2.13}$$

X_t is a vector of covariates that entails dummies for the occupational class, as described in the sections above, and socio-demographic controls. γ_t^v represents the marginal effect of X on $v(F_t)$. Finally, it is possible to decompose the difference of earnings v in the Oaxaca-Blinder traditional manner:

$$\Delta^v = \bar{X}_1(\widehat{\gamma}_1^v - \widehat{\gamma}_0^v) + (\bar{X}_1 - \bar{X}_0)\widehat{\gamma}_1^v \tag{2.14}$$

In the specific case of quantiles, RIF is defined as:[32]

[32] See Firpo et al. (2018) for more detailed information about RIF estimation of quantiles.

$$RIF(t; q_t^p) = q_t^p + \frac{p - I[y \leq q_t^p]}{f_Y(q_t^p)} \tag{2.15}$$

$$E[RIF(y_t, q_t, F_t)|T = 1] = \frac{1}{f_Y(q_t^p)} Pr[Y > q_t^p | X = x] + (q_t^p - \frac{1-p}{f_Y(q_t^p)}) \tag{2.16}$$

$$= c_{1,p} Pr[Y > q_t^p | X = x] + c_{2,p} \tag{2.17}$$

In the above equations, q_t^p is the value of the p-quantiles of Y and $f_Y(q_t^p)$ is the estimated kernel density evaluated in q_t^p. Thus, RIF can be seen more intuitively as the estimation of a conditional probability model of being below or above the quantile q_t^p, re-scaled by a factor $c_{1,p}$, to reflect the relative importance of the quantile to the distribution, and re-centered by a constant $c_{2,p}$. A detailed discussion about RIF for P-shares can be found in Davies et al. (2017).

2.9.2 Auxiliary Tables and Figures

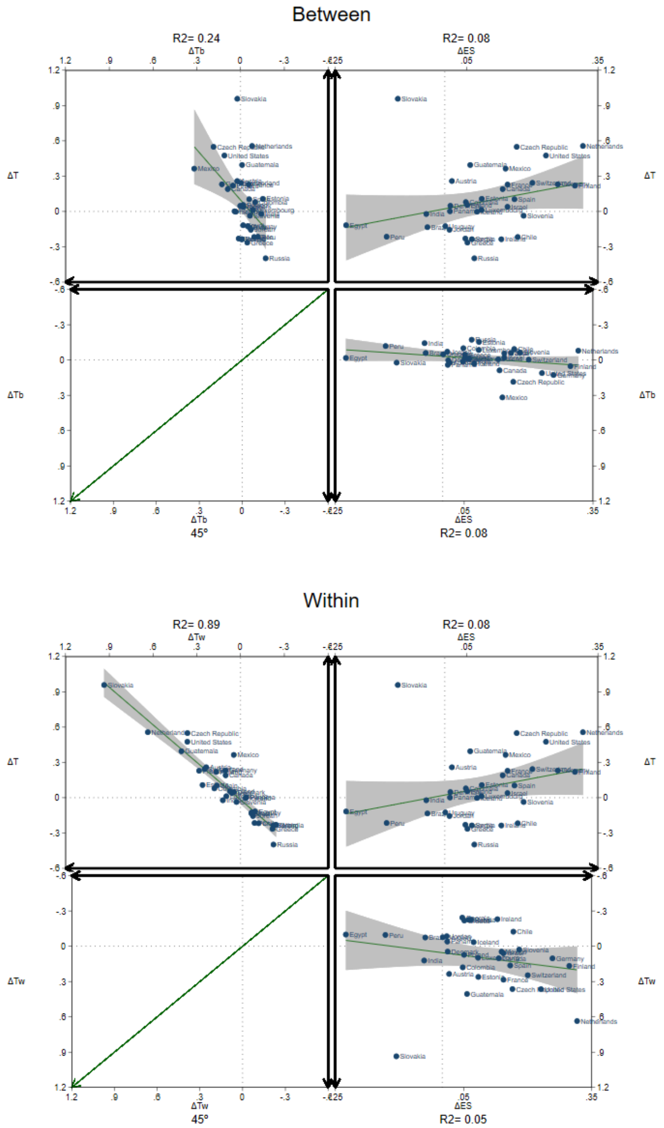
Table 2.4 summarizes the results of our analysis considering the job-polarization hypothesis. The last column reports value of the change in the shares of workers employed in Routine occupations between the indicated time span. Specifically, these values are $-\Delta ES^{Routine}$ explained in Section 2.4.2.

Table 2.4. Summary Results for De-routinization of Jobs

Country	***Time Span***		***Δ Employment Share of the Routine Class (%)***
		De-routinization	
Austria	*2007*	*2004*	-1,6%
Belgium	*2000*	*1995*	-5,8%
Canada	*2010*	*1994*	-13,2%
Chile	*2015*	*1992*	-16,7%
Colombia	*2013*	*2004*	-4,8%
Czech Republic	*2010*	*1992*	-16,4%
Denmark	*2007*	*2004*	-1,3%
Estonia	*2010*	*2007*	-8,5%
Finland	*2010*	*1991*	-29,7%
France	*2010*	*1994*	-14,3%
Georgia	*2016*	*2010*	-4,7%
Germany	*2011*	*1991*	-25,8%
Greece	*2010*	*2007*	-5,2%
Guatemala	*2011*	*2006*	-5,8%
Iceland	*2010*	*2004*	-7,4%
Ireland	*2010*	*2004*	-12,9%
Israel	*2012*	*2007*	-14,3%
Jordan	*2008*	*2002*	-1,1%
Luxembourg	*2010*	*2004*	-8,4%
Mexico	*2012*	*1992*	-13,9%
Netherlands	*2010*	*1990*	-31,6%
Panama	*2013*	*2007*	-1,2%
Poland	*2010*	*2004*	-5,1%
Russia	*2010*	*2000*	-6,7%
Serbia	*2013*	*2006*	-6,2%
Slovenia	*2010*	*1997*	-18,0%
Spain	*2004*	*1990*	-15,9%
Switzerland	*2007*	*1992*	-20,0%
United States	*2016*	*1991*	-23,1%
Uruguay	*2010*	*2004*	-0,1%
		No De-routinization	
Brazil	*2013*	*2006*	4,0%
Egypt	*2010*	*1999*	22,6%
India	*2011*	*2004*	4,2%
Peru	*2013*	*2004*	13,3%
Slovakia	*2013*	*1992*	10,7%

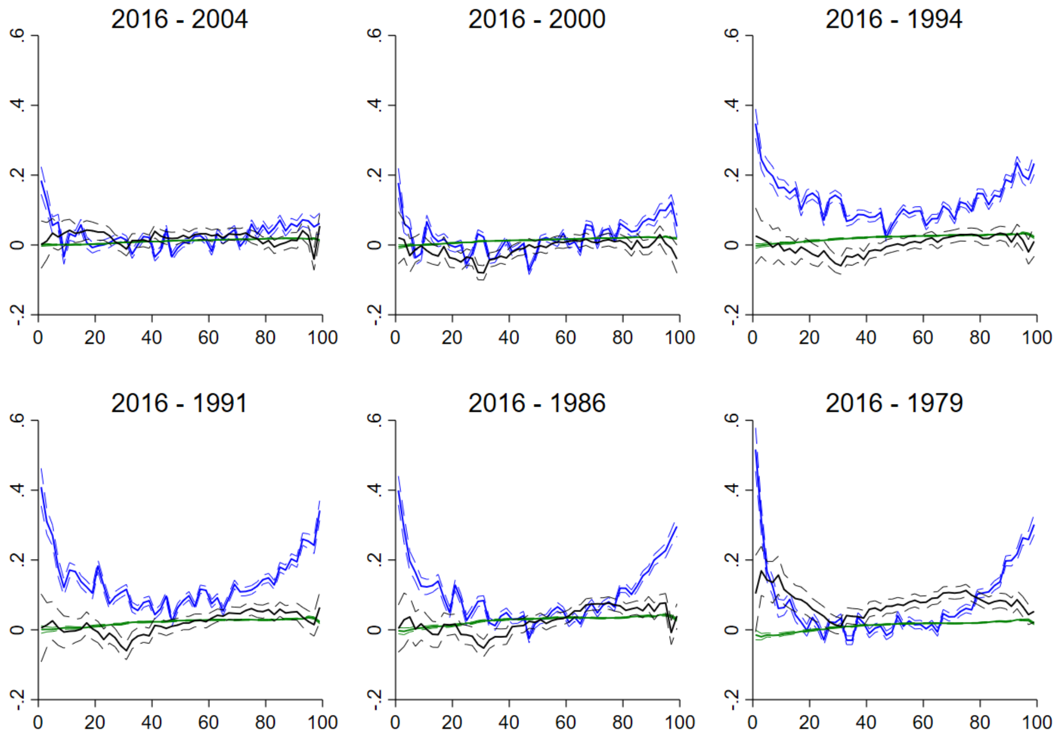
Note. Compiled by authors based on LIS data for prime-age, employed population. The table presents the changes of the employment shares of the routine class over the corresponding time span for each country of our working sample.

Here we present results for the complementary analysis on workers employed in routine and abstract occupations.



Note. Compiled by authors based on LIS data for prime-aged employed population. The construction of the figure is described in detailed in Section2.4.2 and relates changes of the employment share of workers employed in routine occupations (x-axis in the upper right and bottom right panel), with changes in the overall Theil index (y-axis in the upper right and left panels) and in its within-occupations component (y-axis in the lower right panel and x-axis in the upper left panel). Confidence intervals are reported at the 95% confidence level. R^2 are calculated regressing the y-variable on the x-variable in each graph.

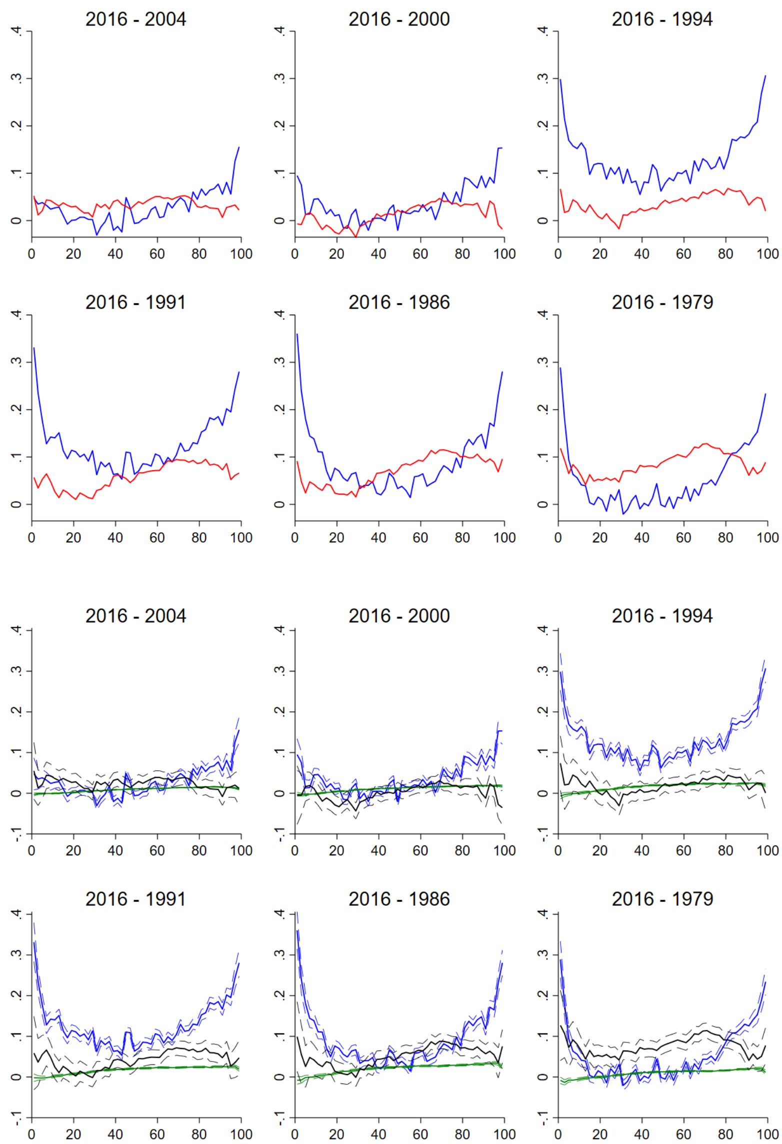
Figure 2.12. Linking H-JP and H-EP: Abstract and Routine Sub-Population



Note. Compiled by authors based on LIS data for prime-age, employed population. The figure shows the total percentile earnings growth (*Total Change*, blue line) and the *Occupational Effect* has been decomposed in *Composition* (in green) and *Coefficient Effects* (in black) for the US based on RIF quantiles decomposition explained in Section 2.4.3. Confidence intervals are based on robust standard errors and are provided at the 95% significance level. The base group is represented by male workers, with a HS diploma, working in routine occupations in manufacturing, mining, or quarrying industry, aged between 35 and 39 years. Confidence intervals are provided at the 95% significance level.

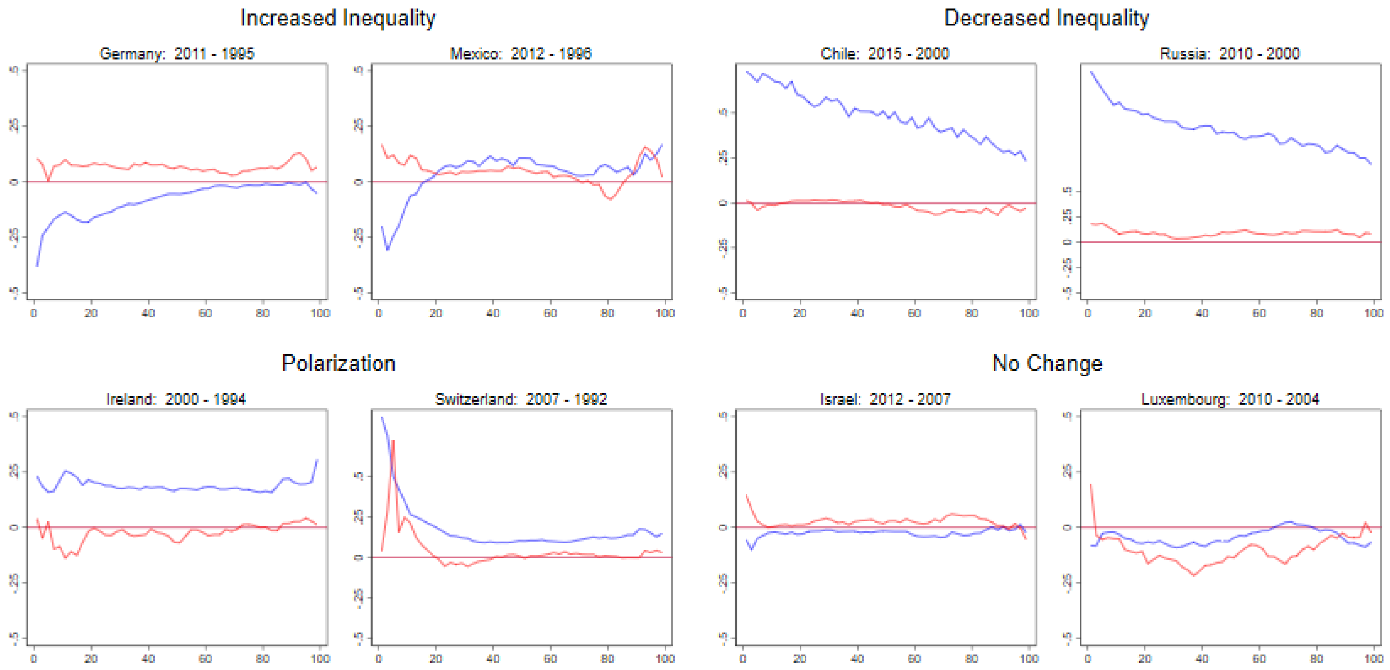
Figure 2.13. Detailed Quantile Decompositions Results for the US

2.9.3 Wage Polarization



Note. Compiled by authors based on LIS data for prime-aged, employed population. The upper panel shows the total percentile earnings growth (*Total Change*, blue line) and the *Occupational Effect* (red line) for the US based on RIF quantiles decomposition explained in Section 2.4.3. The lower panel decomposes the *Occupational Effect* in *Composition* (in green) and *Coefficient Effects* (in black). Confidence intervals are based on robust standard errors and are provided at the 95% significance level. The base group is represented by male workers, with a HS diploma, working in routine occupations in manufacturing, mining, or quarrying industry, aged between 35 and 39 years.

Figure 2.14. Detailed Quantile Decomposition Results for the US - Wages



Note. Compiled by authors based on LIS data for prime-age, employed population. The figure shows the total percentile wage growth (blue line) and the Occupational Effect (red line) for selected countries based on RIF quantiles decomposition explained in Section 2.4. The base group is represented by male workers, with HS diploma, working in routine occupations in in manufacturing, mining, and quarrying industries, aged between 35 and 59 years old.

Figure 2.15. Percentile Growth and Occupational Effect in Selected Countries - Wages

Table 2.5. P-shares decomposition - All Countries - Wages

Country	1-10		10-25		50-75		75-99	
	TC	OE	TC	OE	TC	OE	TC	OE
				Increased Inequality				
Austria: 2007 - 2004	-1.25	-0.39	-1.67	-0.74	-1.10	-0.30	4.01	1.42
Czech Rep.: 2010 - 1996	-0.66	-0.01	-0.53	-0.15	0.09	-0.77	1.09	0.93
Estonia: 2010 - 2007	-0.45	0	-0.41	-0.13	-0.35	-0.85	1.21	0.97
Germany: 2011 - 1995	-0.39	0.12	-0.75	0.16	0.69	-0.29	0.46	0.01
Mexico: 2012 - 1996	-0.28	0.14	-0.27	0.08	-0.34	-0.77	0.89	0.55
Netherlands: 2010 - 1990	-1.17	-0.40	-1.08	-0.54	1.22	-0.51	1.04	1.45
				Decreased Inequality				
Brazil: 2013 - 2006	0.12	-0.04	-0.01	0.11	-2.74	0.42	2.63	-0.50
Chile: 2015 - 2000	0.66	-0.14	1.20	-0.01	2.33	0.47	-4.19	-0.31
Colombia: 2013 - 2004	0.16	-0.17	0.60	0.19	1.85	-0.43	-2.60	0.41
Guatemala: 2011 - 2006	0.18	0.17	0.72	0.16	0.98	-0.68	-1.87	0.35
Russia: 2010 - 2000	0.99	0.03	1.62	-0.17	3.25	-0.78	-5.87	0.92
Uruguay: 2010 - 2004	0.11	-0.17	0.35	-0.07	1.24	0.84	-1.70	-0.59
				Polarization				
Belgium: 2000 - 1995	0.32	-0.20	0.68	0.05	-1.35	1.04	0.35	-0.89
Canada: 2010 - 1994	-0.19	-0.34	-0.36	-0.30	-1.03	0.75	1.57	-0.10
Ireland: 2000 - 1994	-0.01	-0.22	0.12	-0.45	-0.88	-0.45	0.77	1.12
Switzerland: 2007 - 1992	0.83	0.34	0.53	0.01	-1.78	-0.90	0.41	0.55
United States: 2016 - 1991	0.11	-0.08	-0.12	-0.28	-1.51	0.33	1.52	0.03
				No Change				
Greece: 2010 - 2007	0.53	-0.19	1.45	-0.01	2.65	0.47	-4.63	-0.27
Iceland: 2010 - 2004	0.11	-0.04	0.58	-0.41	1.90	0.70	-2.58	-0.25
Israel: 2012 - 2007	-0.17	0.02	-0.06	-0.10	-0.37	0.28	0.60	-0.20
Luxembourg: 2010 - 2004	0.04	-0.01	-0.12	-0.17	0.45	-1.93	-0.36	2.11

Note. Compiled by authors based on LIS data for prime-age, employed population. The table presents detailed result for the P-share decomposition, as explained in Section 2.4.3, with the estimates of four wage bins: lower segment (between the 1st and 10th percentiles), lower-middle segment (between the 10th and 25th percentiles) middle segment (between the 25th and 75th percentiles), and the upper segment (between the 75th and 99th percentiles). TC columns in black report estimates of *Total Change* in four wage bins considered. OE columns in light gray report estimates for *Occupational Effect*. Coefficients are multiplied by 100.

2.9.4 Detailed Country Specific Results

The current Appendix presents country specific results for all the main analysis. Results are based on the LIS-ERF joint dataset and harmonized following to the guidelines explained in Section 2.3. Employment and income shares, Decomposition results for unconditional quantile regressions are reported in country-specific tables and figures are analogous to the ones in the main analysis. The P-share decomposition are provided for three earnings bins, i.e., below the 15th percentile, between the 15th and 85th percentile. Note that Russia, Serbia, Slovakia, Switzerland do not provide industry information. Therefore, we computed RIF decompositions without controlling for industry dummies.

The following notes hold for the graphs and tables of all countries, respectively:

Figures: Employment and Income Shares by Occupational Class

Note. Compiled by authors based on LIS data for the prime-aged, employed population. . The left panels show the change of the employment share for each occupational class over time. The right panels depict average log earnings over time. Dotted lines indicate waves that incur methodological changes in the main variables.

Figures: Theil Decomposition

Note. Compiled by authors based on LIS data for prime-aged employed population. The figures show the relative composition of Theil index once decomposed in its between (light gray bar) and within (dark grey bar) occupations components. Different clusters of occupations are considered. The panels on the left consider 4 main occupational classes (non-routine service, routine manual, routine abstract and non-routine abstract). The panels in the middle decompose the Theil index in the 24 ISCO-88 occupation categories. The panels on the right uses 4-digits occupational codes.

Figurse: Quantile RIF Decomposition

Note. Compiled by authors based on LIS data for prime-age, employed population. The figures show the total percentile earnings growth (*Total Change*, blue line) and the counterfactual earnings growth (*Occupational Effect*, red line) for the respective country based on RIF quantiles decomposition explained in Section 2.4.3. The base group is represented by male workers, with a HS diploma, working in routine occupations in manufacturing, mining, or quarrying industry, aged between 35 and 39 years.

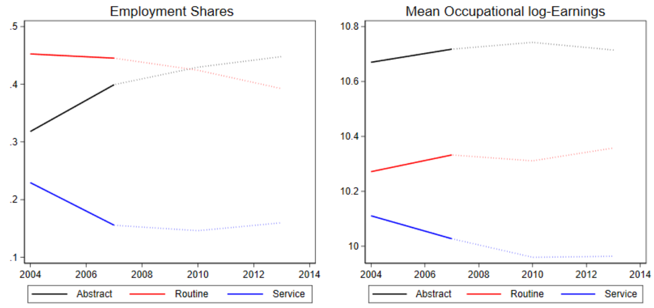
Figures: Occupational Classes Composition and Returns

Note. Compiled by authors based on LIS data for prime-age, employed population. The left panels provide the change of earnings shares by occupational class for the quartiles of the earnings distribution over time. The central panels depict the changes in occupational composition along the ventiles of the earnings distribution. The right panel shows the changes of occupational returns using the routine occupation as baseline category. Dashed lines indicate the estimates in the base year.

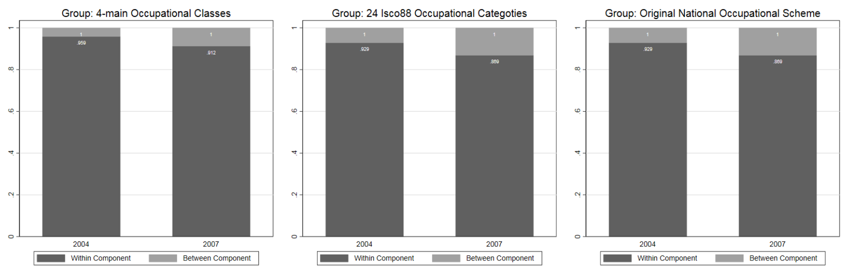
Tables: P-Shares Decomposition

Note. Compiled by authors based on LIS data for prime-age, employed population. The tables presents detailed result for the P-share decomposition, as explained in Section 2.4.3. The delta provides the estimate of the *Total change*. Moreover, we provide the composition and coefficient effect for the *Occupational Effect*, education, females, age, and industry.

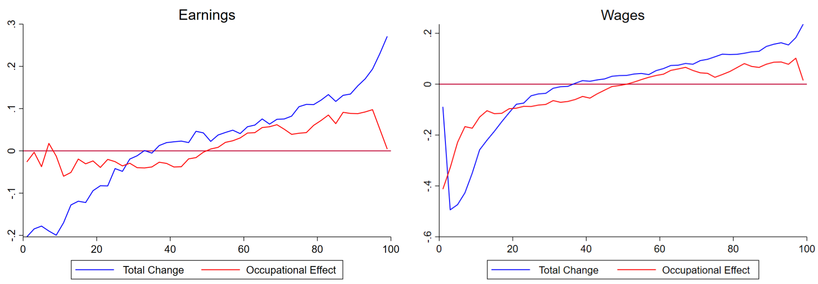
Austria: 2007-2004



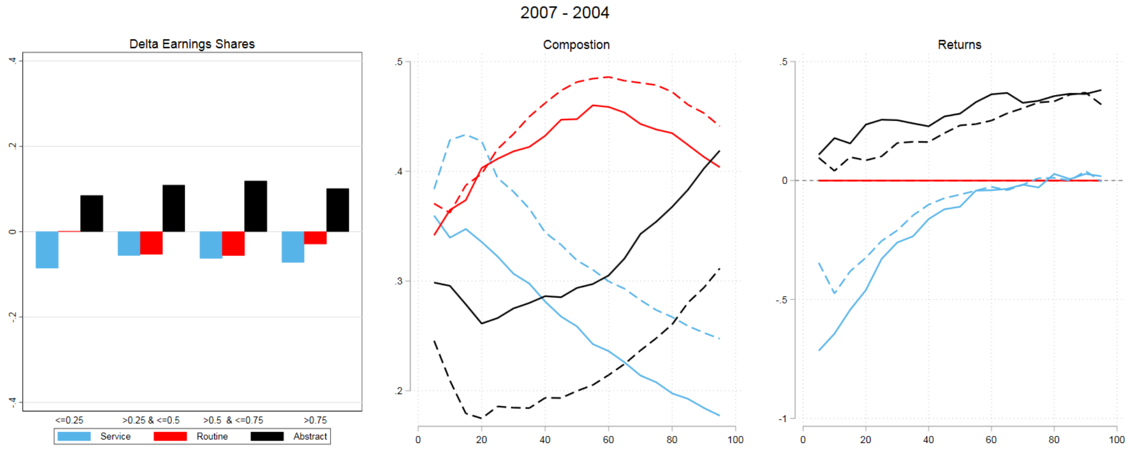
Employment and Income Shares by Occupational Class



Theil Decompostion



Quantile RIF Decompostion

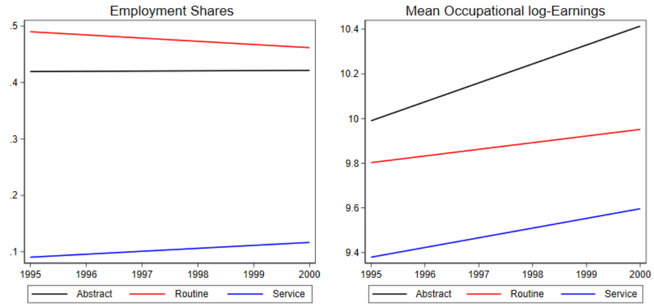


Occupational Classes Compostion and Returns

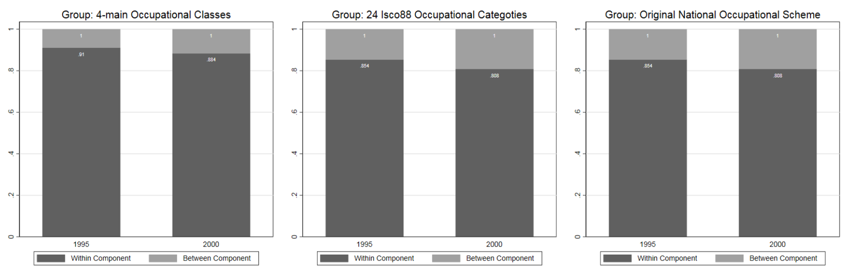
P-Shares Decomposition

Austria: 2007 - 2004			
	<15	15-85	>85
Δ	-0.00289***	0.000570	0.00232***
Specification Error	3.00e-05	1.00e-05	-4.00e-05
Composition Effect			
Occ	0.000130	-0.000180	5.00e-05
Educ	5.00e-05	-1.00e-05	-0.00004*
Female	3.00e-05	0	-0.00003*
Age	5.00e-05	-0.00011*	0.00006*
Ind	-7.00e-05	3.00e-05	4.00e-05
Reweighting Error	-9.00e-05	1.00e-05	8.00e-05
Coefficent Effect			
Occ	-0.000570	-0.000310	0.00088*
Educ	0.00100*	-0.00095*	-5.00e-05
Female	-0.000960	0.000410	0.000550
Age	-0.000210	-8.00e-05	0.000300
Ind	-0.00194	0.000370	0.00157
Constant	-0.000340	0.00140	-0.00105

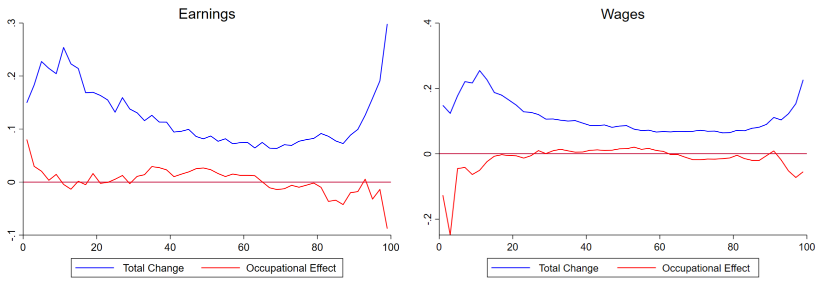
Belgium: 2000-1995



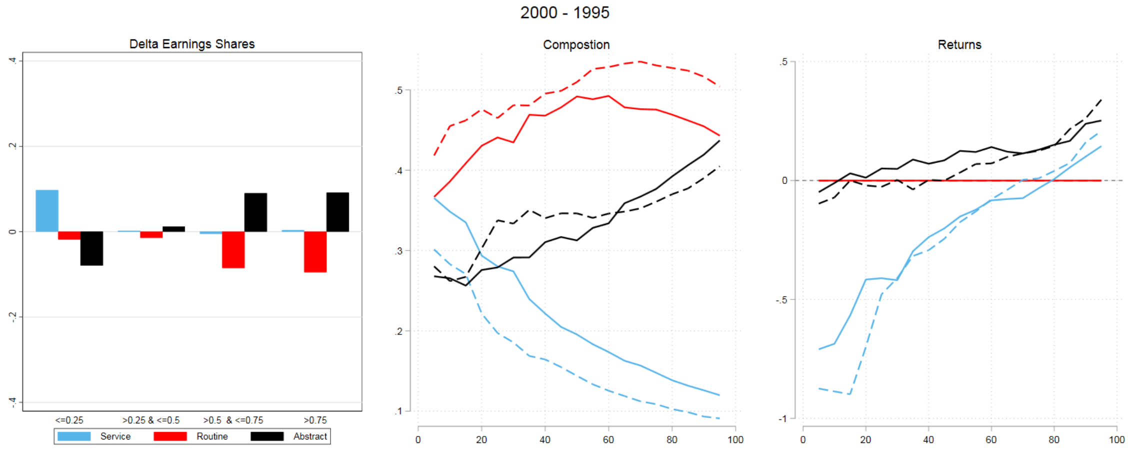
Employment and Income Shares by Occupational Class



Theil Decompostion



Quantile RIF Decompostion

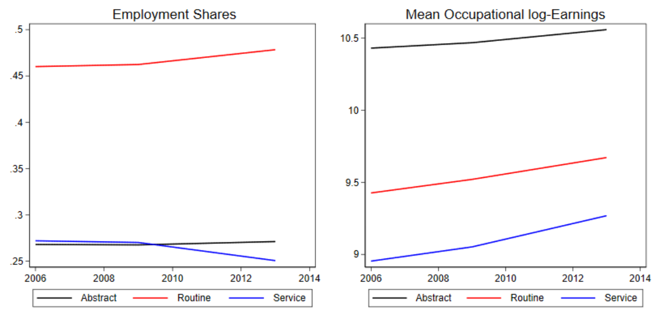


Occupational Classes Compostion and Returns

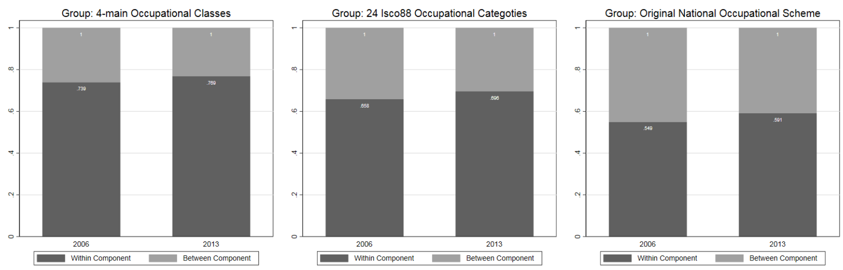
P-Shares Decomposition

Belgium: 2000 - 1995			
	<15	15-85	>85
Δ	0.00162***	-0.00166***	4.00e-05
Specification Error	0	0	0
Composition Effect			
Occ	-0.00022***	1.00e-05	0.00021***
Educ	1.00e-05	-3.00e-05	1.00e-05
Female	-1.00e-05	1.00e-05	1.00e-05
Age	2.00e-05	-6.00e-05	4.00e-05
Ind	0.00014*	-6.00e-05	-9.00e-05
Reweighting Error	-5.00e-05	1.00e-05	4.00e-05
Coefficent Effect			
Occ	0.000410	0.000310	-0.000720
Educ	0.000440	-0.000680	0.000250
Female	0.000180	0.000490	-0.000670
Age	-0.000850	0.000450	0.000400
Ind	0.00217	0.00281	-0.00498
Constant	-0.000620	-0.00492	0.00554

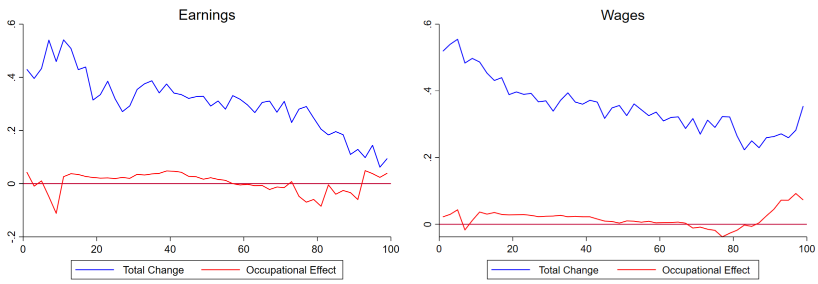
Brazil: 2013-2006



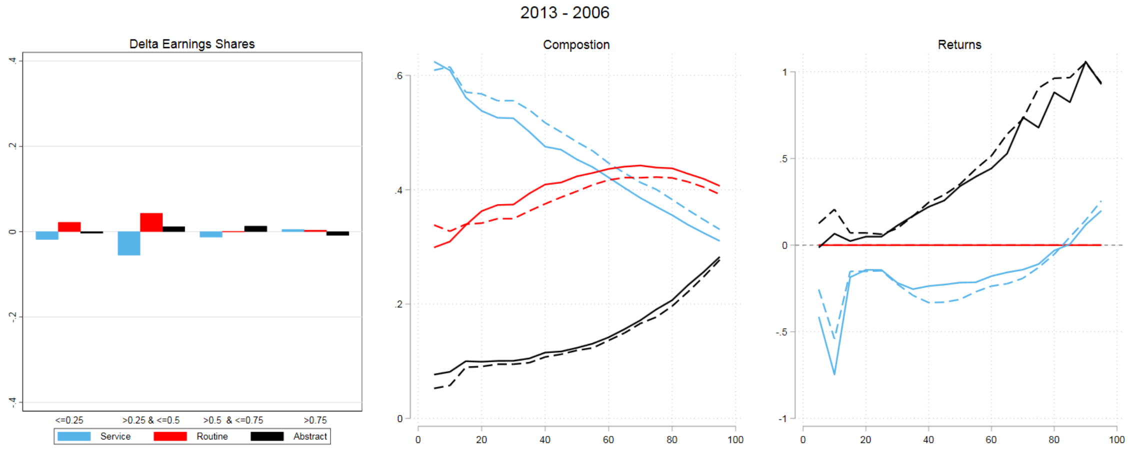
Employment and Income Shares by Occupational Class



Theil Decompostion



Quantile RIF Decompostion

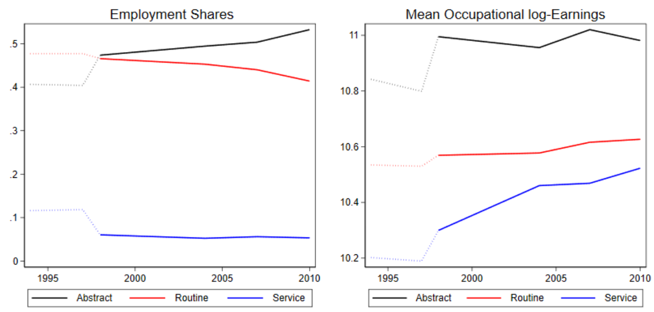


Occupational Classes Compostion and Returns

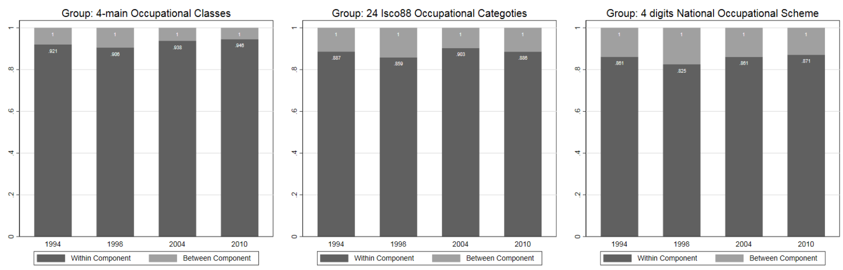
P-Shares Decomposition

Brazil: 2013 - 2006			
	<15	15-85	>85
Δ	0.00313***	0.00024*	-0.00336***
Specification Error	0.00009***	0.00025***	-0.00033***
Composition Effect			
Occ	0.00039***	-0.00009***	-0.00030***
Educ	-0.00014***	-0.00054***	0.00068***
Female	-0.00003***	0.00002***	0
Age	-0.00007***	-0.00001*	0.00008***
Ind	0.00020***	-0.00009***	-0.00011***
Reweighting Error	-5.00e-05	0.00005*	-1.00e-05
Coefficent Effect			
Occ	-0.000240	0.000160	8.00e-05
Educ	0.000190	0.00049**	-0.00068***
Female	0.00049***	-0.00038**	-0.000110
Age	0.000240	0.000140	-0.000380
Ind	3.00e-05	0.00812***	-0.00815***
Constant	0.00201***	-0.00787***	0.00586***

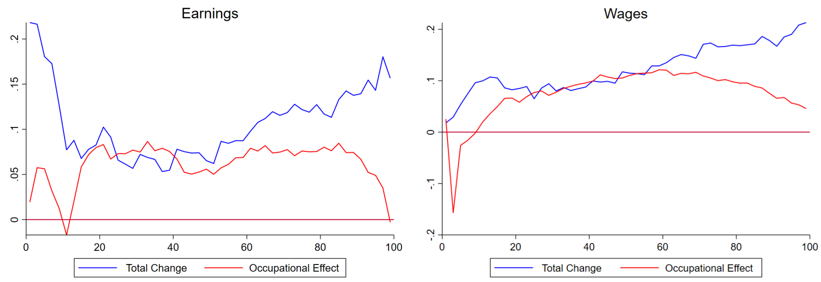
Canada: 2010-1994



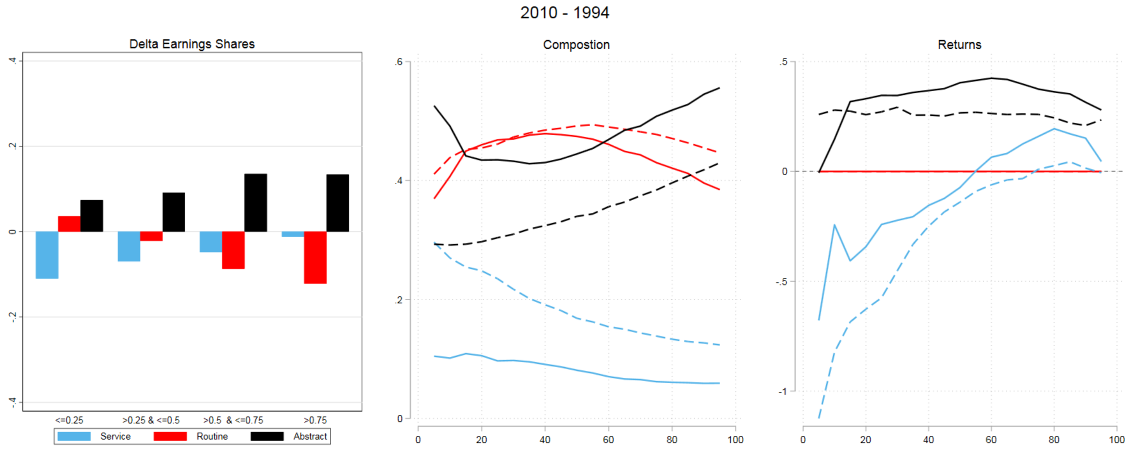
Employment and Income Shares by Occupational Class



Theil Decompostion



Quantile RIF Decompostion

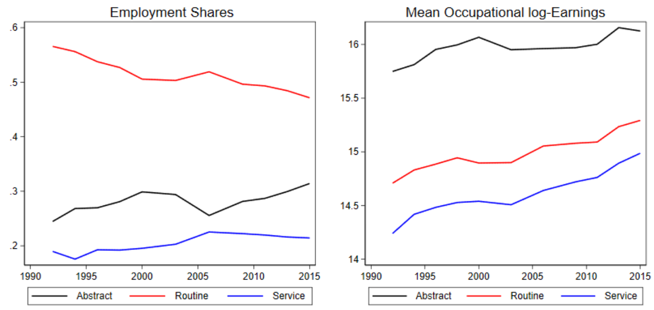


Occupational Classes Compostion and Returns

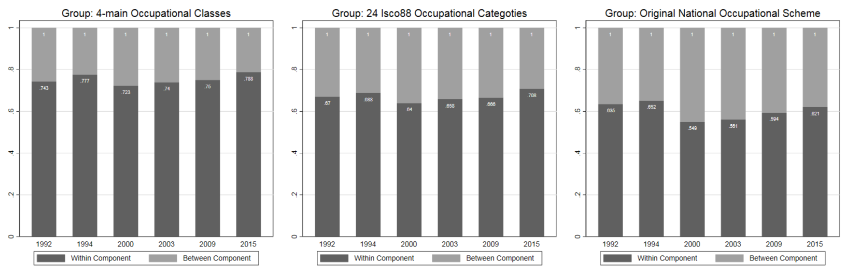
P-Shares Decomposition

Canada: 2010 - 1994			
	<15	15-85	>85
Δ	0.000470	-0.00112***	0.00066***
Specification Error	8.00e-05	0	-0.00007***
Composition Effect			
Occ	0.00032***	-4.00e-05	-0.00027***
Educ	-3.00e-05	0.00018***	-0.00015***
Female	-0.00001**	-0.00001*	0.00002**
Age	0	2.00e-05	-2.00e-05
Ind	-2.00e-05	-1.00e-05	0.00003**
Reweighting Error	-1.00e-05	-1.00e-05	3.00e-05
Coefficent Effect			
Occ	-0.00102	0.00105*	-3.00e-05
Educ	0.00115	-0.000820	-0.000330
Female	0.00110*	-0.000470	-0.00064**
Age	0.00236	-0.00181	-0.000540
Ind	-0.000900	0	0.00090*
Constant	-0.00254	0.000800	0.00174*

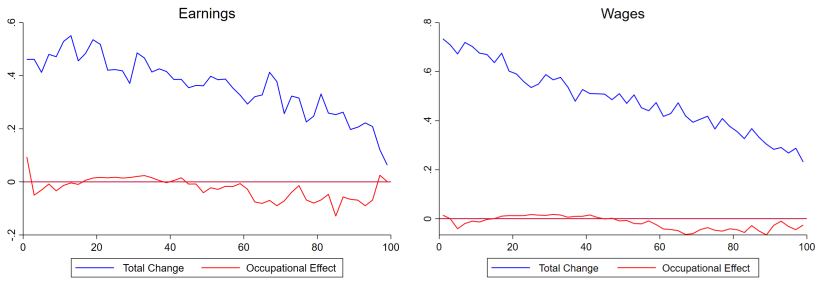
Chile: 2015-1992



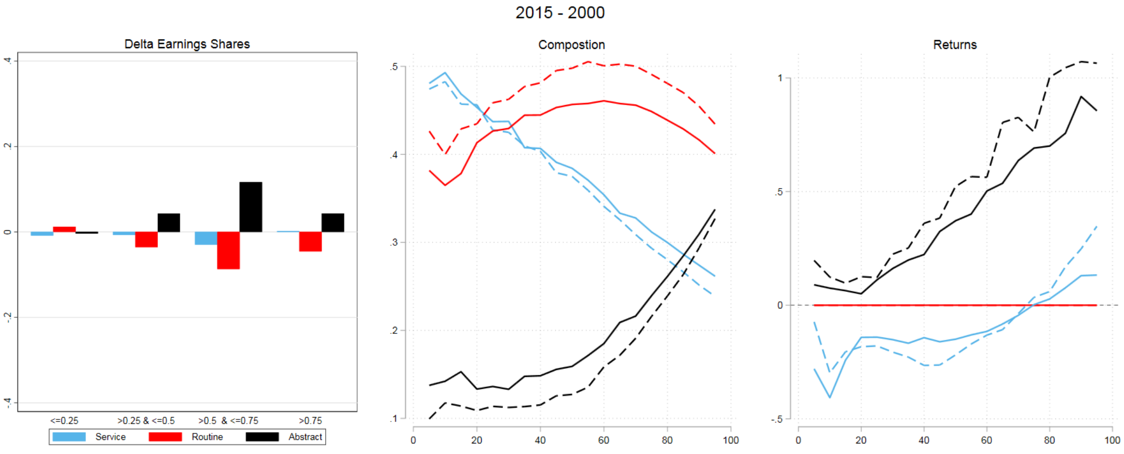
Employment and Income Shares by Occupational Class



Theil Decompostion



Quantile RIF Decompostion

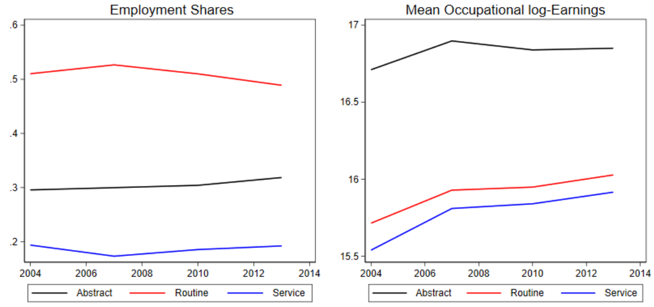


Occupational Classes Compostion and Returns

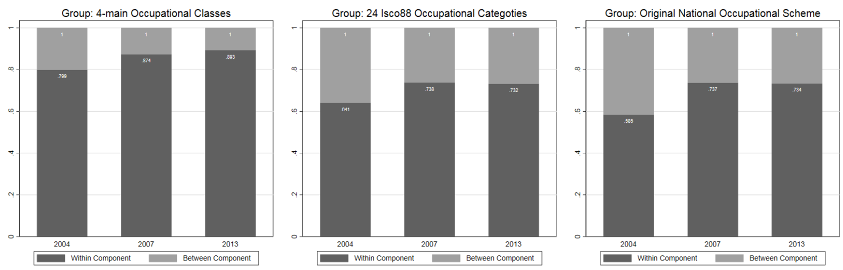
P-Shares Decomposition

Chile: 2015 - 2000			
	<15	15-85	>85
Δ	0.00162***	0.00056***	-0.00219***
Specification Error	0.00004*	6.00e-05	-0.00011**
Composition Effect			
Occ	-0.00007***	-3.00e-05	0.00010***
Educ	-0.00014***	2.00e-05	0.00012**
Female	-0.00002***	0.00004***	-0.00002**
Age	-0.00011***	3.00e-05	0.00007**
Ind	0.00023***	0	-0.00022***
Reweighting Error	-4.00e-05	0.00010***	-6.00e-05
Coefficent Effect			
Occ	0.000240	0.000100	-0.000330
Educ	0.000220	-3.00e-05	-0.000180
Female	-0.00046***	-0.000120	0.00058***
Age	9.00e-05	0.000620	-0.000700
Ind	-0.00120**	0.00235**	-0.00115
Constant	0.00286***	-0.00256*	-0.000300

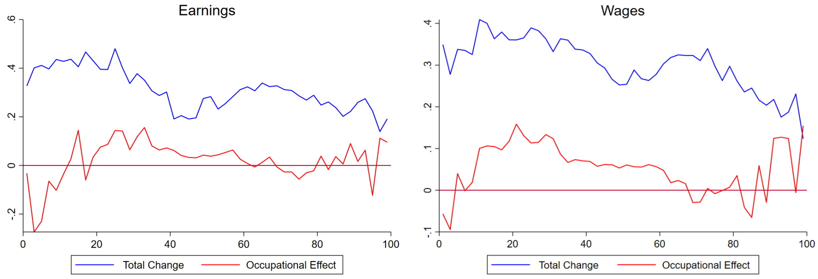
Colombia: 2013-2004



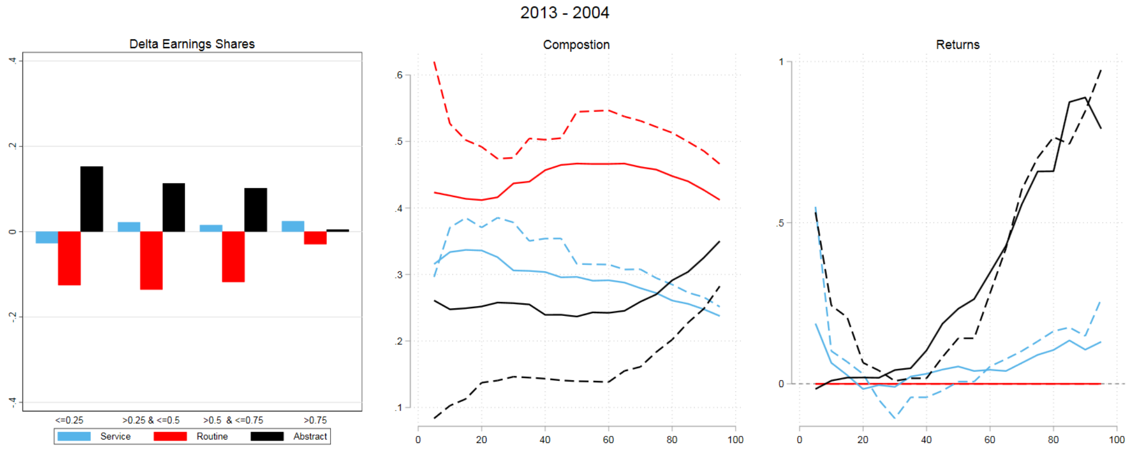
Employment and Income Shares by Occupational Class



Theil Decompostion



Quantile RIF Decompostion

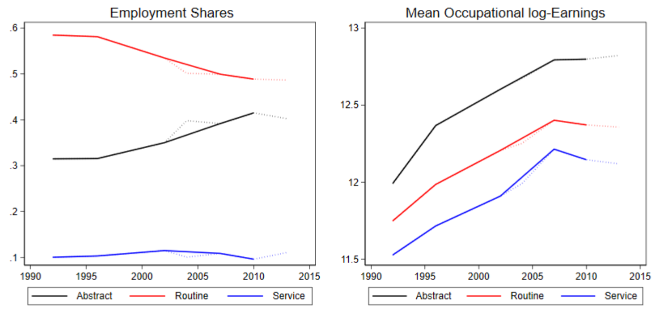


Occupational Classes Compostion and Returns

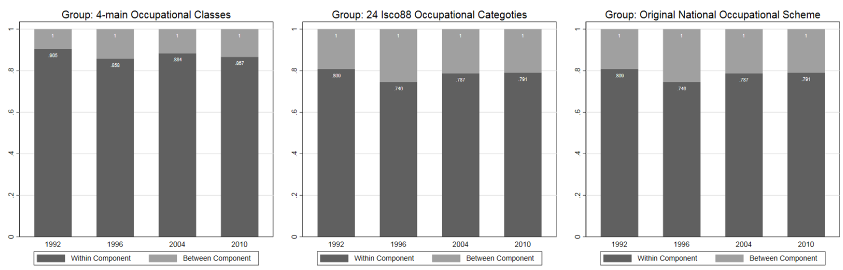
P-Shares Decomposition

Colombia: 2013 - 2004			
	<15	15-85	>85
Δ	0.00100	-3.00e-05	-0.000970
Specification Error	1.00e-05	0.00031*	-0.00032**
Composition Effect			
Occ	-0.000100	4.00e-05	7.00e-05
Educ	-0.00020***	-0.000100	0.00030***
Female	1.00e-05	-1.00e-05	0
Age	-0.00008*	2.00e-05	0.00005*
Ind	0.00026*	-0.00028**	2.00e-05
Reweighting Error	5.00e-05	-4.00e-05	-2.00e-05
Coefficent Effect			
Occ	-0.000650	0.000570	8.00e-05
Educ	0.000260	-0.000230	-3.00e-05
Female	-1.00e-05	-0.000510	0.000510
Age	-0.000180	-0.000300	0.000480
Ind	0.00212	-0.00163	-0.000490
Constant	-0.000500	0.00212	-0.00162

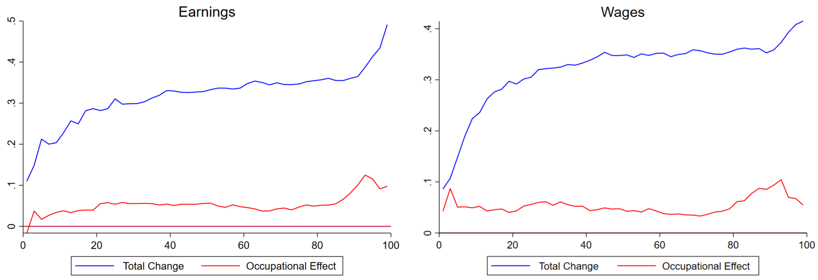
Czech Republic: 2010-1996



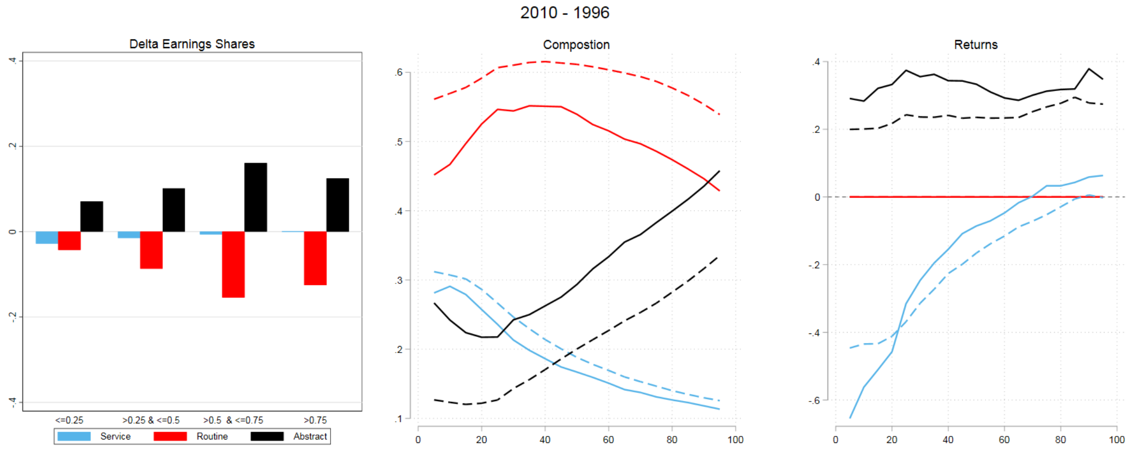
Employment and Income Shares by Occupational Class



Theil Decompostion



Quantile RIF Decompostion

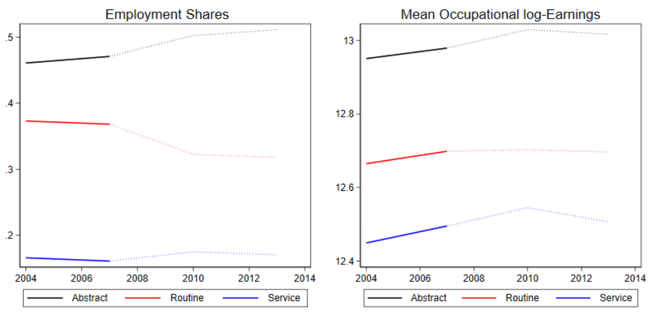


Occupational Classes Compostion and Returns

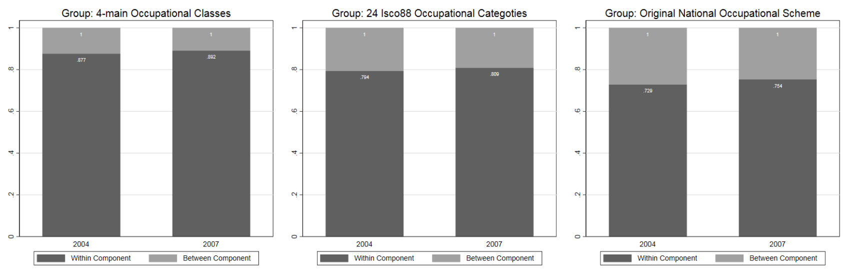
P-Shares Decomposition

Czech Republic: 2010 - 1996			
	<15	15-85	>85
Δ	-0.00125**	0.000490	0.00076*
Specification Error	0.00006**	0.00008**	-0.00014***
Composition Effect			
Occ	-0.00004*	-1.00e-05	0.00005**
Educ	-0.00006*	-0.00014***	0.00020***
Female	0	0.00002***	-0.00002***
Age	-0.00005***	2.00e-05	0.00003**
Ind	-0.00014***	-4.00e-05	0.00017***
Reweighting Error	1.00e-05	1.00e-05	-2.00e-05
Coefficent Effect			
Occ	-0.000240	-0.000200	0.000440
Educ	9.00e-05	-0.000210	0.000120
Female	0.000120	0.000210	-0.000330
Age	-0.000110	-0.000130	0.000240
Ind	-0.00346	0.00193	0.00153
Constant	0.00257	-0.00105	-0.00152

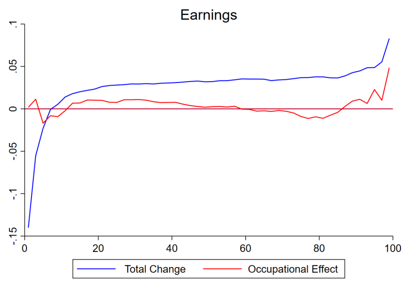
Denmark: 2007-2004



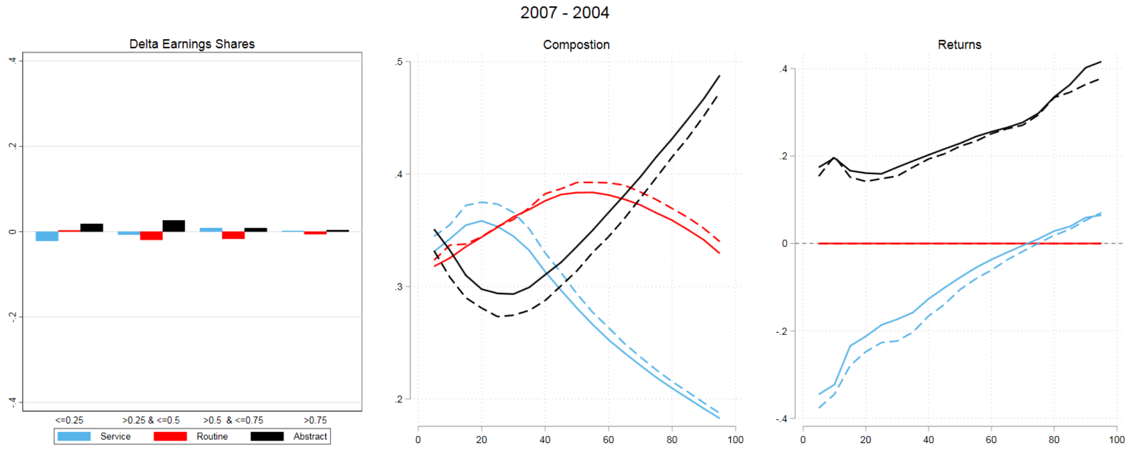
Employment and Income Shares by Occupational Class



Theil Decompostion



Quantile RIF Decompostion

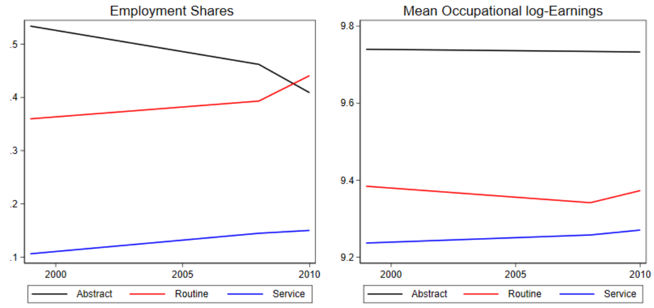


Occupational Classes Compostion and Returns

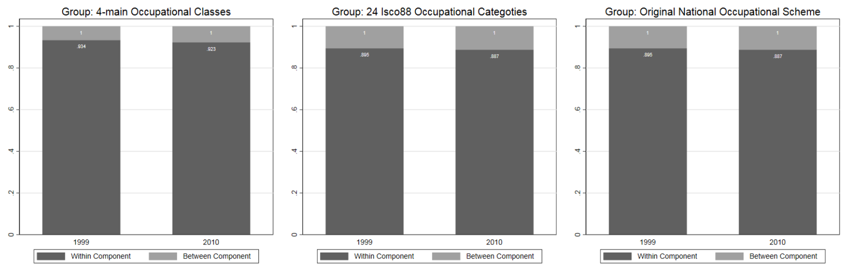
P-Shares Decomposition

Denmark: 2007 - 2004			
	<15	15-85	>85
Δ	-0.00057***	0.00030***	0.00027***
Specification Error	0	1.00e-05	0
Composition Effect			
Occ	0	0	-0.00000*
Educ	-0.00002***	0.00000*	0.00001***
Female	0	0	0.00001***
Age	0.00007***	-0.00005***	-0.00001***
Ind	-0.00004***	-0.00002**	0.00006***
Reweighting Error	0	0	0
Coefficent Effect			
Occ	-7.00e-05	-6.00e-05	0.000130
Educ	-1.00e-05	5.00e-05	-4.00e-05
Female	-5.00e-05	9.00e-05	-4.00e-05
Age	0.000250	-0.00036**	0.000110
Ind	0.00530***	-0.00368***	-0.00162**
Constant	-0.00599***	0.00433***	0.00167**

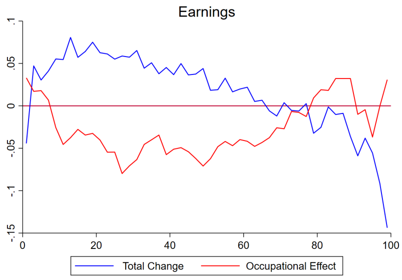
Egypt: 2010-1999



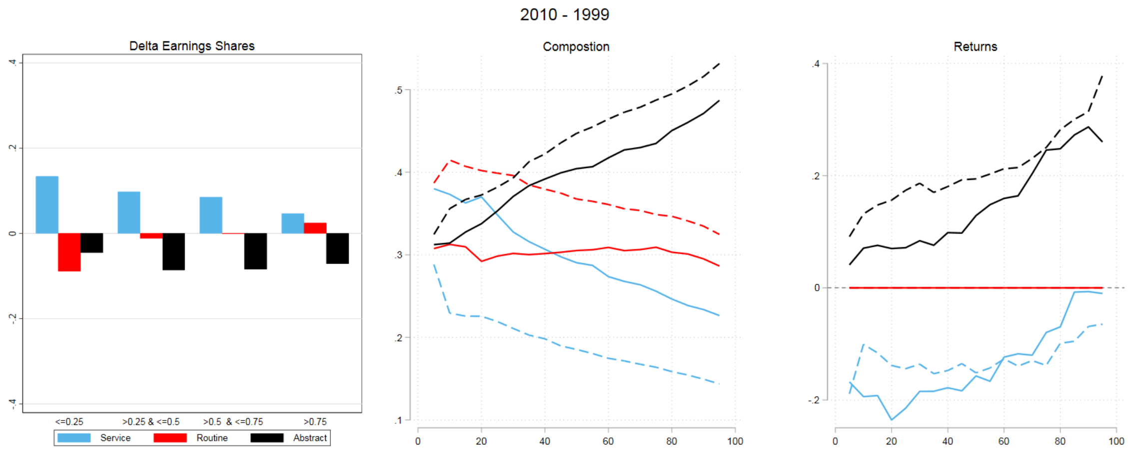
Employment and Income Shares by Occupational Class



Theil Decompostion



Quantile RIF Decompostion

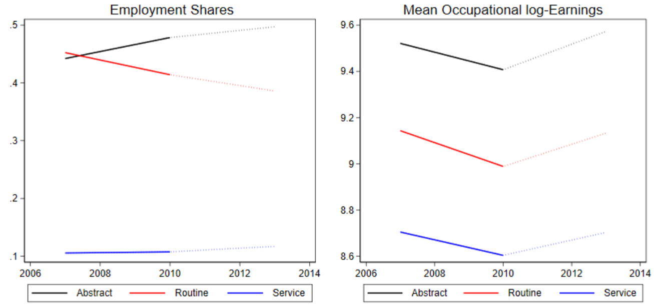


Occupational Classes Compostion and Returns

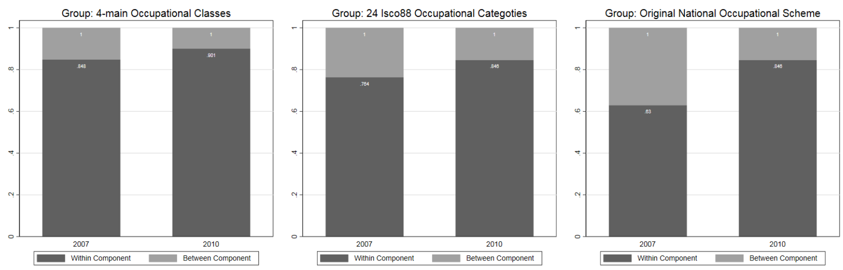
P-Shares Decomposition

Egypt: 2010 - 1999			
	<15	>85	15-85
Δ	0.000200	-0.0012*	0.00100
Specification Error	0	0	0
Composition Effect			
Occ	0.0003***	-0.0003***	-0.000100
Educ	0.0001**	-0.0001**	0
Female	0.0003***	-0.0001***	-0.0002***
Age	-0.0001***	0.0000*	0.0001*
Ind	-0.0001*	0.0001***	0
Reweighting Error	-0.000200	0.000100	0.000100
Coefficent Effect			
Occ	-0.000100	0.000900	-0.000800
Educ	0.000400	0.000100	-0.000500
Female	-0.000200	0	0.000200
Age	0.00170	-0.000500	-0.00110
Ind	-0.000700	-0.00100	0.00180
Constant	-0.00120	-0.000400	0.00150

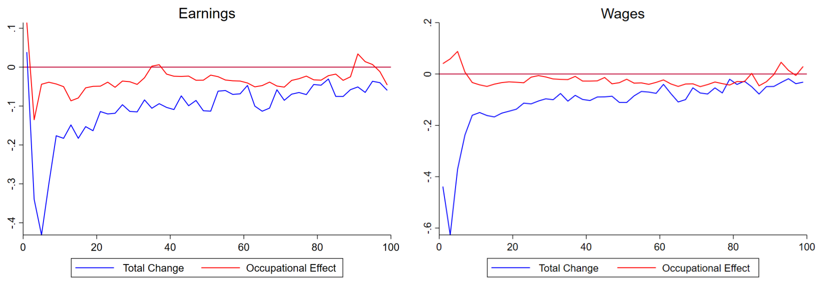
Estonia: 2010-2007



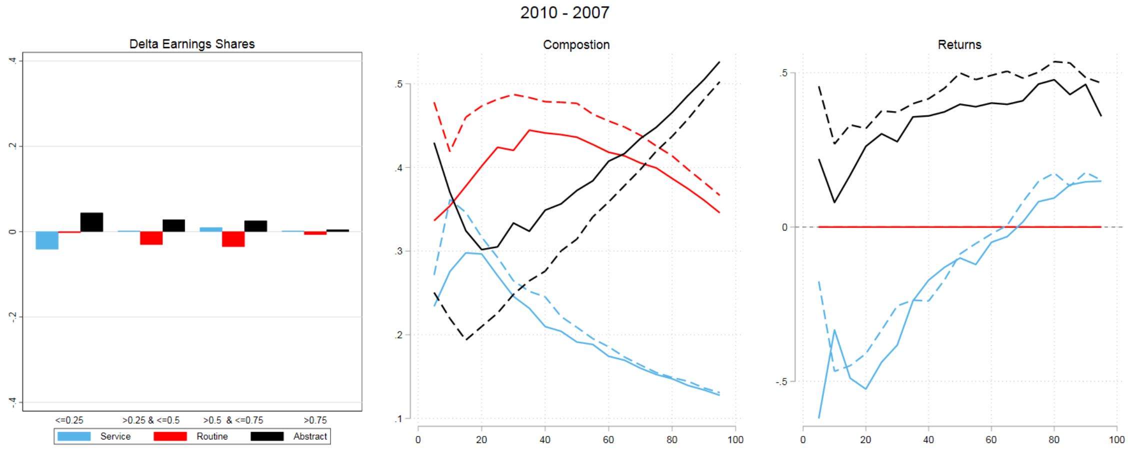
Employment and Income Shares by Occupational Class



Theil Decompostion



Quantile RIF Decompostion

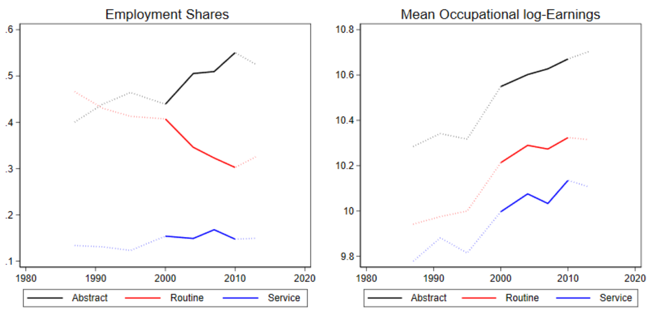


Occupational Classes Compostion and Returns

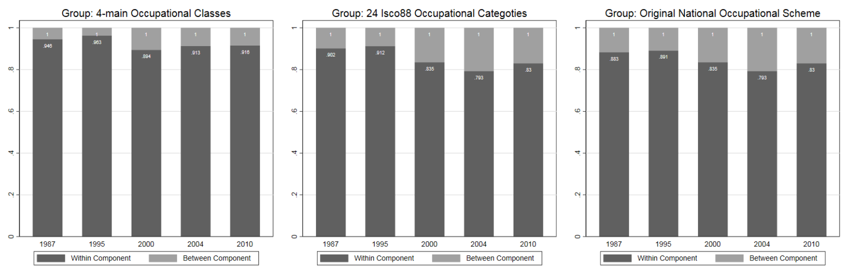
P-Shares Decomposition

Estonia: 2010 - 2007			
	<15	15-85	>85
Δ	-0.00234**	0.00127*	0.00107*
Specification Error	2.00e-05	-4.00e-05	1.00e-05
Composition Effect			
Occ	0	2.00e-05	-2.00e-05
Educ	-0.00014*	0.00022**	-7.00e-05
Female	0.00004*	-0.00006**	0.00002*
Age	-1.00e-05	2.00e-05	-2.00e-05
Ind	0.000110	-6.00e-05	-5.00e-05
Reweighting Error	-2.00e-05	8.00e-05	-5.00e-05
Coefficent Effect			
Occ	-0.000490	-0.000280	0.000780
Educ	0	-0.000790	0.000800
Female	-0.00130	0.00105	0.000240
Age	-8.00e-05	-0.000560	0.000630
Ind	-0.00128	-0.000600	0.00188
Constant	0.000820	0.00226	-0.00307

Finland: 2010-1991

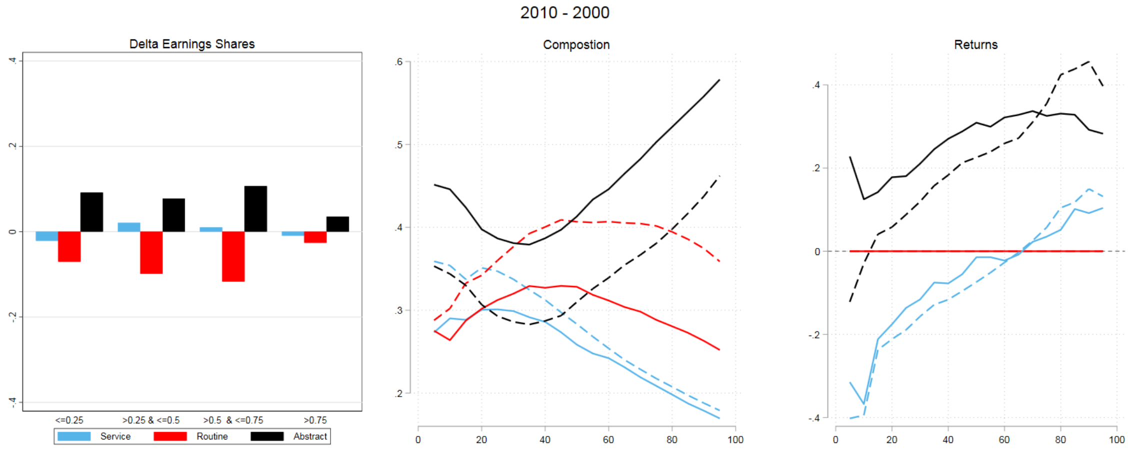


Employment and Income Shares by Occupational Class



Theil Decompostion

Quantile RIF Decompostion

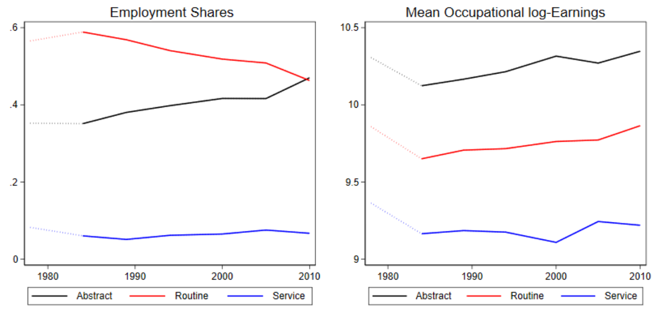


Occupational Classes Compostion and Returns

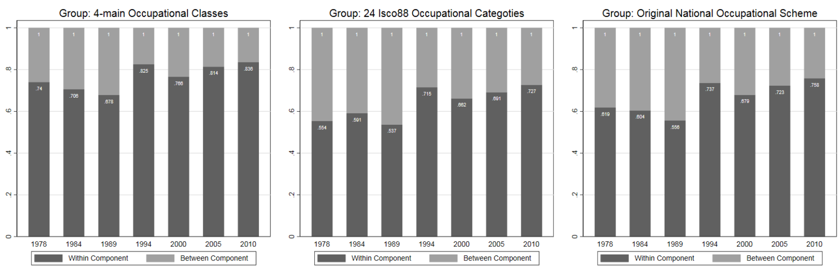
P-Shares Decomposition

Finland: 2010 - 2000			
	<15	15-85	>85
Δ	-0.00250***	0.00205***	0.00045*
Specification Error	4.00e-05	3.00e-05	-0.00007*
Composition Effect			
Occ	-0.00025***	7.00e-05	0.00018***
Educ	0.00018*	-0.00019**	2.00e-05
Female	2.00e-05	-1.00e-05	-1.00e-05
Age	2.00e-05	1.00e-05	-0.00003*
Ind	0.00011*	-0.00010**	-1.00e-05
Reweighting Error	-1.00e-05	0	1.00e-05
Coefficent Effect			
Occ	0.00138	-0.000780	-0.000600
Educ	-0.000640	0.000290	0.000360
Female	-0.000180	-1.00e-05	0.000200
Age	0.00108	-0.00113	5.00e-05
Ind	-0.000100	0.00173	-0.00163
Constant	-0.00414	0.00216	0.00197

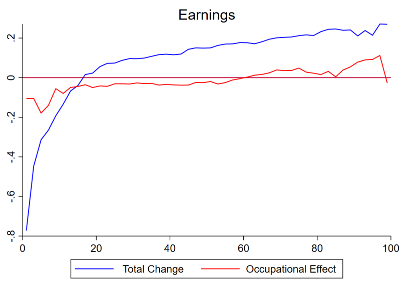
France: 2010-1994



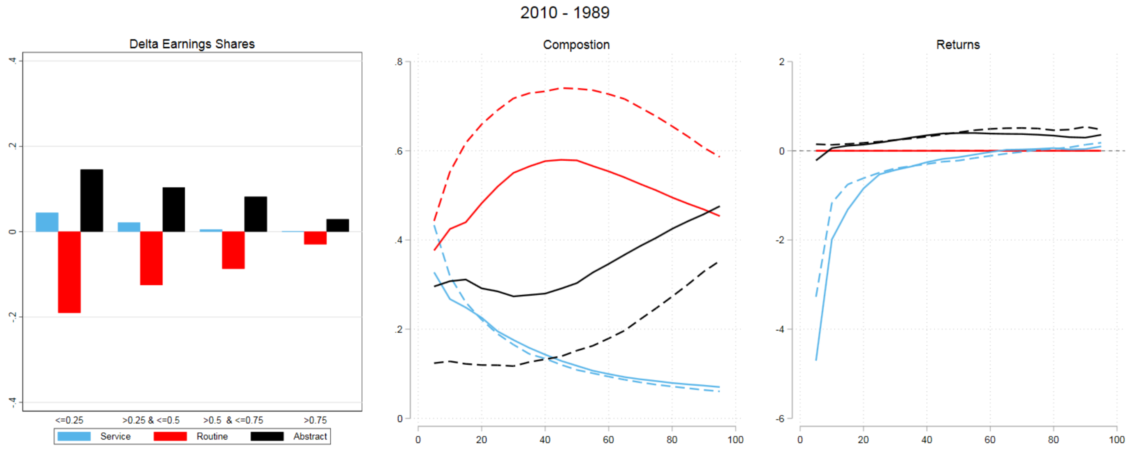
Employment and Income Shares by Occupational Class



Theil Decompostion



Quantile RIF Decompostion

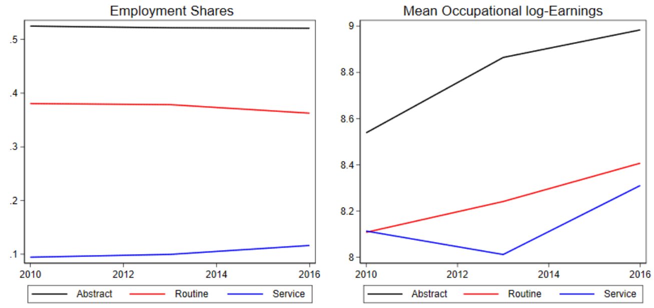


Occupational Classes Compostion and Returns

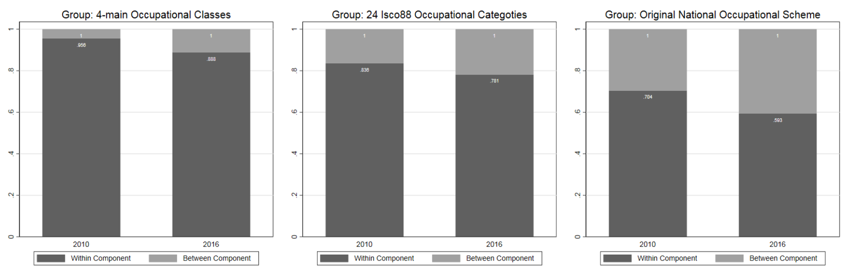
P-Shares Decomposition

France: 2010 - 1989			
	<15	15-85	>85
Δ	-0.00589***	0.00368***	0.00220***
Specification Error	1.00e-05	0.00070*	-0.00071*
Composition Effect			
Occ	-0.00060***	0.00032**	0.00028**
Educ	-0.000190	-0.00108***	0.00127***
Female	4.00e-05	-5.00e-05	0
Age	-0.000140	-0.000120	0.00026**
Ind	-0.00041*	9.00e-05	0.00033**
Reweighting Error	0.000130	-0.000130	-1.00e-05
Coefficent Effect			
Occ	-0.000860	-9.00e-05	0.000950
Educ	0.00201*	-0.000800	-0.00121*
Female	-9.00e-05	-0.000450	0.000540
Age	0.00210	-0.00188	-0.000220
Ind	-1.00e-05	0.00166	-0.00165
Constant	-0.00787*	0.00549	0.00237

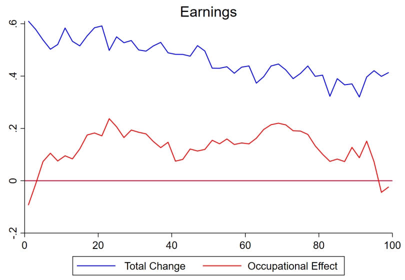
Georgia: 2016-2010



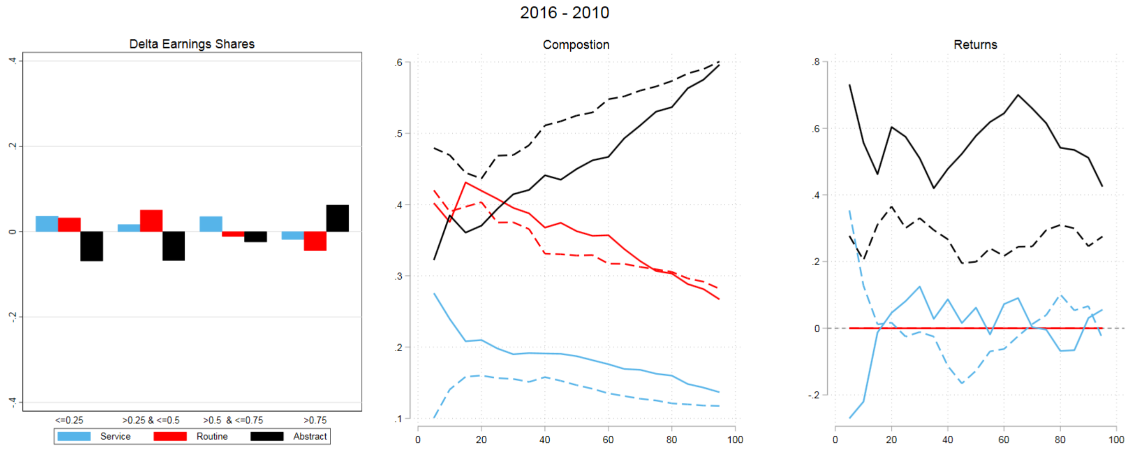
Employment and Income Shares by Occupational Class



Theil Decompostion



Quantile RIF Decompostion

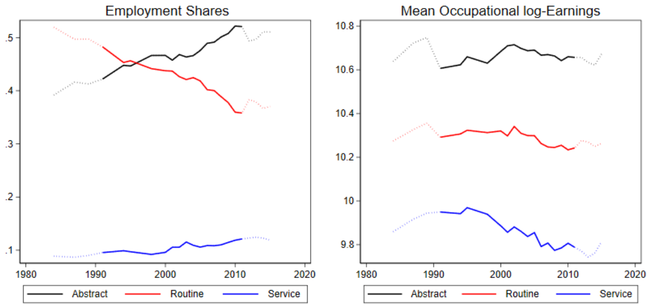


Occupational Classes Compostion and Returns

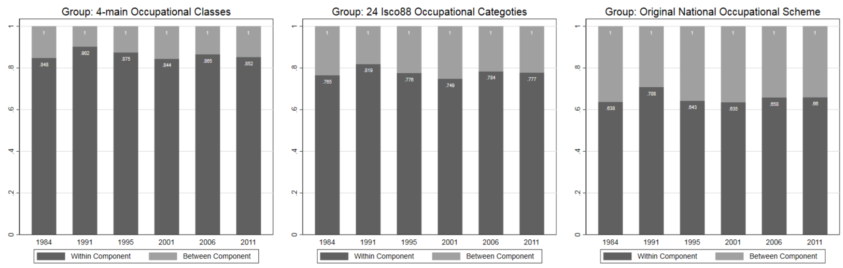
P-Shares Decomposition

Georgia: 2016 - 2010			
	<15	15-85	>85
Δ	0.00337**	-0.000250	-0.00312***
Specification Error	5.00e-05	0	-5.00e-05
Composition Effect			
Occ	0.00034**	-0.000130	-0.00021**
Educ	7.00e-05	5.00e-05	-0.00012**
Female	-1.00e-05	0.00005**	-0.00004**
Age	2.00e-05	-1.00e-05	-1.00e-05
Ind	-5.00e-05	0.000180	-0.000130
Reweighting Error	6.00e-05	-3.00e-05	-3.00e-05
Coefficent Effect			
Occ	-0.000750	0.00248	-0.00173
Educ	-0.00247	0.00121	0.00126
Female	-0.00179	0.000880	0.000910
Age	-0.00117	0.00130	-0.000130
Ind	0.00496	0.00412	-0.00908*
Constant	0.00412	-0.01036*	0.00624

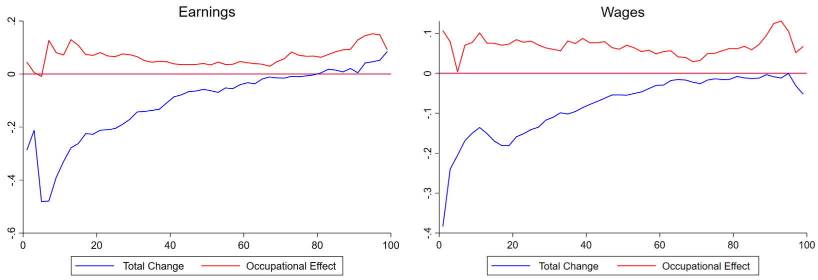
Germany: 2011-1991



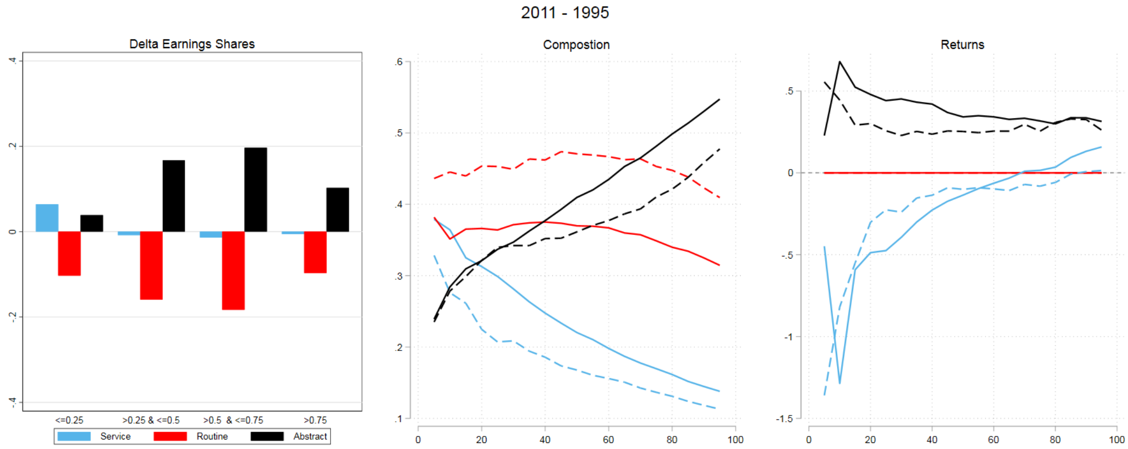
Employment and Income Shares by Occupational Class



Theil Decompostion



Quantile RIF Decompostion

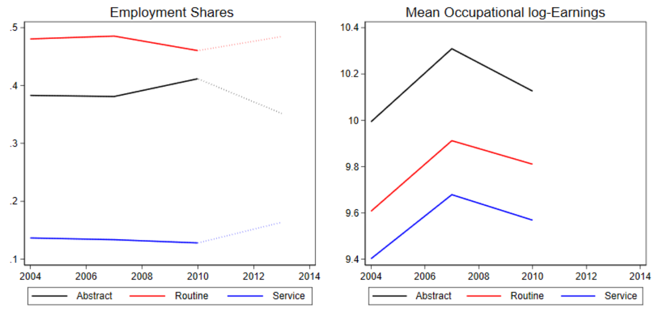


Occupational Classes Compostion and Returns

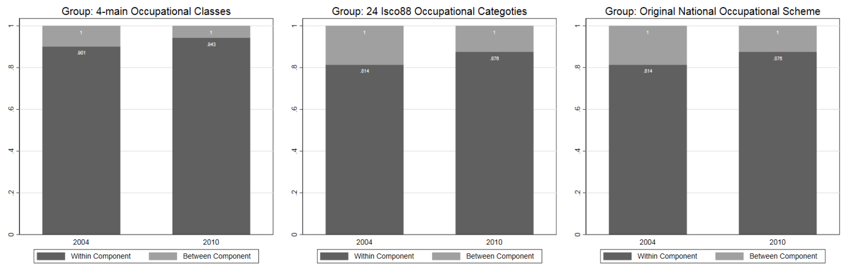
P-Shares Decomposition

Germany: 2011 - 1995			
	<15	15-85	>85
Δ	-0.00390***	0.00163**	0.00227***
Specification Error	5.00e-05	5.00e-05	-0.000100
Composition Effect			
Occ	-0.000150	0	0.00014**
Educ	-2.00e-05	-3.00e-05	4.00e-05
Female	-6.00e-05	3.00e-05	3.00e-05
Age	-0.00060***	0.00042**	0.000180
Ind	2.00e-05	0	-2.00e-05
Reweighting Error	-0.00046**	0.000190	0.00027**
Coefficent Effect			
Occ	0.000270	-0.000820	0.000550
Educ	0.000630	-0.000590	-4.00e-05
Female	0.000750	-0.000680	-7.00e-05
Age	-2.00e-05	-0.000560	0.000590
Ind	-0.00482	0.00394	0.000890
Constant	0.000520	-0.000330	-0.000200

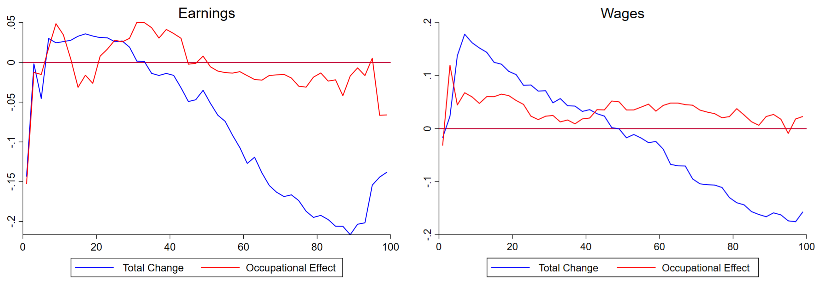
Greece: 2010-2007



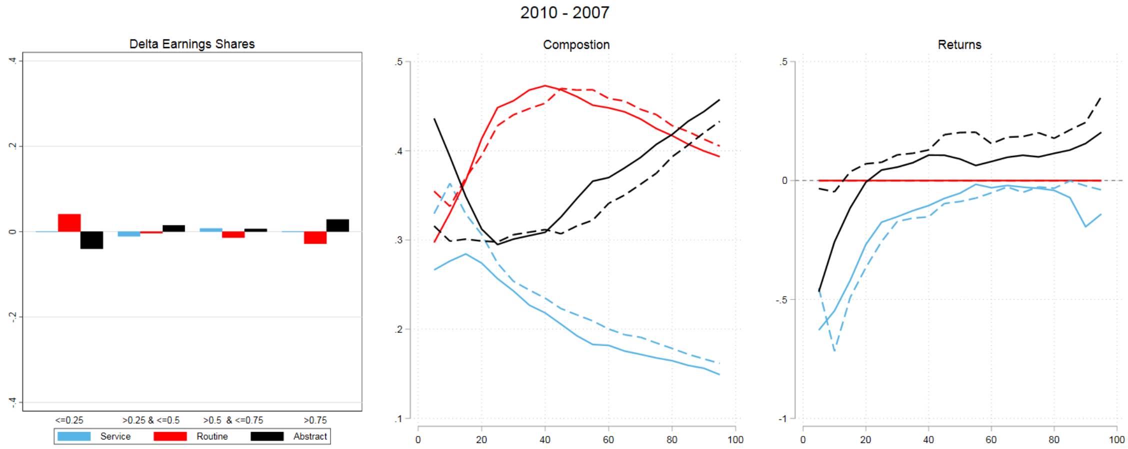
Employment and Income Shares by Occupational Class



Theil Decompostion



Quantile RIF Decompostion

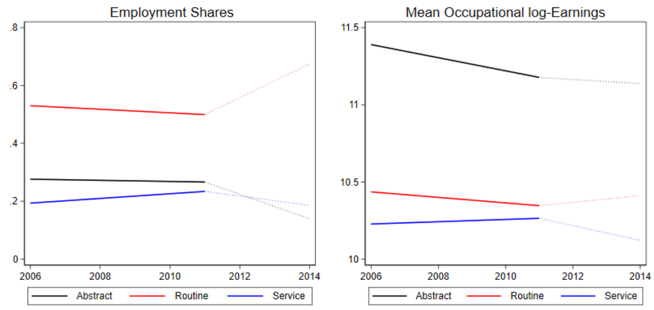


Occupational Classes Compostion and Returns

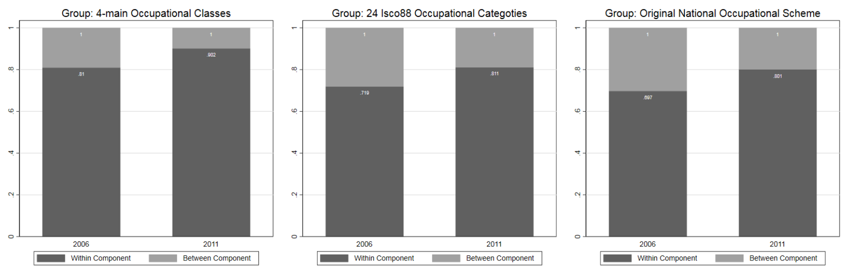
P-Shares Decomposition

Greece: 2010 - 2007			
	<15	15-85	>85
Δ	0.000970	0.000550	-0.00152***
Specification Error	2.00e-05	0	-2.00e-05
Composition Effect			
Occ	5.00e-05	-5.00e-05	-1.00e-05
Educ	0.000130	-7.00e-05	-5.00e-05
Female	-0.00003**	0.00002*	0.00001*
Age	-4.00e-05	0.00008**	-4.00e-05
Ind	-0.00015*	8.00e-05	7.00e-05
Reweighting Error	1.00e-05	1.00e-05	-1.00e-05
Coefficent Effect			
Occ	-7.00e-05	0.000460	-0.000400
Educ	0.00110	-0.000880	-0.000220
Female	-0.000200	0.000760	-0.000550
Age	8.00e-05	-0.00104	0.000960
Ind	-0.00196	0.00773**	-0.00575
Constant	0.00203	-0.00653*	0.00449

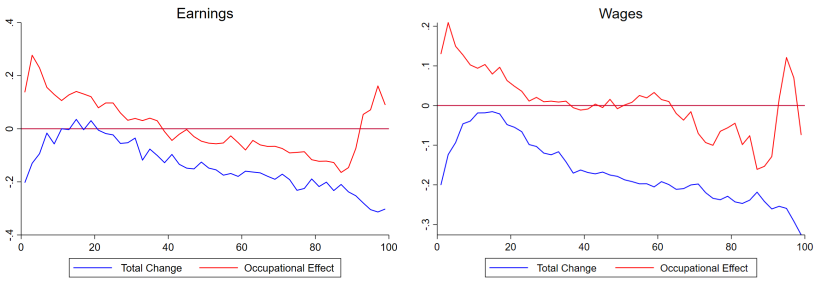
Guatemala: 2011-2006



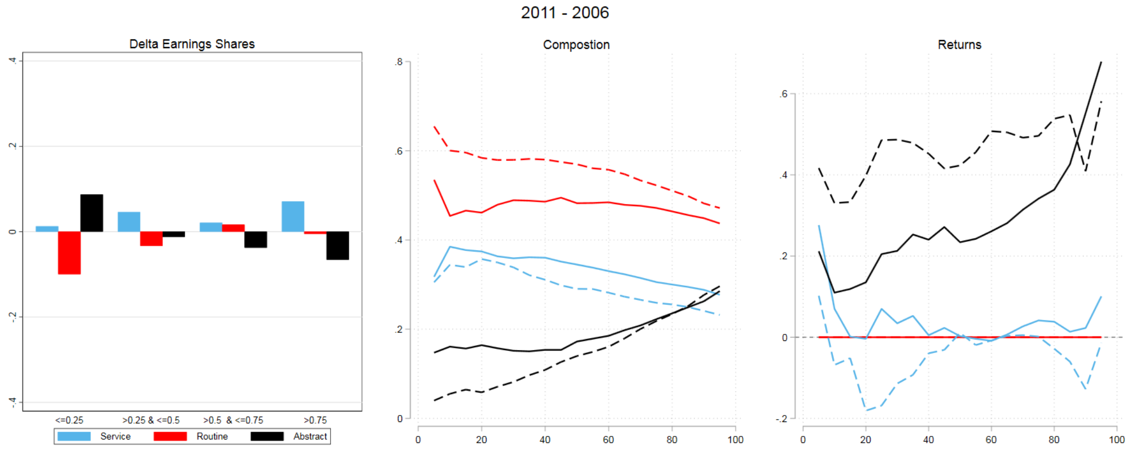
Employment and Income Shares by Occupational Class



Theil Decompostion



Quantile RIF Decompostion

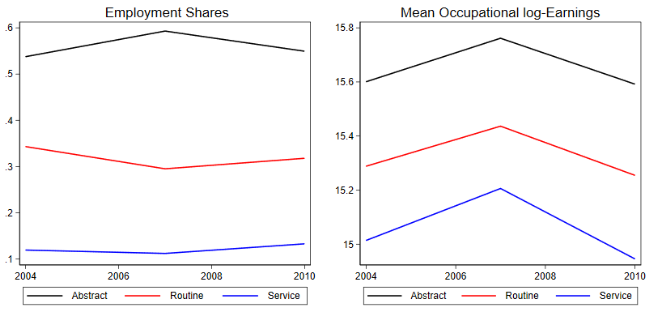


Occupational Classes Compostion and Returns

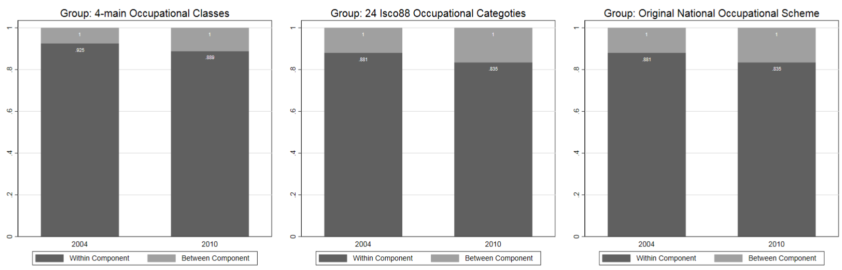
P-Shares Decomposition

Guatemala: 2011 - 2006			
	<15	15-85	>85
Δ	0.000790	0.000770	-0.00155***
Specification Error	1.00e-05	-2.00e-05	1.00e-05
Composition Effect			
Occ	0	-5.00e-05	4.00e-05
Educ	4.00e-05	0.000100	-0.000140
Female	0.00031***	-9.00e-05	-0.00022***
Age	2.00e-05	2.00e-05	-5.00e-05
Ind	-0.00065***	0.000290	0.000360
Reweighting Error	-0.00058***	0.000100	0.00049*
Coefficent Effect			
Occ	0.00257	-0.00214	-0.000440
Educ	-0.000590	-0.000990	0.00158
Female	-0.000730	0.000370	0.000350
Age	-0.00243**	0.000750	0.00169
Ind	-0.000680	-0.00342	0.00411
Constant	0.00350	0.00585	-0.00935*

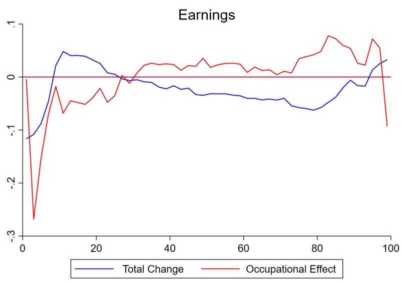
Iceland: 2010-2004



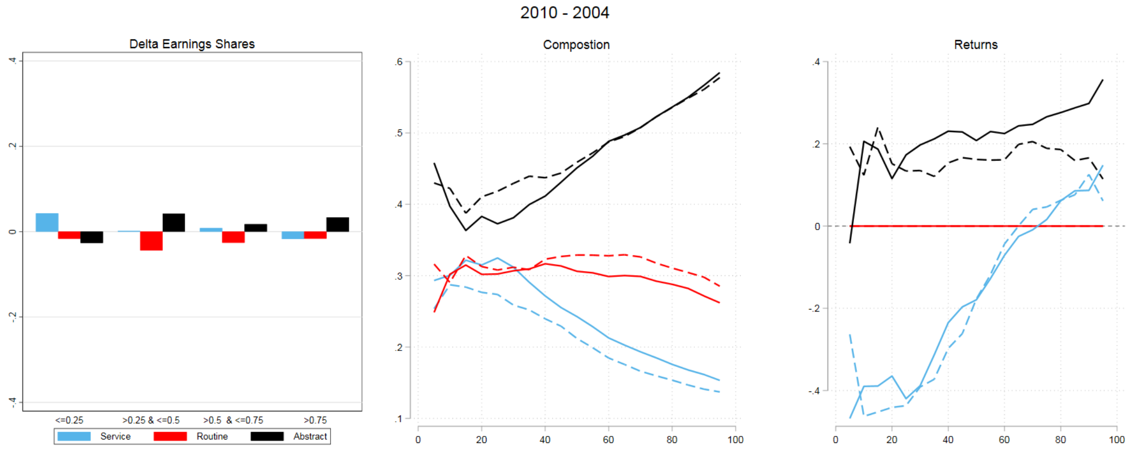
Employment and Income Shares by Occupational Class



Theil Decompostion



Quantile RIF Decompostion

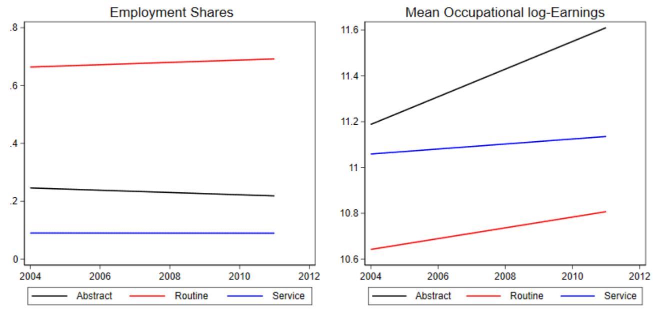


Occupational Classes Compostion and Returns

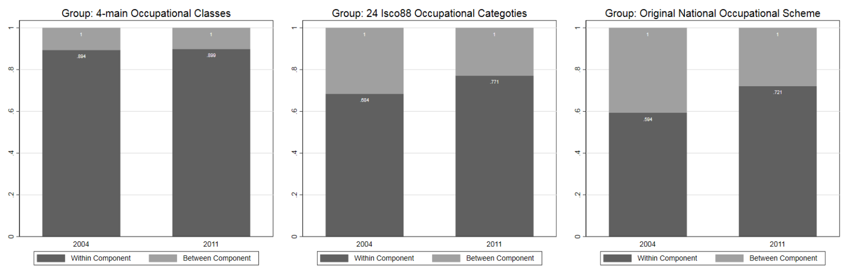
P-Shares Decomposition

Iceland: 2010 - 2004			
	<15	15-85	>85
Δ	-9.00e-05	-0.000120	0.000210
Specification Error	2.00e-05	0	-2.00e-05
Composition Effect			
Occ	-5.00e-05	-2.00e-05	0.00007***
Educ	-7.00e-05	4.00e-05	3.00e-05
Female	0	-1.00e-05	1.00e-05
Age	6.00e-05	-2.00e-05	-0.00004*
Ind	0.000100	-0.000110	1.00e-05
Reweighting Error	1.00e-05	-1.00e-05	0
Coefficent Effect			
Occ	-0.00113	0.000810	0.000320
Educ	0.000190	0.000140	-0.000330
Female	-0.000310	0.000250	6.00e-05
Age	-0.00104	-0.000310	0.00135*
Ind	0.00453*	-0.00370	-0.000830
Constant	-0.00240	0.00281	-0.000420

India: 2011-2004

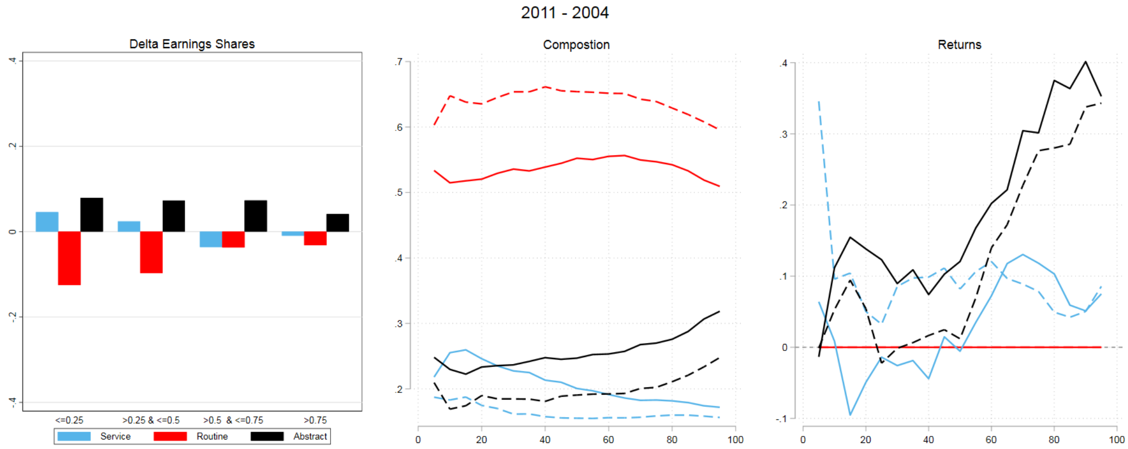


Employment and Income Shares by Occupational Class



Theil Decompostion

Quantile RIF Decompostion

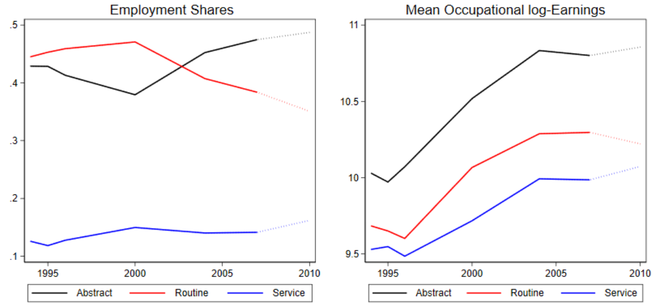


Occupational Classes Compostion and Returns

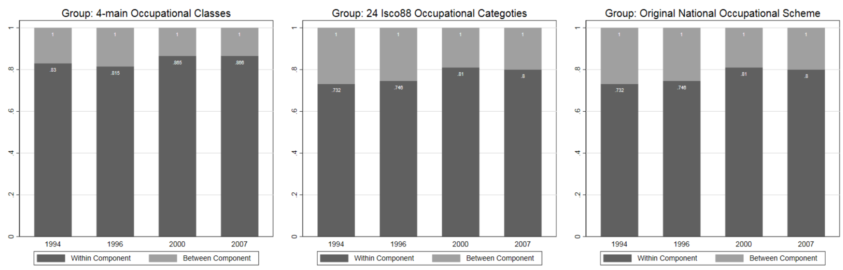
P-Shares Decomposition

India: 2011 - 2004			
	<15	15-85	>85
Δ	-0.00426***	0.00461***	-0.000350
Specification Error	7.00e-05	-5.00e-05	-2.00e-05
Composition Effect			
Occ	0	0.00028***	-0.00028***
Educ	-0.00011***	0	0.00011***
Female	0.00002**	0.00011***	-0.00012***
Age	-0.00008***	-0.00004**	0.00012***
Ind	0.000130	-5.00e-05	-8.00e-05
Reweighting Error	-1.00e-05	-0.000160	0.000180
Coefficent Effect			
Occ	4.00e-05	0.00223	-0.00227*
Educ	-0.000450	0.00355***	-0.00310***
Female	0.000340	7.00e-05	-0.00041***
Age	-0.000900	0.000210	0.00069*
Ind	-0.00355*	0.00460*	-0.00104
Constant	0.000260	-0.00614**	0.00588**

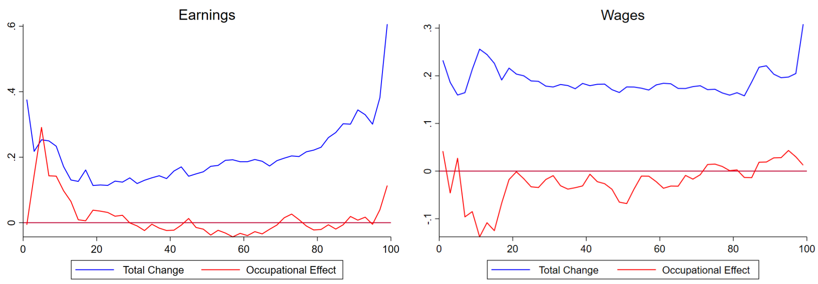
Ireland: 2010-2004



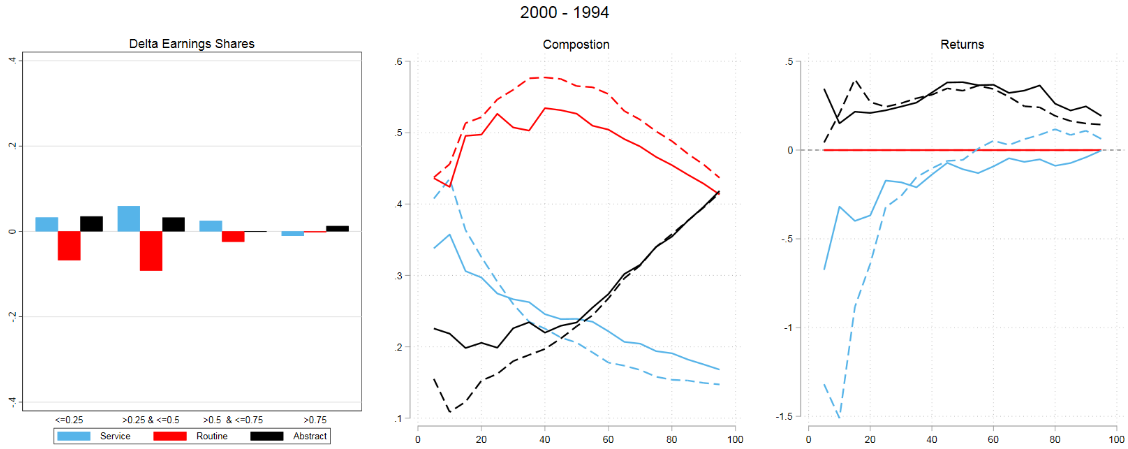
Employment and Income Shares by Occupational Class



Theil Decompostion



Quantile RIF Decompostion

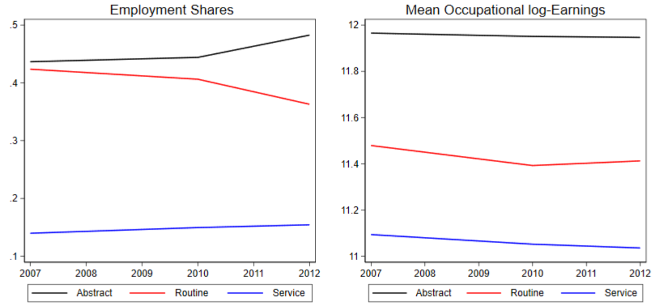


Occupational Classes Compostion and Returns

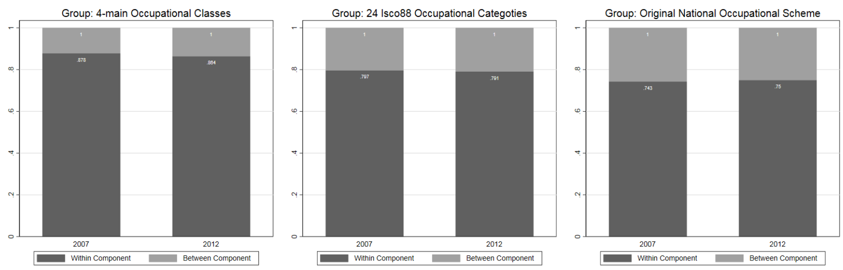
P-Shares Decomposition

Ireland: 2000 - 1994			
	<15	15-85	>85
Δ	0.000690	-0.00284***	0.00215***
Specification Error	4.00e-05	-3.00e-05	-1.00e-05
Composition Effect			
Occ	-7.00e-05	-6.00e-05	0.000130
Educ	-0.000120	0.000120	0
Female	0.00030***	-0.00014**	-0.00016***
Age	-1.00e-05	-7.00e-05	0.00008*
Ind	-0.00112**	0.00050*	0.00062***
Reweighting Error	-0.000320	0.000170	0.000140
Coefficent Effect			
Occ	0.00144	-0.00132	-0.000120
Educ	0.00166	-0.000570	-0.00109
Female	0.00234*	-0.00183*	-0.000510
Age	-0.00395	0.00334*	0.000610
Ind	-0.02253**	0.0105	0.01202***
Constant	0.02304**	-0.01350*	-0.00956*

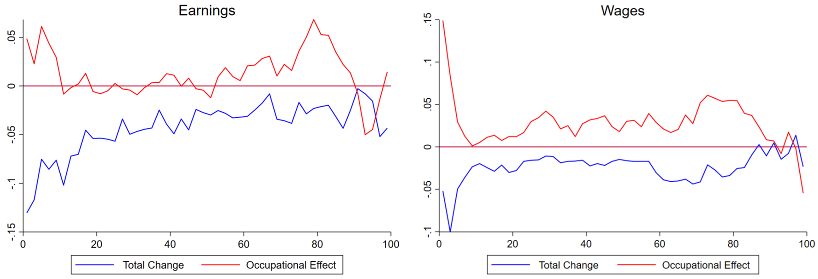
Israel: 2012-2007



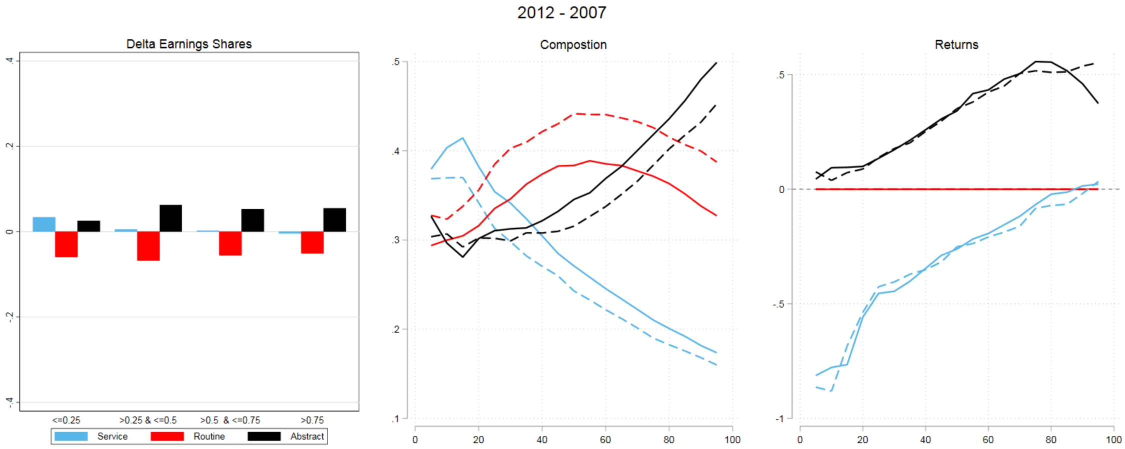
Employment and Income Shares by Occupational Class



Theil Decompostion



Quantile RIF Decompostion

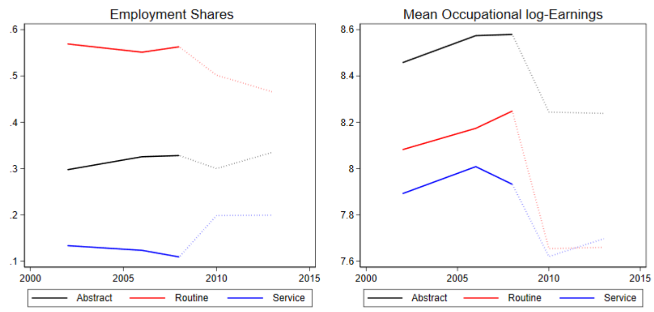


Occupational Classes Compostion and Returns

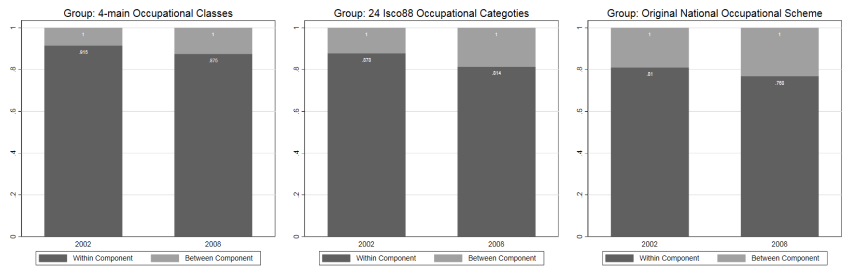
P-Shares Decomposition

Israel: 2012 - 2007			
	<15	15-85	>85
Δ	-0.00081*	0.000590	0.000220
Specification Error	0	2.00e-05	-2.00e-05
Composition Effect			
Occ	-0.00021***	0.00007*	0.00014***
Educ	4.00e-05	1.00e-05	-0.00005**
Female	-1.00e-05	0	0
Age	-2.00e-05	2.00e-05	0
Ind	-4.00e-05	1.00e-05	0.00004*
Reweighting Error	-1.00e-05	0	0
Coefficent Effect			
Occ	0.000420	-0.000100	-0.000320
Educ	0.000590	-0.000260	-0.000340
Female	0.000120	0	-0.000110
Age	-0.000340	0.000510	-0.000180
Ind	-0.00492*	-7.00e-05	0.00499**
Constant	0.00356	0.000370	-0.00393*

Jordan: 2008-2002

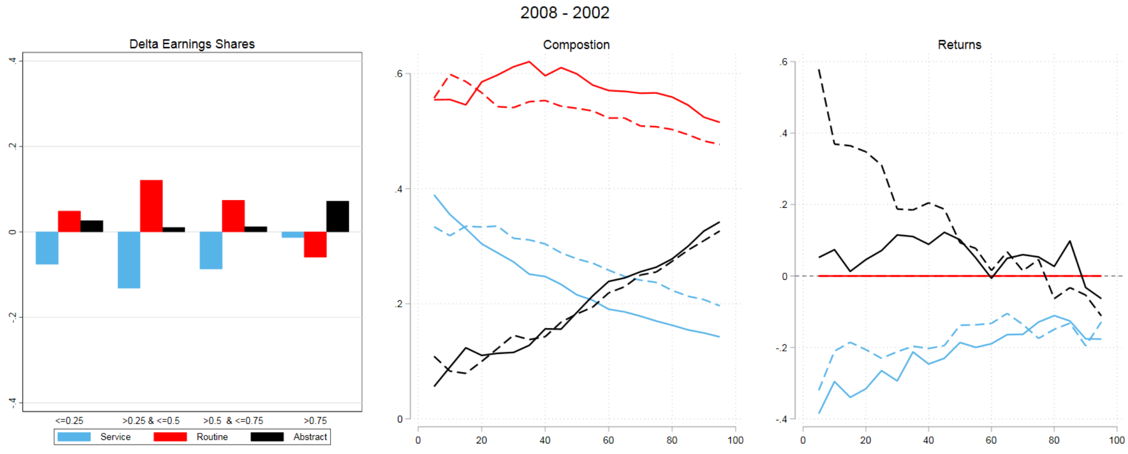


Employment and Income Shares by Occupational Class



Theil Decompostion

Quantile RIF Decompostion

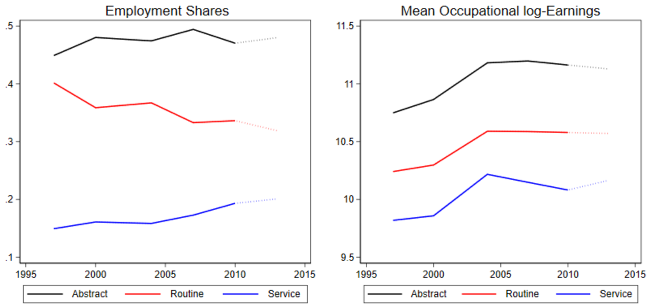


Occupational Classes Compostion and Returns

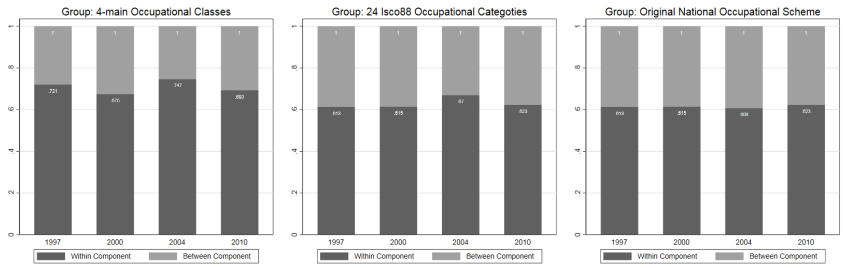
P-Shares Decomposition

Jordan: 2008 - 2002			
	<15	>85	15-85
Δ	0.0074***	-0.0022*	-0.0052***
Specification Error	-0.000100	-0.000200	0.000300
Composition Effect			
Occ	0.000300	0	-0.000300
Educ	-0.000100	0.0002***	-0.000100
Female	0	0	0
Age	0	0	-0.000100
Ind	-0.000300	0.0002*	0.000100
Reweighting Error	0.000400	0	-0.000400
Coefficent Effect			
Occ	-0.000400	-0.000300	0.000700
Educ	0.00120	-0.00220	0.00100
Female	-0.000100	0	0
Age	0.00110	-0.00120	0.000100
Ind	0.000300	-0.000400	0.000100
Constant	0.00510	0.00160	-0.00670

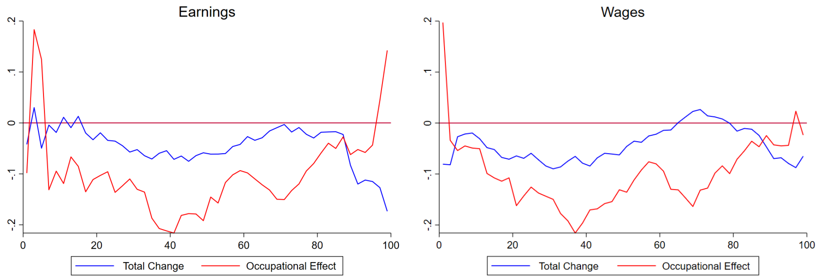
Luxembourg: 2010-2004



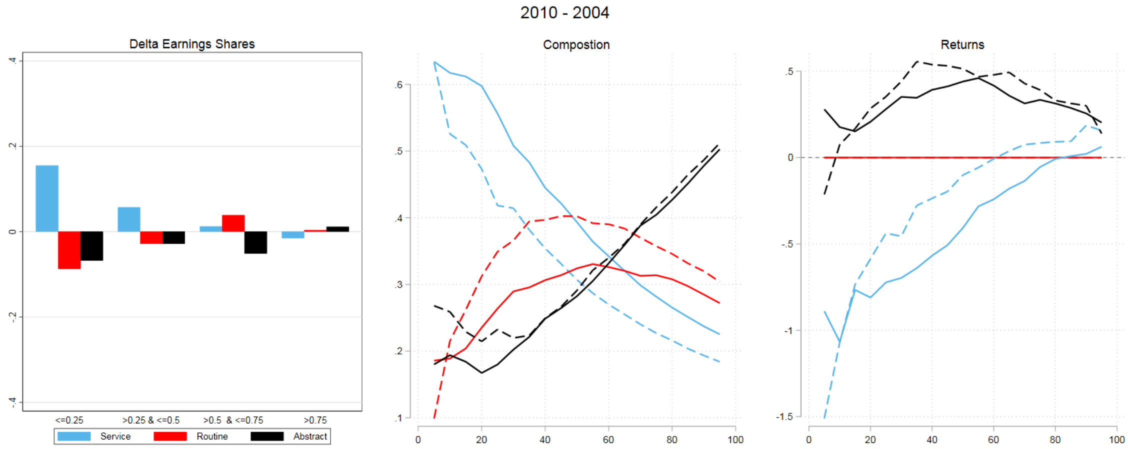
Employment and Income Shares by Occupational Class



Theil Decompostion



Quantile RIF Decompostion

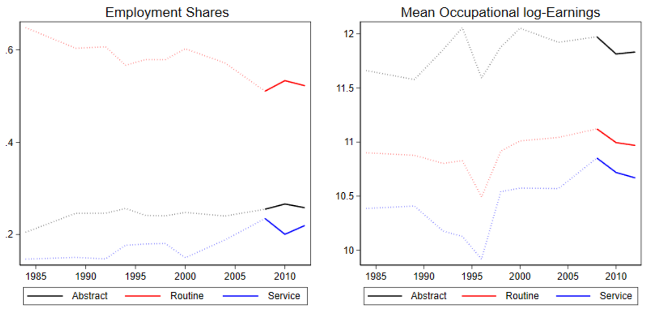


Occupational Classes Compostion and Returns

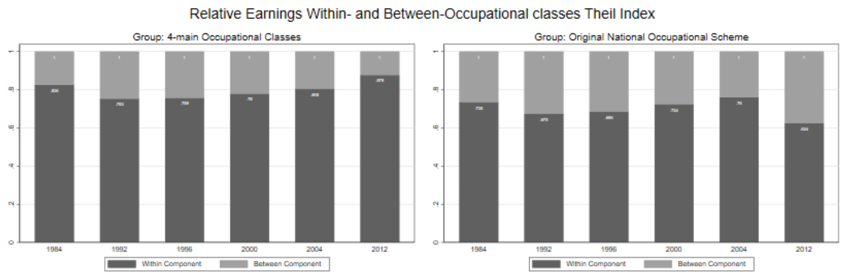
P-Shares Decomposition

Luxembourg: 2010 - 2004			
	<15	15-85	>85
Δ	0.000460	0.000210	-0.000660
Specification Error	3.00e-05	2.00e-05	-5.00e-05
Composition Effect			
Occ	-0.00033***	0.00012*	0.00021***
Educ	-1.00e-05	6.00e-05	-5.00e-05
Female	3.00e-05	-2.00e-05	-1.00e-05
Age	2.00e-05	-2.00e-05	1.00e-05
Ind	0.00026*	-0.000180	-8.00e-05
Reweighting Error	-0.000290	0.000150	0.000140
Coefficent Effect			
Occ	0.00127	-0.00215*	0.000880
Educ	-0.000170	0.000440	-0.000270
Female	0.000560	-0.000380	-0.000180
Age	-0.000570	-0.000160	0.000720
Ind	-2.00e-05	-0.000500	0.000520
Constant	-0.000310	0.00282	-0.00251

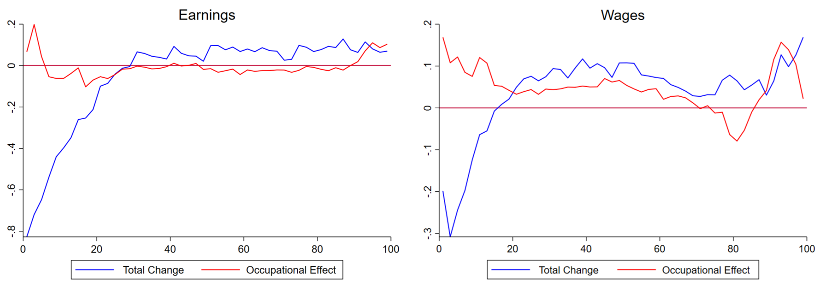
Mexico: 2012-1992



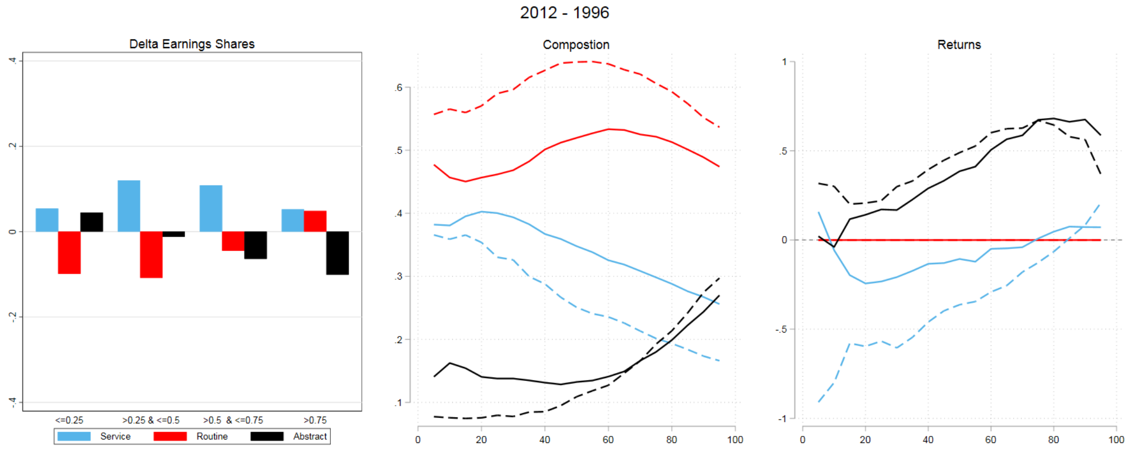
Employment and Income Shares by Occupational Class



Theil Decompostion



Quantile RIF Decompostion

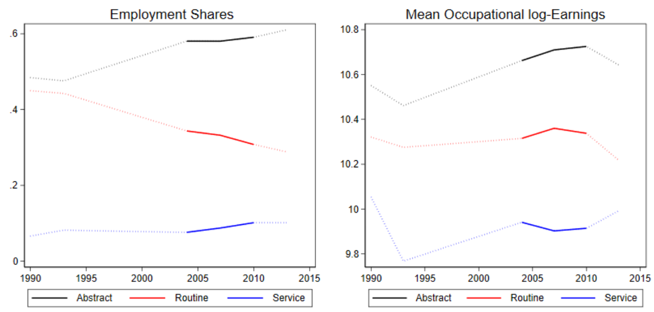


Occupational Classes Compostion and Returns

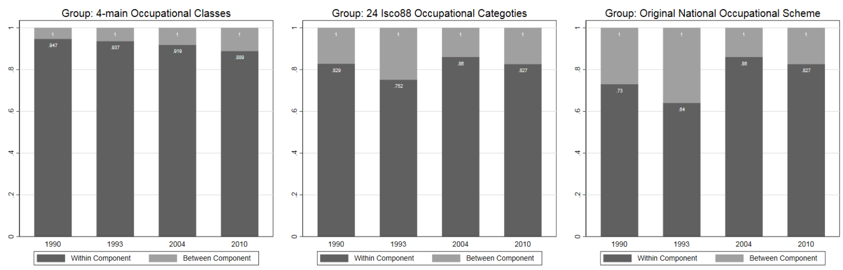
P-Shares Decomposition

Mexico: 2012 - 1996			
	<15	15-85	>85
Δ	-0.00701***	0.00512***	0.00189***
Specification Error	6.00e-05	-5.00e-05	0
Composition Effect			
Occ	-0.00023***	-0.00021***	0.00044***
Educ	-0.00024***	-0.00021***	0.00045***
Female	-0.00005**	0.00004**	1.00e-05
Age	-0.00034***	6.00e-05	0.00028***
Ind	0.00011***	0.00010***	-0.00020***
Reweighting Error	-0.000260	0.00026**	0
Coefficent Effect			
Occ	0.000450	-0.000740	0.000290
Educ	-0.00304***	0.00229*	0.000750
Female	-0.00338***	0.000810	0.00258***
Age	-0.00168	0.00159	9.00e-05
Ind	-0.00170	-0.000320	0.00202**
Constant	0.00331	0.00149	-0.00481**

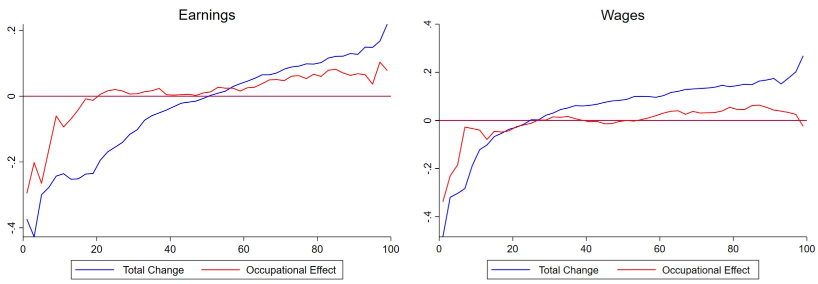
Netherlands: 2010-1990



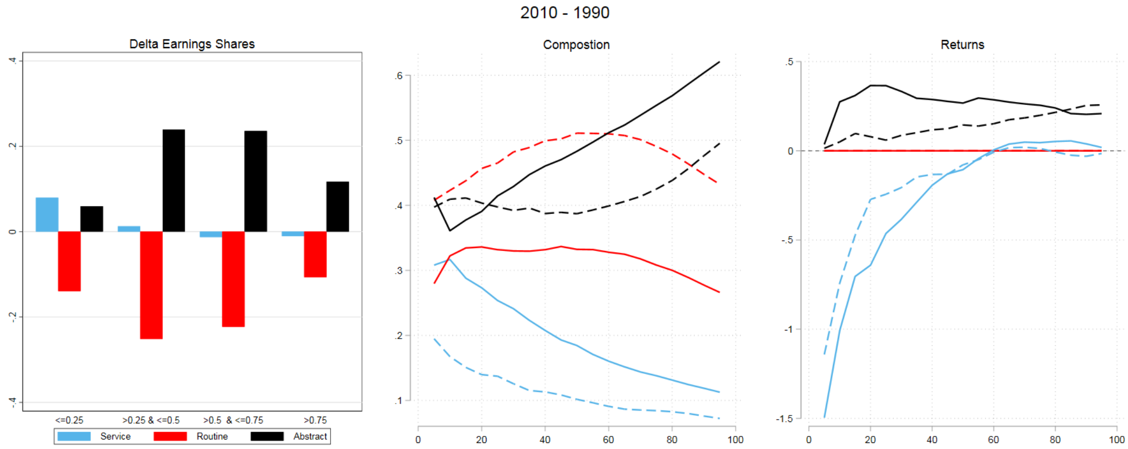
Employment and Income Shares by Occupational Class



Theil Decompostion



Quantile RIF Decompostion

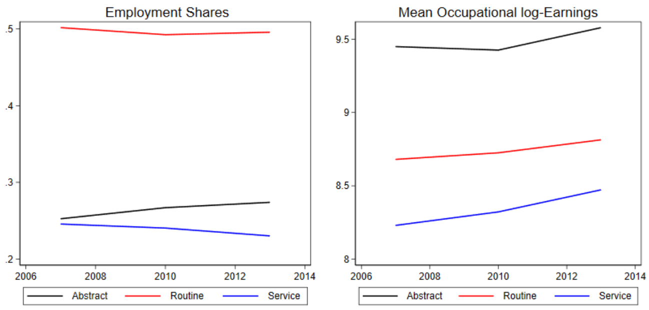


Occupational Classes Compostion and Returns

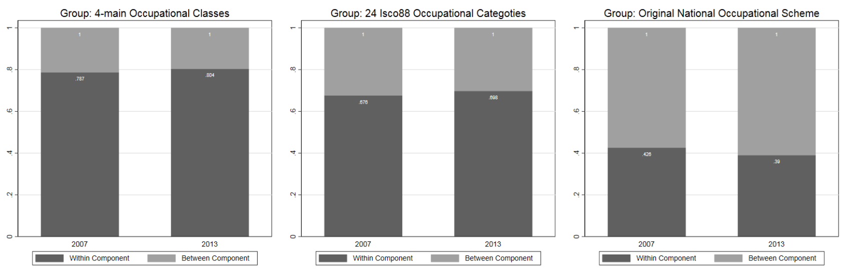
P-Shares Decomposition

Netherlands: 2010 - 1990			
	<15	15-85	>85
Δ	-0.00382***	0.00112	0.00270***
Specification Error	0.000140	0.000200	-0.00035*
Composition Effect			
Occ	-0.000370	0.000140	0.00023*
Educ	0.000170	-0.000260	9.00e-05
Female	0.00019**	-0.00011*	-0.00008**
Age	-0.00058***	0.00030*	0.00028***
Ind	0.000240	-0.000100	-0.000150
Reweighting Error	-0.00242***	0.00134***	0.00108***
Coefficent Effect			
Occ	-0.00192	0.00125	0.000670
Educ	-0.00152	0.00137	0.000150
Female	0.00503***	-0.00357***	-0.00146***
Age	-0.00197	0.00217	-0.000200
Ind	0.00605	-0.00371	-0.00234
Constant	-0.00686	0.00209	0.00477*

Panama: 2013-2007

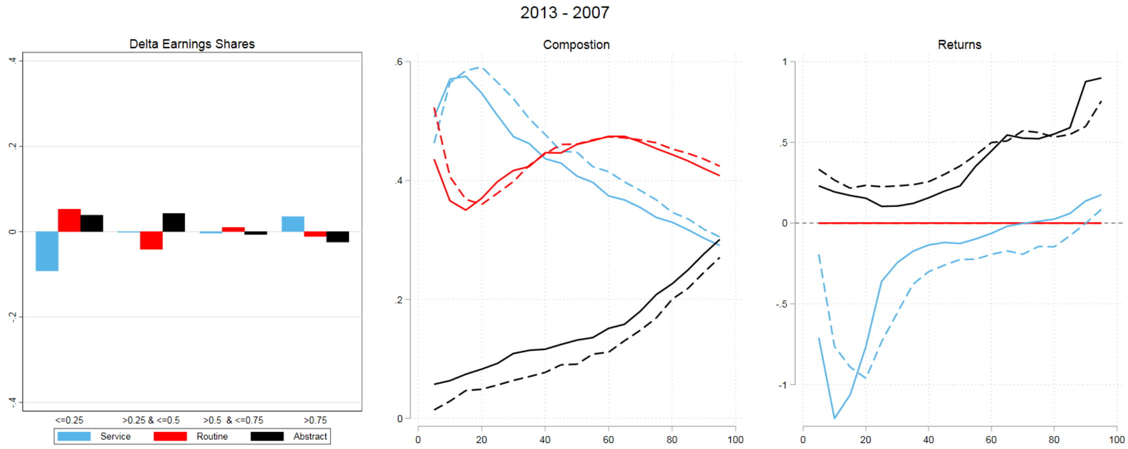


Employment and Income Shares by Occupational Class



Theil Decompostion

Quantile RIF Decompostion

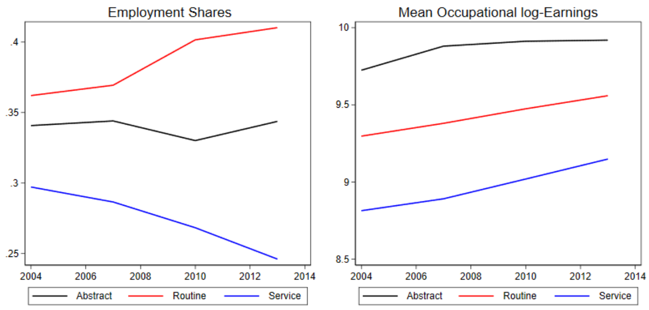


Occupational Classes Compostion and Returns

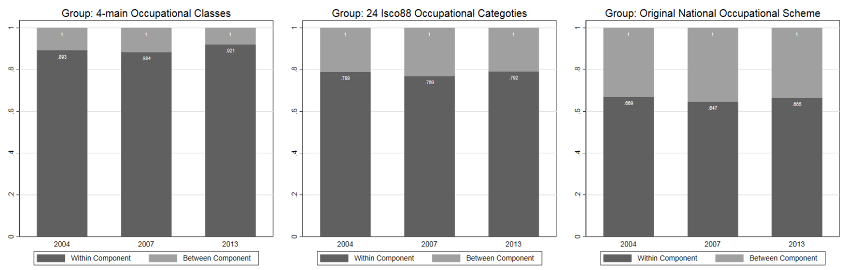
P-Shares Decomposition

Panama: 2013 - 2007			
	<15	15-85	>85
Δ	0.000780	0.000480	-0.00126**
Specification Error	0.00027***	-0.000110	-0.00016**
Composition Effect			
Occ	0.00022**	0.00013*	-0.00035***
Educ	5.00e-05	-0.00034***	0.00029***
Female	-0.00003**	0.00001**	0.00002**
Age	-0.00004*	-1.00e-05	0.00006***
Ind	0.00025**	6.00e-05	-0.00031***
Reweighting Error	-0.000180	9.00e-05	9.00e-05
Coefficent Effect			
Occ	-0.00193**	0.000260	0.00167**
Educ	-1.00e-05	0.000470	-0.000460
Female	0.000110	0.000770	-0.00089*
Age	0.00100	-0.000130	-0.000880
Ind	0.00662	-0.00611*	-0.000510
Constant	-0.00558	0.00540	0.000180

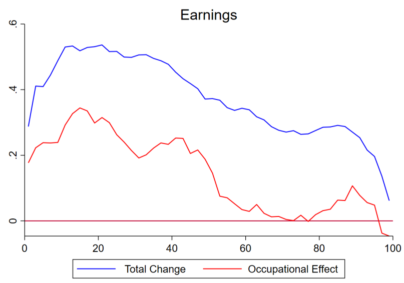
Peru: 2013-2004



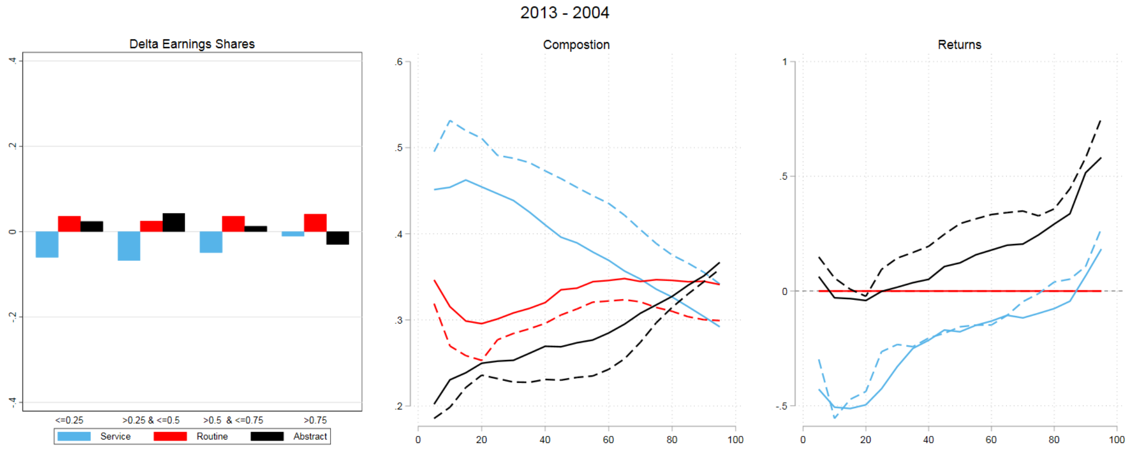
Employment and Income Shares by Occupational Class



Theil Decompostion



Quantile RIF Decompostion

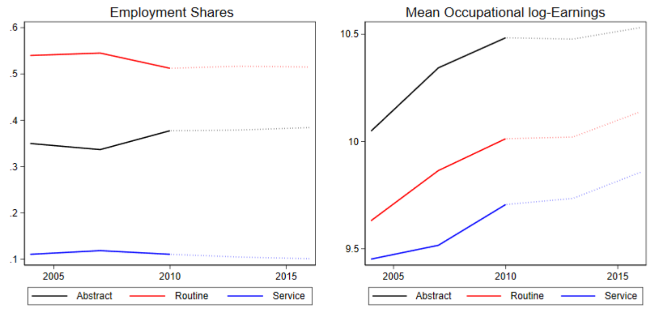


Occupational Classes Compostion and Returns

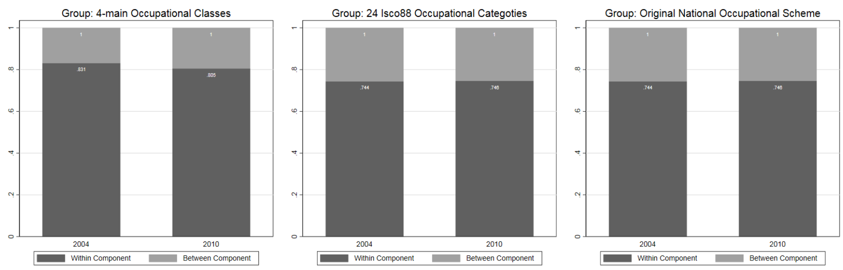
P-Shares Decomposition

Peru: 2013 - 2004			
	<15	15-85	>85
Δ	0.00260***	0.00115**	-0.00375***
Specification Error	0.00031**	-0.00026**	-5.00e-05
Composition Effect			
Occ	0.00084***	0.000240	-0.00108***
Educ	5.00e-05	1.00e-05	-6.00e-05
Female	0.00008***	-0.00002*	-0.00006***
Age	-0.00015***	-6.00e-05	0.00021***
Ind	6.00e-05	-5.00e-05	-1.00e-05
Reweighting Error	-0.00049***	0.00020*	0.00029*
Coefficent Effect			
Occ	0.00164*	-0.000570	-0.00108
Educ	-0.000500	0.000850	-0.000340
Female	-0.000360	-6.00e-05	0.000420
Age	0.000630	0.000960	-0.00160*
Ind	0.00137	-0.00427**	0.00290
Constant	-0.000880	0.00418*	-0.00329

Poland: 2010-2004

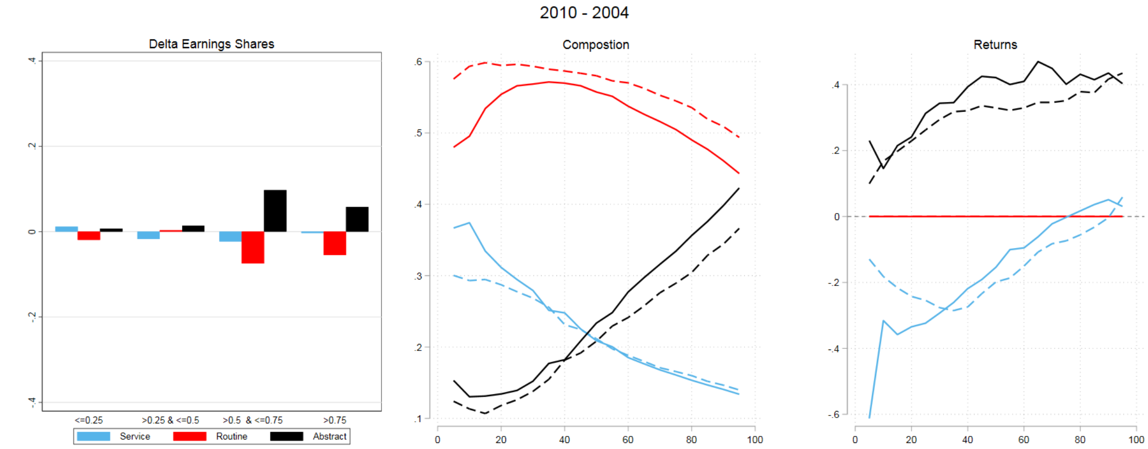


Employment and Income Shares by Occupational Class



Theil Decompostion

Quantile RIF Decompostion

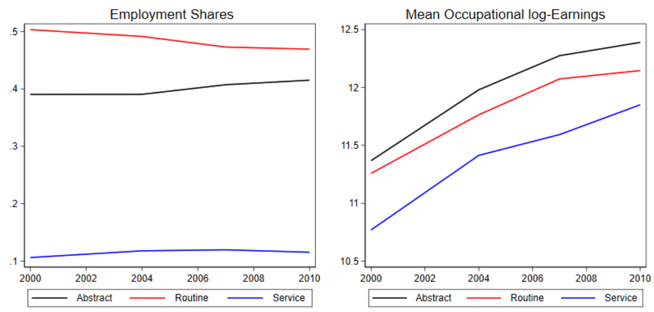


Occupational Classes Compostion and Returns

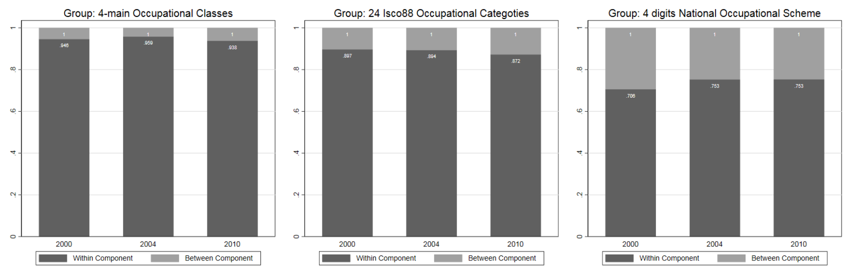
P-Shares Decomposition

Poland: 2010 - 2004			
	<15	15-85	>85
Δ	-0.00148***	0.00130***	0.000180
Specification Error	0	5.00e-05	-5.00e-05
Composition Effect			
Occ	-0.00008***	0.00010***	-0.00002**
Educ	-0.00020***	-2.00e-05	0.00022***
Female	-0.00002***	0.00002***	0
Age	-0.00001*	-2.00e-05	0.00003***
Ind	-3.00e-05	-0.00008***	0.00011***
Reweighting Error	-0.00016**	1.00e-05	0.00014*
Coefficent Effect			
Occ	-0.00059***	0.00035*	0.000240
Educ	0.000170	-1.00e-05	-0.000160
Female	-8.00e-05	7.00e-05	1.00e-05
Age	-0.000100	0.000310	-0.000210
Ind	-0.000800	0.00195	-0.00115
Constant	0.000430	-0.00145	0.00102

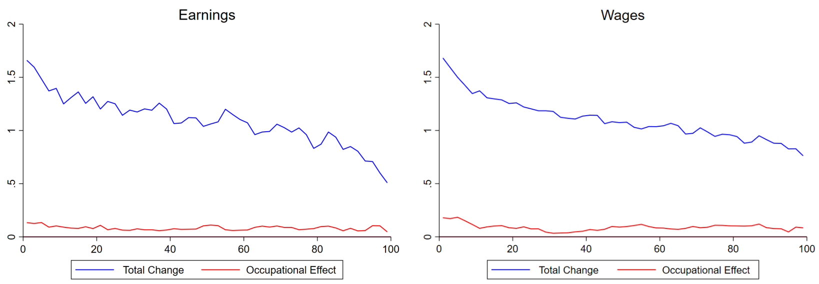
Russia: 2010-2000



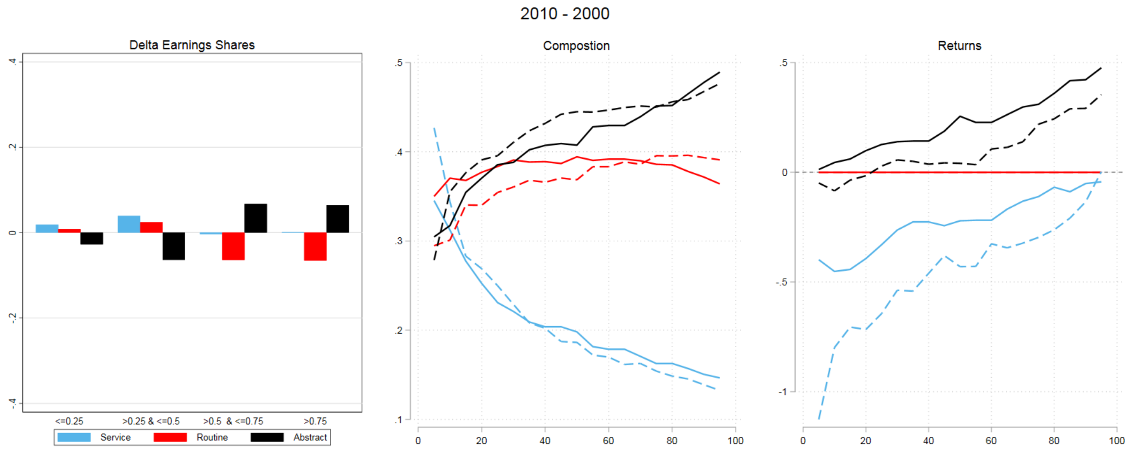
Employment and Income Shares by Occupational Class



Theil Decompostion



Quantile RIF Decompostion

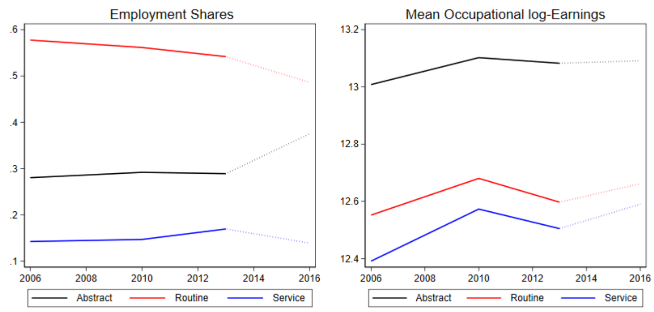


Occupational Classes Compostion and Returns

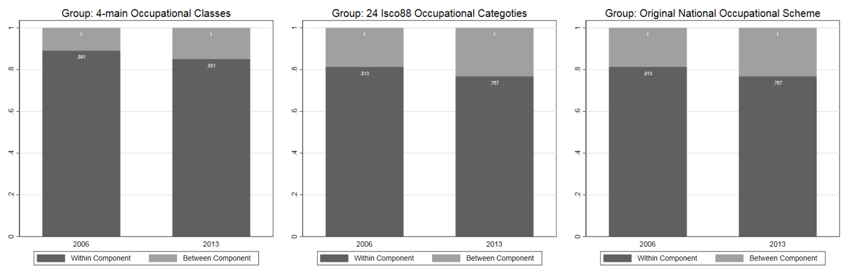
P-Shares Decomposition

Russia: 2010 - 2000			
	<15	15-85	>85
Δ	0.00609***	7.00e-05	-0.00615***
Specification Error	3.00e-05	-1.00e-05	-1.00e-05
Composition Effect			
Occ	9.00e-05	-5.00e-05	-4.00e-05
Educ	0.00012**	6.00e-05	-0.00017***
Female	0.00003*	-0.00004*	1.00e-05
Age	0.000110	-5.00e-05	-6.00e-05
Reweighting Error	2.00e-05	0	-2.00e-05
Coefficent Effect			
Occ	0.000540	-0.000220	-0.000330
Educ	-0.000930	7.00e-05	0.000860
Female	-0.000480	0.000600	-0.000120
Age	-0.00134	1.00e-05	0.00133
Constant	0.00789***	-0.000300	-0.00759***

Serbia: 2013-2006

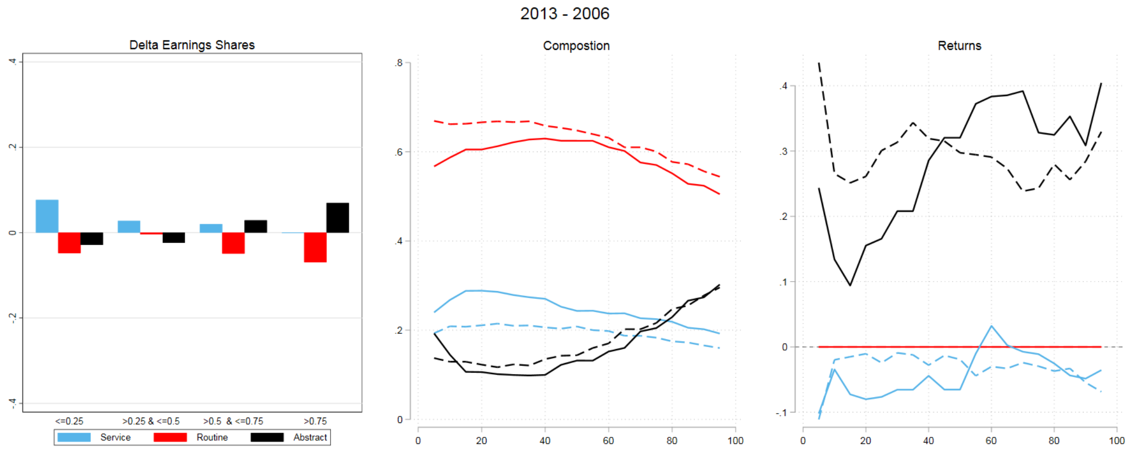


Employment and Income Shares by Occupational Class



Theil Decompostion

Quantile RIF Decompostion

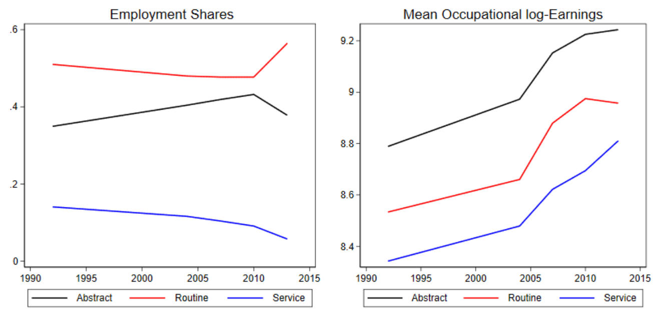


Occupational Classes Compostion and Returns

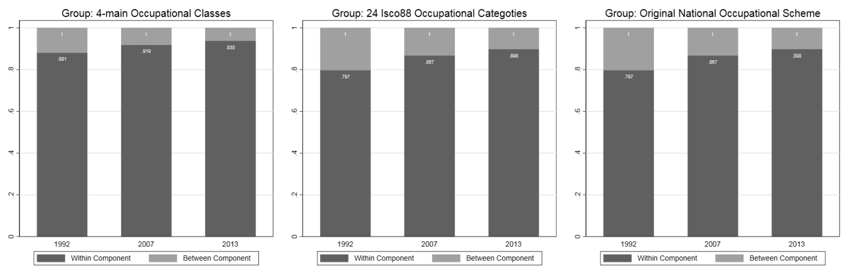
P-Shares Decomposition

Serbia: 2013 - 2006			
	<15	15-85	>85
Δ	0.00262***	-0.00141***	-0.00122***
Specification Error	0	2.00e-05	-1.00e-05
Composition Effect			
Occ	-0.00006*	4.00e-05	2.00e-05
Educ	-0.00012**	4.00e-05	0.00008**
Female	1.00e-05	0	0
Age	3.00e-05	-2.00e-05	0
Reweighting Error	1.00e-05	-1.00e-05	0
Coefficent Effect			
Occ	-0.00091*	0.00068*	0.000230
Educ	0.000290	0.000380	-0.00067*
Female	7.00e-05	-0.000240	0.000170
Age	-0.000270	0.000480	-0.000210
Constant	0.00359**	-0.00277**	-0.000820

Slovakia: 2013-1992



Employment and Income Shares by Occupational Class



Theil Decompostion

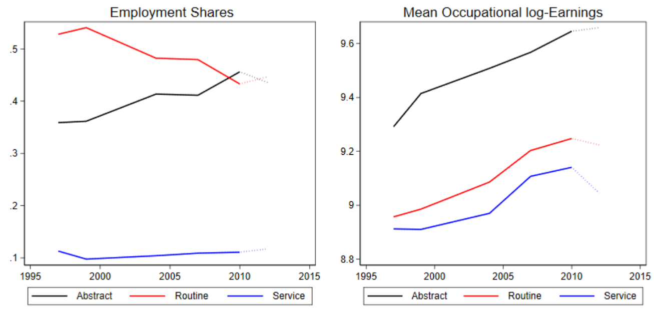
Quantile RIF Decompostion

Occupational Classes Compostion and Returns

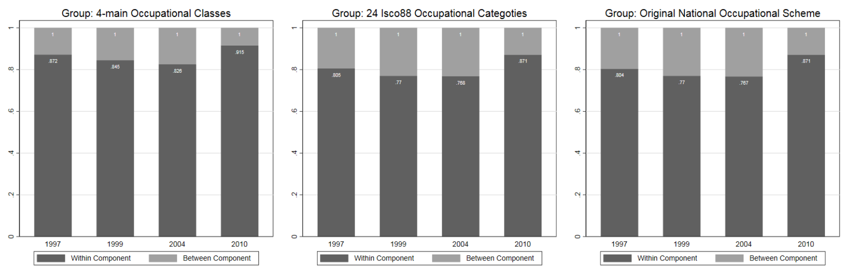
P-Shares Decomposition

Slovakia: 2013 - 1992			
	<15	15-85	>85
Δ	-0.00181**	0.000120	0.00169***
Specification Error	0.00025***	2.00e-05	-0.00027***
Composition Effect			
Occ	0.00042***	-5.00e-05	-0.00036***
Educ	-0.00032***	0	0.00031***
Female	1.00e-05	0.00005***	-0.00006***
Age	3.00e-05	1.00e-05	-0.00003*
Reweighting Error	-0.00012**	1.00e-05	0.00011***
Coefficent Effect			
Occ	-0.000260	-0.000290	0.000540
Educ	0.000130	-0.000110	-3.00e-05
Female	0.000940	-0.000510	-0.000430
Age	0.000610	-0.000260	-0.000350
Constant	-0.00349*	0.00123	0.00226*

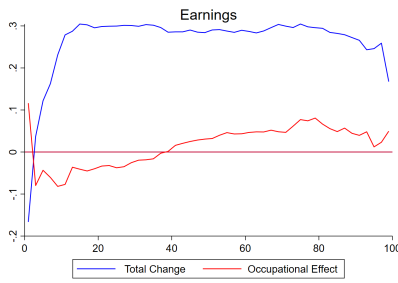
Slovenia: 2010-1997



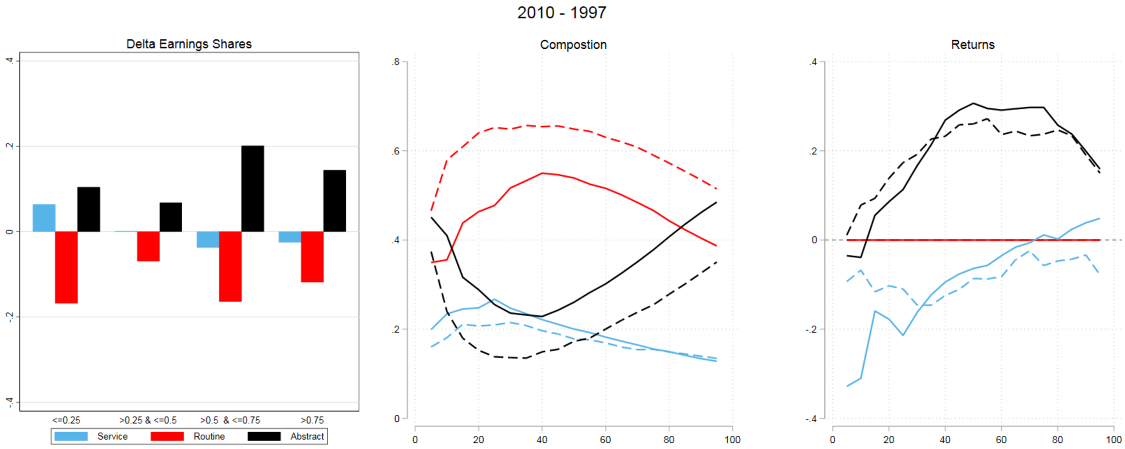
Employment and Income Shares by Occupational Class



Theil Decompostion



Quantile RIF Decompostion

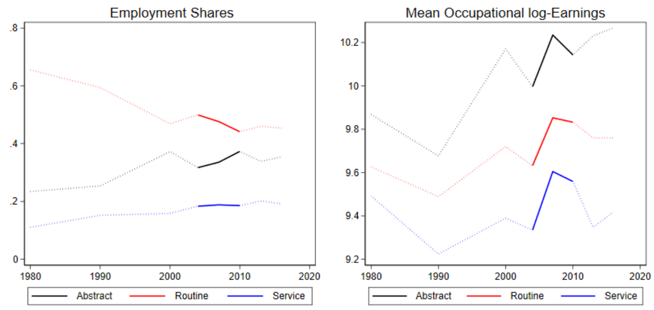


Occupational Classes Compostion and Returns

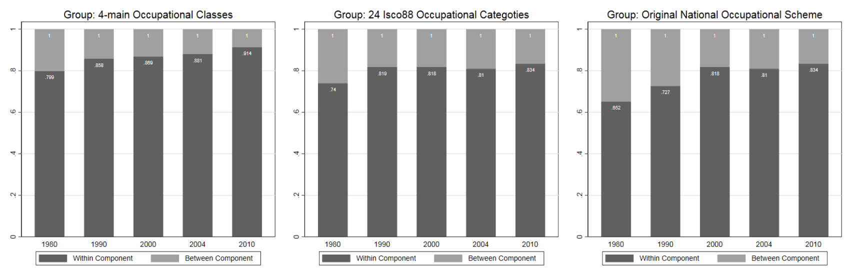
P-Shares Decomposition

Slovenia: 2010 - 1997			
	<15	15-85	>85
Δ	-0.00171***	0.00222***	-0.000510
Specification Error	0.00028*	0.000130	-0.00040**
Composition Effect			
Occ	-0.00024*	0.00023**	1.00e-05
Educ	-0.00061***	9.00e-05	0.00051***
Female	0.00009**	-0.00005*	-0.00003*
Age	-9.00e-05	-3.00e-05	0.000110
Ind	-0.00029**	0.00029**	0
Reweighting Error	-6.00e-05	5.00e-05	1.00e-05
Coefficent Effect			
Occ	-0.000740	0.000360	0.000380
Educ	0.00118	-0.000610	-0.000570
Female	0.000660	-0.000440	-0.000220
Age	0.00121	-0.00108	-0.000130
Ind	0.02084*	-0.01453*	-0.00632*
Constant	-0.02394**	0.01781*	0.00614

Spain: 2004-1990

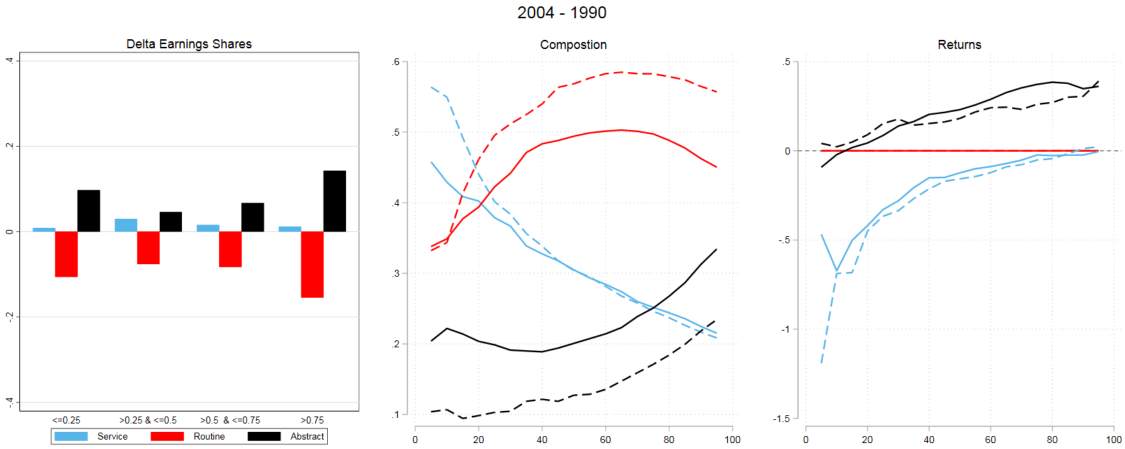


Employment and Income Shares by Occupational Class



Theil Decompostion

Quantile RIF Decompostion

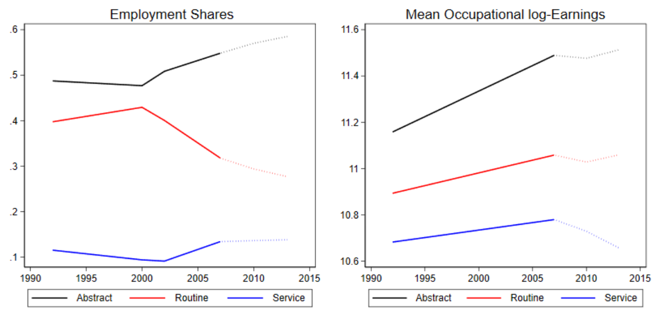


Occupational Classes Compostion and Returns

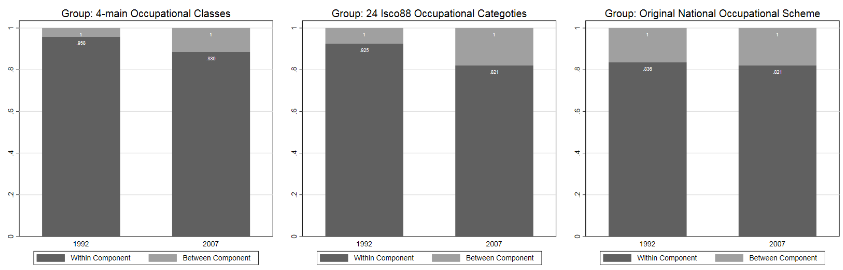
P-Shares Decomposition

Spain: 2004 - 1990			
	<15	15-85	>85
Δ	-0.000450	-0.00073*	0.00118***
Specification Error	2.00e-05	-4.00e-05	2.00e-05
Composition Effect			
Occ	0.00037***	-0.00008***	-0.00029***
Educ	-0.00011***	-0.00002***	0.00012***
Female	0.00065***	-0.00025***	-0.00039***
Age	-1.00e-05	0.00004**	-0.00002**
Ind	0.00016***	-0.00096***	0.00080***
Reweighting Error	-0.00144***	0.00158***	-0.00013*
Coefficent Effect			
Occ	0.00011***	-0.00020***	0.00009***
Educ	0.00029***	-0.00011***	-0.00018***
Female	0.00135***	-0.00068***	-0.00068***
Age	-0.00125***	0.00089***	0.00035***
Ind	0.00271***	0.00249***	-0.00520***
Constant	-0.00331	-0.00339	0.00670

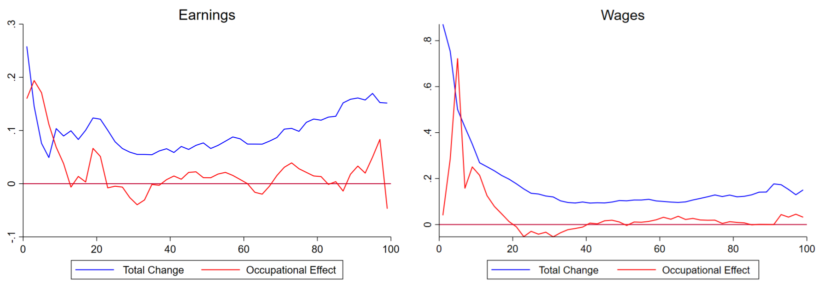
Switzerland: 2007-1992



Employment and Income Shares by Occupational Class



Theil Decompostion



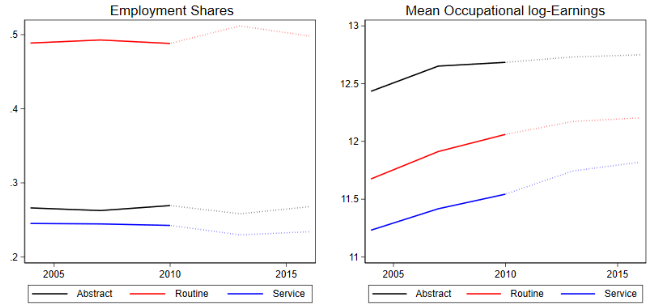
Quantile RIF Decompostion

Occupational Classes Compostion and Returns

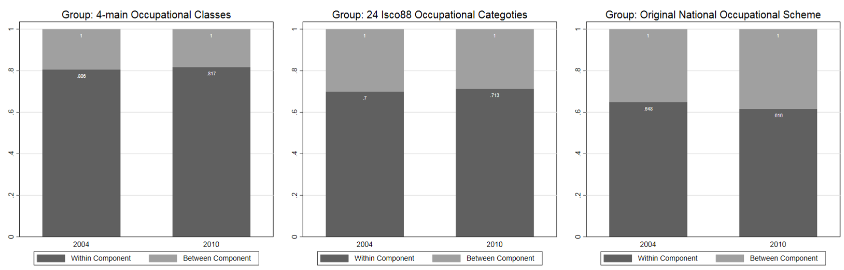
P-Shares Decomposition

Switzerland: 2007 - 1992			
	<15	15-85	>85
Δ	0.000480	-0.00112	0.000640
Specification Error	9.00e-05	7.00e-05	-0.000160
Composition Effect			
Occ	-0.00067***	0.00040***	0.00026***
Educ	-0.00045***	-3.00e-05	0.00048***
Female	0.00010**	-1.00e-05	-0.00009**
Age	-0.00041***	0.00019**	0.00022***
Reweighting Error	-0.000520	4.00e-05	0.00048*
Coefficent Effect			
Occ	0.00156	-0.00123	-0.000320
Educ	-0.000630	0.000810	-0.000180
Female	0.00119	4.00e-05	-0.00123
Age	-0.00263	0.00352	-0.000890
Constant	0.00286	-0.00493	0.00207

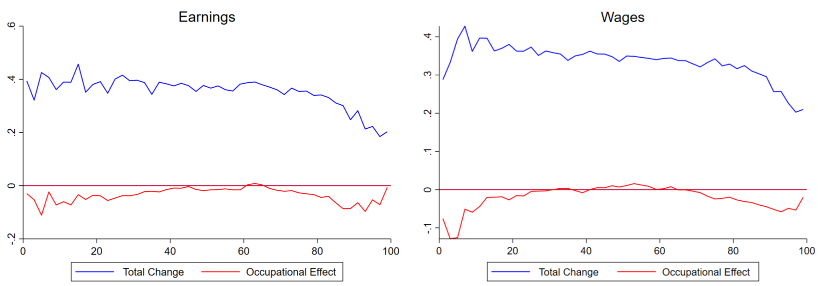
Uruguay: 2010-2004



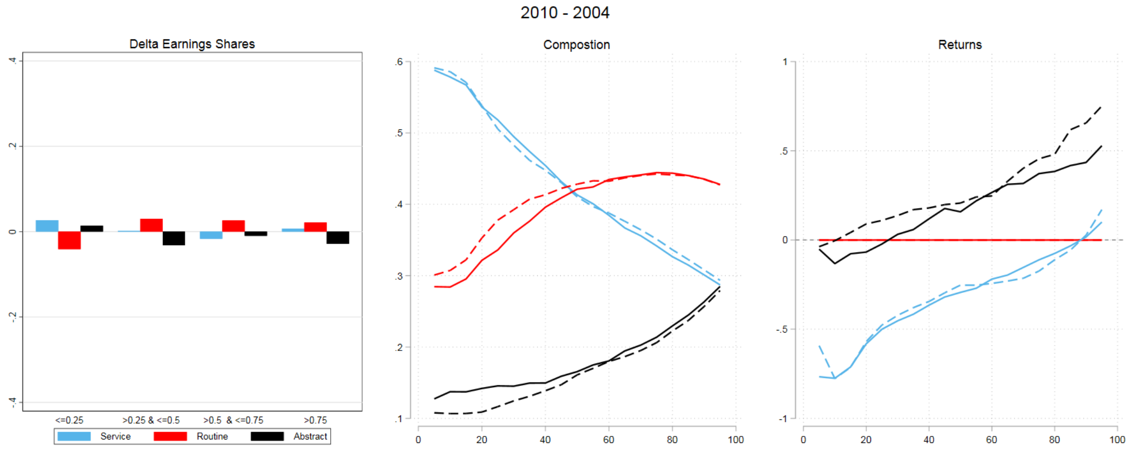
Employment and Income Shares by Occupational Class



Theil Decompostion



Quantile RIF Decompostion

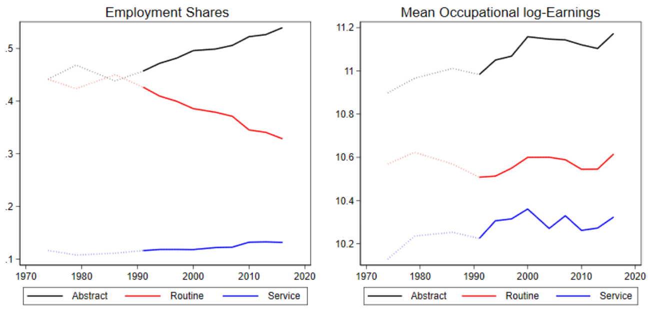


Occupational Classes Compostion and Returns

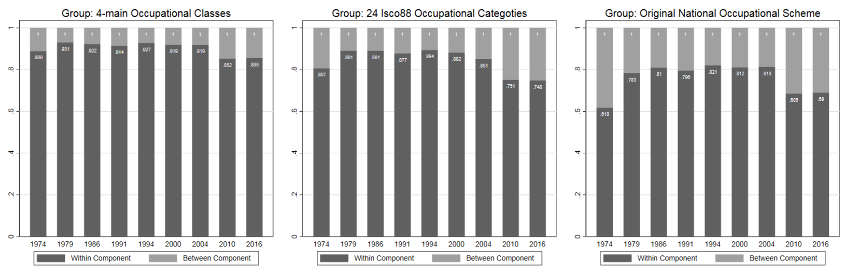
P-Shares Decomposition

Uruguay: 2010 - 2004			
	<15	15-85	>85
Δ	0.00058*	0.00131***	-0.00189***
Specification Error	-1.00e-05	3.00e-05	-3.00e-05
Composition Effect			
Occ	-0.00012*	3.00e-05	0.00010**
Educ	0.00006***	0.00003*	-0.00009***
Female	0.00006***	-0.00004***	-2.00e-05
Age	0	0.00002*	-0.00002**
Ind	0.00025***	-0.00047***	0.00023***
Reweighting Error	-0.00010*	6.00e-05	4.00e-05
Coefficent Effect			
Occ	-0.000420	0.00090*	-0.000480
Educ	-0.000680	-0.000370	0.00105*
Female	-0.000600	-6.00e-05	0.00066**
Age	-0.000710	0.000580	0.000120
Ind	-0.00357***	-0.00237	0.00594*
Constant	0.00643***	0.00296	-0.00939***

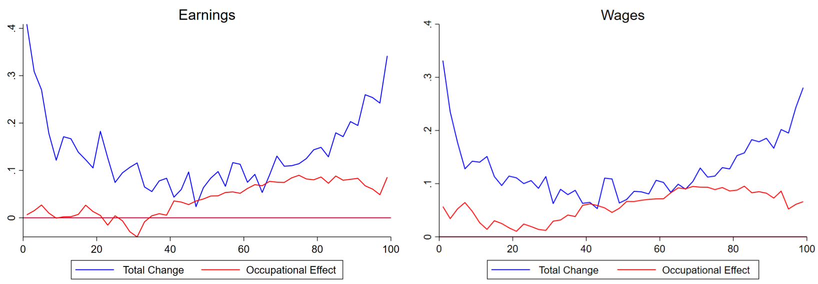
US: 2016 - 1991



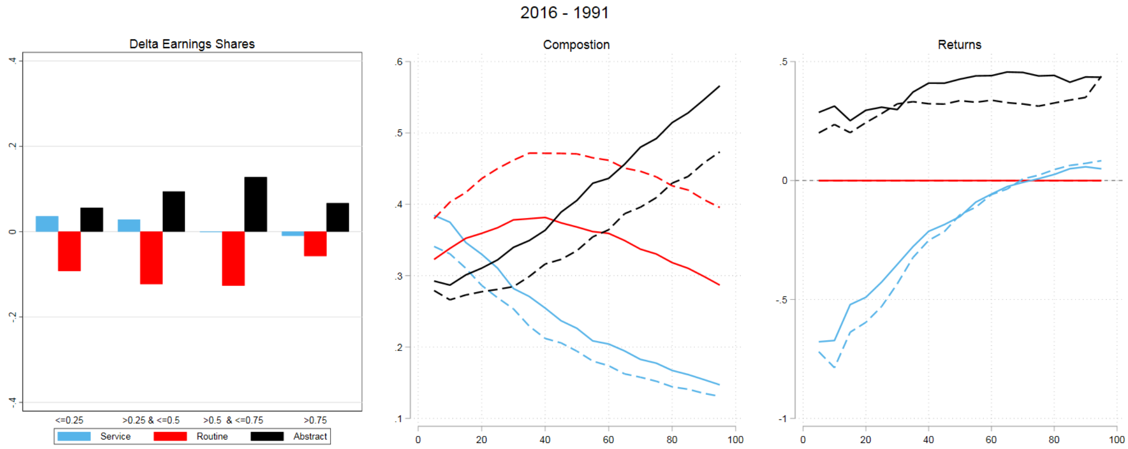
Employment and Income Shares by Occupational Class



Theil Decompostion



Quantile RIF Decompostion



Occupational Classes Compostion and Returns

P-Shares Decomposition

US: 2016 - 1991			
	<15	15-85	>85
Δ	0.00128***	-0.00256***	0.00128***
Specification Error	0.00005*	0.00010***	-0.00014***
Composition Effect			
Occ	-0.00022***	0.00011***	0.00010***
Educ	-0.00025***	0.00018***	0.00008**
Female	-0.00001***	-0.00000*	0.00001***
Age	-0.00006*	-2.00e-05	0.00008***
Ind	-0.00029***	-6.00e-05	0.00035***
Reweighting Error	2.00e-05	-2.00e-05	0
Coefficent Effect			
Occ	-0.000220	2.00e-05	0.000200
Educ	7.00e-05	-3.00e-05	-4.00e-05
Female	0.00048**	9.00e-05	-0.00057***
Age	-0.000520	0.000120	0.000400
Ind	0.00130***	-0.000540	-0.000760
Constant	0.000940	-0.00251***	0.00157*

3 The Distribution of Augmented Wealth in Australia from 2002 to 2018

3.1 Introduction

The analysis of wealth distributions has gained more and more attention around the world throughout the last decade. A highly unequal distribution of wealth can negatively affect growth and innovation (Bagchi and Svejnar, 2015; Berg et al., 2018) and raises concerns about an unequal distribution in welfare (Stiglitz, 2012). Pension wealth challenges the comparability of wealth data across countries. (Bönke et al., 2019; Frick and Headey, 2009; Kuhn, 2020; Wolff, 1996) Private pension wealth is often included in wealth surveys and, therefore, in wealth distributions, while public pension entitlements are not. Several studies showed, however, that public pension wealth is an (imperfect) substitute for net worth (Attanasio and Brugiavini, 2003; Wolff, 2015a).[1] Hence, augmented wealth, i.e., net worth plus private and public pension wealth, contributes to a sincere comparison of international wealth distributions incorporating a more reliable measure of economic well-being (Wolff, 2015b; Bönke et al., 2019). Moreover, pension wealth aggregates are a helpful tool in a panel analysis to understand the distributional effects of pension policies.

This paper explores the evolution of augmented wealth aggregates in Australia between 2002 and 2018. Australia is particularly interesting for an assessment, as its pension system relies on two main schemes, i.e., the means-tested social security pension system, called Age Pension, and the private pension system, called Superannuation. Like many advanced economies, Australia struggled to provide a sustainable and effective pension scheme that secures a commensurate level of income, and, hence, consumption, during retirement. In 1992, the Australian government[2], therefore, introduced compulsory contributions by employers to the employee's Superannuation accounts as an addition to the established, tax-funded Age Pension. Superannuation is mostly a defined distribution scheme for employees and employers, and it is designed to retain and guarantee the standard of living for future retirees (Superannuation (Objective) Bill, 2016).

This analysis provides a deeper understanding about the establishment of the Superannuation scheme. It comprises of two main components: first, I describe

[1] See Bönke et al. (2019) for further discussions.

[2] Henceforth, I will refer to the Australian government as an collective term for the different Coalition and Labor governments over time.

income during retirement. This includes the analysis of dissaving patterns of Superannuation accounts and how they interact with the Age Pension scheme. Second, I analyze augmented wealth in Australia over time by focusing especially on the development of the pension schemes. I use the Household, Income and Labour Dynamics in Australia survey, which provides five wealth modules at the household level between 2002 and 2018 (Watson and Wooden, 2002). Furthermore, it includes exhaustive pension information at the individual level.

I find that Australian households, especially those with a household head aged 50 and older, realized significant wealth gains between 2002 and 2018. Despite large wealth increases, inequality remains stable in Australia. The results show that Age Pension remains the most important source of income for most retried Australians, though income from Superannuation has increased its share. Social security pension wealth is the largest asset category at the lower end of the wealth distribution. Pension wealth has an equalizing effect, as the Gini index in 2018 reduces from 0.66 for net worth to 0.571 for augmented wealth. Comparing my findings to Germany, Switzerland, and the United States (US) (Bönke et al., 2019; Kuhn, 2020), Australia's Gini index of augmented wealth is the second highest after the US. Furthermore, I find evidence for behavioral responsiveness from the interaction between Superannuation and Age Pension, as individuals dissave Superannuation accounts faster when they are also eligible for Age Pension. My results also raise concerns as to whether the retiree population will diverge between those depending on Age Pension and those relying on Superannuation.

Despite the increased data availability, the role of pension wealth in wealth distributions has been marginally studied. Harding (2002) analyzes pension wealth, especially Superannuation accounts in Australia, and finds that the scheme reduces wealth inequality while overall wealth inequality remained constant between the 1980s and the 1990s. Frick and Headey (2009) compare Australian and German augmented wealth for the retired population in 2002 and find that including pension wealth reduces the observed inequality in both countries. More recently, Bönke et al. (2019) analyze augmented wealth in Germany and the US. They estimate that pension wealth accounts for 48 (61) percent of the US (German) household wealth. They also find an equalizing effect of pension wealth with a net worth Gini coefficient of 0.889 (0.755) compared to an augmented wealth Gini coefficient of 0.700 (0.508) in the US (Germany). Kuhn (2020) analyzes augmented wealth in Switzerland and finds similar results as in Germany.

This paper contributes to the existing literature by taking a first look at augmented wealth in Australia. Furthermore, I present a primary inter-temporal analysis of social security pension wealth. To the best of my knowledge, I am also the first who empirically assesses the dissaving behavior of public pension scheme after retirement. Moreover, this paper also contributes to the understanding of the enrollment phase of a new private pension scheme, and the lessons to be learned for other economies.

The remainder of this paper is as follows: Section 3.2 describes the Australian pension system in detail, Section 3.3 provides the applied methodology, Section 3.4 depicts the data, Section 3.5 presents the empirical findings. Section 3.6 discusses several caveats of my analysis. Section 3.7 provides a general discussion and concludes.

3.2 Australian Pension System

This chapter provides a detailed overview of the Australian pension system. It builds on two main pillars: the social security scheme Age Pension and the private Superannuation scheme.[3] The Australian pension system contains several unique features. First, the means-tested social security pension is tax-funded, which does not rely on individuals' employment history, but on an income and asset test at retirement. Second, the private occupational pension scheme Superannuation became compulsory in 1992 to provide additional incentives to build up private pension wealth (Australian Government, 1992). Third, Age Pension and Superannuation are not independent, but interact with each other. Wealth and returns from Superannuation alter an individual's eligibility to Age Pension.

In this section, I describe the institutional settings of the different schemes and what has changed between 2002 and 2018.[4]

3.2.1 Age Pension

General information The Age Pension scheme is a tax-funded, means-tested social security scheme. Individuals or households, who pass the income and asset test, are eligible to Age Pension at the age of 65 to 67. Additional to the full basic rate, which was $629[5] for singles and $834.40 for couples per fortnight in 2018, the Australian government provides several further payments, i.e., Rent Assistance, Energy Supplement, and Pension Supplement.[6] Moreover, individuals are entitled to the Pensioner Concession Health Card, providing higher refunds for health care costs. Depending on the state, local councils also offer additional discounts on property and water rates, public transport fares or motor vehicle registration.

[3]Some sources also report three pillars, meaning Age Pension, compulsory contributions to Superannuation funds and voluntary savings, which includes additional investment into the Superannuation accounts or other investment vehicles. In the analysis, I focus on the two major schemes. Superannuation henceforth includes both private and voluntary contributions.

[4]The information presented in this section is based on several documents and websites provided by the Australian Government (2018, 2021b).

[5]1 AUD = 0.7405 USD = 0.6325 EUR in 30th June 2018.

[6]In 2018, Rent Assistance added up to $128.00 ($135.80) for singles (couples) per fortnight, Energy Supplement another 10.60 ($14.10) , and Pension Supplement $51.10 ($67.80) . For more information, the Table 3.13 in the Appendix.

Eligibility conditions The first hurdle for being eligible to Age Pension is the retirement age and the residency in Australia. In 2018, the retirement age was 65 for both genders, i.e., cohorts born in 1953 or earlier. The residency rules demand that an individual has been living in Australia for at least ten years in total, with five of these years without a break (Australian Government, 2018).

Income and asset test Singles or couples, who reached the retirement age and meet the residency rules, are eligible to Age Pension, if they pass the income and asset test. The income test includes gross labor earnings, pensions, received gifts, rental income, as well as financial gains. Retirees are allowed to work during retirement, but labor income above the work bonus, i.e., $250 per fortnight, is included in the income test. The income test does not consider realized, financial returns, but imputed, "deemed" returns on the basis of total financial assets (Australian Government, 2021b).[7] The Australian government provides low and high interest rates[8], which are multiplied with the value of the financial assets below and above the threshold, respectively. The deeming rates are adjusted on a non-regular basis by the Australian tax authorities (Australian Government, 2021a). The asset tests inquire the level of net worth, including real estate, business assets, financial wealth, Superannuation accounts and valuables. Real estate does not include the principal home, however different thresholds apply for homeowners and non-homeowners, providing a higher allowance for the latter. In 2018, the total wealth threshold for the full pension for singles (couples), without owning their home, was at $465,500 ($594,500), whereas home-owners were allowed to own $258,500 ($387,500).[9]

Full vs partial pension Individuals and couples can qualify for a full or partial pension, depending on their level of income and wealth. Once they earn more than the first threshold, i.e., $172 ($304) per fortnight in 2018, every additional dollar reduces the payment by a reduction rate, which was for singles (couples) 0.5 (0.25) in 2018. The wealth thresholds for the partial pension are considerably higher than those for the full pension, which were $771,000 ($1,055,000) for non-homeowners, and were $564,000 ($848,000) for homeowners. There are strong incentives to be at least partially eligible to Age Pension, even if the payment is low, as they are eligible to the benefits of the Pensioner Concession Health Card.

[7]The corresponding asset base includes the market value of savings accounts and term deposits, managed investments, loans, and debentures, listed shares and securities, gifts, and Superannuation accounts.

[8]Also called deeming rates: lower rate was 1,75 percent, higher rate was 3.25 percent per year in 2018.

[9]This shows that the scheme considers sharing resources of couples, and therefore provides lower payment and lower threshold, compared to two single individuals.

Table 3.1. Eligibility

Type	2002	2006	2010	2014	2018
			Eligibility Testing		
Income Test					
single threshold full	116	128	146	160	172
couple threshold full	204	228	256	284	304
single threshold partly	1,185	1,391.75	1,578.20	1,868.60	2,004.60
couple threshold partly	1,979	2,328.50	2,415.20	2,860	3,066.80
work bonus	0	0	250	250	250
reduction rate single	0.4	0.4	0.5	0.5	0.5
reduction rate couple	0.2	0.2	0.25	0.25	0.25
Deeming					
cash free	500	500	500	500	500
interest rate low	0.025	0.03	0.03	0.02	0.0175
interest rate high	0.04	0.05	0.045	0.035	0.0325
single threshold	34,400	38,400	43,200	48,000	51,200
couple threshold	57,400	63,800	72,000	79,600	85,000
Asset Test					
single w. house full	145,250	161,500	181,750	202,000	258,500
single no house full	249,750	278,500	313,250	348,500	465,500
couple w. house full	206,500	229,000	258,000	286,500	387,500
couple no house full	311,000	346,000	389,500	433,000	594,500
single w. house partly	288,000	330,000	659,250	771,750	564,000
single no house partly	392,500	447,000	790,750	918,250	771,000
couple w. house partly	443,500	509,500	978,000	1,145,500	848,000
couple no house partly	548,000	626,500	1,109,500	1,292,000	1,055,000
reduction rate	3 per 1,000	3 per 1,000	3 per 1,000	3 per 1,000	3 per 1,000
Year of Birth for Eligibility					
Female	1940	1944	1948	1949	1953
Male	1937	1941	1945	1949	1953
			Payments per fortnight		
Age Pension					
pension single p. ft.	352.10	402.40	496.30	585.50	629.00
pension couple p. ft.	421.80	478.50	658.40	776.70	834.40

Note. Information is provided by the Australian Government (2018, 2021b). Table provides thresholds for the income and asset test, deeming, and payments for Age Pension. Further payments are provided in Appendix 3.8.1. All values are provided in AUD. As the numbers represent the thresholds of the respective years, they are not adjusted for inflation. Following the official approach, most values, apart from the Asset Test and Deeming, are calculated per fortnight.

Changes between 2002 and 2018 There were several policy changes and adjustments in the Age Pension scheme. The Age Pension benefits, as well as the income and asset test thresholds, are adjusted bi-annually.[10] Table 3.1 depicts the changes of rates and thresholds valid in July of the years included in my analysis. The threshold for income tests increased steadily over the years, however, those for the partial pensions nearly doubled for singles, whereas those for couples increased by around two-thirds. The asset thresholds increased relatively steadily too, but were actively adjusted by the Federal Government in 2017: the full pension thresholds were increased, while the thresholds for the partial pension were reduced, explaining the difference between 2014 and 2018. Between 2006 and 2010, the reduction rate for the partial pension increased from 0.4 to 0.5 (0.2 to 0.25 for couples). The Age Pension payments were additionally increased in 2009 (Australian Government, 2009).

At the beginning of the observed period, in 2002, women could qualify earlier for Age Pension than men. The transition of the retirement age for women increased from 60 to 65 by 2014. Another transition for both women and men started in 2018, which will gradually raise the retirement age to 67 by 2024. The 2018 wave is partially affected, as the retirement age for those born after July 1952 and before 1954 can retire at the age of 65 and a half.

3.2.2 Superannuation

General information The Superannuation scheme represents, for most Australians, an investment in an accumulation fund.[11] The fund can be managed by financial institutions (retail funds), by the employing company (corporate funds) or industry (industry funds), by the public sector for civil servants (public sector funds) and by the individuals themselves (self-managed funds).[12] The total wealth in Superannuation accounts was \$2.9 trillion in 2019 (ASFA, 2021), i.e., 1.3 times the annual Australian GDP.

The choice of the Superannuation fund is not always subject to the employee, since enterprise agreements can specify the fund type.[13] The Australian government sets standards for contributions by employees and compulsory monthly contribu-

[10] Until 2007, this was done on the basis of the Consumer Price Index (CPI). After the introduction of the Pensioner and Beneficiary Living Cost Index (PBLCI), the rates are increased by whichever index is greater.

[11] Even though Australians potentially possess several Superannuation funds, I refer to them in singular.

[12] Industry and retail funds include more than 11 million members each out of a total 27.4 million accounts, and are the most dominant fund categories in 2019 (Australian Prudential Regulation Authority (APRA), 2021).

[13] This affects around 30 percent of all receivers (Australian Government, 2021b). Theoretically, employees could set up their own fund and transfer the money from the default fund. However, this is costly. I gratefully thank Roger Wilkins for this remark.

tions for employers, i.e., the Superannuation Guarantee, which is at least 9.5 percent of the monthly wage in 2018.[14] Additionally, individuals can invest further savings from their gross or net income. The withdrawal after the preservation age, i.e., 60 in 2018, is unlimited and free from income taxes.[15] Low income earners qualify to a governmental supplement contribution match of 50 percent, capped at $500 per year.

Superannuation and taxation Superannuation savings are liable to tax (Australian Tax Office, 2021). The tax rate depends on the type of contribution. Most commonly, contributions are taxed lump-sum at 15 percent. These are "concessional" contributions including the compulsory payments by the employer[16] and additional private investments up to $25,000 per year from the gross income. The "Division 293 tax" raises the tax rate to 30 percent for individuals earning more than $250,000. Additional "Non-concessional" contributions are not taxed, as they stem from net-income. However, they are capped at $100,000 per year in 2018 and only allowed for those with less than $ 1.6 million wealth in Superannuation accounts. Any contribution above is taxed at the marginal income tax rate.[17] Except in rare circumstances, e.g., due to medical conditions, individuals cannot access their Superannuation accounts before they reach the preservation age, without paying marginal income tax rates. Self-managed funds can also borrow a loan against its Superannuation before preservation age is reached. Generally, returns on Superannuation investments are taxed at a 15 % tax rate[18], while returns on the first $1.6 million are not taxed if they are realized in the retirement phase.

The reduced taxation rate is the main vehicle to incentivize savings throughout the accumulation period. The taxation of Superannuation investments and returns is, at 15 percent, considerably lower than the marginal income tax rates, which includes income from capital gains. In Australia, there is no separate tax rate on capital gains, as it is added to labor income and other income sources. The total sum defines

[14] The employer has to contribute quarterly. Delayed payments are taxed with the 10 percent Super Guarantee Charge. The term "Guarantee" can be misleading, as it may imply a defined benefit later in retirement. However, it describes the monthly contribution to the Superannuation scheme and does not represent a "guaranteed" income flow during retirement.

[15] Exceptions apply for individuals holding an untaxed super fund (contributions are not taxed), which can occur in the case of a public sector fund. However, the tax rate does not depend on the withdrawal sum for retired Australians.

[16] The contribution is deductible for employers at the end of financial year.

[17] These tax rates are calculated at the end of the financial year. As income taxes are normally paid directly, this can delay tax payments. Therefore, the Australian government introduced the Excess concessional contribution charge rates, which adds an extra rate on the marginal income tax (4.96% in 2018).

[18] Australian Tax Office applies a dividend imputation system, meaning that tax payments, e.g., by the share emitting company, can be used by the shareholder to offset their own tax liabilities. It potentially decreases the tax rate for the Superannuation owner.

Table 3.2. Marginal Income Tax Rates

2002		2018	
Tax bracket	Tax	Tax bracket	Tax
\$1 – \$6,000	Nil	\$1 – \$18,200	Nil
\$6,001 – \$20,000	17 %	\$18,201 – \$37,000	19 %
\$20,001 – \$50,000	30 %	\$37,001 – \$87,000	32,5 %
\$50,001 – \$60,000	42 %	\$87,001 – \$180,000	37 %
\$60,001 and over	47 %	\$180,001 and over	45 %

Note. Information is taken from Australian Tax Office (2021). The table provides the marginal income tax rates for 2002 and 2018. For foreign residents, different tax rates apply and they are not depicted here. Thresholds are in AUD and not adjusted for inflation.

the individual tax rate. Table 3.2 provides the tax brackets and the corresponding tax rates for the years 2002 and 2018. The brackets increased significantly in the considered period and the tax rates were raised at the lower end and decreased at the top. As the Division 293 Tax does not apply for individuals with an income below \$250,000, the tax advantage from Superannuation savings is substantially higher at the upper end of the income distribution.

Changes between 2002 and 2018 In the beginning year of my analysis, the compulsory Superannuation scheme was ten years old and therefore relatively new. Several policy changes and adjustments have been made in the successive years, of which several were of a larger scale. This includes the raise of the preservation age, changes to the concessional contributions cap, the abolition of Reasonable Benefit Limits, the end of the Superannuation surcharge in 2005, and the introduction of the Division 293 tax.

In 1999, the preservation age was gradually increased from age 55 to 60. Those born before 1960 could access their Superannuation savings at 55, however the preservation age increased year by year to 60 for those who were born after June 1964. This means that from 2015 onward, individuals access their accounts a year later. An even more substantial change could be the reduction of the concessional contribution cap. As shown in Table 3.3, the cap was reduced from \$100.000 to \$25.000 in the 2010s. Moreover, the age specific caps were abolished. As a consequence, the potential tax advantages per year for Superannuation savings are significantly lower compared to end of the previous decade. On the other hand, the Reasonable Benefit Limits were abolished in 2007, which conditioned the concessional tax rate limits to wealth levels in Superannuation accounts.[19] Another considerable change was the abolition of the Superannuation Surcharge in 2005, which increased the tax rate by 12.5 percent, applied to contributions from individuals earning more than \$121,075.

[19]The non-concessional, tax-free tax rate is still limited at \$1.6 million.

Nonetheless, in 2012, a new type of surcharge tax was introduced, the Division 293 tax. It increases the Superannuation tax rate by 15 percent for individuals earning more than \$250,000 (\$300,000 until 2017). In July 2017, the Australian Tax Office introduced the \$1.6 million cap to the tax-exempt status, with a 15 percent tax rate for the amount above.[20]

Table 3.3. Concessional Contributions Cap

Age	2009	Age	2010	Age	2014	Age	2018
<50	50,000	<50	25,000	<59	25,000	all	25,000
>50	100,000	>50	50,000	>59	35,000	ages	

Note. Information is taken from Australian Tax Office (2021). The table shows the change of the concessional contribution cap over time. Values in earlier years were the same as in 2009. All monetary values are in AUD and not adjusted for inflation.

There were smaller adjustments in the 2002 to 2018 period. The introduction of the excess concessional contributions charge was introduced in 2013, which declined gradually from 5.82% to 4.96% until 2018. The governmental supplement contribution started in the same year.[21] The Superannuation Guarantee increased moderately from 9 percent to 9.25 percent in 2013, and to 9.5 in 2014.

3.2.3 Interactions between Superannuation and Age Pension

The Superannuation system was introduced to address the aging population in Australia, as in most advanced economies. An aging society with increasing live expectancy could bring a singular tax-based pension scheme as Age Pension to its limits. Superannuation was set as an additional pillar to support private wealth accumulation (Australian Government, 1992). Consequently, Superannuation wealth affects Age Pension eligibility twofold: deemed income from Superannuation wealth is included in the income test and the total wealth in Superannuation accounts is included in the asset test. In 2016, the Australian government projected in the Superannuation (Objective) Bill (2016) that the proportion of retirees receiving no Age Pension will remain stable until 2050 at 20 percent of the population. Due to the maturity of the Superannuation scheme, they expect a shift away from full to partial pensions.

These interactions are highly relevant for understanding how the introduction of the private pension scheme affects its public counterpart. Individuals face new inter-temporal consumption decisions through their accumulation and retirement

[20] I, again, thank Roger Wilkins for this remark.
[21] Originally, the contribution per invested \$ 1 was \$1.5, but reduced to \$0.5 in 2018.

phase. Phasing out regulated Age Pension payments by individually chosen private annuities from Superannuation causes new dissaving decisions, which are affected by the incentives set by the two pension schemes. An assessment of this behavioral responsiveness is of interest for policy implications beyond the Australian case, as policy makers in other countries may learn from it. This paper provides a first analysis of these dissaving decisions.

3.2.4 Other Schemes

Besides Age Pension, the Australian government provides several other pension schemes.[22] The Disability Pension is available to compensate veterans and their partners and/or descendants for injuries or diseases caused or aggravated by war service or certain defense service. Disability Support Pension is for people aged over 16 and below retirement age with a physical, intellectual, or psychiatric impairment that prevents them from working, or being re-skilled to work. Mature Age Allowance is a bridging income support payment for individuals of at least 60 years of age until they reach the retirement age. A Service Pension is paid to veterans at the same level, but five years earlier than, the Age Pension. The War Widow's/Widower's pension is paid to widowed partners and dependents of veterans. Widow Allowance is a means-tested benefit for women (born on or before 1 July 1955) widowed, divorced, or separated after turning 40, working less than 20 hours per week. Wife pension used to apply to female partners of recipients of Age Pension where those partners were not eligible for another pension. Given a relatively high immigration rate, Australians potentially receive pensions from other governments.[23]

Table 3.4 depicts conditional means, medians, and number of observations of each pension scheme taken from the HILDA dataset in 2002 and 2018 respectively. The statistics are based on individuals, who are at least 55 years old, and retired. The total number is calculated by multiplying the number of observations in the dataset with their personal population weight. The total number represented in both years increased from around 2.9 million to 4.0 million. This is in line with the figures provided by the Australian Bureau of Statistics (2021d).

Age Pension is the most important pension in both years with 1.5 (2.1) million individuals receiving Age Pension in 2002 (2018). Australian retirees who were eligible for Age Pension, received on average $12,664 per year in 2002 and $16,608 in 2018 with a slightly higher median in both years. The number of individuals

[22] Additional information on pensions are taken from Australian Government (2018). A detailed overview is provided by Harmer (2008).

[23] Not included here is the Widow B Pension, which used to be paid to widowed, divorced, or separated women aged 50 years and over, as it is not observed in HILDA, probably because it stopped for new entrances in 1997. It potentially is still included under the "Other Pension" category in Table 3.4. Further, the Mature Age Partner Allowance is not included, as it existed only from 1994 to 1996.

Table 3.4. Pensions in Australia: Annuities

	2002			2018		
Pension Type	N (in '000)	Mean (in $)	Median (in $)	N (in '000)	Mean (in $)	Median (in $)
Age Pension	1,472	12,664	13,380	2,048	16,608	17,542
Superannuation	579	22,783	17,644	1,367	28,561	20,000
Disability Pension	96	14,134	11,469	56	21,619	20,000
Disability Sup Pension	219	12,811	13,233	222	19,175	20,000
Mature Age Allowance	53	12,163	12,807	0	0	0
Service Pension	176	12,773	13,380	88	19,498	17,784
War Widow Pension	83	19,279	21,026	40	27,866	29,000
Wife Pension	27	11,433	12,577	0	0	0
Widow Allowance	16	13,607	14,704	15	11,290	14,000
Foreign Pension	246	7,567	5,293	282	5,744	4,000
Other Pension	10	3,816	662	2	4,655	2,000
Total N	2,894			3,961		

Note. Own calculation based HILDA Survey wave 18. The total number reflects the total sum of the cross-sectional population weight for individuals, who are retired and at least 55 years of age. Means and medians are in AUD and set to 2018 prices on the basis of the Consumer Price Index (The World Bank, 2021) and refer to the recipients of the respective pension type. The Mature Age Allowance and the Wife Pension were phrased out by 2018. The Widow Allowance was stopped for new entrances but still paid to individuals in 2018.

withdrawing annuities from Superannuation accounts increased from around 0.6 million to 1.4 million in the same time span. The amount is considerably higher on average in both years, with $22,783 and $28,561 respectively. The median is lower than the mean in both years, indicating a slightly left-skewed distribution. Further important pension schemes in 2002 are the Disability Support Pension, the Service Pension, and foreign pensions. The payments of the first two are around the same level as the Age Pension, whereas those of foreign pensions are on average lower. Pensions related to individuals serving for the Australian military, i.e., Disability Pension, Service Pension, and War Widow Pension, decreased over the years, which is potentially connected with the decrease of the WWII generation in this period. The Mature Age Allowance, the Widow Allowance, and the Wife Pension were stopped for new entrances in 2003 (phased out 2008), 2018, and 1995, respectively. Those eligible before these years still received their payments until the schemes were phased out.

3.3 Methodology

In this Section, I provide the definition of augmented wealth and its underlying aggregates following Bönke et al. (2019), and Wolff (2015b). I apply the accrual method, i.e., basing pension entitlements on the household's socio-economic characteristic at the observed point in time. Finally, I present the methodology for analyzing the dissaving patterns of Superannuation in retirement.

3.3.1 Wealth Aggregates

The augmented wealth definition applied here is closely related to the definition established by Bönke et al. (2019). I define the same 15 wealth aggregates as listed in Table 3.5. The gross wealth (w6) is the sum of w1 to w5, including owner occupied property (w1), other additional real estate (w2), tangible assets (w3), business assets (w4), and financial assets (w5). Subtracting debts, i.e., debts from owner-occupied property (w7), from other real estate (w8), and consumer debts (w9), from gross wealth, I receive net worth (w10). Statutory pension wealth without - (w11), and from survivor benefits (w11s), and after dissaving pension wealth (w11d)[24] add up to the social security pension wealth (w12). Pension wealth (w14) is, therefore, the sum of social security pension wealth (w12) and occupational and private pension wealth (w13). Augmented wealth (w15) is the sum of net worth (w10) and pension wealth (w14).

While aggregates w1 to w10 are standard for the distributional analysis of wealth, aggregates w11 to w15 allow for a broader perspective on wealth endowments. Private pension wealth, w13, represents wealth in Superannuation funds, which potentially is included in financial assets in standard wealth analysis. The w12 aggregate represents the present value of discounted income flows from social security pensions, listed in Table 3.4 above. In other words, the value considered as social security wealth is the actuarially fair, discounted price to which an individual would sell their social security pension claims on the complete capital market. Hence, the measure incorporates social security entitlement as the present value of pension p for individual i in year y, and is defined as

$$PV^{p}_{i,y} = \sum_{t=0}^{T}[\frac{1}{(1+r)^t}\sum_{p}\sum_{t} d^{P}_{t,i,y} \times pension^{p}_{t,i,y} \times \sigma_{t,g,c,y}], \tag{3.1}$$

where T is the "end-of-life" period, when the individual reaches the age of 100, r is a constant discount rate, i.e., 2 percent. $d^{P}_{t,i,y}$ is equal to 1 if individual i is eligible for pension p in period t. $pension^{p}_{t,i,y}$ represents the pension entitlement and $\sigma_{t,g,c,y}$ is the probability of staying alive in period t depending on gender g in cohort c in year

[24]This is the only deviation from Bönke et al. (2019) and it is explained later in this section.

Table 3.5. Wealth Aggregates

Acronym	Variable
w1	Owner-occupied property
w2	Other real estate
w3	Tangible assets (collectibles)
w4	Business assets
w5	Financial assets
w6	**Total gross wealth (sum up w1 to w5)**
w7	Mortgage debts - owner-occupied property
w8	Mortgage debts - Other real estate
w9	Consumer debts
w10	**Net worth (w6-(w7 + w8 + w9))**
w11	Statutory pension wealth without survivor benefits
w11s	Statutory pension wealth from survivors benefits
w11d	Statutory pension wealth after dissaving Superannuation accounts
w12	**Social security pension wealth (w11 + w11s + w11d)**
w13	**Occupational and private pension wealth**
w14	**Pension Wealth (w12 + w13)**
w15	**Augmented wealth (w10 + w14)**

Note. Description of the 15 wealth aggregates according to Bönke et al. (2019), p.12. This analysis adds w11d, which includes statutory pension wealth, which households may be eligible to after dissaving their Superannuation account.

y. The survival probability is provided by Australian Bureau of Statistics (2021b) for the waves under analysis.

I assume that individuals who receive social security pension in year y remain eligible throughout the rest of their life cycle. This assumption is realistic for individuals receiving Disability Pension, Disability Support Pension, Service Pension, or War Widow Pension. Individuals receiving Age Pension could potentially lose their eligibility, from one period to the next, if they failed the income and asset test, e.g., by starting to work or having increased capital gains. This is however rare.[25] A reasonable concern could be that individuals lose their Age Pension eligibility due to policy adjustments. As the accrual method relies on the expected value of futures pension schemes, it does not include future policy changes in year y. Once these changes are introduced, they affect the present value calculation.

The statutory pension wealth from survivors pension (w11s) includes the Widow Allowance, the only scheme where the payment $pension^{p}_{t,i,y}$ depends on the male partner's survival probability. Eligibility is conditioned on being female, born on or before 1 July 1955, being widowed, divorced, or separated since turning 40. Women have to meet the requirements of the income and assets test and meet residence rules, i.e., living in Australia for at least 10 years. In equation 3.1, this means that

[25] I provide more evidence for this in Section 3.4 and 3.5

the survival probability is, therefore, $(1-\sigma_{t,m,c,y}) \times \sigma_{t,f,c,y}$.[26] The economic relevance remains small, as only 0.2 percent of the whole population was eligible in 2014, and the program stopped in 2018.

Including social security pension wealth and its effect on Australian wealth inequality builds the basis of the contributions of this paper. As social security pension wealth is financed with taxes, the contribution to the tax system is indirectly included in the classical wealth analysis. Social security pension wealth increases tax rates, which potentially reduces household net incomes and, eventually, hinders wealth accumulation compared to a situation without a tax-based pension scheme. It potentially affects the wealth aggregate directly, as it reduces net-wealth. Incorporating social security pension wealth, therefore, provides a more accurate wealth measure as it includes the benefits of the pension system.

Even though my methodology is closely related to the ones by (Bönke et al., 2019; Kuhn, 2020), the peculiarities of the Australian pension system compared to Germany, Switzerland, and US cause some deviation. The social security schemes in these three countries are pay-as-you-go schemes relying on individual labor income histories and associated pension contributions. Hence, they can calculate the present value of their pension contributions at any point in their life cycle. In these countries, tax-funded social security pensions exist as a basic income support for those who received a lower lifetime income. As these payments are not per se a pension scheme, they are not included in the present value calculation by (Bönke et al., 2019; Kuhn, 2020). However, in my analysis, I include them in the present value calculation, because Age Pension is the major social security scheme in Australia. As I cannot observe whether cohorts below the retirement age will qualify for Age Pension, I only include their savings in Superannuation accounts.

Calculating dissaving rates and aggregate w11d Another central aspect is the interaction between Age Pension eligibility and Superannuation wealth. Once individuals reach the preservation age, the government lets individuals choose how much they retrieve per year. As Superannuation is included in the Age Pension income and asset test, there can be an incentive for those who are slightly above the income and wealth thresholds, to dissave Superannuation wealth at a higher rate. Contrary to other financial assets, this would not affect the individual tax rate[27] and one could potentially fulfill the income and asset test requirements. For this reason, I run an artificial income and asset test for those who reached the retirement age

[26] I refrain from including divorce rates and focus on survival probabilities.
[27] e.g., returns from selling shares falls under the capital gain tax.

and meet the residency rules. I calculate the average Superannuation dissaving rate $\overline{v_{c,y}}$ for each cohort c in year y:

$$\overline{v_{c,y}} = \frac{1}{N_{c,y}} \sum_{i=1}^{N} \frac{y_{i,c,y}^{s}}{w_{i,c,y}^{s}} \tag{3.2}$$

with $y_{i,c,y}^{s}$ describing the annual annuity retrieved from Superannuation wealth by individual i, and $w_{i,c,y}^{s}$ representing the individuals total Superannuation wealth. $N_{c,y}$ is the total size of cohort c in year y. I assign the average dissaving rate to those individuals who would pass the artificial income and asset test by reducing their Superannuation in each period of the present value calculation in Equation 3.1. Aggregate w11d is then calculated as the present value similar to the other pensions in Equation 3.1. As soon as individuals are eligible, the dummy variable $d_{t,i,y}^{P}$ switches from 0 to 1 and the individuals receive $pension_{t,i,y}^{p,imputed}$ from this period onward.[28]

3.3.2 Analysis of Dissaving Rates

I shed more light on the interaction between the two main pension schemes by scrutinizing dissaving rates. While saving for retirement has been addressed exhaustively in empirical analysis, dissaving dynamics during retirement has been covered far less. Dissaving decisions play a vital role for consumption smoothing. Retired households have to take the probability of their own life expectancy into account. Moreover, they potentially face new risk types, e.g., health risk, which could affect their income and, consequently, consumption.

An interesting feature of the Superannuation scheme is the flexibility once an individual reaches the preservation age. While many rules apply during the accumulation phase, individuals are free to choose their income stream in retirement. Moreover, Australians can retrieve the full amount at once without any financial losses. Furthermore, the tax advantages from the Superannuation scheme leave little incentives to transfer wealth away to other financial investments. Analyzing these private dissaving rates helps to understand how they vary across several socio-economic characteristics, e.g., age, household types, education, and wealth endowments. Moreover, I show that households, who become eligible for Age Pension at one point in retirement, dissave more.

I estimate a pooled fractional probit model. The advantage of this, is that compared to a binary probit model, I can take the intensive margin of the continuous dissaving rate into account. The fractional probit model was introduced by Papke

[28]This affects 7.91 percent of households with a retired household head in my working sample.

and Wooldridge (1996), who analyze aggregated employee participation rates in 401(k) pension plans in the US. The model has the following form:

$$E(\nu_i|X_i) = \Phi(X_i\beta_j) \tag{3.3}$$

where ν_i represents the dissaving rate, and X_h represents a set of the covariates of individual i and the intercept. $\Phi()$ represents the standard normal cumulative distribution function.

3.4 Data

The main source of my analysis is the Household and Income Dynamics in Australia (HILDA) Survey (Watson and Wooden, 2002). Additional data on the income and asset testing, as well as payments of the Age Pension scheme is taken from the Australian Government (2018). This section describes the HILDA dataset, especially in regard to the wealth modules. Furthermore, I define the working sample for the analysis of Superannuation dissaving rates.

Wealth data in HILDA The HILDA survey is representative for the Australian population, and it has been conducted annually by the Melbourne Institute since 2001. I use the HILDA survey, as it includes, first, broad information on socio-economic characteristics and a detailed wealth module on the household level for the years 2002, 2006, 2010, 2014, and 2018. Second, it includes information on surrender pension values $pension^p_{t,i,y}$ from all Australian pension schemes on the individual level. Third, the panel survey allows me to track wealth pensions over time. Fourth, neither the wealth module, nor the methods of data collection have been changed over time, so I can rely on consistent information over a 16-year time period.

Table 3.6 depicts the numbers of observation and mean values for selected variables in the HILDA survey over time. In 2002, 7,063 households are surveyed. With a top up in 2011, more than 9,000 households are included since then. The household member answering the questions related to the household is defined as household head. The average age of the household head is close to 50, and the average number of individuals in a household is constant and around 2.6 in all waves, and 46 to 48 percent of the household heads are female. All monetary values in my analysis are provided in 2018 prices. The mean of the households' equivalent income increases in the 2000s from \$52,449 to \$63,502, and ends up at \$67,247 in 2018. The mean net worth follows the definition of the HILDA Survey which includes Superannuation wealth (Wilkins et al., 2020) increases between 2002 and 2006 from \$598,160 to \$810,683, then decreases to \$792,315 in 2014, and rises to \$933,664 in 2018.

Table 3.6. Means of Selected Characteristics: HILDA

	2002	2006	2010	2014	2018
Age HH Head	48.08	48.70	49.07	50.00	50.54
	(0.13)	(0.14)	(0.14)	(0.14)	(0.13)
Numb HH	2.60	2.60	2.59	2.58	2.56
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Female HH Head (%)	45.92	46.31	48.39	47.64	47.57
	(0.39)	(0.41)	(0.42)	(0.40)	(0.42)
Yearly Eq. Inc HH	52,449	58,617	63,502	65,684	67,247
	(340)	(375)	(415)	(434)	(478)
Net Worth (Survey)	598,160	810,683	801,490	792,315	933,664
	(7,016)	(10,715)	(10,599)	(9,062)	(10,342)
Obs	7,063	7,003	7,193	9,363	9,486

Note. Own calculation based on the full HILDA Survey wave 18.. Weighted monetary mean values are in AUD and set to 2018 prices on the basis of the Consumer Price Index (The World Bank, 2021). Income is equalized by using the modified OECD scale (Hagenaars et al., 1994). Net worth provided here follows the definition of the HILDA Survey which includes Superannuation wealth (Wilkins et al., 2020). All statistics are based on imputed values. Bootstrapped standard errors in brackets using 1000 replica weights (Efron, 1979).

Wealth data based on survey information comes with the caveat, that the top one percent are difficult to capture adequately (Eckerstorfer et al., 2016; Kennickell and McManus, 1993). Oversampling of the rich potentially addresses this problem (Kennickell, 2008), but this is not provided by the HILDA survey. Moreover, wealth and income variables are top-coded, which affects 30 to 40 observations in each survey year.[29] Therefore, the results of my analysis do not incorporate changes for the top of the wealth distribution. Nevertheless, the HILDA survey provides reliable wealth data for large parts of the population, which is surveyed consistently over time.

Working sample for the analysis of dissaving rates Data requirements potentially put limits on an analysis of dissaving rates. Long-term panel data with wealth modules are scarce. Even if panel wealth data is available, dissaving patterns are difficult to obtain, as the extraction values from private pension accounts are commonly not provided. The HILDA dataset allows me to observe both: I can observe the total Superannuation in five waves between 2002 and 2018 and, for every year, observe the annuity retrieved from this account.

[29]The observations hold an imputed, mean preserving value. This affects the top 0.5 percent.

For the estimation of Equation 3.3, I pool all waves with the wealth module from between 2002 and 2018. I rely on individual information, as data on Superannuation accounts is one of the few wealth items which is provided on the personal level in the HILDA dataset. I restrict my dataset to retired individuals over 55 years of age, who hold at least some wealth in their Superannuation accounts. Table 3.7 compares the mean values of the full retired population in all considered waves and the mean of the working sample. There are significant differences in means, which does not come by surprise, as the working sample is selected on the Superannuation wealth variable.

Table 3.7. Selected Means: Retired Population

Variable	Full Sample	WS Regression
Annual Dissaving Rate ν	0.182	0.207
	(0.001)	(0.003)
Receive Age Pension in %	32.3	52.9
	(0.274)	(0.458)
Annuity Superannuation	6,280	15,442
	(30)	(223)
w13 Superannuation	99,892	226,405
	(944)	(3,868)
w10 Net Worth	402,752	674,070
	(3,449)	(8,094)
Female in %	56.6	46.8
	(0.340)	(0.453)
Single Households in %	29.1	25.3
	(0.299)	(0.396)
Age	71.5	69.6
	(0.066)	(0.075)
Medium Education in %	31.4	41.2
	(0.321)	(0.443)
High education in %	11.1	20.8
	(0.147)	(0.200)
Obs	15,019	5,679

Note. Own calculation based on HILDA Survey wave 18. Pooled dataset include the waves 2002, 2006, 2010, 2014, and 2018. *Full Sample* describes all retired individuals who are at least 55 years. *The Working Sample* includes all retired individuals who are at least 55 years old, and who hold at least some Superannuation wealth. All statistics are based on imputed values. Monetary values are in AUD. Bootstrapped standard errors in brackets using 1000 replica weights.

The mean annual dissaving rate in the working sample is slightly higher than the one in the full sample. This is due to the fact that the observations outside of the working sample have primarily missing -, and some zero values. The latter is the case, if individuals claim zero values for both income from Superannuation and wealth in Superannuation accounts. More than thirty percent of the full retired population receive income from Superannuation and Age Pension during their retirement. This increases to 55 percent in the working sample, as those who only receive Age Pension and do not hold any Superannuation wealth, drop out. The remaining covariates show that individuals owning Superannuation wealth hold generally more net wealth, are proportionally more male, and more educated. I control for all these dimensions in the regression analysis.

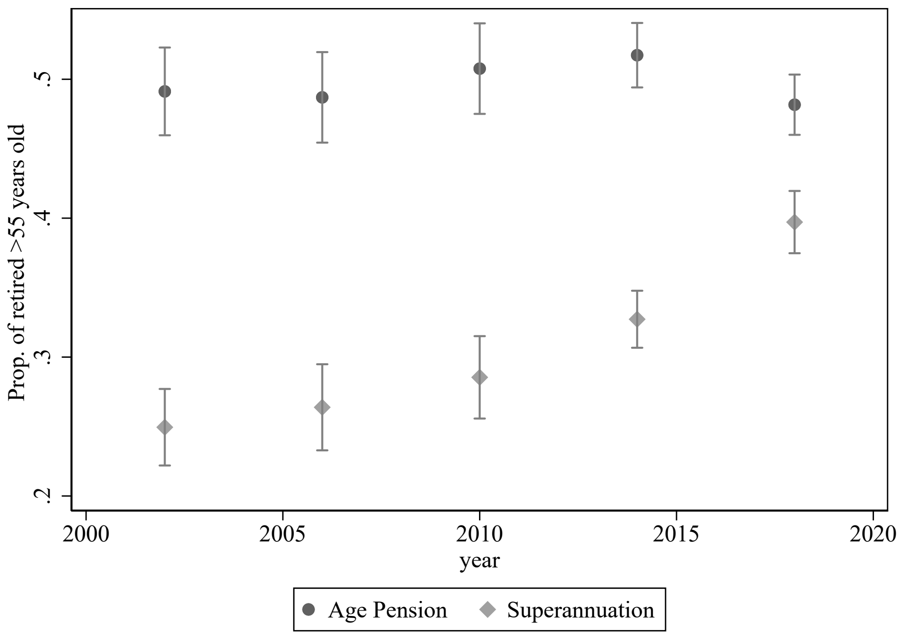
3.5 Empirical Findings

This section provides the results of my analysis and it is divided into two major subsections scrutinizing income during retirement and augmented wealth inequality. The first subsection provides an overview of the main income sources during retirement over time. I also provide the results of my inquiry on Superannuation dissaving rates. Hence, the first subsection focuses on the retired population only. The second subsection analyzes augmented wealth inequality, focusing on the contribution of the pension schemes and describing differences across age cohorts, and sets the results within the context of augmented wealth in Germany, the US (Bönke et al., 2019), and Switzerland (Kuhn, 2020).

3.5.1 Income during Retirement

This subsection sheds light on the general ramifications of the main pension schemes' relevance for the retired population. The relative importance of the Superannuation scheme rises significantly during the 2002 to 2018 time span. Figure 3.1 provides the proportion of retired households with a household head at 55 years of age or higher, who receive at least some Age Pension or withdraw some positive annuity from their Superannuation account. The proportion of recipients of Age Pension is relatively constant at around 50 percent throughout the considered years. The proportion of those withdrawing Superannuation annuities, however, increases from around 26 percent to 40 percent. The dominant factor for the increase is the establishment of the scheme: retired individuals in 2018 were enrolled for a longer time span throughout their life-cycle in comparison to those in earlier years. Although the establishment of the scheme was affected by the general financial crisis (GFC) at the end of the 2000s, it did not stop its growing relevance.

Income and Gross Wealth Even though the Age Pension and Superannuation are the main source of income in retirement, income from private investments, rent and, to some extent employment, are further sources of income. Nonetheless, they contribute differently to households' income along the wealth distribution. I present the income means and ratios to overall income along the gross wealth distribution in Figure 3.2. The panels include households with a retired household head. For the sake of clarity, I compare the first and last observed wave of the considered



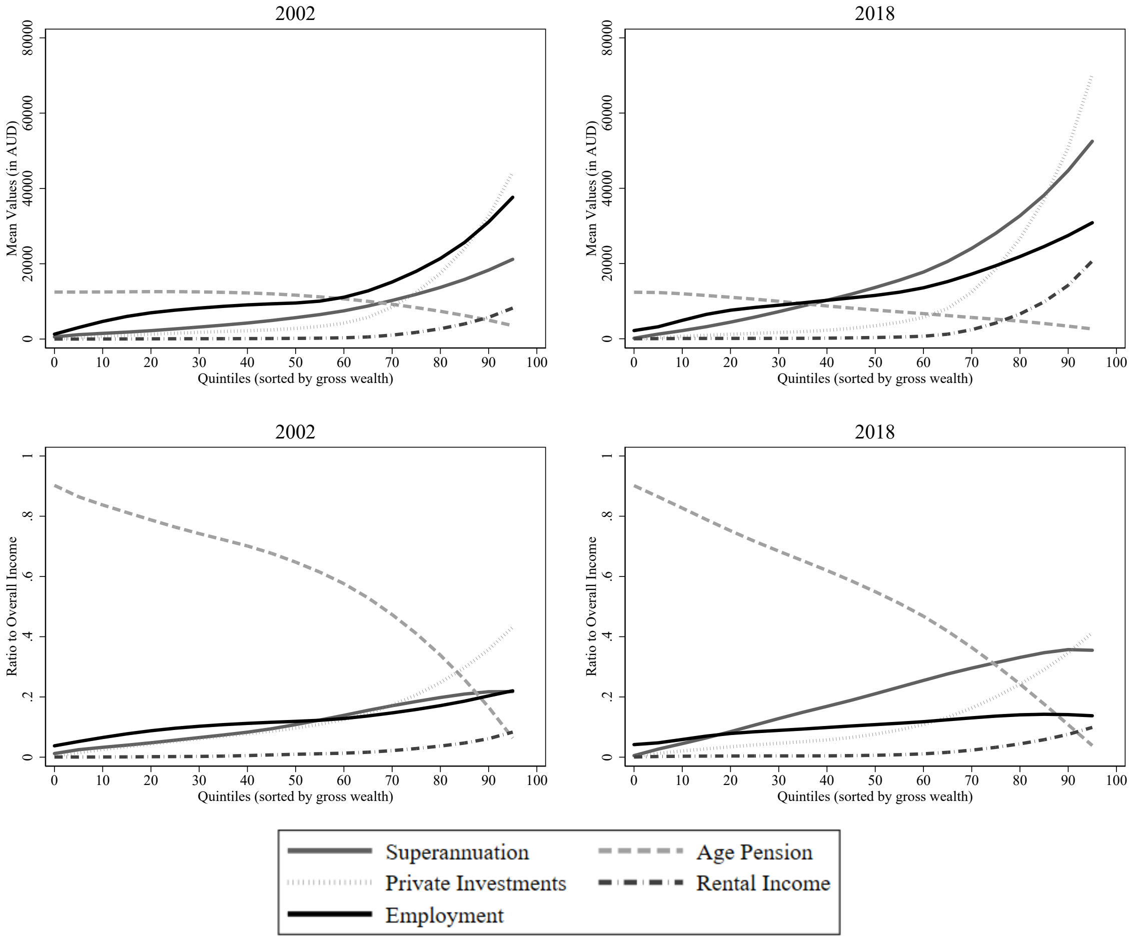
Note. Own calculation based on HILDA Survey wave 18. Numbers represent the proportion of households with a retired household head over the age of 55 receiving some positive Age Pension and/or some positive annuity from their Superannuation accounts. Brackets represent bootstrapped 95 percent confidence intervals. All estimates are weighted with cross-sectional household weights and based on imputed data.

Figure 3.1. Age Pension and Superannuation Accounts in Australia by Wave

time span, that is 2002 and 2018. These panels contribute to the understanding of pension wealth aggregates as they are closely linked: social security pension wealth is the present value of future income flows of public transfers, i.e., predominantly Age Pension. Moreover, Superannuation annuities are used to calculate the dissaving rates, which are analyzed later in this subsection.

The upper two panels describe the absolute mean values for income from public transfers, private investments, rental income, and employment. Age Pension payments remained relatively stable while Superannuation annuities increased considerably over the 16 years. The mean income from private investments increased only for the upper 30 percent of the wealth distribution. Rental income stayed relatively constant and is equal to 0 for most households and, again, is more important for the upper 30 percent. This is not surprising, as additional housing wealth investments for rental purpose mainly appear at the top of the wealth distribution.

Income from employment, which is earned by non-retired partners or children in the household, is relevant but it slightly decreased over time.[30]



Note. Own calculation based on HILDA Survey wave 18. Sample includes households with a retired household head over 55 years. The weighted mean of 5 percentiles along the gross wealth-distribution and smoothed with LOWESS. Monetary values are are in AUD and set to 2018 prices on the basis of the Consumer Price Index (The World Bank, 2021). All statistics are based on imputed values.

Figure 3.2. Income in Retirement and Gross Wealth

The lower two panels of Figure 3.2 provide the income ratios of the same income types. In both years, public transfers represent around 90 percent of the total income for households at the lower end of the gross-wealth distribution. This share steadily decreases until income from private investments (Superannuation) are more important from the 85th (75th) gross-wealth percentile in 2002 (2018). Income from

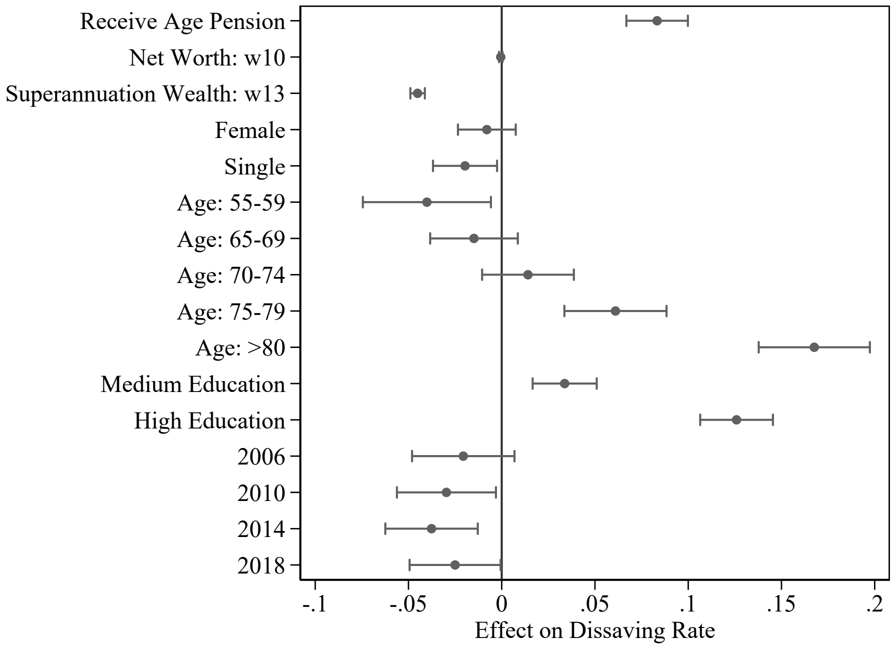
[30]The population shown here focuses on household heads who retired from the work force. Potentially, they could be employed part-time and receive e.g., Age Pension, as long as the total amount remains below the allowance.

Superannuation became more important for nearly all households along the gross-wealth distribution between 2002 and 2018. On the contrary, income from private investments is relatively less important in 2018. Nevertheless it remains, together with Superannuation, the most important source of income for the top 15 percent. Income from employment contributes around 10 percent for most parts of the gross-wealth distribution to the overall income, which is less than the approximately 20 percent in 2002. Reasons for this are difficult to pin down, as this income is mainly provided by other household members. One pension policy that could contribute to this result, is the increase of the deduction rate on income at the end of the 2000s, encouraging a reduction in earnings to keep up the eligibility for the Age Pension scheme.

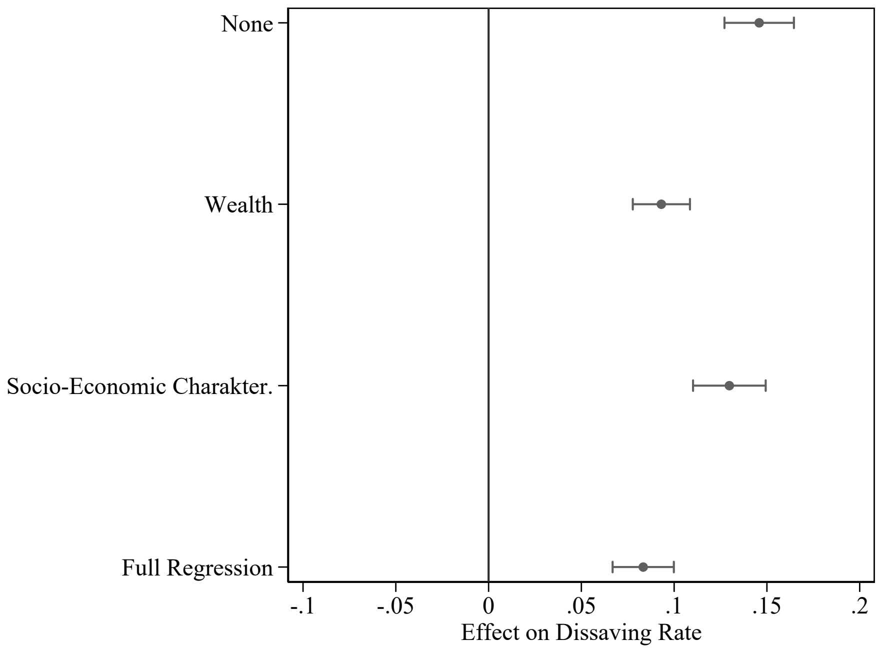
I can conclude from this that Age Pension is the most important source of income in 2002 and 2018. Even if income from Superannuation accounts and private investments are added up as overall private pension income, public transfers are still more important for more than 50 percent of the gross-wealth distribution. The relevance of Superannuation increased between 2002 and 2018, potentially replaced income from private investments. This does not come as a surprise, as the scheme started in the 1990s and the retired population in 2018 had to participate in the scheme for large parts of their working life compared to the retired population in 2002. Additionally, tax incentives may support the transfer of private investments into the scheme. Figure 3.2 also indicates that the upper half of the wealth distribution has started replacing age pension, while the lower half has not. This development hints at a two-tiered society of retirees: those who rely on Age Pension and those that rely on Superannuation. I find more implications for this in my analysis of dissaving rates.

Dissaving Superannuation in Retirement In this subsection I present the results of the pooled fractional probit model. As discussed in the methodology section, the working sample is based on 5,679 retired individuals in five waves, who hold at least some wealth in their Superannuation account. I present the marginal effects at the mean (MEM) of the covariates on the Superannuation dissaving rate with the 95 percent confidence interval in Figure 3.3. The MEM describes the partial effect of the independent variable, when all other covariates are set at their means. The full model, as well as OLS estimates, are provided in Appendix 3.8.2. The upper panel of Figure 3.3 provides the MEMs for all covariates used in the regression, the lower panel demonstrates results for several model specifications for the binary variable, which is equal to 1 if individuals receive Age Pension.

The regression results provide several insights into dissaving at retirement. Individuals receiving Age Pension choose higher dissaving rates. The effect size is not trivial: for the average individual, the dissaving rate is 8.33 percentage points higher per year. A back-of-the-envelope calculation would suggest that, all else equal,



(a) Marginal Effects at the Mean from Covariates on the Dissaving Rate



(b) Effects of the "Receive Age Pension Dummy' on Dissaving Rates for Different Model Specifications

Note. Own calculation based on HILDA Survey wave 18. The working sample includes retired individuals over 55 years of age. Panels provide the marginal effects at the mean for several covariates. The confidence intervals are specified at the 95% level. The upper panel provides the full set of the main regression. The lower panel shows the marginal effect at the mean from those receiving Age Pension for several different model specifications. The y-axis describes the set of covariates included in the regression. All statistics are based on imputed values.

Figure 3.3. Regression Results of the Fractional Probit: Marginal Effects at the Mean

a fund of $100,000 would reduce to $24,106 in 5 years, instead of the $35,891 at the mean dissaving rate (21.6, see table above) of the working sample. Australian retirees, hence, seem to react to the interplay between Superannuation and Age Pension and dissave their accounts faster, potentially to qualify for (higher) Age Pension rates. The methodology applied in this analysis does not verify a causal relationship, but given the benefits for being only partially eligible, it is likely to be one central driver of this behavior. Aside from this objective, those individuals may also have to use comparably more wealth from Superannuation accounts to guarantee their consumption levels. The effect is robust to several specifications of the fractional probit model. The lower panel provides the results of those receiving Age Pension with different model specifications. The estimate of the full regression is the lower bound:[31] adding only wealth or socio-economic characteristics as covariates to the Receive Age Pension dummy variable does not change the significance of the result.

Net worth and Superannuation wealth are divided by $100,000. The MEM of net worth is statistically and economically zero. Wealth levels of financial assets and housing investments are included in net worth and they do not seem to drive the individual dissaving rate decision in this specification.[32] Superannuation wealth, however, is negatively associated with the saving rate. An increase of $100,000 decreases the dissaving rate at the means by 4.51 percentage points. One reason for this is potentially the rising importance of other income sources at the higher end of the wealth distribution, as shown above. Moreover, the tax advantage for holding wealth in Superannuation accounts is much stronger for individuals in higher marginal tax brackets. Thus, this can be reflected in the lower dissaving rate.

The socio-economic characteristics show that women do not dissave significantly differently from men. Single households extract slightly less than those living in coupled households, but the estimate is only significant at the five percent level. A negative coefficient could indicate higher exposure to income risk, as household pooling is not possible, therefore choosing more prudent rates. The age dummies reveal a clear increase in age, significant from the 75-79 bracket. A credible factor would be the decreasing insecurity over an individual's life expectancy. A bequest motive could limit, however, the excess of the dissaving rate, as Superannuation wealth is not subject to inheritance or other additional tax that affects the transition to the closer kin. Individuals with a medium level of education, here classified as at least twelve years of schooling and less than a bachelor degree, as well as those with a high level of education, i.e., bachelor degree and above, dissave relatively more. It

[31] The OLS regression provides a similar effect size, which is provided in Appendix 3.8.2

[32] This result comes with the caveat, that net worth is measured on the household level. By sharing resources, many components could affect the individuals dissaving decision, e.g., joint consumption decisions or tax-considerations. Therefore, the true net worth effect is difficult to classify from this model specification. If only financial wealth or additional housing wealth is included, instead of net wealth as a covariate, the coefficient remains around zero. Regression result are available upon request.

is not straightforward to qualify the result, as a higher dissaving rate is per se not more or less efficient.

I also include year dummies in the regression, showing that the year-specific conditional mean of the dissaving rate reduces over time. This is may be explained, again, by the growing maturity of the Superannuation scheme associated with increasing life expectancy. Naturally these are long-term trends and, therefore, probably do not dominate the choice of the decision rate.

From this subsection, I conclude that Age Pension remains by far the most important payment for individuals and households, with income from Superannuation accounts following second. Further, I focus on the dissaving rate of Superannuation wealth and I find suggestive evidence for individuals reactions due to the provision of Superannuation wealth and income in the Age Pension eligibility test. The income and asset tests can be considered as an indirect tax on Superannuation wealth, as it reduces or hinders social security payments, but only for some, i.e., for those who would receive the full or partial pension without Superannuation wealth in the picture. This bears three major concerns: first, minimizing the tax burden, the policy may force individuals to choose inefficient dissaving rates, in regards to consumption smoothing or insurance decisions, especially at the wake of increasing health risks with age. This could induce welfare losses. Second, the policy may encourage individuals who are potentially eligible for Age Pension to reduce contribution to their Superannuation scheme before they reach the retirement age. As the contribution is linked to the employment status, this could lead to distortions, i.e., lower supply at the labor market. Third, in combination with the nearly unlimited access after the preservation age, this could contribute to tax evasion strategies, transferring wealth into less traceable means, i.e., cash or wealth in accounts overseas.

3.5.2 Augmented Wealth Inequality

This subsection describes augmented wealth inequality from several angles. Aside from basic statistics of the wealth aggregates, I show portfolio shares along the wealth distribution over time, Gini coefficients and percentile ratios, and illustrate wealth levels along several household characteristics, e.g., age and household types.

Descriptive Statistics Table 3.8 provides the mean values of the wealth aggregates w1, w5, w6, w10, w12, w13, and w15. All values are provided in Australian dollars and in 2018 prices. I observe a substantial increase in all wealth aggregates between 2002 and 2006, except for social security pension wealth. Mean gross wealth increases between 2002 and 2010, slightly decreases in 2014 and increases significantly to $896,743 in 2018. Net worth provides similar patterns. The estimates suggest that the home owner residence value (w1) drives much of that development, as a comparably high ownership rate, i.e., around 65 percent in the total population

(Wilkins, 2016) coincides with a steep increase in housing prices in Australia (Knoll et al., 2017; Wilkins et al., 2020). The mean of the financial wealth (w5) aggregate rises by around 40 percent during the 16 years, only stagnating between 2006 and 2010 – the time of the GFC. Social security pension wealth remained relatively constant over time, with an increase in 2014 followed by a decrease close to the 2010 level at $75,805. I also include wealth from survivor benefits (w11s), which depend on widow allowance payments. Only very few households receive pensions from this scheme and it stopped in 2018. The mean pension wealth from Age Pension after dissaving their Superannuation account (w11d) is higher than those from aggregate w11s, but not also not relevant for the overall wealth distribution.[33] The mean of Superannuation (w13) doubled between 2002 and 2018, which can be explained by the advanced maturity of the scheme in combination with a high take-up rate. The family home represents, on average, the most valuable asset in 2018, followed by Superannuation wealth, financial assets, and eventually social security pension wealth.

Australian households, on average, yield very high augmented wealth positions (w15). The aggregate increases greatly between 2002 and 2006, plateaus between 2006 and 2014, and increases again in 2018. During the considered 16 years, augmented wealth has increased by 51 percent. Increasing housing (households' main residence, HMR) wealth and Superannuation drive these results.[34] Table 3.8 also shows the relevance of pension wealth, as the mean of augmented wealth is 45 percent higher than net wealth in 2018.

The pension schemes were affected by several policy adjustments and they may help to explain the changes over time. The rise of the retirement age for women and the bi-annual adjustments of the thresholds and pension payments contribute in keeping the social security pension wealth values stable, even though the numbers of retirees has increased. The relatively steep upturn between 2010 and 2014 is likely due to a composition effect: the retirement age for women was successively increased from 60 to 65 until 2014. Therefore, several women had to postpone their retirement and become eligible in that year. The decrease of the mean from 2014 to 2018 might be a consequence of the rather large threshold adjustment by the government and the first step of increasing the retirement age to 67 for men and women. Major changes of the Superannuation scheme, like the reduction of the concession cap in the early 2010s, almost certainly reduced the Superannuation wealth growth. However, the results do not provide a counterfactual scenario, which would also be difficult to disentangle from the aftermath of the GFC at that time. The increase of the preservation age between 2014 and 2018 possibly increases wealth in Superannuation accounts, as the accumulation phase is prolonged. Again, which is

[33]Henceforth, w11s and w11d are not presented separately.
[34]This is also shown by Wilkins et al. (2020).

Table 3.8. Selected Weighted Mean Values based on HILDA

	2002	2006	2010	2014	2018
w1: HMR Value	307,381	407,842	443,434	420,146	508,951
	(2,834)	(3,884)	(4,224)	(3,674)	(5,145)
w5. Financial Wealth	101,038	125,351	121,199	134,375	147,243
	(2,142)	(3,252)	(3,093)	(2,850)	(3,225)
w6: Gross Wealth	575,738	805,250	809,938	775,967	896,743
	(6,692)	(10,869)	(10,249)	(8,296)	(10,053)
w10: Net Worth	474,784	650,131	624,835	589,465	692,618
	(6,002)	(9,672)	(8,933)	(7,073)	(8,668)
w11s: PW from Survivor Benefits	22	225	629	676	0
	(7)	(7)	(19)	(25)	(0)
w11d: PW after Dissaving Super	1,102	769	1,315	2,056	3,403
	(132)	(105)	(139)	(194)	(257)
w12 Social Security PW	70,242	73,415	71,800	81395	75,805
	(1,219)	(1,457)	(1,375)	(1,563)	(1,478)
w13: Superannuation	123,280	160,876	177,978	203,086	240,726
	(1,979)	(2,911)	(3,245)	(2,761)	(3,135)
w15: Aug. Wealth	668,306	884,422	874,613	873,946	1,009,149
	(6,865)	(10,999)	(10,979)	(8,778)	(10,296)

Note. Own calculation based on the full sample of the HILDA Survey wave 18.. Weighted mean values are in AUD and set to 2018 prices on the basis of the Consumer Price Index (The World Bank, 2021). All statistics are based on imputed values. Bootstrapped standard errors in brackets using 1000 replica weights.

difficult to disconnect from the developments of international stock markets, that significantly rose during that period.

Looking beyond the mean, Table 3.9 provides the mean, median, 25th-, 75th-, and 90th percentiles of selected wealth aggregates in 2018. Typical for net worth is having a much lower median than the mean, indicating a highly rightly-skewed distribution, with \$35,000 at the 25th –, \$858,900 at the 75th –, and \$1,718,000 at the 90th percentile, respectively. It also provides the ratio of households which hold a positive amount in the wealth aggregate. For net worth in 2018, these were 90.51 percent of all households. Social security pension wealth applies for 21.07 percent of Australian households. Superannuation is much more dominant in comparison to the social security pensions, but its distribution is rightly-skewed as well. 84.88

Table 3.9. Descriptive Statistics Wealth Aggregates: 2018

Aggregates	Mean	p10	p25	p50	p75	p90	Frac>0
w10: Net Worth	692,618 (10,053)	200	35,000	357,797	858,900	1,718,000	90.51 (0.33)
w12 Social Sec PW	75,805 (1,478)	0	0	0	0	357,809	21.07 (0.44)
w13 Super	240,726 (3,135)	0	16,628	95,000	275,000	600,000	84.88 (0.12)
w15 Aug. Wealth	1,009,149 (10,296)	28,000	165,830	601,000	1,307,299	2,325,522	96.20 (0.14)

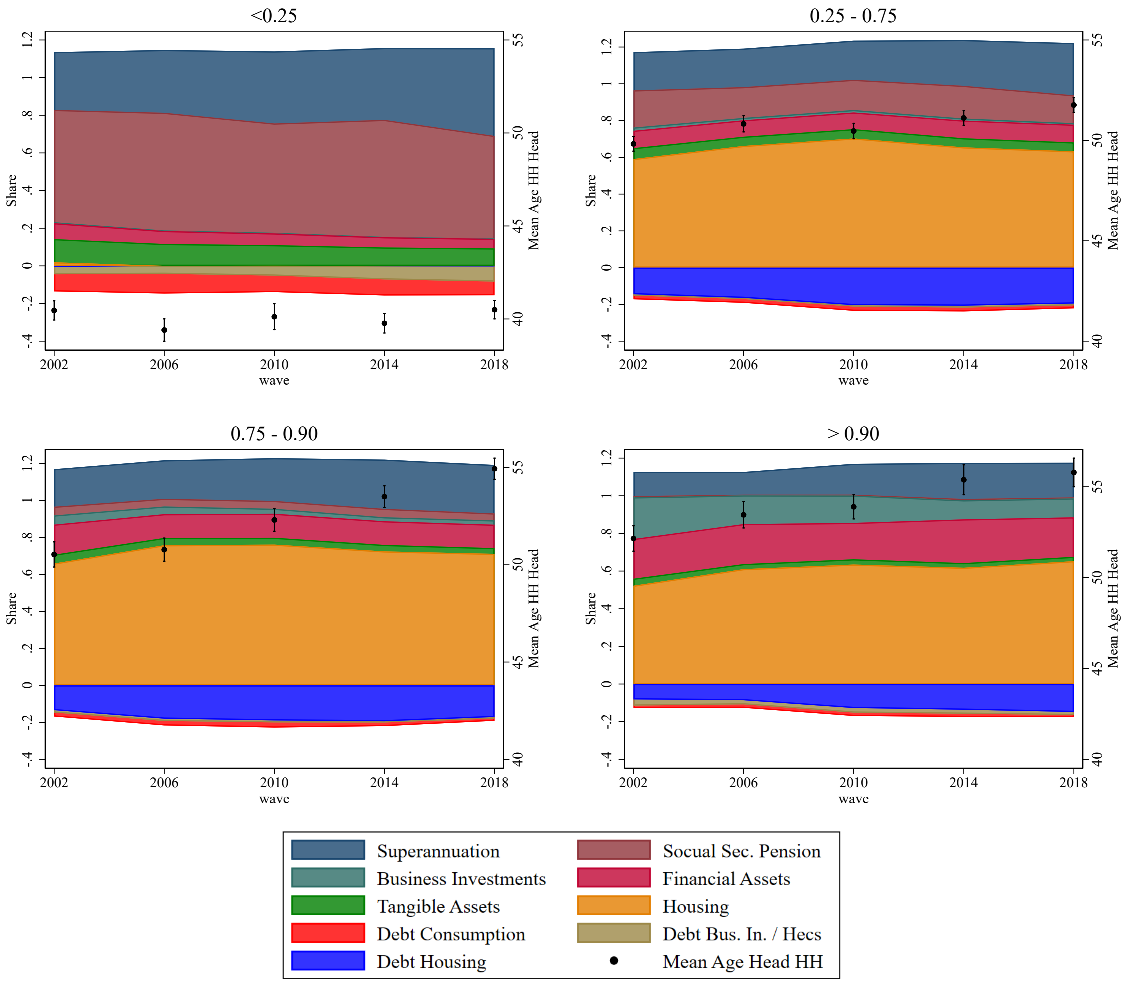
Note. Own calculation based on the full sample of the HILDA Survey wave 18. Monetary values are in AUD. All statistics are based on imputed values. Bootstrapped standard errors in brackets using 1000 replica weights.

percent of Australian households hold at least some wealth in Superannuation accounts. Augmented wealth is considerably higher than net worth for all statistics shown in Table 3.9 , and only 3.8 percent of all households do not possess a positive amount of augmented wealth.

Wealth Portfolios The analysis of wealth portfolios shows the relative importance of different wealth components along the gross wealth distribution over time. Figure 3.4 depicts four graphs plotting the mean wealth portfolio shares for households below the 25th percentile of the gross wealth distribution ("the poor"), for those between the 25th and the 75th percentiles ("the middle"), for those between the 75th and the 90th percentile ("the upper middle") and those above the 90th percentile ("the rich"), respectively. For the sake of clarity, all components are divided by the year and group mean of augmented wealth.[35] As debts are included, the shares can be larger than 1. Moreover, the mean age of the household head for each group at each wave is provided at the right y-scale.

The figure reveals that the portfolio composition and the relative importance of pension wealth vary along the distribution. The first panel shows that pension wealth is the most important wealth component for "the poor" and this finding is consistent over time. The Superannuation wealth share increases from 35 percent to 46 percent for "the poor". Social security is the most important asset at the lower end of the wealth distribution share, i.e., 56 percent of augmented wealth in 2002, decreasing slightly to 53 percent in 2018. This changes for "the middle

[35]Tables and graphs with monetary values are provided in Appendix 3.8.3.



Note. Own calculation based on the full sample of the HILDA Survey wave 18. The four groups are defined by percentiles in the gross wealth distribution. The share is calculated as the ratio between the weighted average of the wealth aggregate and the weighted average of augmented wealth in the respective subgroup. The right scale provides the mean age of the household head in the respective group and year. All estimates are based on multiple imputations.

Figure 3.4. Portfolio Shares Along the Gross-Wealth Distribution

class" group in the second panel, where housing becomes the most important wealth component and the relative importance of pension wealth decreases relatively to "the poor". Financial assets play a bigger role from here, but are still less important than pension wealth for "the middle class". Housing wealth remains by far the most significant wealth aggregate for "the upper middle" and "the rich", the relevance of social security pension diminishes. Furthermore, the proportion of financial assets increase in the portfolio. The relative importance of Superannuation reduces for "the rich" category, where its share decreases while business investments and financial assets increase their share, respectively.[36]

The portfolios provide consumption -, business/HECS[37] -, and housing debts. Consumption debts are the main source of debts for the lower end of the wealth distribution, which is replaced by housing debts for the wealthier groups. Business/HECS debts are of minor relevance for all groups, compared to the other debts and wealth sources.

The four panels also show the age gradient for each group in each wave. "The poor" household heads are on average around 40 years, hence, considerably younger than those in the upper parts of the gross-wealth distribution. As wealth accumulation continues throughout the working life, some households in "the poor" group potentially end up in one of the groups above in the following years. However, as social security pension wealth is only provided for retirees, there is also a considerable fraction of older household heads in "the poor" group. Furthermore, the age difference is relatively small between "the middle", "the upper middle", and "the rich". Life cycle accumulation patterns do not seem to be a main determinant as to whether a household belongs in one of these three groups.

These portfolio patterns contribute to the understanding how relevant social security pension wealth is for the lower end of the distribution, especially compared to Superannuation wealth for "the poor". Nonetheless, this could change in the years to come, as Superannuation wealth is constantly increasing its share. This offsets some of the distributional consequences of the housing boom, as they are not captured by this group. Even in "the middle", social security pension wealth plays a vital role. Financial assets and business investments seem to be an aggregate rather for "the upper middle" and "the rich". The Superannuation scheme has increased wealth for all four groups, which points to the success of the scheme during the considered periods. However, it is still open to debate whether these gains would have occurred in other financial assets if the scheme had not been introduced.

Wealth Modules and Inequality I continue with the description of the wealth aggregate distributions over time. The distribution of wealth in Australia has been studied before. Using data of the Australian Bureau of Statistics (ABS), Harding (2002) finds a net wealth Gini coefficient at 0.64 in 1986 and in 1998, which remained constant due to the equalizing effect of the Superannuation accounts. Later studies reveal estimates between 0.6 and 0.65 using ABS data (Kaplan et al., 2018a) or HILDA data (Headey et al., 2005; Frick and Headey, 2009; Sila and Dugain, 2019; Wilkins, 2016). The trends over time are, however, controversial as Kaplan et al. (2018a) describe increasing inequality, while Sila and Dugain (2019) and Wilkins

[36] Similar patterns were found in the US by Kuhn et al. (2017).

[37] Stands for "Higher Education Contribution Scheme" and includes tertiary education fees

(2016) find stable inequality patterns. This paper contributes to the discussion by adding the, thus far, not included social security pension wealth aggregate.

Table 3.10. Gini Coefficients of Wealth Aggregates

Aggregates	2002	2006	2010	2014	2018
w10: Net Worth	0.644	0.661	0.649	0.661	0.664
	(0.0035)	(0.0040)	(0.0040)	(0.0032)	(0.0033)
w10 + w11: Personal Ent.	0.602	0.624	0.615	0.622	0.628
	(0.0035)	(0.0040)	(0.0040)	(0.0033)	(0.0034)
w10 + w12: Social Sec PW	0.602	0.624	0.615	0.622	0.628
	(0.0035)	(0.0040)	(0.0040)	(0.0044)	(0.0034)
w10 + w13: Superannuation	0.619	0.632	0.625	0.628	0.624
	(0.0035)	(0.0037)	(0.0039)	(0.0031)	(0.0030)
w15: Augmented Wealth	0.577	0.597	0.593	0.592	0.592
	(0.0033)	(0.0037)	(0.0039)	(0.0032)	(0.0030)
Net worth <= 0 (%)	7.66	9.41	9.75	10.93	11.10

Note. Own calculation based on the full sample of the HILDA Survey wave 18. All estimates are based on imputed values and weighted with household population weights. Bootstrapped standard errors in brackets using 1000 replica weights.

Table 3.10 presents the Gini coefficients of net worth, adding up social security pension wealth and Superannuation to augmented wealth in the considered waves. Based on my estimates, the Gini index of net worth inequality increased slightly over time. Bottom coding or censoring does not alter the estimates veritably, as only a relatively small proportion of the distribution holds negative net worth.[38] My estimates of the Gini coefficient correspond to previous estimations with HILDA data in Headey et al. (2005), Frick and Headey (2009), and Wilkins (2016). Aggregate "w10 + w13" comes closest to the net wealth definition in the survey, as it includes wealth from Superannuation accounts.

Adding pension wealth to the net worth distribution decreases the measured inequality, as augmented wealth exhibits lower Gini estimates than net worth in all waves. The Gini coefficient of augmented wealth increases slightly over time, i.e., from 0.577 in 2002 to 0.592 in 2018. The aggregates w11s and w11d have no distributional relevance, which is shown by the same estimates of "w10+w11" and "w10+w12". In 2002, net worth and social security pension wealth are slightly more equally distributed than net worth combined with wealth in Superannuation accounts. However, this changes by 2018, when both pension aggregates exhibit about the same level of inequality.

Relatively constant Gini coefficients are not what one would necessarily expect in the Australian case. Economic growth can be observed constantly between 2002 and

[38] Corresponding estimates are provided in Appendix 3.8.4

2018, and it is, in advanced economies, associated with growing wealth inequality during the last decades (Islam and McGillivray, 2019; Stiglitz, 2012), as economic growth is more beneficial to high-income-earners, and they choose higher saving rates (Saez and Zucman, 2016). This does not seem to hold for Australian households. The share of households in the sample which possess zero net worth or less has, however, increased from 7.66 to 11.10 percent. Hence, the economic growth does not seem to coincide with a broader accumulation of wealth.[39]

Going beyond the Gini coefficient, I provide percentile ratios to investigate the tails of the net worth and augmented wealth distribution. Figure 3.5 depicts the 20/50 and 90/50 percentile of the two aggregates between 2002 and 2018.[40]. The first panel shows that households held less than 10 percent of the median net worth at the 20th percentile in 2002, and around 5 percent in 2018, while they hold around 18 percent of the median augmented wealth. The second panel shows that households held 4.1 times the median net worth at the 90th percentile in 2002, increasing to 4.8 in 2018, at the same time holding 3.3 (3.8) time the median augmented wealth in 2002 (2018). As shown with the Gini coefficient above, percentile ratios of augmented wealth exhibit less inequality than those of net worth. However, the differences are more distinct at the top tail of the distribution, i.e., the 90/50 percentile ratio. Table 3.17 in Appendix 3.8.6 shows that this is fully attributable to social security pension wealth. Finally, the 90/50 ratio reveals an increase of inequality on both wealth aggregates in the 2010s.

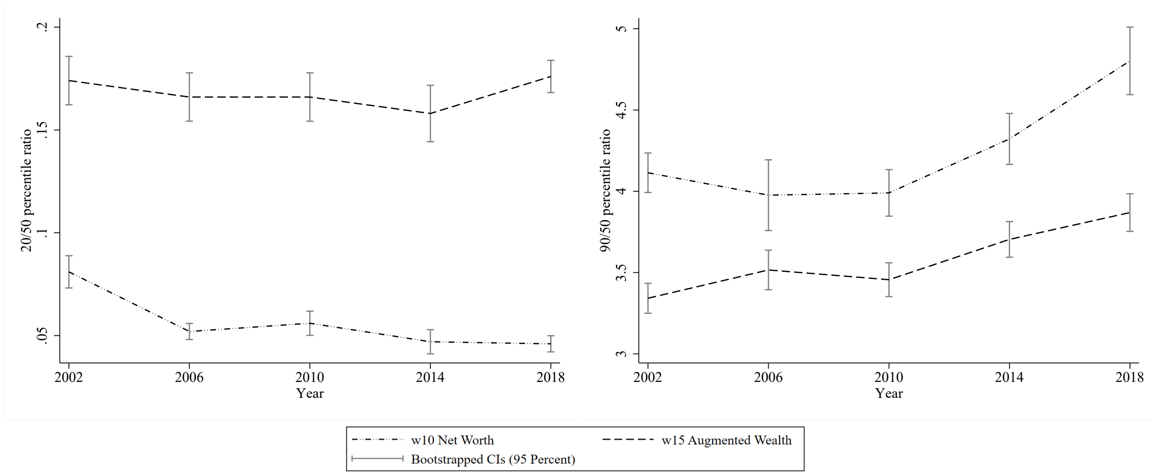
Effective distributional consequences from policy changes of the Age Pension or Superannuation scheme cannot be clearly identified in these figures. This, however, does not mean that they do not occur, as the estimates presented here assess overall wealth inequality. Adjustments of the pension scheme may take more than the observed years to reveal long-term consequences.

It is important to keep in mind that the applied wealth data is top coded and survey data potentially does not capture the top adequately. This may explain why other datasets show a more significant increase of inequality in the same period (Australian Bureau of Statistics, 2021a). The 90/50 percentile ratio also hints towards a larger variation at the top. It also shows that including social security pension wealth affects the level, but not trend of inequality, as its contribution is relatively constant over time and plays no role at the top of the wealth distribution.

The assessment in this subsection shows, that Superannuation has not hitherto increased overall wealth inequality in the considered periods and its compulsory

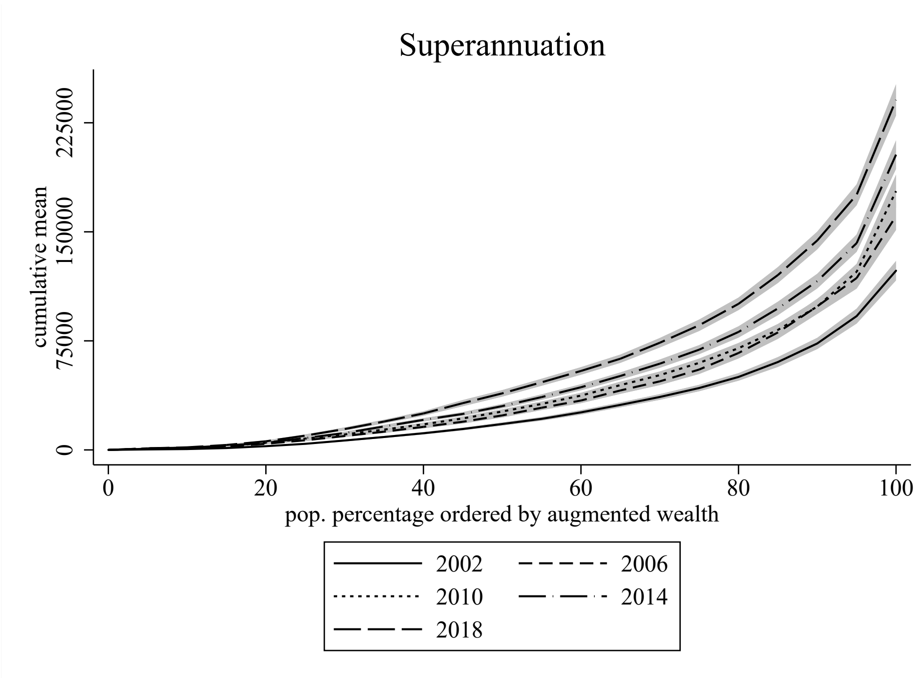
[39] To further investigate the contribution of net worth and pension wealth to the overall augmented wealth inequality, I apply a factor decomposition of the wealth aggregates in Appendix 3.8.5. The decomposition originates from Lerman and Yitzhaki (1985) and is applied in the same context by Bönke et al. (2019).

[40] I choose those percentiles as they guarantee that individuals hold some positive wealth at the lower end. Table 3.17 in Appendix 3.8.6 shows the corresponding numbers for all pension wealth aggregates.



Note. Own calculation based on the full sample of the HILDA Survey wave 18.. The left (right) panel provides the 20/50 (90/50) percentile ratio over time. All estimates are based on multiple imputations. Bootstrapped standard errors are based on 1000 replica weights.

Figure 3.5. Percentile Ratios: 20/50 and 90/50



Note. Own calculation based on the full sample of the HILDA Survey wave 18. Generalized concentration curves of Superannuation from 2002 to 2018. Values are in AUD and set to 2018 prices on the basis of the Consumer Price Index (The World Bank, 2021). Grey area represents 95 % confidence intervals.

Figure 3.6. Generalized Concentration Curves of Superannuation and Augmented Wealth: 2002 to 2018

component supported a broad range of households to accumulate wealth for retirement. There is some indication that this could change in the future. Figure 3.6 provides generalized concentration curves of Superannuation wealth along the population ordered by augmented wealth. The cumulative mean of Superannuation increased more at the upper end of the augmented wealth distribution, while over 15 percent did not accumulate any Superannuation wealth at all. This pattern is not surprising, as the Superannuation Guarantee depends on monthly labor income, which is highly correlated with wealth. Moreover, the tax incentives are higher for those at the upper end. Even though the maximum contribution cap was reduced in the 2010s, this did not stop these dynamics, as there is a significant increase between 2014 and 2018. The up to $500 annual support for low income earners does not seem to induce essential wealth gains for the lower end of the distribution.

Despite these absolute differences, households at the top did not increase their overall share of Superannuation wealth. In Appendix 3.8.7, I provide the normalized concentration curves for 2002 and 2018. It shows that the concentration curves are not significantly different in 2018 compared to 2002. Nevertheless, the top 20 percent hold nearly 60 percent of the overall Superannuation wealth. Given that this can be transferred from one generation to the next, these differences can accumulate over time and may contribute more to inequality in the future. I discuss this in more detail below.

Life Cycle Patterns I now turn to wealth aggregates along age patterns. As wealth from the Age Pension scheme affects retired households, the first part focuses on wealth levels and inequalities for different birth cohorts of the retired population. The second part takes the whole population into account. The individual age determines the stage of wealth accumulation over the life cycle. Following the neoclassical theory, individuals choose their saving rate to smooth their consumption over time. Previous research shows an inverted u-shaped pattern of wealth accumulation throughout an individual's life, with an increase throughout the working life and a decline, once an individual retires (Atkinson, 1971; Davies and Shorrocks, 2000). There are many factors that shape the wealth accumulation patterns, e.g., saving rates, especially precautionary saving (Cagetti, 2003; De Nardi and Fella, 2017), investment behavior (Calvet et al., 2007a; Benhabib et al., 2011; Lusardi et al., 2017; Longmuir and König, 2019), and bequest motives, either accidental (De Nardi, 2004) or due to the "warm glow" (Andreoni, 1989; De Nardi and Fella, 2017).

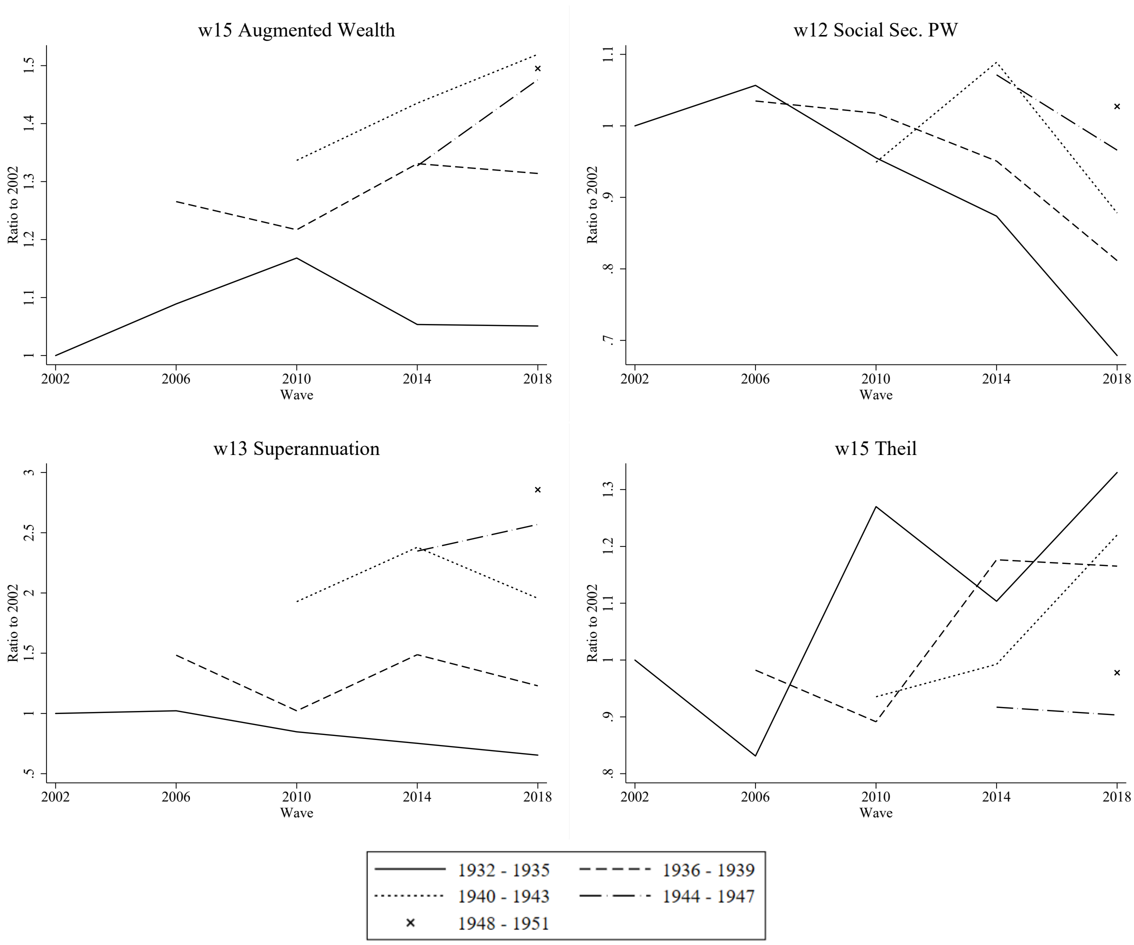
Retiring population The HILDA panel data allows me to analyze the retired population from another angle, by following different birth cohorts over time. A comparison of households at the start of their retirement phase is interesting for several reasons. First, a comparison of augmented wealth levels between wealth cohorts can indicate long-term trends. Second, at that point in time in the life cycle, pension wealth is normally at its peak. As the present value of social security pension wealth depends on the individual life expectancy, it is by construction, the highest at the very beginning of retirement. Third, Superannuation accounts can then be accessed and dissaved.

Three panels in Figure 3.7 show the mean values of wealth aggregates for different cohorts. The forth compares inequality between the cohorts over time. I pool four birth cohorts at the age of 67 to 70 in each year. Thus, the first birth cohort in 2002 includes retired household heads, born between 1932 and 1935, and the last cohort in 2018 covers those who are born between 1948 and 1951. I choose these birth cohorts, as they are in their early years of retirement. Moreover, it is likely that retirees have applied for Age Pension at that point in time.[41] To facilitate comparison across time and cohorts, I divide each value by the estimate of the 1932 to 1935 cohort in 2002.

Starting at the upper left panel, the mean augmented wealth increased between 2002 and 2006 by more than 20 percent for the succeeding cohort. It increases again in 2010, stagnates for those entering retirement in 2014 and increases again in 2018. Therefore, the augmented wealth levels of the cohorts after 1935 are consistently higher than those of the initial cohort. The upper right panel shows that the social security pension wealth level decreases slightly by each retiring cohort after 2002, rises back to the 2002 level in 2014 and remains fairly stable in 2018. Due to the present value estimation of the social security pension wealth aggregate, it decreases with each cohort over time, because the individual life expectancy declines. The panel also shows that social security pension wealth increases for the 1940 to 1943 cohort in 2014. As described above, this period is affected by the raise of the female retirement age. Hence, women had to postpone their retirement and proportionately more of them retired in 2014. The lower left panel provides the ratios of Superannuation wealth. For retiring households, it nearly tripled during the 16-year period. In the lower right panel, I depict some distributional insights of the augmented wealth aggregate. As I am analyzing subgroups, I use the Theil index, which is decomposable by subgroups, to describe change of augmented wealth inequality between the birth cohorts over time. Interestingly, the initial inequality remains relatively stable in each wave. There does not seem to be an increase in inequality over time for those entering the retirement phase. During retirement, the inequality in each subgroup, however, rises. This effect is not surprising: while social security pension wealth declines, its equalizing effect on augmented wealth also diminishes gradually.

I conclude that retiring households realized significant wealth gains between 2002 and 2018. This is associated with increasing wealth Superannuation accounts.

[41] Even though Superannuation can be accessed much earlier, the comparison should still accurately capture differences between cohorts.



Note. Own calculation based on different birth cohorts of the retired population from the HILDA Survey wave 18. The upper -, and the lower left panel shows the four-year birth values by birth cohort. The lower left panel provides the Theil index of augmented wealth by birth cohort. All estimates are normalized by the the level of the birth cohorts 1932 to 1935 in 2002. Birth cohort refers to HH Head. All estimates are based on imputed values. For the sake of clarity, bootstrapped standard errors are not reported here and available upon request.

Figure 3.7. Retiring Population Over Time

Full population Broadening the perspective to the full population, I provide augmented wealth over age profiles in Figure 3.8, again with age referring to the household head. I calculate the mean wealth of 21 age cohorts. Each age cohort includes three years of age, for instance, household heads at the age 22 to 24 in the second age cohort.[42] The last age cohort comprises of all households with a household head at the age 80 or older. I plot the results for each wave using LOWESS regressions.[43] In the four panels, I depict augmented –, social security pension –, Superannuation –, and housing wealth.

Augmented wealth, in the upper left panel, follows the same pattern in all considered waves, starting at a level close to 0 at the beginning of the life cycle and then a steady increase until the 60s, after which mean wealth values start to decline again. The estimates show, that the large increase of wealth between 2002 and 2006 seems to evaluate the wealth of those in their 50s and above. In the years 2010 and 2014, augmented wealth stagnated at the 2006 levels for all age cohorts. In 2018 augmented wealth increases again for older age cohorts, while the difference for those below their 50s is small.

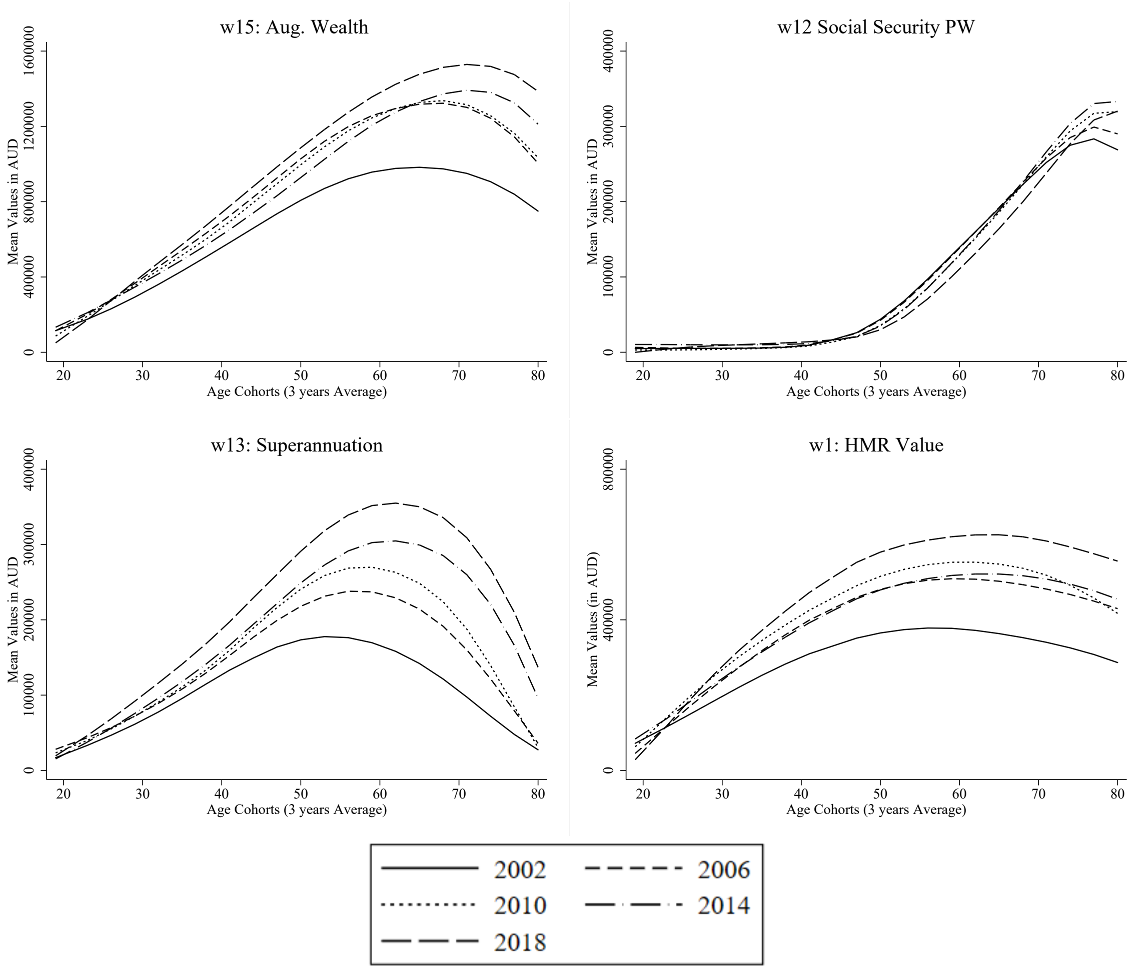
The patterns of social security pension wealth, in the upper right panel, are very persistent over time. Pension wealth from social security schemes starts to grow at the age of 50 in all years and then continuously increases almost linearly until the late 70s cohorts. There are other social security pension schemes included, e.g., the Service Pension or Disability Pension, which explains the take off before the age of 67.[44] Social security wealth in 2018 appears to be slightly lower then in the cohorts before, but the difference is small. Even though the present value of social security pension wealth decreases with age for birth cohort, as seen in the figure above, the mean values increase with age cohorts. Some households seem to become eligible to Age Pension later during retirement, increasing the mean of their age cohort.

The lower panels confirm the previous finding that Superannuation wealth and housing wealth drive the increase of augmented wealth in these periods. Superannuation wealth follows an inverted u-shaped pattern across the age cohorts and grows in every wave, but less during the time of the GFC. The increase of Superannuation wealth between 2014 and 2018 is interesting, as several changes and economic factors interact in this period. The rise of the preservation age potentially increased wealth levels, as the accumulation phase was prolonged. The reduced concessional rate, established in the 2010s, however, could have reduced accumulation patterns. Well performing financial markets in that period potentially offset effects on the accumulation rates and raised wealth for all cohorts. In the cross section, mean wealth reduces for households with a household head above the retirement age.

[42]The first cohort includes 4 years, aged 18-21.

[43]Scatter plots with bootstrapped standard error for the years 2002 and 2018 are provided in Figure 3.14 in Appendix 3.8.8.

[44]Also older household members can contribute to this finding.



Note. Own calculation based on the full sample of the HILDA Survey wave 18. 3 year average values by age cohort, smoothed with LOWESS. Age refers to HH Head. All values are in AUD and set to 2018 prices on the basis of the Consumer Price Index (The World Bank, 2021).

Figure 3.8. Life Cycle Wealth Profiles

Mean values for household heads 80 or older get closer to zero for the waves 2002, 2006, and 2010, but remain at above \$100,000 in the last two waves. Housing wealth increases throughout the life cycle as home-ownership becomes more and more relevant. Home-ownership did not change significantly across age cohorts[45]. The increase is, therefore, fully attributable to increasing housing values.

Superannuation and inequality The evolution of Superannuation wealth so far shows an increase of the wealth endowment for many Australians during the 2002 to 2018 period. As the compulsory component of the scheme was 26 years old in 2018, nearly a whole generation has participated in the scheme so far. While the observed *intra*-generational period of the scheme seems to affect overall inequality only to a small extent, my results incorporate concerns about the distributional consequences of *inter*-generational transfers in the future.

I observe higher wealth levels at the end of the life cycle in 2014 and 2018. In the decades to come, inheritances from Superannuation accounts will play a vital role for younger age cohorts. This may lead to new implications for the contribution of the scheme to overall inequality. My analysis above shows that higher amounts of Superannuation wealth are associated with lower dissaving rates. This can indicate higher remainder values for larger Superannuation accounts. Then, the effect on inequality depends on the dispersion of Superannuation wealth across the augmented wealth distribution. The concentration curve in Figure 3.6 above shows that Superannuation wealth is higher at the upper end of the augmented wealth distribution. Even though this does not increase the share of Superannuation wealth between 2002 and 2018, inter-generational transmission may change these patterns and could potentially lead to an increasing concentration of Superannuation wealth.

Another concern is the divergence in tax advantages along the distribution. As discussed above, the tax advantage is at four percent for the lowest tax rate, and 30 percent at the highest margin. As indicated in this paper, retirees at the lower end of the wealth distribution dissave Superannuation faster and potentially drop out of the scheme completely. This will ultimately increase the dispersion between retirees receiving Age Pension and those using Superannuation accounts.

Naturally, my analysis is not sufficient to provide clear predictions. Even though I find suggestive evidence for that the Superannuation scheme successfully supported *intra*-generational wealth accumulation for many Australians, these considerations on *inter*-generational transfers and taxation may lead to increasing inequality in the future.

Pension wealth by characteristics Focusing on 2018, I provide pension wealth by household characteristics along age cohorts following the same definition as above. This adds to the general understanding of pension wealth dynamics in the

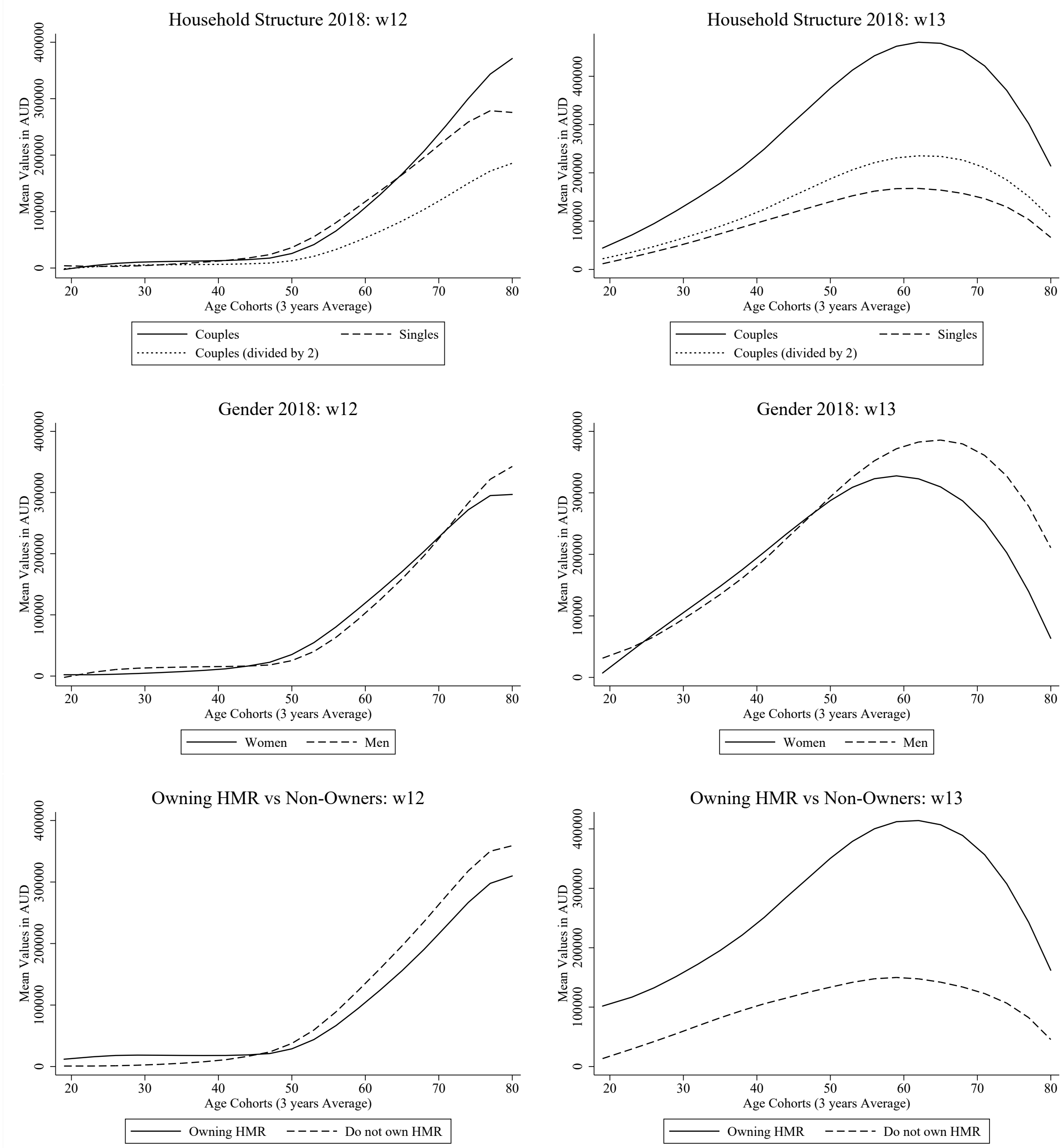
[45] A graph is provided in Appendix 3.8.8.

population. The results are shown in Figure 3.9. The left panels depict social security pension wealth, the right panels represent Superannuation wealth in 2018, respectively. I start with differentiating between single and couple households in the upper two panels. To make wealth levels more comparable, I also divide the couples' wealth by two. The panels show that social security pension wealth is relatively similar for coupled households, which means in regard to the per head perspective, couples hold considerably less social security pension wealth. It reflects the proportionally lower payments and stricter thresholds of the Age Pension scheme for couples compared to two single households. In terms of wealth in Superannuation accounts, couples hold more Superannuation wealth per head than single couples. I am not able to disentangle the dynamics behind it, whether they receive less pension due to higher amounts of Superannuation, or save more, as they expect a lower pension. Nevertheless, it potentially indicates another interaction between the two schemes.

The two middle panels show differences between genders. Social security pension wealth is held at equal levels by female and male household heads. Superannuation accounts show no difference during the accumulation period, but start do diverge in the 60s. Both are interesting findings. This panel cannot detect child birth "penalties" for women, as this would imply a lower mean earlier in the life cycle. One would not necessarily expect this results, because contributions are conditioned on labor market participation. This does not mean that the penalties do not exist. As this is a household perspective, these effects are potentially mitigated by partners in the household. The divergence during retirement age cannot be explained by higher dissaving rates, as I could not find a significant difference between men and women in my analysis above. Men seem to accumulate Superannuation wealth several years longer than women and, therefore, hold more wealth in their accounts.

Finally I compare pension wealth between households who own their main household residence (HMR) with those who do not. Those who do not own their HMR hold more social security pension wealth, but the difference is small. In terms of wealth in Superannuation accounts, HMR owners hold much more wealth than their counterparts. The result is striking: those households who do not own their primary home are also considerably worse off in terms of pension wealth.[46] This raises concerns about how housing wealth is treated in the Age Pension asset test. Those who do not own their home potentially do not benefit from comparably higher thresholds. Reforming Age Pension in that matter could offset some of the differences in retirement.

[46] It directly relates to the generalized concentration curves in Figure 3.6 above, as housing wealth is the main component of augmented wealth.



Note. Own calculation based on the respective subgroup taken from HILDA Survey wave 18. 3 year average values by age cohort, smoothed with LOWESS. Left panels present social security pension wealth, right panels depict wealth in Superannuation accounts. Age refers to HH Head. All panels refer to the year 2018.

Figure 3.9. Pension Wealth by Characteristics

3.5.3 Augmented Wealth in an International Comparison

Augmented wealth is a helpful tool to enhance the comparability of wealth data between countries, as it reduces the bias from different social security pension schemes. This allows me to compare Australian wealth inequality with Germany, the US (Bönke et al., 2019), and Switzerland (Kuhn, 2020). The pension system in Germany, Switzerland, and the US are all based on public, private and occupational pension schemes. The public pension in these countries are pay-as-you-go schemes, hence, the present value of the future benefits can be assigned to an individual at any age in their life cycle. As I discussed above, Australia's pension scheme differs, as it is not conditional on income during the worker's life, but on income and wealth endowment starting at the retirement age, and is, therefore, a special case in this group of countries. Furthermore, only the Australian Superannuation scheme and the occupational pension scheme[47] in Switzerland have a compulsory component. The Riester-Scheme in Germany, and the 401k-plan in the US are voluntary schemes.

In the following analysis, all values are in US Dollars and purchasing-power-parity-adjusted (ppp). For Australia, one Australian Dollar is equal to 0.90 US Dollars, i.e., the average exchange rate in 2014, and adjusted with the 1.452 ppp-conversion-rate provided by the OECD (2021). For Switzerland, I transform the estimates from Kuhn (2020), where one Swiss Franc is equal to 1.041 US Dollars and the ppp-conversion-rate is 1.235 in 2015. For a better comparison to Bönke et al. (2019), I exclude vehicles from the Australian wealth aggregates, as it is not asked in Germany and excluded from the US.[48]

Comparison of Descriptive Statistics I start with the comparison of mean values and several quintiles of the wealth aggregates w10, w12, w13, and w15 in Table 3.11. Regarding the mean net worth in Australia, at 324,664 USD, it is considerably higher than in Germany (182,329 USD) and Switzerland (223,525 USD) and slightly below the US (337,570 USD). However, mean social security pension wealth is 4.2 times higher in Germany, 2.6 in Switzerland, and 3.4 times higher in the US. The mean value in Superannuation accounts is higher than those of the occupational and private pension schemes in Germany or Switzerland but lower than the one in the US. Consequently augmented wealth in Australia sits at a mean value of 487,822 USD, which is above the one of Germany (472,401 USD) and Switzerland (451,294 USD) and below the US (652,504 USD). Augmented wealth in Australia is lower at the lower end of the wealth distribution, i.e., at the 25th percentile, than in the other countries.

[47] In German: "Berufliche Vorsorge"

[48] Kuhn (2020) does not provide information on this, but the SILC survey asks generally for overall wealth, which potentially includes vehicles. Hence, the estimates from Switzerland potentially include slightly more wealth types than the other countries.

Table 3.11. Descriptive Statistics Wealth Aggregates across Countries

Aggregates	Mean	p25	p50	p75	frac >0
Australia					
w10: Net Worth	324,664	8,368	171,257	406,970	84.53
	(3,850)				(0.29)
w12 Social Security	45,552	0	0	0	24.28
	(921)				(0.33)
w13 Superannuation	117,606	6,198	40,289	123,967	84.33
	(1,686)				(0.27)
w15 Augmented Wealth	487,822	66,474	296,876	636,399	93.84
	(4,953)				(0.19)
Germany					
69 w10: Net Worth	182,329	0	49,623	228,528	71.64
	(2,287)				(0.23)
w12 Social Security	200,424	68,620	162,780	296,048	93.17
	(923)				(0.21)
w13 Occupational and Private Pensions	89,648	0	13,059	78,352	64.24
	(1,116)				(0.15)
w15 Augmented Wealth	472,401	149,128	326,990	630,784	98.38
	(2,761)				(0.07)
Switzerland					
w10: Net Worth	223,525	12,074	66,870	513,221	n/a
	(7,414)				
w12 Social Security	123,868	52,298	88,915	262,859	n/a
	(688)				
w13 Occupational and Private Pensions	103,901	16,801	51,223	268,932	n/a
	(1,116)				
w15 Augmented Wealth	451,294	116,364	259,388	967,464	n/a
	(7,889)				
United States					
w10: Net Worth	337,570	0	40,001	198,800	73.14
	(5,351)				(0.28)
w12 Social Security	161,481	64,486	124,938	227,458	96.49
	(806)				(0.13)
w13 Occupational and Private Pensions	153,453	0	13,000	140,000	61.68
	(2,227)				(0.4)
w15 Augmented Wealth	652,504	86,311	246,663	608,473	95.83
	(6,710)				(0.14)

Note. German and US estimates by Bönke et al. (2019) based on SOEP v30/v31 and SCF 2013, respectively, Switzerland results are based on EU-Silc. Results from Australia are based on own calculations from the HILDA Survey wave 18. Australian estimates are transformed into USD (1 AUD= 0.90 USD, average exchange rate in 2014) and ppp-adjusted, with the factor 1.452 provided by the OECD (2021) Estimates from Germany and the US are taken from Bönke et al. (2019). Estimates from Switzerland are taken from Kuhn (2020) and transformed ppp-adjusted USD by using the 1 CHF= 1.041 USD and the 1.235 ppp-conversion-rate (OECD, 2021). All statistics are based on imputed values. Bootstrapped standard errors in brackets using 1000 replica weights.

The statistics reveal the differences between the social security pension schemes. As it is only accessible for the retired population in Australia, it accounts for only 24.28 percent of the overall population in 2014. As described above, this is mainly due to the different pension concepts. Two-thirds of the Australian Pension wealth comes, on average, from Superannuation accounts, with 84.33 percent of the households holding some wealth here. The coverage is considerably higher than in the other countries and can be explained by the compulsory nature of the Superannuation scheme.[49] Comparing the values of pension wealth, it shows that pension wealth of Australian households is proportionally, and in absolute terms, lower than in the other three countries depicted.

The Gini indices of the different wealth aggregates in Table 3.12 provide further insights into the distributional differences of the wealth aggregates. The net worth Gini coefficient in Australia lies at 0.661, which is lower than the Gini coefficient of 0.755 in Germany, 0.750 in Switzerland, and 0.889 in the US. Adding personal entitlements of social security pension wealth to net worth, the Gini coefficient reduces in Australia with 5.90 percent less than in the other two countries, i.e., 32.84 percent in Germany, 24 percent in Switzerland, and 20.13 percent in the US. Net worth, plus occupational and private pension wealth, reduces the Gini index slightly in all countries. Eventually, the Gini index of augmented wealth in Australia is with 0.592 higher than Germany with 0.508 and Switzerland with 0.55, but remains below the one in the US with 0.700. Including augmented wealth, therefore, leads to a different ranking of Australia in terms of inequality.

3.6 Qualification and Extensions

This section provides caveats and a critical classification of my results. Moreover, I discuss potential extensions of this paper.

I fist discuss the analysis of income during retirement, and especially the analysis of the dissaving behavior of Superannuation accounts. The fractional probit model aims to describe different dissaving behavior between retirees, who would be potentially eligible to Age Pension after dissaving their Superannuation accounts, to the rest of the retirees. My approach does not allow for causal interpretation. Moreover, the effects measured are the marginal effects at the mean. Thus, they only hold for the average retired household and the marginal effects can differ for other parts of the distribution. One could explore this in more detail and extend the analysis by setting up a causal framework by comparing retirees who are just below the Age Pension threshold to those who are just above it.

There are several limitations in the applied accrual method in the context of the Australian pension scheme. The aim of the method is to estimate the value of

[49]Coverage is not provided by Kuhn (2020) for the other compulsory scheme in Switzerland.

Table 3.12. Wealth Aggregates: Gini Coefficients

	Australia	Germany	Switzerland	United States
w10: Net Worth	0.661	0.755	0.750	0.889
	(0.003)	(0.036)	(0.007)	(0.029)
w10 + w12: Social Sec PW	0.622	0.507	0.570	0.585
	(0.003)	(0.037)	(0.004)	(0.035)
w10 + w13: Occ. / Private PW	0.628	0.705	0.650	0.826
	(0.003)	(0.034)	(0.005)	(0.031)
w15: Augmented Wealth	0.592	0.508	0.550	0.700
	(0.003)	(0.034)	(0.008)	(0.033)

Note. German and US estimates are based on SOEP v30/v31 and SCF 2013, respectively, Switzerland results are based on EU-Silc. Results from Australia are based on own calculations from the HILDA Survey wave 18. German and US estimates are taken from Bönke et al. (2019). Estimates from Switzerland are taken from Kuhn (2020). For Australia, all statistics are based on imputed values. Bootstrapped standard errors in brackets using 1000 replica weights.

the pension plans based on the individual's work history (Bönke et al., 2019). In Australia, this can be easily provided for the Superannuation scheme, as the to-date value is available in individuals pension accounts and provided in the HILDA survey on the individual level. Contrarily, the social security scheme Age Pension is means-tested annually for individuals in retirement and, therefore, difficult to evaluate. Individuals below the Age Pension retirement age will face an income and asset test later in their life cycle which incorporates a high uncertainty in determining the to-date present value of the scheme. I avert from estimating future income and wealth endowments by calculating present values only for those who actually receive social security pensions in the respective year. This underestimates the expected value of the Age Pension scheme, as it is only calculated for retired households, but it avoids modeling assumptions on the development of households' income and wealth.

What does this mean for the estimates of augmented wealth in my analysis? Some individuals below the retirement age may almost certainly receive Age Pension as soon as they reach the retirement age, as they may experience long unemployment spells or receive other kinds of persistent economic shocks during their life cycle. Including this "unobserved" pension wealth into the social security pension wealth aggregate (w12) would potentially reduce inequality, as it would increase wealth for those at the lower end of the wealth distribution.[50]

[50] In earlier stages of this paper, I calculated several scenarios where the present value of social security wealth for individuals below the retirement age was calculated. As several arbitrary

My applied methodology on calculating the present value of social security schemes requires several assumptions: I assume that if an individual is eligible to Age Pension in a certain period, the eligibility in retirement does not change throughout the period-specific present value calculation. However, one could lose their eligibility from one period to the other, as eligibility is checked annually by the Australian government. Given that wealth decreases constantly during retirement, this should not occur on a large scale. Another concern is that the approach neglects changes in household formations. As eligibility differs between singles and married couples, this potentially includes a wrong measurement of the present value for some. I argue, that would affect my estimates only slightly, as a relatively small number of retirees got married, i.e., one percent, and the divorce rate among retirees was around 0.8 percent in Australia in 2019 (Australian Bureau of Statistics, 2021c).

As discussed in this paper, Bönke et al. (2019) and Kuhn (2020) do not include means-tested schemes. Therefore, the international comparison needs to be addressed with caution. The conclusion on re-ranking between the four countries, however, is not necessarily affected by the difference of the scheme. If means-tested social security pension would be excluded from the Australian analysis, meaning that the aggregate w15 is set equal to w10+w13, the re-ranking in terms of overall inequality would remain the same. An extension of this paper could include an international comparison of additional means of well-being during retirement, e.g., consumption and income.

3.7 Discussion and Conclusion

The analysis of augmented wealth in Australia between 2002 and 2018 provides several insights about pension wealth dynamics. It evaluates the two main pension schemes, i.e., Age Pension and Superannuation. This section discusses the question what other countries can learn from the Australian retirement system, and summarizes my results.

Discussion My analysis opens a discussion about the distributional consequences of the two pension schemes. The Superannuation scheme helps to generally increase wealth endowment for the retired population and, thus, there is a high take-up rate, also due to its compulsory nature. Even though the Superannuation wealth contribution to overall augmented wealth inequality increased only slightly between 2002 and 2018, *inter*-generational transmission and dispersed tax incentives along the wealth distribution may further affect inequality patterns in the future. Age pension payments remained remarkably stable over the considered period, but

assumptions where required, I did not include this in my final version. The implications for the estimation of overall inequality where, however, small.

it is not certain if this trend will prevail. The long-term plan to relieve the Age Pension scheme can entail several risks, which affect those especially at the lower end of the wealth distribution. Individuals who cannot participate regularly in the workforce during their life cycle face lower income streams, if Age Pension eligibility is hampered. There is a risk of a two-tiered retirement system: those depending on Superannuation and those depending on Age Pension.

Australia's Superannuation scheme shows that compulsory contributions are key for a high take-up rate in the population and, consequently, for the success of its broad coverage. There are, however, several consequences to consider for the policy planner. The consideration of Superannuation savings on the eligibility of the Age Pension scheme sets potentially adverse incentives for individuals saving and, as provided in my analysis, their dissaving plans. When introducing an additional pension scheme, policy planners need to set a solid saving and extraction framework, so that the tax schedule and the scheme itself does not set adverse incentives. Simultaneously, the planners need to minimize potential deadweight loss effects, i.e., lessen the amount of individuals attending the scheme for tax advantages without adjusting their actual ceteris paribus saving behavior. The set of rules for a new scheme also needs to carefully consider its distributional consequences, as long-term imbalances are difficult to address with ex post policies.

Policy planners also need to be aware that Superannuation imposes further risk on the population, i.e., financial market – and institutional risk. As Superannuation accounts invest in the stock market, individuals in the maturing Superannuation scheme face more financial market risk in their household portfolios. For retirement wealth, this can have severe consequences when individuals, that are close to retirement, face a downward trend on the financial markets, e.g., as seen lately in the COVID-19 crisis. The government could provide aid in these situations to prepare individuals to re-balance their portfolios before retirement. Over the years, the Superannuation scheme has been transformed and adjusted regularly. Naturally, the scheme is shaped by partisan policies, e.g., the initial plans of one administration for significant increases of the Superannuation Guarantee were reduced and postponed by the successor. While fundamental parameters, like the compulsory nature of the contributions, are not affected, the scheme bears institutional risk.

Conclusion This article focuses on two main topics: income during retirement and augmented wealth inequality. Through the analysis of income during retirement, I show that Age Pension is the most important source of income for retirees whilst simultaneously being replaced by Superannuation at the top end of the gross wealth distribution. It also appears that individuals react to the eligibility requirements of the Age Pension scheme and dissave their Superannuation accounts at a higher rate to avoid penalties on their Age Pension payments. Moreover, dissaving rates are associated negatively with higher levels of Superannuation wealth.

I find that the levels of net worth and augmented wealth inequality in Australia are persistent over time. This coincides with large wealth gains during the 2000s, which were driven by increases in housing values, accompanied by a relatively high home-ownership, and large gains in Superannuation accounts. Wealth gains are the most profound for households with a household head over 50 years of age. Entitlements to social security pension wealth are the most important wealth position for households at the lower end of the gross wealth distribution. In regard to the overall wealth portfolio, the relative importance of Superannuation reduces for the top 10 percent, where financial assets and business investments become more important. Conditioning wealth on several household characteristics, I also find that couples hold considerably less social security pension wealth and hold more Superannuation wealth. Women seem to hold less Superannuation wealth in retirement, and home owners are also better off in terms of pension wealth.

In an international comparison, Australia exhibits relatively high values of net worth, but relatively low values of pension wealth. Net worth is more equally distributed than in Germany, Switzerland or the US. However, adding pension wealth reduces the Gini coefficient less than in the other countries, so that augmented wealth in Australia is less equally distributed than in Germany and Switzerland. The main reason for this, is the means-tested social security pension wealth, which covers only retired Australian households and is not an asset for those still in employment. This also shows the limits of the accrual method in Australia. Moreover, one could include other factors of individual welfare at retirement, as some countries provide considerably more public goods for retirees than others. In conclusion, this article represents a first analysis of the Australian pension schemes and their interaction with augmented wealth, leaving room for further assessments in the future.

3.8 Appendix

3.8.1 Additional Payments and Eligibility

Table 3.13 provides the payments per fortnight in AUD for each wave. The maximum Age pension payment has nearly doubled during the observed period for both, singles and couples. The government also provides further support, i.e., Rent Assistance, Energy and Pension Supplements. Rent Assistance is paid for singles (couples) who pay at least 121.2 (196.20) AUD per fortnight in 2018. The payment is linear to the paid rent. Individuals receive the maximum payment of 135.8 (128 per person) AUD if they pay more or equal to 302.27 (366.87). Energy Supplement was introduced in 2013 to protect retired individuals from increasing energy costs. The payment is 14.10 (10.60) per fortnight, but only if income is below 1,197 (2,201.68) AUD. The Pension Supplement is an extra payment to ensure that individuals can cover their utility, phone, internet, and medicine costs. The payments of 67.80 (51.10) AUD are paid to every eligible recipient.

Table 3.13. Eligibility and Payments

Type	2002	2006	2010	2014	2018
			Payments per fortnight		
Age Pension					
pension single	352.10	402.40	496.30	585.50	629.00
pension couple	421.80	478.50	658.40	776.70	834.40
Rent Assistance					
payment single	90.60	100.6	115.20	127.60	135.80
payment couple	85.40	95	108.60	120	128.00
red. rate	0.25	0.25	0.25	0.25	0.25
min. rent single	80.40	89.6	102.40	113.2	121.20
max. rent single	201.20	223.73	256	283.3	302.27
max. rent couple	131.00	145.8	166.80	184.20,	196.20
max. rent couple	244.87	272.47	311.60	344.20	366.87
Energy Sup.					
payment single	0	0	0	14.10	14.10
payment couple	0	0	0	10.60	10.60
single max.	0	0	0	1197	1197
couple max.	0	0	0	2201.68	2201.68
Pension Sup.					
payment single	0	17.80	57.70	63.50	67.80
payment couple	0	14.80	87.00/2	95.80/2	51.10

Note. Information is taken from Australian Government (2018, 2021b). Table provides thresholds for the income and asset test, deeming, and payments for Age Pension. Further payments, i.e., Rent Assistance, CDEP Participant Supplement (CPS), Energy and Pension Supplements are potentially paid es well. All values are provided in AUD, as they are directly taken from the public authority. Following the official approach, most values, apart from the Asset Test and Deeming, are calculated per fortnight.

3.8.2 Dissaving Regression Results

Table 3.14. Regression Results

	(1)	(2)	(3)
Receive Age Pension	0.0936***	0.3575***	0.0833***
	(0.0096)	(0.0364)	(0.0084)
Net Worth: w10	-0.0005	-0.0022	-0.0005
	(0.0004)	(0.0021)	(0.0005)
Superannuation Wealth: w13	-0.0151***	-0.1935***	-0.0451***
	(0.0010)	(0.0091)	(0.0020)
Female	-0.0031	-0.0343	-0.0080
	(0.0091)	(0.0339)	(0.0079)
Single	-0.0183	-0.0844*	-0.0197*
	(0.0104)	(0.0376)	(0.0088)
Age: 55–59	-0.0288	-0.1721*	-0.0401*
	(0.0157)	(0.0753)	(0.0175)
Age: 65–69	-0.0124	-0.0638	-0.0149
	(0.0121)	(0.0515)	(0.0120)
Age: 70–74	0.0222	0.0604	0.0141
	(0.0136)	(0.0538)	(0.0126)
Age: 75–79	0.0945***	0.2616***	0.0610***
	(0.0166)	(0.0597)	(0.0140)
Age: >80	0.2786***	0.7190***	0.1676***
	(0.0203)	(0.0634)	(0.0152)
Medium Education	0.0294**	0.1448***	0.0338***
	(0.0101)	(0.0377)	(0.0088)
High Education	0.1150***	0.5402***	0.1259***
	(0.0122)	(0.0434)	(0.0100)
2006	-0.0360*	-0.0884	-0.0206
	(0.0182)	(0.0601)	(0.0140)
2010	-0.0441*	-0.1272*	-0.0297*
	(0.0175)	(0.0581)	(0.0135)
2014	-0.0643***	-0.1614**	-0.0376**
	(0.0160)	(0.0541)	(0.0127)
2018	-0.0554***	-0.1072*	-0.0250*
	(0.0159)	(0.0533)	(0.0125)
R^2	0.1557	0.1593	
N	5,679	5,679	5,679

Note. Regression are based on the working sample described in the main text, based on the HILDA Survey wave 18. The table provides the regression results of the OLS regression (1), the estimates of the Fractional Probit Model (2), and the marginal effect at means (3). Wealth variables are divided by 100,000. Robust standard errors are provided in brackets. The R^2 in (2) represents the pseudo-R^2 of the fractional probit model. The significance levels are reported with * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$.

3.8.3 Portfolios

Table 3.15 provides the estimates for the mean wealth portfolios. These the estimates are the basis of the four panels of Figure 3.4 in the main analysis.

Table 3.15. Portfolio: Mean Values

wave	dist	w15	w1 + w2	w3	w4	w5	w12	super	w7 + 8	debt busin hecs	debt other	Age	s.e. Age	N
2002	1	84,587	0.019	0.123	0.085	0.005	0.597	0.307	-0.007	-0.038	-0.091	40.46	0.263	1,765
	2	462,165	0.591	0.059	0.094	0.018	0.202	0.208	-0.145	-0.008	-0.020	49.83	0.185	3,532
	3	987,395	0.659	0.050	0.163	0.050	0.047	0.203	-0.134	-0.013	-0.022	50.53	0.330	1,059
	4	2,622,084	0.522	0.037	0.211	0.221	0.007	0.129	-0.082	-0.033	-0.013	52.15	0.356	707
2006	1	86,071	0.001	0.116	0.070	0.003	0.625	0.334	0	-0.042	-0.105	39.41	0.306	1,750
	2	569,656	0.662	0.050	0.090	0.013	0.167	0.209	-0.165	-0.007	-0.020	50.82	0.205	3,502
	3	1,225,613	0.758	0.039	0.128	0.042	0.042	0.208	-0.180	-0.013	-0.024	50.78	0.303	1,050
	4	3,692,633	0.611	0.027	0.212	0.152	0.004	0.121	-0.086	-0.022	-0.018	53.46	0.370	701
2010	1	85,888	0.001	0.109	0.062	0.003	0.581	0.383	-0.001	-0.051	-0.088	40.12	0.356	1,798
	2	572,665	0.703	0.051	0.091	0.013	0.164	0.213	-0.204	-0.008	-0.023	50.47	0.196	3,596
	3	1,291,714	0.761	0.037	0.130	0.027	0.042	0.231	-0.190	-0.011	-0.028	52.31	0.296	1,079
	4	3,388,828	0.635	0.028	0.193	0.145	0.005	0.164	-0.128	-0.024	-0.018	53.90	0.341	720
2014	1	88,932	0.000	0.098	0.054	0.002	0.622	0.382	-0.002	-0.070	-0.085	39.77	0.266	2,340
	2	575,659	0.655	0.049	0.097	0.011	0.178	0.249	-0.207	-0.010	-0.021	51.12	0.186	4,682
	3	1,351,735	0.724	0.035	0.128	0.022	0.045	0.266	-0.194	-0.007	-0.020	53.50	0.282	1,404
	4	3,282,191	0.618	0.025	0.232	0.101	0.007	0.192	-0.137	-0.021	-0.017	55.38	0.420	937
2018	1	96,564	0.001	0.093	0.051	0.002	0.545	0.466	-0.003	-0.081	-0.073	40.50	0.253	2,371
	2	640,421	0.633	0.049	0.096	0.008	0.151	0.284	-0.195	-0.010	-0.016	51.76	0.191	4,743
	3	1,556,088	0.711	0.031	0.127	0.023	0.038	0.262	-0.171	-0.008	-0.012	54.94	0.277	1,423
	4	3,720,798	0.653	0.023	0.210	0.101	0.005	0.184	-0.149	-0.016	-0.011	55.79	0.402	949

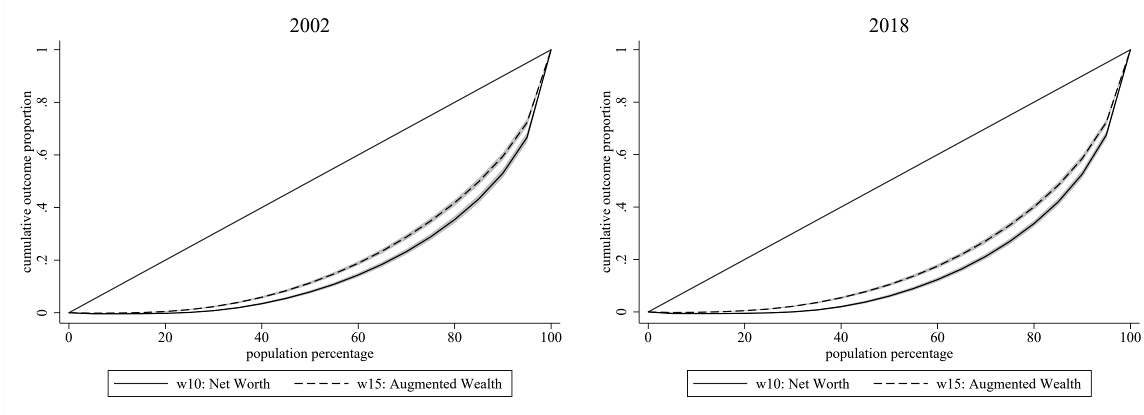
Note. Own calculation based on Hilda Wave 18. The table provides the estimates of Figure 3.4. Age refers to HH Head. All monetary values are in 2018 AUD.

3.8.4 Ginis

Table 3.16. Gini Coefficients of Wealth Aggregates

Aggregates	2002	2006	2010	2014	2018
Bottom coded at 0					
w10: Net Worth	0.638	0.653	0.636	0.646	0.654
	(0.0039)	(0.0044)	(0.0031)	(0.0035)	(0.0018)
w10 + w12: Social Sec PW	0.597	0.617	0.604	0.609	0.619
	(0.0042)	(0.0044)	(0.0035)	(0.0043)	(0.0021)
w10 + w13: Superannuation	0.615	0.626	0.616	0.618	0.617
	(0.0036)	(0.0041)	(0.0023)	(0.0036)	(0.0018)
w15: Augmented Wealth	0.577	0.597	0.593	0.592	0.592
	(0.0038)	(0.0042)	(0.0029)	(0.0046)	(0.0019)
Bottom censored at 0					
w10: Net Worth	0.609	0.619	0.601	0.608	0.618
	(0.0044)	(0.0044)	(0.0045)	(0.0052)	(0.0028)
w10 + w12: Social Sec PW	0.567	0.582	0.568	0.570	0.582
	(0.0037)	(0.0050)	(0.0045)	(0.0064)	(0.0028)
w10 + w13: Superannuation	0.589	0.597	0.586	0.587	0.587
	(0.0039)	(0.0040)	(0.0023)	(0.0038)	(0.0012)
w15: Augmented Wealth	0.556	0.577	0.570	0.568	0.572
	(0.0036)	(0.0042)	(0.0022)	(0.0051)	(0.0012)

Note. Own calculation based on HILDA Survey wave 18. Table provides the bottom coded and bottom censored estimates of the Gini index. All estimates are based on imputed values. Bootstrapped standard errors in brackets using 1000 replica weights.



Note. Own calculation based on HILDA Survey wave 18. Figures show the Lorenz curves of net worth and augmented wealth in 2002 and 2018, respectively. Gray area represents the 95 percent confidence intervals.

Figure 3.10. Lorenz Curves: Net Worth and Augmented Wealth

3.8.5 Gini Decomposition

This analysis comprises a factor decomposition to study the contribution of net worth and the pension wealth components to the overall inequality of augmented wealth in every survey year. This decomposition allows me to evaluate the interaction between the changes in pension wealth and overall augmented wealth inequality over time. Following Lerman and Yitzhaki (1985); Bönke et al. (2019), I decompose the Gini index as follows:

$$Gini_y = \sum_{a=0}^{A} \rho_{a,y} \times Gini_{a,y} \times s_{a,y} = \sum_{a=0}^{A} O_{a,y}, \tag{3.4}$$

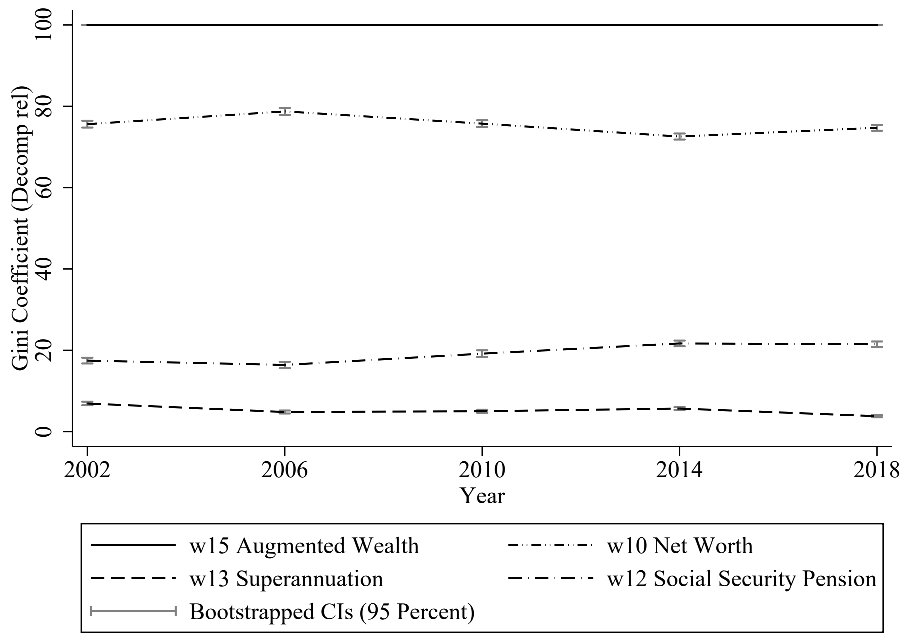
where the $Gini_y$ represents the Gini index of augmented wealth in year y, $\rho_{a,y}$ denotes the Gini correlation[51] between wealth aggregate $w_{a,y}$, with $a = 1, ..., A$, and augmented wealth. $Gini_{a,y}$ is the Gini index of wealth aggregate a and $s_{a,y}$ the share of wealth aggregate a in augmented wealth. The product $O_{a,y}$ is the absolute contribution of a wealth aggregate to overall inequality of augmented wealth. In the main analysis, I provide the relative contribution $o_{a,y}$=$O_{a,y}/Gini_y$.

I present the results for the relative contribution of each aggregate in Figure 3.11. The x-axis represents the contribution of the aggregates to the augmented wealth Gini coefficient as a percentage and the y-axis depicts the years. Net worth is the highest contributor to overall augmented wealth in Australia, with a share consistently over 75 percent. The steep increase in housing wealth may explain the

[51] Its properties are a mixture of Pearson and Spearman correlations, see Lerman and Yitzhaki (1985); Schröder et al. (2014) for more details.

upward trend of the net worth aggregate between 2002 and 2006 but the change is small and not persistent.

The relative contribution from social security pension wealth reduced over the years from 5.5 percent in 2002 to 1.3 percent in 2018. On the contrary, the contribution of Superannuation wealth increased from 18.22 percent to 22.55 percent in 2018. The relative decrease of the contribution of social security pension wealth stems from a decrease in its total wealth share, showing that it did not grow as much as the other wealth aggregates. The Superannuation scheme matures throughout the 16-years time frame and increases, therefore, its distributional relevance.



Note. Own calculations based on the full sample of the HILDA Survey wave 18. The relative contribution of net worth, Social Security Pension wealth and Superannuation to overall augmented wealth inequality. All estimates are based on multiple imputations. Bootstrapped standard errors are based on 1000 replica weights.

Figure 3.11. Gini Decomposition

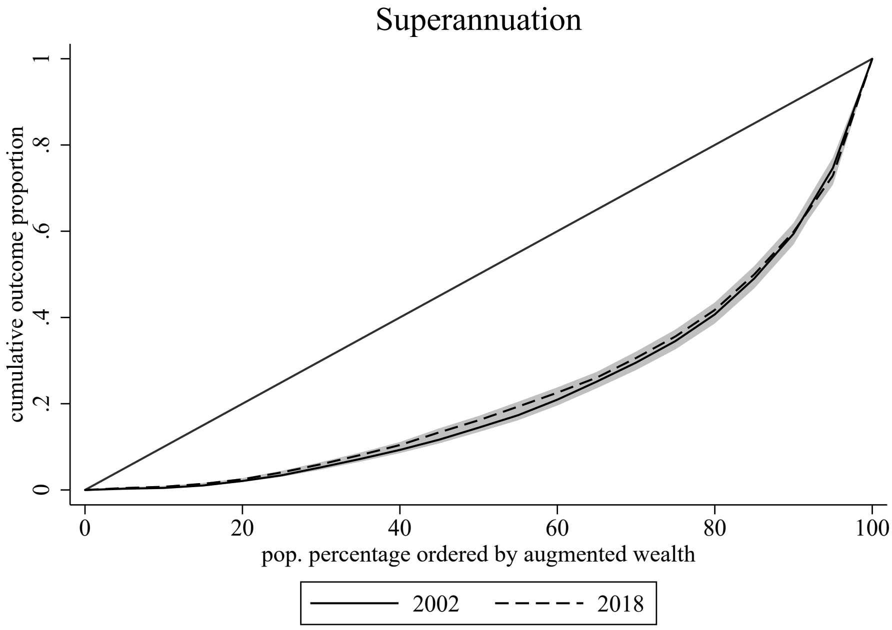
3.8.6 Percentile Ratios

Table 3.17. Percentile Ratios

	2002	2006	2010	2014	2018
20/50 ratio					
w10: Net Worth	0.081	0.052	0.056	0.047	0.046
	(0.004)	(0.003)	(0.003)	(0.003)	(0.002)
w10 + w11: Personal Ent.	0.105	0.0790	0.0750	0.0650	0.0630
	(0.004)	(0.005)	(0.004)	(0.004)	(0.003)
w10 + w12: Social Sec PW	0.105	0.0790	0.0750	0.0650	0.0630
	(0.004)	(0.005)	(0.004)	(0.004)	(0.003)
w10 + w13: Superannuation	0.147	0.130	0.140	0.138	0.144
	(0.006)	(0.005)	(0.004)	(0.004)	(0.005)
w15: Augmented Wealth	0.174	0.166	0.166	0.158	0.176
	(0.006)	(0.006)	(0.005)	(0.007)	(0.005)
90/50 ratio					
w10: Net Worth	4.114	3.976	3.990	4.322	4.802
	(0.067)	(0.113)	(0.073)	(0.084)	(0.105)
w10 + w11: Personal Ent.	3.445	3.582	3.555	3.819	4.029
	(0.055)	(0.062)	(0.057)	(0.061)	(0.067)
w10 + w12: Social Sec PW	3.463	3.586	3.580	3.856	4.085
	(0.054)	(0.060)	(0.058)	(0.060)	(0.068)
w10 + w13: Superannuation	4.137	4.133	4.036	4.411	4.488
	(0.068)	(0.091)	(0.075)	(0.071)	(0.065)
w15: Augmented Wealth	3.342	3.516	3.456	3.704	3.869
	(0.050)	(0.061)	(0.049)	(0.048)	(0.059)

Note. Own calculation. All estimates are based on imputed values. Bootstrapped standard errors in brackets using 1000 replica weights.
Source: HILDA Survey wave 18.

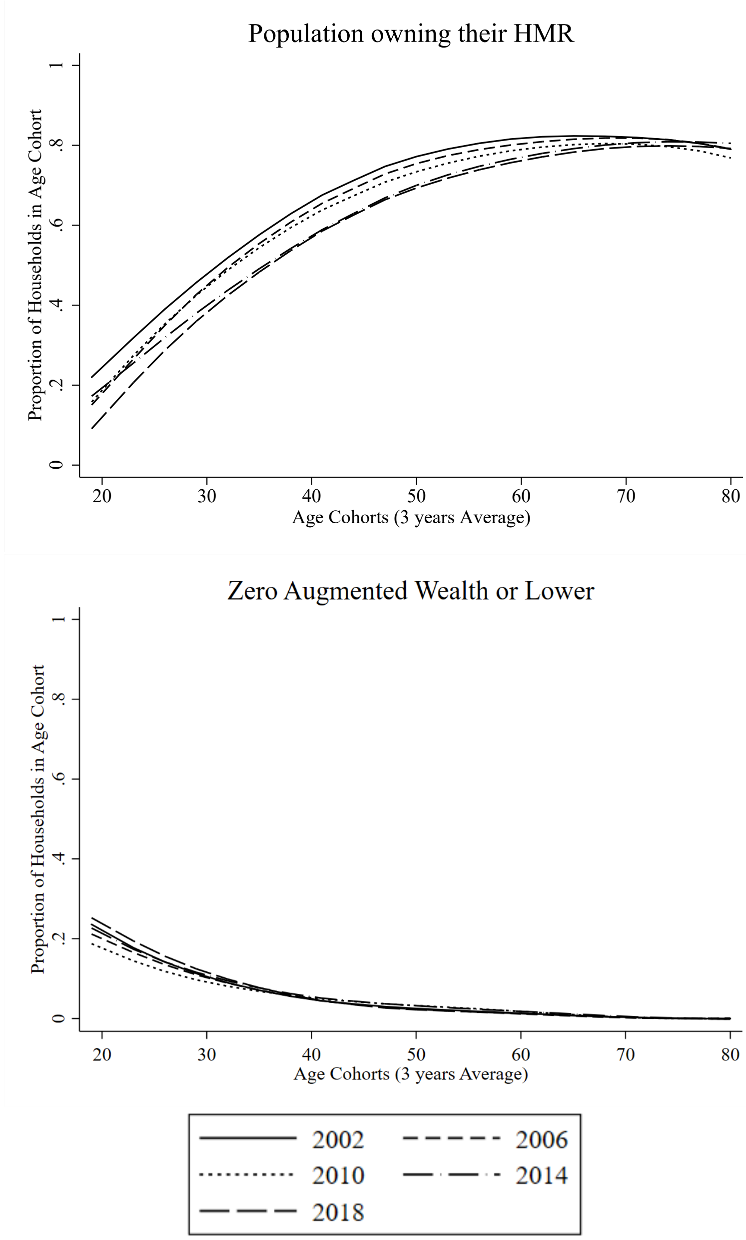
3.8.7 Concentration Curves



Note. Own calculations based on HILDA Survey wave 18. Generalized concentration curves of Superannuation along the population ordered by augmented wealth for 2002 and 2018. Grey area represents 95 % confidence intervals.

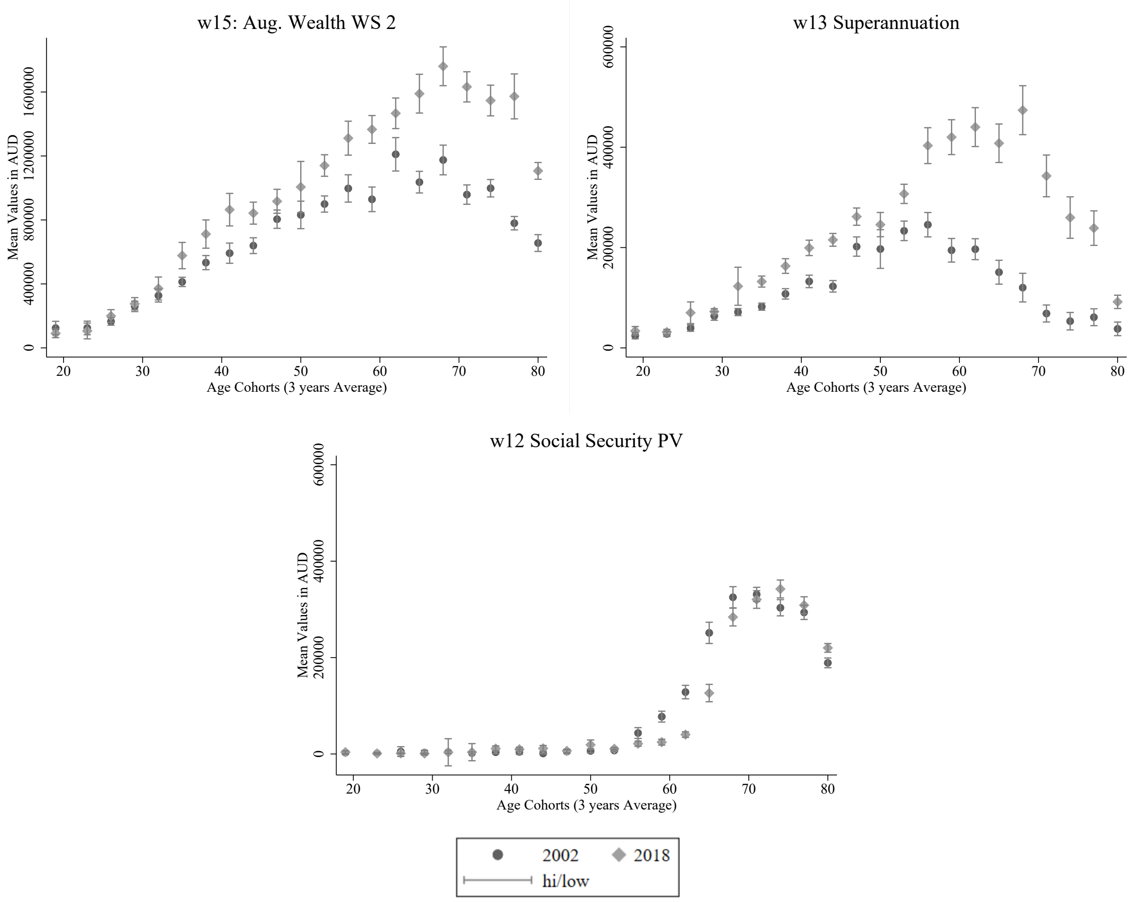
Figure 3.12. Concentration Curves of Superannuation and Augmented Wealth

3.8.8 Life Cycle Patterns



Note. Own calculations based on HILDA Survey wave 18. 3-year average values by age cohort, smoothed with LOWESS. Age refers to HH Head. Upper panel provides the proportion of households owning their housing main residence. The lower panel shows the proportion of households holding zero augmented wealth or lower.

Figure 3.13. House-Ownership and Non-Positive Augmented Wealth



Note. Own calculations based on HILDA Survey wave 18. 3-year average values by age cohort. Age refers to HH Head. All values are in 2018 prices. 95 percent confidence intervals are based on bootstrapped standard errors in brackets using 1000 replica weights. This figure replicates Figure 3.8 in the main analysis, providing bootstrapped standard errors for 2002 and 2018.

Figure 3.14. Life Cycle Means with Bootstrapped Standard Errors

4 Wage Risk and Portfolio Choice: The (Ir)relevance of Correlated Returns

4.1 Introduction

Unpredictable and uninsurable shocks to labor income affect households' portfolio choice. In standard portfolio choice models, the optimal risky portfolio share reduces when households face higher degrees of labor income risk (Heaton and Lucas, 2000; Cocco et al., 2005; Guiso and Paiella, 2008; Cardak and Wilkins, 2009; Betermier et al., 2012; Fagereng et al., 2018). These standard models typically assume that the correlation between labor income risk and asset returns risk is (close to) zero (Guiso and Sodini, 2013). This assumption is not innocuous. For example, when risky asset returns are negatively correlated with labor income risk, investing in risky assets hedges labor income risk. Thus, it is crucial for individuals to consider the correlation when minimizing overall risk exposure. Ultimately, whether the theory holds, is an empirical question. Hence, in this paper, we quantify the influence of wage risk and its correlation with financial market returns on German investors' financial portfolio shares.

Labor income risk is one of the prime components of background risk (Guiso and Sodini, 2013). Investigating background risk is important for our understanding of individuals' investment behavior, and, thus, the dynamics of the capital stock. Previous research has concluded that individuals do not follow standard portfolio choice theory and invest too little in risky assets (Calvet et al., 2007a). Background risk serves as one of the explanations to reconcile this discrepancy between empirics and theory. From a welfare economics perspective, it is important to recognize that background risk is beyond the control of the individual (Eeckhoudt et al., 1996). Thus, alleviating individuals' exposure to background risk, either by offering insurance on the labor market, or improving individuals' portfolio choice, may also induce welfare gains. Finally, since labor income risk impacts individuals disparately, it may not only lead to income inequality, but through its influence on portfolio choice, drive wealth inequality (Benhabib et al., 2017).

Both the precise definition and the quantification of background risk are demanding. In most applications authors restrict their view to yearly labor income risk, which appears to be the prime quantity for uninsurable risk. Even when we restrict our attention to labor income risk, two important concerns are in the way of identifying an effect on portfolio choice: an omitted variable and a measurement error bias. The omitted variable bias arises, because risk preferences are heterogeneous in the population and affect not only portfolio choice but also various decisions in the labor market, among them job choice or educational attainment (Brown et al., 2006; Bonin et al., 2007; Dohmen et al., 2011). The measurement error bias emerges, since only a part of the risk that the econometrician measures is due to exogenous shocks that the individual cannot control. To address the omitted variable bias, the literature has relied on specifications with fixed effects. To address measurement error bias, instrumental variables (IV) have been used (Fagereng et al., 2018).

We follow the literature and address these problems by estimating a fixed effects IV regression. Concerning the measurement error bias, we use hourly wages and not labor income to quantify risk. Idiosyncratic wage fluctuations are generally not attributable to individual choice compared to variations in labor income (Blundell et al., 2016). Hence, we tackle part of the measurement error problem immediately. Wages are also analogous to asset returns since they represent the marginal payoff from allocating one additional hour of one's time budget to work. Therefore, they are the ideal quantity to calculate the appropriate correlation with financial market returns.

To thoroughly treat the measurement error problem, we construct a group-IV based on occupational, time, and fine-grained regional information (Blundell et al., 1998; Blau and Kahn, 2006; Burns and Ziliak, 2017). We interact dummies of the occupational, time, and regional information to construct well-differentiated groups. This IV captures variation in local labor market conditions. For example, take the case of an automobile plant opening in a specific region. Due to the plant opening, demand for certain occupations, for example mechanics, will rise, changing their wages. These types of shocks are beyond the control of the individual and thus identify true wage risk. In this sense, the variation we use for our IV is very similar to the IV established in Fagereng et al. (2018), which relies on firm-specific productivity shocks and links them to individuals working in these firms. Their IV also behaves like a group IV, since all workers in a certain firm experience the productivity shock. Our IV is bound to react to a broader array of shocks, for example, changes in bargaining power, firm closures and openings, or differences in development trends of regional infrastructure, however, the general idea of the identification strategy is the same.

First, we find that the coefficient of the influence of wage risk on the financial portfolio share falls from about -0.03 in the fixed effects specification to about -0.09 in the IV fixed effects specification; a change by a factor of 3. This supports the case for adequately dealing with measurement error, because any attenuation bias resulting from (classical) measurement error would shift coefficient estimates toward zero. Second, the correlation of wage risk with financial asset returns is never statistically significant in our regressions. This is in contradiction with standard portfolio choice models, which would predict that higher correlation with financial markets leads to a smaller share invested.[1]

We suspect that we cannot find an effect, because of a lack of salience with respect to the correlation. As a test of this hypothesis, we regress a subjective measure of economic worries on the wage risk and the correlation both in OLS and IV specifications. As in the portfolio choice regressions, wage risk significantly increases economic worries, while the correlation between wage risk and asset returns is not significant.

To facilitate the analysis, we rely on the German Social Economic Panel (SOEP), which contains long-running panel data on hourly wages and wealth portfolios. Further, the SOEP offers the labor market characteristics of individuals and the fine-grained regional information that we need to construct the IV. Additionally, we use financial market return data from the Thomson Reuters Eikon database to construct the correlation.

Our paper makes three major contributions to the literature: first, we quantify the influence of hourly wage risk on portfolio choice and find that the size of its effect is of comparable magnitude to those found in the existing literature. Second, using a new IV-strategy we show the considerable influence of wage risk on portfolio choice in an economically important country with considerable heterogeneity in both asset holdings and wage risk, i.e., Germany. Third, to the best of our knowledge, we are the first to empirically quantify the influence of the correlation between financial market returns and wage risk on portfolio choice and we give a plausible explanation as to why this influence is not statistically significant.

Background risk, especially income risk and its effect on portfolio holdings, has been studied extensively in the literature. Heaton and Lucas (1999, 2000) show with a life cycle model of consumption and portfolio choice with non-tradable labor income, that households with higher background risk hold less financial assets. Cocco et al. (2005) include various correlations between income variance and asset returns in their calibrated decision-theoretic models, showing that even small, positive correlations reduce the risky portfolio share. Buraschi et al. (2010) provide a multivariate model framework to show that risk correlations affect the optimal portfolio choice.

Several studies provide empirical evidence for the importance of background risk. Guiso and Paiella (2008), using Italian survey data, show that background

[1] Our finding could be due to small correlations between wage risk and financial market returns, which has been argued in the literature (Guiso and Sodini, 2013). However, while the average correlation is close to zero, there is much individual heterogeneity. Most of the evidence regarding the correlation of labor income and financial market returns rests on aggregate and not individual level correlations. Davis and Willen (2014) find large and significant correlations of residual occupation-specific wages with a portfolio sorted on firm size. Bottazzi et al. (1996) use a VAR model to derive the correlation between human capital and financial return innovation and finding them to be negative on average (about -0.4).

risk and borrowing restrictions shape consumers risk aversion and, therefore, background risk decreases the willingness to take risk on financial markets. Cardak and Wilkins (2009) find a significant correlation between income risk and risky portfolio holdings in Australia, while Betermier et al. (2012) find a causal effect between the increase of income risk and the decrease of risky portfolio holdings in Sweden. Recently, Fagereng et al. (2017a) include uninsurable labor income as a fundamental component to estimate portfolio choice over the life cycle in Norway. Fagereng et al. (2018) link individual workers to their firms to use the variability in the profitability of the firm as a measure of labor market risk. Their findings underline the significance of background risk for portfolio choice.

Our findings, beyond quantifying the influence of wage risk on portfolio choice and testing the influence of the correlation with financial market returns, have important implications in the areas of household finance and welfare economics. Our study implies that individuals make considerable investment mistakes by disregarding the correlation of their asset portfolio returns and wage risk. Individuals could improve their investment strategy by receiving information on this correlation structure and thus decrease their overall risk exposure, and increasing their welfare.

The rest of the paper is organized as follows. Section 4.2 presents a simple model of portfolio choice with correlated labor income and financial market returns. Section 4.3 presents our two data sources: the Socio-Economic Panel and the Thomson Reuters Eikon database. Section 4.4 shows how we construct our variables for analysis and the specifications we implement. Section 4.5 presents our results. Section 4.6 provides several robustness checks. Section 4.7 qualifies our assumptions and Section 4.8 concludes.

4.2 A Two-Period Model of Portfolio-Choice

To start, we provide relevant theoretical insight into labor market risk, especially the variance of labor income and its correlation with asset returns. Then we show how both risks affect portfolio choice, i.e., the optimal risky portfolio share. We show simulations from a two-period model, where an individual receives some wealth and risk-free income in the first period and risky labor income as well as asset returns in the second period. The individual chooses consumption, and hence saving, in the first period as well as the share invested in a risky asset. Then, we introduce correlation between the risky asset returns and risky labor income.

The Model The formulation and parametrization of the model follows Cocco et al. (2005).[2] Apart from not featuring a longer time-horizon, we also disregard the bequest motive and labor supply. This reduction in complexity is acceptable, since

[2]See page 501, Table 4.

the model is only intended to illustrate individuals' reactions to changes in risk and the correlation of risk factors, which are qualitatively unaffected by this reduction in complexity.[3]

An individual's utility exhibits constant relative risk aversion and thus the individual maximizes:

$$\max_{c_1, c_2, \omega} E\left[\frac{c_1^{1-\sigma}}{1-\sigma} + \beta \frac{c_2^{1-\sigma}}{1-\sigma}\right], \tag{4.1}$$

where c_1 and c_2 denote first and second period consumption respectively and the parameter σ determines relative risk aversion. The maximization is subject to the inter-temporal budget constraint given by

$$c_2 = w_2 + R\omega(w_1 - c_1 + a_1) + (1+r)(1-\omega)(w_1 - c_1 + a_1), \tag{4.2}$$

where w_1 and w_2 is labor income, a_1 is the stock of assets in the first period, $(1+r)$ is the non-risky return on the safe asset, R is the risky return, and ω is the share invested in the risky asset (risky share). Unlike w_1, w_2 is risky and, together with R, they are log-normally distributed with the underlying normal distribution:

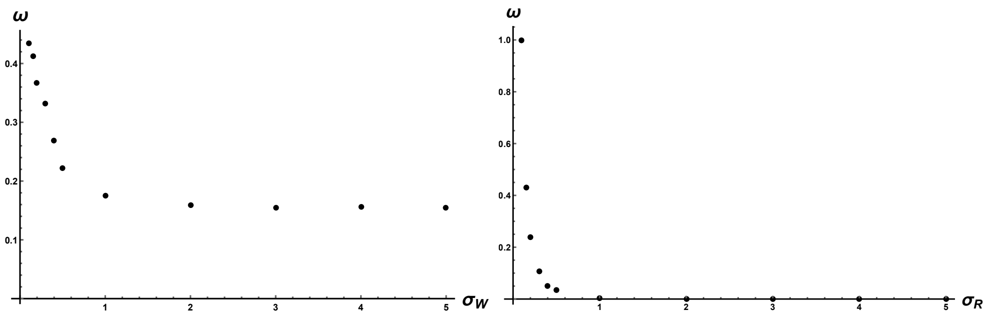
$$\{\log w_2, \log R\} \sim N\left(\begin{pmatrix} \log w_1 - \frac{\sigma_w^2}{2} \\ \log(1+r+\tilde{e}) - \frac{\sigma_R^2}{2} \end{pmatrix}, \begin{pmatrix} \sigma_w^2 & \rho\sigma_w\sigma_R \\ \rho\sigma_w\sigma_R & \sigma_R^2 \end{pmatrix}\right), \tag{4.3}$$

where the mean of w_2 is w_1 and the mean of R is given by $1+r+\tilde{e}$, i.e., the safe return plus an expected excess return. The standard deviation of the log of w_2 and the log of R are σ_w and σ_R, respectively, and the correlation $\log w_2$ and $\log R$ is ρ. In the following exercises, we will vary σ_w, σ_R, and ρ to illustrate their influence on the choice of ω. The baseline calibration of the model is given in Table 4.1.

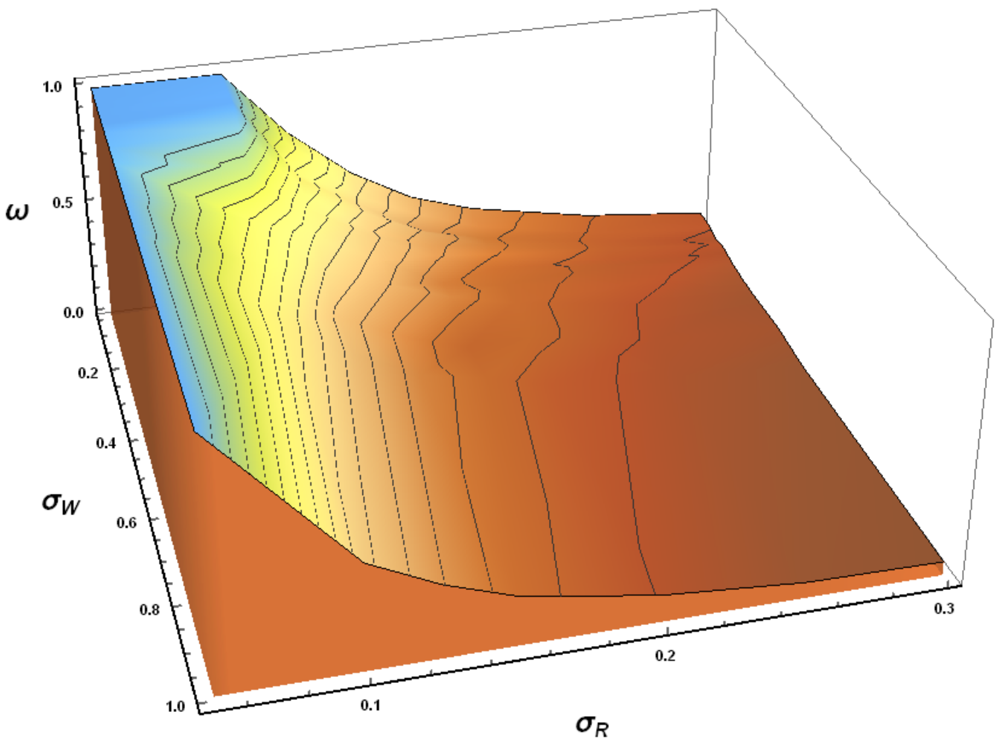
Optimal Portfolio-Choice We solve the model numerically using the Nelder-Mead algorithm implemented in Mathematica 11. First, in the top panel of Figure 4.1, we vary labor income risk and asset return risk separately at ρ equaling zero. The risky share ω declines as the economic environment becomes more risky, be it from labor income risk σ_w or return risk σ_R. However, ω declines much faster when return risk σ_R rises. Second, in the bottom panel, we vary both dimensions of risk jointly, yet still keep ρ at zero. The plot demonstrates that, in the end, it is total risk, that determines portfolio choice.

In Figure 4.2 we illustrate the importance of correlation between the two sources of risk. Although we have calibrated the correlation ρ to two relatively moderate

[3] For example, one can think of this model as looking at a decision between two periods in a model with more than two periods. Further, regarding labor supply, as long as individuals still have incentive to work, they will still partially be subject to labor income risk.



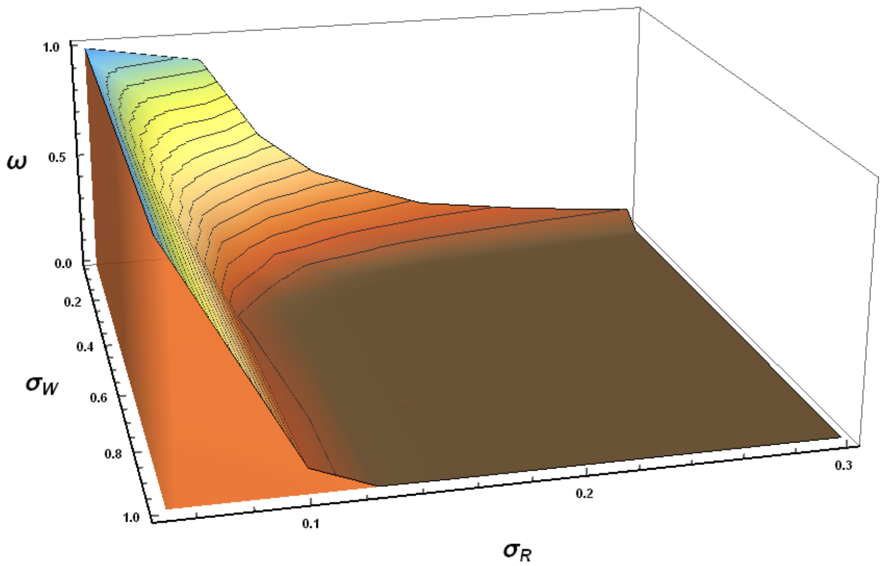
(a) Vary Risks Separately



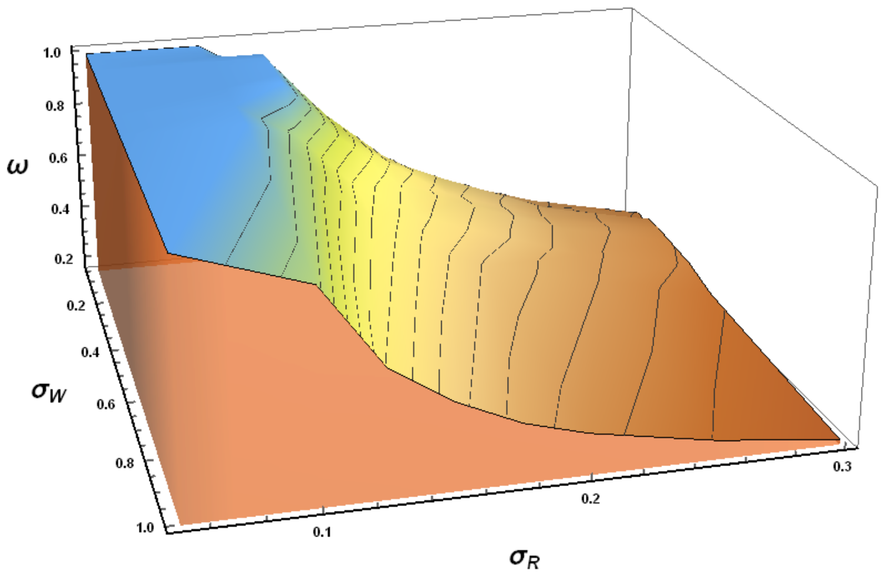
(b) Vary Risks Jointly

Note. Authors' calculations using Mathematica 11. We plot the optimal share of risky assets for different σ_w and σ_R at $\rho = 0$.

Figure 4.1. Optimal Risky Share ω Varying Non-Asset Income Risk σ_w and Return Risk σ_R



(a) $\rho = 0.2$



(b) $\rho = -0.2$

Note. Authors' calculations using Mathematica 11. We plot the optimal share of risky assets for different σ_w and σ_R at ρ equal to -0.2 or 0.2.

Figure 4.2. Optimal Risky Share Varying Non-Asset Income Risk σ_w and Return Risk σ_R with Correlation ρ

Table 4.1. Baseline Calibration

Description	Variable	Value
CRRA parameter	σ	10
first period assets	a_1	25000
first period labor income	w_1	25000
s.d. of risky income	σ_w	0.15
discount factor	β	0.96
non-risky return	r	0.02
expected excess return	$\tilde{e}$	0.04
s.d. risky asset	σ_R	0.157
correlation between $\log w_2$ and $\log R$	ρ	0.0

Note. Model specification taken from Cocco et al. (2005).

values, i.e., 0.2 and -0.2[4], the change from the base case in Figure 4.1 is remarkable. Since now labor income risk and return risk are linked, the individual chooses far smaller risky portfolio shares at even moderate levels of risk when the correlation is positive. Conversely, the risks serve to hedge each other, which leads to a much higher risky share, if the correlation is negative.

In conclusion, we note that the two risk sources have non-trivial interactions that lead to diverse consumer behavior. Even at a correlation of zero, both risks contribute to the total risk and thus influence the choice of the risky portfolio share. When the correlation between the two risk sources is non-zero, the consumer's response is even more pronounced: at a moderate positive correlation, consumers completely eliminate risky assets from their portfolio. This occurs even at values of risk, that would have resulted in a positive risky share had the correlation been zero. Contrarily, a moderate negative correlation strongly incentivizes risky asset holdings since they offer insurance against labor market risk. These results point out that labor income risk and its co-movement with return risk are crucial for portfolio choice.

4.3 Data

For our analysis we use two main data sources: the German Socio-Economic Panel and the Thomson Reuters Eikon database.

German Socio-Economic Panel We use the German Socio-Economic Panel (SOEP) as our primary source of data. The SOEP is a nationally representative panel study with data running from 1984 to 2018 (Goebel et al., 2019; Schröder et al., 2020b).

[4] Same values as chosen by Cocco et al. (2005).

Data on assets, collected on the individual level, are available in 2002, 2007, 2012, and 2017, while data on labor market outcomes are available in every year.

Table 4.2. Observations: From the Full - to the Working Sample

Description	2002	2007	2012	2017	Sum
Full Dataset	23,892	20,869	27,940	32,397	105,098
Between 18 and 65	-4,066	-4,435	-5,296	-5,110	-18,907
Positive net wealth	-5,100	-4,166	-7,745	-12,612	-29,623
Full Labor Market Information	-10,029	-5,097	-9,421	-6,314	-30,861
Characteristics	-288	-374	-308	-657	-1,627
Working sample	4,409	6,797	5,170	7,704	24,080

Note. Compiled by the authors using SOEP v35. The table provides number of observations of full dataset and the working sample by year. The table also shows, how several sample restrictions reduce the numbers of individuals in the dataset. The numbers of observation are based on imputed data.

Sample Definition We restrict the sample to the working population aged 18 to 65, who hold positive net wealth, with sufficient observations to construct the outcome variable, the regressors, and the instrumental variable. We include four waves since data on assets is only available for 2002, 2007, 2012, and 2017. We provide an overview of the numbers of observations in Table 4.2. The restriction on the working population reduces the dataset by 18,907 observations. A large reduction is imposed by including only individuals with positive net wealth, eliminating another 29,623 observations. We choose to exclude those observations for two reasons: first, our analysis focuses on the financial share, which can only be observed for those who own some assets. Second, we exclude those with negative net wealth, mainly to keep our results comparable to the previous literature (Fagereng et al., 2017b, 2018).[5] Focusing on individuals with full labor market information reduces another 30,861 observations from our sample. This includes individuals, which are employed long enough to provide information on wage volatility[6] and its correlation with financial market returns. Moreover, for the calculation of the instrument, we need information on their occupation (isco 88), and their residency (NUTS2). As we control for a set of socio-economic characteristics in our regression analysis, we lose some observations due to missing values and limited panel information. Our basic working sample includes 24,080 observations.

[5]This affects in total about 3,200 observations with financial assets. Including those observations, however, does not affect our results.

[6]For the construction of wage risk we use the years from 1998 to 2017.

Mean values for some of the labor market variables as well as socioeconomic characteristics are shown in Table 4.7 in Appendix 4.9.1 for both our working sample and the overall dataset in the considered years. Our sample selects on those who own at least some financial wealth and are in the labor force. Individuals in our working sample are more educated. Moreover, there are proportionally less women in the sample. The average number of children and the average age is around the same level in both samples. There are more married individuals, as our definitions potentially restricts younger singles, who do not hold financial wealth, less individuals with a migration background, and slightly more individuals who live in former East Germany. Individuals in our working sample earn higher levels of gross labor income, and have higher net and financial wealth. There are also more individuals who own their own home. We also provide the financial share, with and without housing wealth.

Financial Market Returns We retrieve annual excess returns from the Thomson Reuters Eikon database. The returns are calculated as mean total return of the German HDAX index, which combines the main German indices, i.e., DAX, MDAX and TecDAX. To retrieve the excess return, we subtract mean annual interest rates of 10-year German sovereign bonds from the mean total HDAX returns. These excess return should capture the main volatility in German equity markets. The mean annual excess return is 6.01% with a standard deviation of 24.34% over the 1998-2017 period. A graph depicting mean annual returns per year is depicted in Appendix 4.9.4 in Figure 4.7.

Required Variables Three types of information from the SOEP are central to our analysis: hourly wages, to construct the idiosyncratic variance, wealth portfolios, to construct the financial portfolio share, and fine-grained regional information, to construct the instrumental variable.

Wages — To compute hourly wages, we use reported monthly labor income and contractual working hours per week. A common concern with hourly wages constructed from survey data is measurement error. While it is not clear how measurement error would affect our estimates, we can allay these concerns by pointing out that measurement error of hourly wages has been studied in the SOEP before.[7] Caliendo et al. (2018) provide a cross-validation of SOEP-based distributions of working hours, monthly labor income, and hourly wages with the cross-sectional Structure of Earnings Survey (SES). SES is a firm-level survey dataset providing payroll information including income and contractual working hours. Caliendo et al.

[7] Since we are ultimately only interested in the effect of idiosyncratic wage variance on the financial portfolio share, it is not clear that measurement error, if it were present to a noticeable extent, will affect this relationship in a meaningful way. Since we use an IV-strategy for our main estimates, measurement error should not be a problem in any case.

(2018) show that SOEP-based and SES-based distributions of income, working hours, and hourly wages are very similar, thus suggesting no relevant differences.

Wealth Portfolios — The SOEP surveys a total of eight types of assets in 2017:

1. owner-occupied residential property,
2. miscellaneous property ownership (including undeveloped land and holiday and weekend homes),
3. financial assets (savings accounts, savings bonds, corporate stocks, and fund shares),
4. assets from private insurance policies (life and private pension insurance including Riester pensions),
5. balance on savings account with a building and loan association,
6. business assets (ownership of sole proprietorships and participation in partnerships or corporations, net operating liabilities),
7. tangible assets in the form of valuables such as gold, jewelry, coins, or artwork,
8. value of vehicles.

SOEP reports four types of liabilities

9. mortgage loans on owner-occupied property,
10. mortgage loans on miscellaneous property,
11. consumer loans,
12. student loans.

Figures on the value of vehicles and the balance of student loans were not collected between 2002 and 2012. To produce internally consistent wealth concepts, we exclude these items from the analysis.[8] Deducting the liabilities from the assets results in the total net wealth. Our financial portfolio share is constructed from gross financial assets divided by overall gross wealth excluding housing wealth.

Regional Information —The SOEP records the current location of an individual's residence at different levels of aggregation down to the "Kreis"-level, comparable to a district. We use the level of aggregation above the "Kreis"-level, the NUTS2 code (Nomenclature of Territorial Units for Statistics). This code is based on 38 "Regierungsbezirke", which translates to governmental districts.

[8] Vehicles can be argued not to be relevant to wealth, since they are a consumption good and act as a poor store of value. Education loans are almost completely irrelevant in Germany (Schröder et al., 2020a).

4.4 Methodology

In the following we detail how we calculate idiosyncratic wage variances and their correlation with financial market returns on the individual level, then how we define the financial share, and finally how we construct our instrumental variable.

Residual Wage Variance and Correlation To remove variation in wages that is predictable by the individual, we run the OLS regression:

$$w_{it} = \alpha + \rho w_{it-1} + \gamma' \tilde{\mathbf{X}}_{it} + e_{it}, \tag{4.4}$$

where w_{it} are log hourly wages in year t of individual i, $\tilde{\mathbf{X}}_{it}$ is a set of control variables containing a quadratic in age and dummies for the survey year, gender, company size, German federal states, employment history (unemployment, part-time, full-time), migration background, and completed education.[9] The term e_{it} contains idiosyncratic wage innovations. To appropriately employ the panel data of the SOEP, we run this regression for the years 1998 to 2017. The estimates of the unexplained residuals, $\hat{e}_{it}$, serve to calculate the idiosyncratic variance as

$$V_{it} = \frac{1}{4} \sum_{n=t-4}^{t} \hat{e}_{in}^2. \tag{4.5}$$

Thus, V_{it} are four-year rolling-window means of squared differences for each individual following Fagereng et al. (2018).

To calculate individual-level correlations, we use the excess returns series from the German HDAX:

$$C_{it} = \frac{\frac{1}{4}\sum_{n=t-4}^{t} (\hat{e}_{in}) \times \left(ER_n - \frac{1}{4}\sum_{k=t-4}^{t} ER_k \right)}{\sqrt{V_{it}}\sqrt{Var(ER_t)}}, \tag{4.6}$$

where $Var(ER_t)$ is an analogously defined rolling-window variance of excess returns from the German HDAX.

It is important to note that by restricting our attention to wage variance, instead of all of labor income, we are not threatening, but rather enhancing identification. Our primary focus is on identifying the transmission parameter that tells us how changes in background risk affect the financial portfolio share. By focusing on

[9] This specification follows Fagereng et al. (2018) in the choice of controls. Unlike them, we include the lagged dependent directly in the specification and do not restrict ρ to one.

hourly wages, we are eliminating one important source of potentially contaminating variation not due to background risk, i.e., variation in hours of work.

Financial Portfolio Share The SOEP does not ask individuals directly for what they consider to be risky investments. However, we can calculate the share of the portfolio invested in financial assets. Thus, we write the financial portfolio share as:

$$fs_{it} = \frac{fa_{it}}{gw_{it}}, \tag{4.7}$$

where fs_{it} is the stock of financial assets, and gw_{it} is gross wealth, that is the sum of all assets, excluding housing wealth, i.e., owner-occupied and other real-estate. By choosing the (overall) financial portfolio share, instead of the risky financial share, which is the more frequently used concept in the literature, we may encounter some discrepancies compared to other studies: a decrease (increase) in the overall financial portfolio share may occur due to a reduction (increase) in risky assets or a reduction (increase) in non-risky assets. Thus, we cannot be completely certain that we are detecting the same effects as the literature based on the risky financial share. However, our results on subjective economic worries support the argument that people are reacting on the risky financial margin. Further, it appears that the SOEP question on financial assets is formulated in a way to reduce the likelihood that individuals report wealth in sight accounts or other non-risky investments. The questionnaire specifically asks for "Geldanlagen", i.e., financial investments, which suggests that these types of assets to be recorded are at least to some extend risky.

We exclude housing because it is generally not only an investment decision, but rather a joint consumption and investment decision. Further, housing is not easily liquidated. The macroeconomic literature has recognized that there are "wealthy hand-to-mouth" households that cannot easily liquidate their wealth in reaction to shocks to smooth consumption (Weidner, 2014). The major illiquid asset for this group is housing. In the same sense, we expect that housing is not easily liquidated or adjusted to react to changes in labor income risk. The return on housing wealth, especially in Germany, is far less risky than the return on equity. Data from Jordà et al. (2019) and a calculation using this data in König et al. (2020) support this fact.[10] Thus, when we include housing wealth as a non-risky/non-financial asset and reproduce our main results in Appendix 4.9.5, our main conclusions are qualitatively unaffected.

[10]Granted, the risk exposure for the individual is possibly far greater because the investment is indivisible and diversification within housing is not possible for most individuals.

Empirical Strategy In what follows, we wish to measure the effect of idiosyncratic wage risk on the financial portfolio share. Following Fagereng et al. (2018), we model the relationship as

$$fs_{it} = \alpha + \beta V_{it} + \gamma' \mathbf{X}_{it} + \nu_i + \epsilon_{it}, \tag{4.8}$$

where the set of control variables $\mathbf{X}_{it}$ contains a quadratic in age, dummies for household composition and marriage, survey year, home-ownership, as well as the log of net wealth. Further, ν_i is a fixed effect, which will capture time-invariant differences in portfolio choice, i.e., individual-level heterogeneity in risk preference. ϵ_{it} is an error that captures time-varying unexplained changes in the financial share. The coefficient of interest is β.

When we wish to measure the effect of the correlation of wage risk with financial market returns on the financial share, the specification takes the form

$$fs_{it} = \alpha + \beta_1 V_{it} + \beta_2 C_{it} + \gamma' \mathbf{X}_{it} + \nu_i + \epsilon_{it}, \tag{4.9}$$

where now the coefficients of interest are β_1 and β_2. For both equations we run regressions in fixed effects and fixed effects instrumental variables specifications.

Instrumental Variables After eliminating the omitted variable problem in estimating Eq. (4.8) using fixed effects, only the problem of measurement error in the idiosyncratic variance remains. The concern is that some of the variance will remain due to individuals' choices and is actually under their control (Fagereng et al., 2018). Thus, the wage risk is measured with error and, as Fagereng et al. (2018) show, this attenuates its effect on portfolio choice.

A valid instrument captures a part of the variance that individuals cannot meaningfully influence. Our IV-strategy exploits the interactions of fine-grained regional, time, and occupational information (2-digit ISCO code) to form cell-specific deviations from the overall variance, i.e., a group IV estimator (Blundell et al., 1998; Blau and Kahn, 2006; Burns and Ziliak, 2017). This IV captures, among other things, variation in local labor market conditions, such as plant closures and openings or different developments in regional infrastructure. These types of shocks are beyond the control of the individual and, thus, are due to true wage risk. In this sense, the variation we use for our IV is comparable to the IV established in Fagereng et al. (2018), which relies on firm-specific productivity shocks and links them to individuals working in these firms. To check whether the IV is relevant, we report F-tests of the first stage in the IV regressions.

The regional classification uses the SOEP's NUTS2 variable, which is one level above the finest regional information, i.e., the "Kreis"-level (NUTS3). We choose this level to achieve a balance between granularity and thus more variation on the one hand, and sample size within each of the groups on the other. We then interact these regional dummies with time dummies and occupational dummies to obtain our final dummy set. The number of groups used in the IV is 3,427 and the mean group size is 15.44 (s.d. 13.06).[11]

One could be concerned that at low group size, the cell mean is driven by one or a couple of observations with a particularly high wage variance. However, we are estimating a fixed effects IV regression, so that persistent individual level differences are controlled for. Thus, the dynamics of the cell means drive the variation of the IV.

To construct the instrument for the variance, we run a regression of V_{it} on the group indicators and predict the IV from this regression. Similarly, since we also need an IV for the correlation with financial market returns, we run an analogous regression and prediction for C_{it}. Thus, we have exactly one instrument for each of the endogenous variables.[12]

4.5 Results

We begin by showing descriptives for our main outcome and independent variables and then present the results from our central estimation.

4.5.1 Descriptives

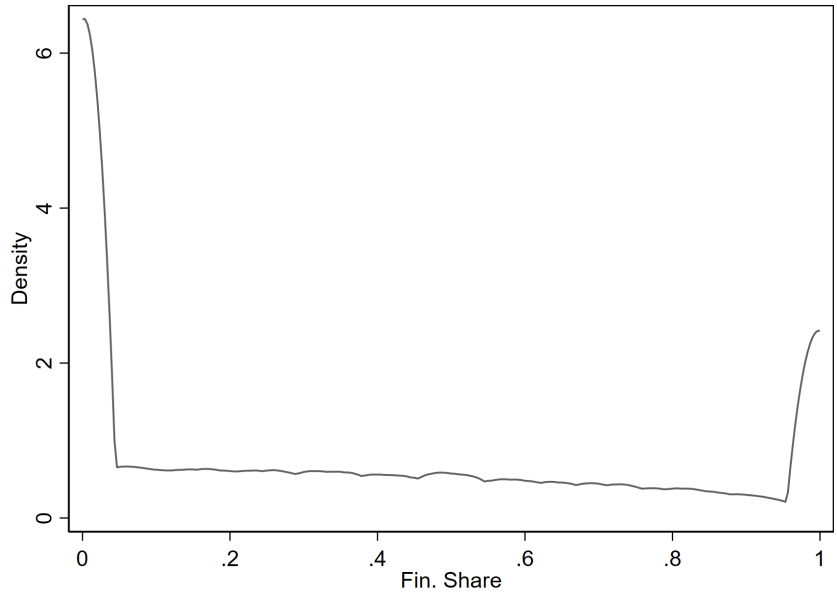
As a first description of our dataset, we illustrate the distribution and dynamics of the financial share and we show its PDF and the movement of its mean over time in Figure 4.3.

The pattern, both in the pooled sample and in the time-series picture, line up closely with what is found for the risky financial portfolio share in the comparable literature (Fagereng et al., 2018). Inspecting the upper panel, we find that a very large fraction of observations have no financial assets at all. For low financial shares smaller than 20% the fraction of observations is still considerable. For financial shares larger than 40% the fraction of observations is very small with a final, considerable hump at 100%.

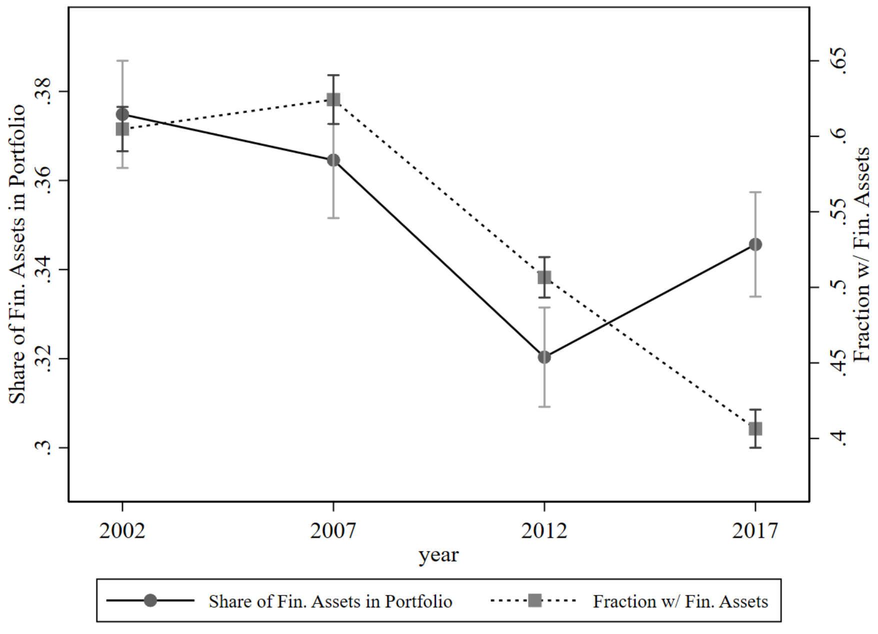
The lower panel shows the share of observations with financial assets, which ranges around 60% to almost 40%. The share dropped considerably after 2007 to levels around 40% in 2017. The conditional financial share dropped since 2002 (about 37%), with a pronounced decrease after the financial crisis of 2008, and a recovery to pre-crisis levels in 2017 (about 34%). The financial share comoves with the number of observations with positive financial asset holdings.

[11] We show more descriptive statistics in Appendix 4.9.2

[12] This means, even though, the variation is based on many cell means, we collect them in one variable. Therefore we have no overidentifying restrictions.



(**a**) Kernel Density of the Financial Share



(**b**) Mean of the Financial Share and Fraction of Observations with some Financial Assets Over Time

Note. Compiled by the authors using SOEP v35. The kernel density was calculated with a bandwidth of 0.02. Unweighted means in panel (b) are provided with 95 percent confidence intervals based on bootstrapped standard errors using 1000 replica weights. Sample is not restricted to working sample.

Figure 4.3. Descriptive Statistics on the Financial Share

Our financial share includes both risky and risk-free investments. The fraction of households with financial assets, hence, does not represent the stock market participation rate. Other studies suggests that the participation rate in Germany lies between 18 percent (Breunig et al., 2019) and 24 percent (Necker and Ziegelmeyer, 2016).

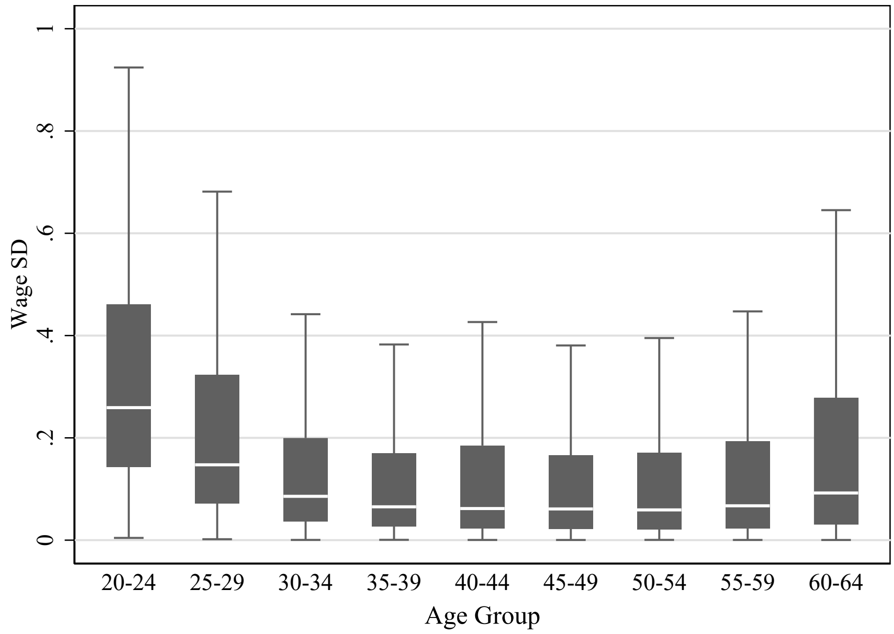
Table 4.3. Descriptives of Focal Variables for the Full Working Sample

Variable	year	mean	SD	min	max
Financial Share	2002	0.339	0.366	0	1
	2007	0.329	0.350	0	1
	2012	0.326	0.351	0	1
	2017	0.316	0.360	0	1
Residual Wage Variance	2002	0.217	0.383	0.0005	5.251
	2007	0.206	0.368	0.0003	5.642
	2012	0.194	0.341	0.0004	4.355
	2017	0.190	0.356	0.0003	4.570
Correlation	2002	-0.049	0.585	-1	1
	2007	-0.049	0.563	-1	1
	2012	0.074	0.534	-1	1
	2017	-0.003	0.557	-1	1

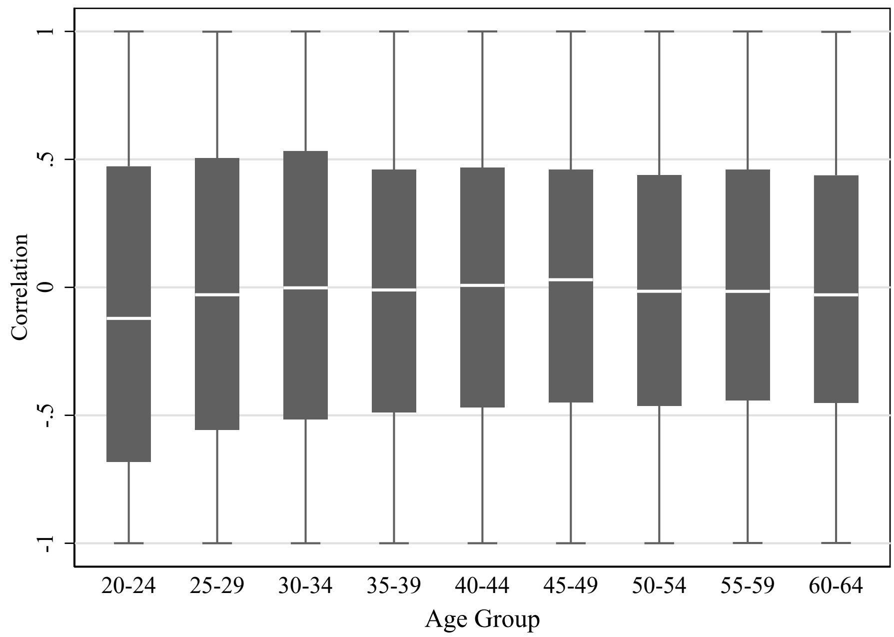
Note. Compiled by the authors using SOEP v35. The table provides unweigthed means of the focal variables for every year of the sample.

In Table 4.3 we show basic descriptives of the focal variables for the working sample. The average financial share in the working sample is about 33%, while the standard deviation is quite large. The average variance of wage risk is fairly small, but seems to have a long tail. The average correlation is very small and negative, except in 2012, and the standard deviation is large. The descriptive statistics of all variables do not vary much of the years, but they reveal plenty of variation in the focal variables in the working sample.

We provide box plots of the standard deviation of wage risk and its correlation with financial market returns for five year age groups in Figure 4.4. The standard deviation is higher for younger age cohorts, remains at a lower level for the middle aged and slightly increases for the groups closer to retirement. This is what we would expect, indicating the individuals face higher unexplained wage volatility at the beginning of their career (Kaplan, 2012). Increasing experience combined with academics entering the labor force at the end of their education reduces the standard deviation during the 20s. More individuals dropping out of the labor force and devaluation of human capital potentially explain the dispersion before retirement. The correlation with HDAX returns, however, does not show any connection to age



(a) Box Plot of Standard Deviation of Wage Risk by Age Cohorts.



(b) Box Plot of Correlation Between HDAX Returns and the Idiosyncratic Wage Risk by Age Cohorts.

Note. Compiled by the authors using SOEP v35. The white bar in the box represents the median. The ends of the boxes represent the 25th and 75th percentile, respectively. The whiskers define the adjacent value, which is the first (third) quartile minus (plus) 1.5 times the interquartile range. Sample is to working sample

Figure 4.4. Life Cycle Pattern of Wage Risk and Correlation

cohorts. The median is about constant and close to zero, while the quartile ranges, as well as the adjacent values, reveal a large dispersion over the full support.

4.5.2 The Influence of Wage Risk and its Correlation with Asset Returns on the Financial Portfolio Share

We discuss the results of our IV estimation in this subsection. Table 4.4 shows the causal effect of wage risk and the correlation of wage risk with financial excess returns on the financial portfolio share.

Table 4.4. Influence of Wage Risk and its Correlation with Asset Returns on the Financial Portfolio Share

	(1) FE	(2) IVFE	(3) FE	(4) IVFE
V_{it}	-0.0297***	-0.0846**	-0.0297***	-0.0849**
	(0.0098)	(0.0335)	(0.0098)	(0.0335)
C_{it}			0.0008	0.0029
			(0.0049)	(0.0132)
age	-0.0300***	-0.0315***	-0.0300***	-0.0315***
	(0.0037)	(0.0038)	(0.0037)	(0.0038)
age^2	0.0003***	0.0004***	0.0003***	0.0004***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
home ownership	-0.0494***	-0.0484***	-0.0494***	-0.0483***
	(0.0121)	(0.0121)	(0.0121)	(0.0121)
log wealth			0.0008	0.0029
	(0.0044)	(0.0044)	(0.0044)	(0.0044)
constant	1.1806***	1.2293***	1.1805***	1.2293***
	(0.0850)	(0.0904)	(0.0850)	(0.0904)
Controls	Yes	Yes	Yes	Yes
Obs.	24,080	24,080	24,080	24,080
Fstat		79.45		79.45

Note. Compiled by the authors using SOEP v35. Robust standard errors in parentheses. FE is a fixed effects specification, IVFE is an instrumental variables fixed effects specification. Log wealth refers to the log of net wealth. Fstat refers to the first stage F-statistic. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Columns (1) and (2) in Table 4.4 regress the financial portfolio share on wage risk showing that the coefficient of the wage variance is negative and statistically significant at least at the 5%-level in all specifications. Note that the coefficient roughly triples in absolute size when switching from the FE to the IVFE specification.

Our baseline estimate is in column 2, the IVFE specification. Taking this value, we find that increasing wage risk by one standard deviation decreases the financial share by roughly 3 percentage points.[13] Comparing this value to the existing literature, we find that the value is on the upper end of recent estimates being close to the coefficient in Betermier et al. (2012) (-0.12). However, as we have detailed in Section 4.4, our dependent variable is the (overall) financial share of the portfolio and not the risky share. Thus, while the financial share is closely aligned with the risky financial share, we expect the effect of wage risk on the (overall) financial share to be a lower bound of the effect of wage risk on the risky financial share.[14]

We choose control variables that are familiar to the literature, but since our dependent variable is the financial portfolio share, coefficients and interpretations are not necessarily closely aligned. Our age effect is u-shaped and reverses around age forty, the prime age for home ownership and family formation. After forty, we find an increasing trend in the financial share since individuals are more likely to save for retirement. The log of net wealth is negatively associated, pointing to the fact that individuals with very high values in wealth do not have a large financial asset share. This is in line with common observations in the SOEP data that very high-wealth individuals hold their assets primarily in their own businesses (Schröder et al., 2020). Home ownership also enters negatively. This corresponds, in a broader sense, with Chetty et al. (2017), who find that less wealth invested in housing would increase liquid wealth. Further, as individuals in Germany often build (financial) wealth to buy a house, it is not surprising that after this purchase their financial share drops.

Columns (3) and (4) of Table 4.4 additionally include the correlation of wage risk with financial excess returns. In the FE-specification the coefficient is small and positive, while it is a lot larger in the IVFE-specification. However, in both specifications the correlation is statistically insignificant. Thus, we do not find any evidence that this correlation influences individuals portfolio choice.

4.5.3 Discussion

Table 4.4 contains two main findings.

1. The effect of wage risk on the financial portfolio share is negative and statistically significant, and the effect sizes are comparable to those found in the existing literature.

2. The effect of the correlation of wage risk with financial excess returns on portfolio choice are economically and statistically insignificant.

[13] Note that this calculation is only ultimately meaningful for behavior if it refers to uninsurable wage variance, which we cannot fully disentangle from insurable wage variance.

[14] Since we control for net wealth, we do not suspect that we are picking up primarily movements in the denominator of the financial portfolio share.

The first result confirms the basic qualitative pattern in the literature: more wage risk leads to less investment in financial assets. The second results is completely novel. The estimated effect of the correlation is close to zero and statistically insignificant. Our suspicion is that, wage risk is understood intuitively by individuals, but the correlation with financial market returns is not. As individuals observe their wage changes, some of them being not in line with their expectations, they can develop an intuitive understanding of their wage risk (Guvenen, 2007). Hence, it seems plausible that this understanding would also influence portfolio choice. For example, if a worker unexpectedly receives lower wages because demand for the product she is producing has dropped, it seems reasonable that she would lower her exposure to financial market risks.

For the correlation, a case for intuitive understanding appears weak. Developing any understanding of the correlation of wage risk with financial market returns is much less natural and cannot develop as individuals observe the realizations of their wages alone. Rather, it would take explicit calculation to know how one's wage moves with the markets. Accordingly, we suspect that the small and insignificant coefficient on the correlation stems from insufficient salience of it.

Our argumentation would fail, if German portfolio investments were rather internationally diversified and, therefore, do not take German equity market volatility into account. This is unlikely, as the literature confirms a persistent home bias in investment decisions (Levy and Levy, 2014). Moreover, the volatility of international financial markets is smaller, but correlates highly with the internationally integrated German equity market[15]: even if investors would diversify their portfolios only internationally, the risk would correlate similarly with their wage risk. Another possibility is that financial market returns and wage risk do not comove enough. The descriptive statistics in Table 4.3 and Figure 4.4 point to considerable variation in the correlation across the working sample. Thus, it seems unlikely that the issue arises due to a lack of variation or a concentration on zero, as presumed in Guiso and Sodini (2013). In the next subsection we take a look at whether wage variance and correlation are perceived differently.

4.5.4 The Influence of Wage Risk on Economic Worries

To see whether wage risk and its correlation with financial asset returns are perceived with different degrees of salience, we exploit a variable in the SOEP that asks people to report their worries about their own financial or economic conditions on a three point Likert scale (1=strongly worried, 2=somewhat worried, 3=not worried). We perform OLS and IV regressions of this index of worries on the wage variance and the correlation with financial asset returns. Again, we use the IV regressions to

[15]See Appendix 4.9.4.

account for the measurement error problem. Moreover, we control for the same socio-economic characteristics as in the analysis above. Table 4.5 shows the results.

Table 4.5. The Influence of Wage Risk and its Correlation with Asset Returns on Economic Worries

	(1) OLS	(2) IV	(3) OLS	(4) IV
V_{it}	-0.1386***	-0.2217***		
	(0.0124)	(0.0294)		
C_{it}			0.0063	0.0154
			(0.0073)	(0.0196)
age	-0.0375***	-0.0406***	-0.0322***	-0.0323***
	(0.0034)	(0.0036)	(0.0034)	(0.0034)
age^2	0.0004***	0.0004***	0.0003***	0.0003***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
home ownership	-0.0799***	-0.0812***	-0.0776***	-0.0777***
	(0.0103)	(0.0103)	(0.0103)	(0.0103)
log wealth	0.1028***	0.1038***	0.1012***	0.1012***
	(0.0032)	(0.0032)	(0.0032)	(0.0032)
constant	1.9906***	2.2232***	1.8626***	2.0252***
	(0.0729)	(0.0766)	(0.0721)	(0.0721)
Controls	Yes	Yes	Yes	Yes
N	24,012	24,012	24,012	24,012
Fstat		73.61		452.67

Note. Compiled by the authors using SOEP v35. Robust standard errors in parentheses. IV is an instrumental variables specification. Log wealth refers to the log of net wealth. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

In both the OLS and IV specifications the wage variance coefficient is sizable, statistically significant and negative, thus pointing to an increase in worries. When the IV specification is run, just like in Table 4.4, the size of the coefficient increases strongly in absolute terms. The coefficient of the correlation with financial asset returns, however, is positive in both specifications and not significant in either one.[16] These outcomes support the hypothesis that individuals perceive wage risk and factor it into their assessment of economic worries, but they do not take the correlation with financial market returns into account.

4.6 Robustness Checks

We provide several robustness checks for our results. They include variations of the primary fixed effects IV model specification and a jackknife procedure to test the sensitivity of the IV estimates to single groups in the group IV.

[16]Note that we lose 68 observations compared to the baseline specification, due to missing values.

Variations of Model Specification and Sample Selection Table 4.6 depicts the coefficients of wage risk V_{it} and the correlation C_{it} for several different specifications of the fixed effects IV model, as provided in Equation 4.9. The full results of our different model types are provided in Appendix 4.9.5. Model (1) provides the model without any controls, model (2) includes age controls only, model (3) presents the result when including wealth controls only, model (4) increases the rolling-window of the wage and return variances from four to five years, model (5) excludes self-employed workers, and model (6) excludes civil servants from the sample.

The table shows that our results are robust to different model specifications. The effect size of the instrumented residual wage variance ranges between −0.112 and −0.073 and generally stays significant at the five percent level. The correlation, however, remains close to zero and non-significant.

Excluding the self-employed in model (5) is an important robustness check for two reasons. First, the self-employed often invest a large share of their wealth in their own business and invest very little in financial markets (Fossen, 2011; Fossen et al., 2020). Second, wage risk is different in levels and dynamics for this group, especially because the self-employed may more freely choose their hours and may choose to forgo labor income in favor of savings in their business (Hurst et al., 2010).

We exclude individuals working as civil servants in model (6), as they are typically less affected by local labor market conditions. The estimated coefficient of the wage variance is with −0.112 more negative, then the one from our baseline model. This is not surprising, as we exclude a group which faces less wage risk.[17] Consequently, we would expect less adjustments to the financial share. By excluding them, we find a more negative overall coefficient.

In Appendix 4.9.5, we also provide the results of including household wealth into the financial share definition. Table 4.16 shows that the IVFE estimates of V_{it} and C_{it} are smaller and become non-significant. As the change of the financial share represents now higher amounts of wealth, the results point to into the same direction, even though they are not significant. As we argue above, we do not think that it is appropriate to consider housing for short-run portfolio choice decisions.

Jackknifing IV Groups We test the sensitivity of our IV estimates to single groups of the instruments. Therefore, we re-run our main regressions while eliminating one group at the time. As we exclude every group once in this procedure, we obtain 3,427 estimates of the instrumented idiosyncratic wage risk, i.e., β_1 from Equation 4.8, and the effect of the correlation of wage risk with financial market returns volatility, β_2 from Equation 4.9, on the financial portfolio share.

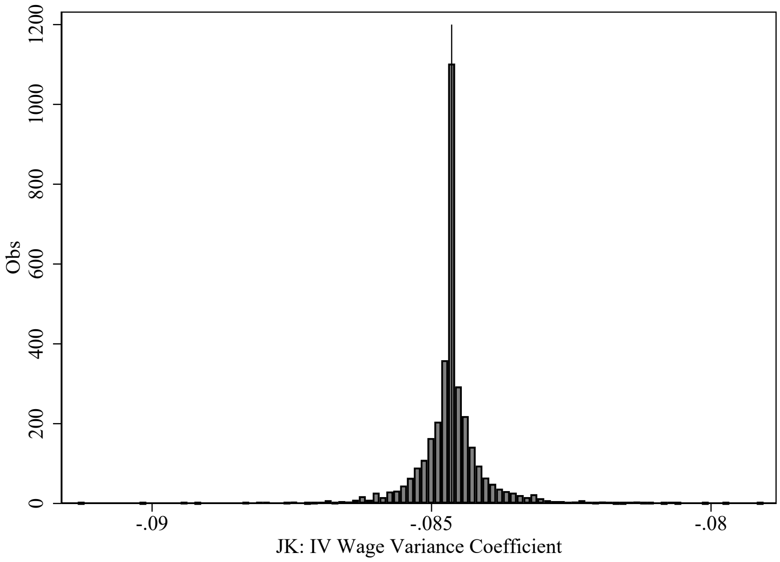
Figure 4.5 provides the histograms of the 3,427 coefficients estimates. The upper panel provides the coefficient of the instrumented residual wage variance on the x–axis and the number of estimations on the y–axis. The histogram shows, that our

[17] The mean value of V_{it} for civil servants is 0.13 compared to 0.21 of the overall working sample.

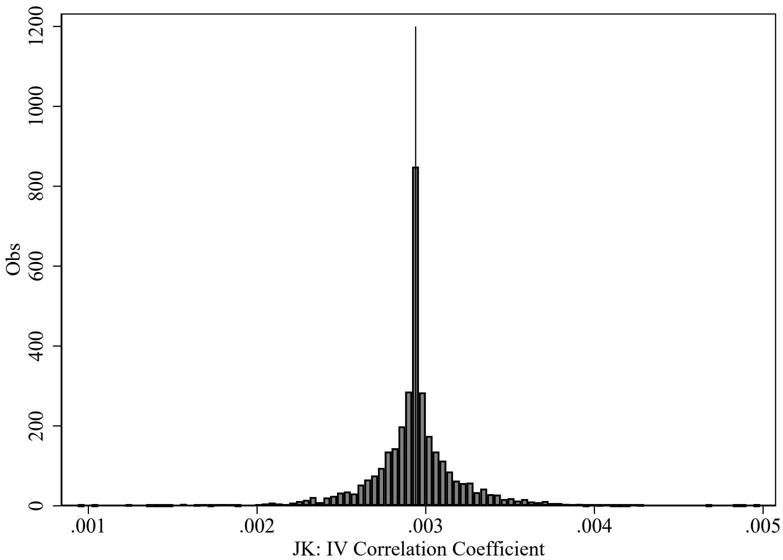
Table 4.6. Robustness Checks: Influence of Wage Risk and its Correlation with Asset Returns on the Financial Portfolio Share

	(1)	(2)	(3)	(4)	(5)	(6)
	No Controls	Age Controls	Wealth Controls	5 years Rolling Window	Excluding Self-Employed	Excluding Civil Servants
V_{it}	-0.0734**	-0.0878***	-0.0797**	-0.0908**	-0.0817*	-0.1119***
	(0.0330)	(0.0329)	(0.0329)	(0.0417)	(0.0458)	(0.0381)
C_{it}	-0.0026	0.0039	0.0021	-0.0014	-0.0001	-0.0005
	(0.0128)	(0.0132)	(0.0128)	(0.0160)	(0.0139)	(0.0160)
Obs.	24,080	24,080	24,080	24,080	22,741	17,611
Fstat	510.59	92.33	259.25	66.95	56.13	56.27

Note. Compiled by the authors using SOEP v35. The table provides the coefficients of the wage variance V_{it} and correlation C_{it} of the fixed-effects IV regression with the financial share as dependent variable. Models (1) to (3) provide the estimates excluding all or some control variables. Model (4) increases the rolling time window to five years, and the models (5) to (6) include the same control variables as the main model in our analysis, but exclude some specific subgroups. Full regressions are provided in Appendix 4.9.5 in Tables 4.11 to 4.17. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$



(a) Jackknife Results: Instrumented Residual Wage Risk



(b) Jackknife Results: Instrumented Correlation

Note: Compiled by the authors using SOEP v35. The panels provide the histograms of the jackknifed estimates of β_1 from Equation 4.8 (upper panel), and β_2 from Equation 4.9 (lower panel). The y-axis depicts the number of estimates. The solid line represents the original point estimate. The size of the bins is 100.

Figure 4.5. Histograms of Jackknifed Results

point estimate of −0.0846, provided as vertical line, is confirmed by more than 1,000 estimates while the variation of the estimates is small. The lower panel provides the coefficient of the instrumented correlation on the ordinate, and, again, the point estimate of 0.0029 does not vary significantly.[18] We conclude from this that our results do not depend on single groups of the instruments.

4.7 Qualification and Extensions

This section discusses caveats of the analysis and extension possibilities. This discussion includes the vocal variables, i.e., the financial share, the wage volatility, the correlation, and the empirical strategy.

A considerate caveat is the lack of information on the individual risky share, which is defined as the ratio of financial wealth invested into risky assets. We are not able to compile the risky share, according to the literature, using the GSOEP data. Our analysis, therefore, focuses on the financial portfolio share. We find suggestive evidence that the financial share captures risky investments, as the GSOEP asks for financial investments (Geldanlagen), the distribution mirrors the risky share provided by Fagereng et al. (2018), and our results on economic worries reflects what we would expect from risky investments. Nevertheless, we cannot proof that our concept approximates the risky share well. [19] Thus, if the overall amount invested remains constant, we cannot rule out that our financial share remains constant, even though re-balancing between risky and less (or non) risky assets has occurred. If this transpires on a large scale, our estimates of wage risk and the correlation are potentially inaccurate.

One could question whether the volatility of unexplained wage residuals is an appropriate measure for labor market risk. The measure would fail, for instance, if the same volatility for two individuals would describe very different scenarios, with large wage increases for one, and large decreases for the other. The literature (Cocco et al., 2005; Cardak and Wilkins, 2009; Betermier et al., 2012; Fagereng et al., 2018) treats wage (and income) risk similarly to a lottery, as those shocks are outside of the individual's control. Therefore, the directions of the shock does not matter, as it only measures the potential variation of the exogenous wage component, which can be positive or negative in future periods. Following this perspective, one could argue that the investors in our analysis are not falling for the *gambler's fallacy* (Tversky and Kahneman, 1971), which would imply that investors hold subjective probabilities on the direction of future shocks. We follow the literature cited above, but this perspective is also contested: Guvenen (2007) argues that individuals may

[18]We provide further descriptive statistics of the jackknifed results in Table 4.18 in Appendix 4.9.5.

[19]Data restrictions in Germany lead to a trade off between sufficient labor market information (provided by the GSOEP) and the accurate conceptualization of the risky share (as found in the Panel of Household Finances).

adjust their expectations based on recent experiences during their life-cycle due to incomplete information. This would mean for our results that we would have to treat positive and negative shocks differently.[20]

A central finding of the paper includes the salience of individuals towards the correlation between financial market risk and wage risk. We show in our theoretical model, that even small levels of correlation affects the optimal risky portfolio share. In our empirical notion of the correlation, we cannot measure individual returns on the financial markets, hence, we use the German stock index HDAX as a proxy. This has two implications: first, individuals may invest in international stocks, which do not necessarily correlate well with the German Index. Second, even if individuals follow a home-bias strategy, as described by (Levy and Levy, 2014), they may invest in single assets with a return structure, which differs from the HDAX. In both cases, our empirical strategy would potentially not capture potential adjustments in the financial share due to variation in the correlation, even if individuals would take it into account. In summary, we argue that even if individuals invest into single assets or mutual funds, we believe that the returns taken from HDAX correlate with these investment and they are, hence, a valid proxy. Apart from model assumptions, one could raise the general question whether it is realistic that this ex-post correlation can be acknowledged by individual investors at all. Surely this would require a detailed knowledge of optimal portfolio choice and certain assumptions about the expected correlation, similar to the expected return. From this perspective, there could be uncertainty about (or even a lack of) information, rather than salience, and the optimal choice for a risk adverse investor may be not to react to changes in the correlation.

As we are interested in the causal effect of wage risk and its correlation with financial risk on the financial portfolio share, the instrument variable needs to be relevant and fulfill the exclusion restriction, i.e., the instrument is not correlated with the error term. Our F-statistics suggest the relevance, but naturally, the exclusion restriction cannot be tested. Our group IV is an interaction between region, time, and occupational information. We argue that it is unlikely to actively select oneself into one of these specific groups. Nevertheless, the empirical estimation can be affected by individuals commuting long distances to work or by small groups. Individuals that work in a different region compared to where thy live, meaning a different region in the NUTS2 scheme, our instrument does not capture local changes in the labor market conditions. Moreover, the IV relies, to some extent, on a small number of observations which can lead to biased estimates. Our Jackknife procedure shows that the result is not driven by one group, but the problem cannot be resolved completely.

Our analysis provides some insights into German portfolio choice. A possible way to extend the analysis would rely on complete information on households'

[20] This matter has not been analyzed in this context so far.

asset portfolios. In this way, one could identify the true correlation between financial market, and labor market risk, without relying on proxies and several strong assumptions.

4.8 Conclusion

We have demonstrated a considerable relationship between wage risk and portfolio choice in Germany. We find that an increase of idiosyncratic wage volatility by one standard deviation reduces the financial portfolio share by three percentage points, while the correlation between financial market returns and idiosyncratic wage volatility is economically and statistically not significant. Individuals with non-zero correlation do not appear to adjust their wealth portfolios taking correlation into account, even though standard portfolio choice theory predicts this. We believe that this is due to a lack of salience and support this claim by regressing an index of individuals' economic worries on idiosyncratic wage volatility and its correlation with financial market risk. This exercise shows that individuals with more wage risk worry more about their economic situation, but the correlation shows no such association.

Given the considerable effect this correlation has on optimal portfolio holdings, there are reasons to be concerned about this result. Since, depending on the sign and size of the correlation, individuals face either more or less risk than they are aware off, large investment mistakes can occur. Improving individual's financial knowledge on correlation structures with their wage risk, could either reduce their risk level for a given return, or help in using the hedging mechanism to maximize returns at a given risk level. In the current framework, our study is limited: we cannot quantify the extent of these investment mistakes, because we do not observe total background risk. Thus, we also cannot quantify the extent to which these mistakes impact individual welfare, which leaves it as a route for promising future research.

4.9 Appendix

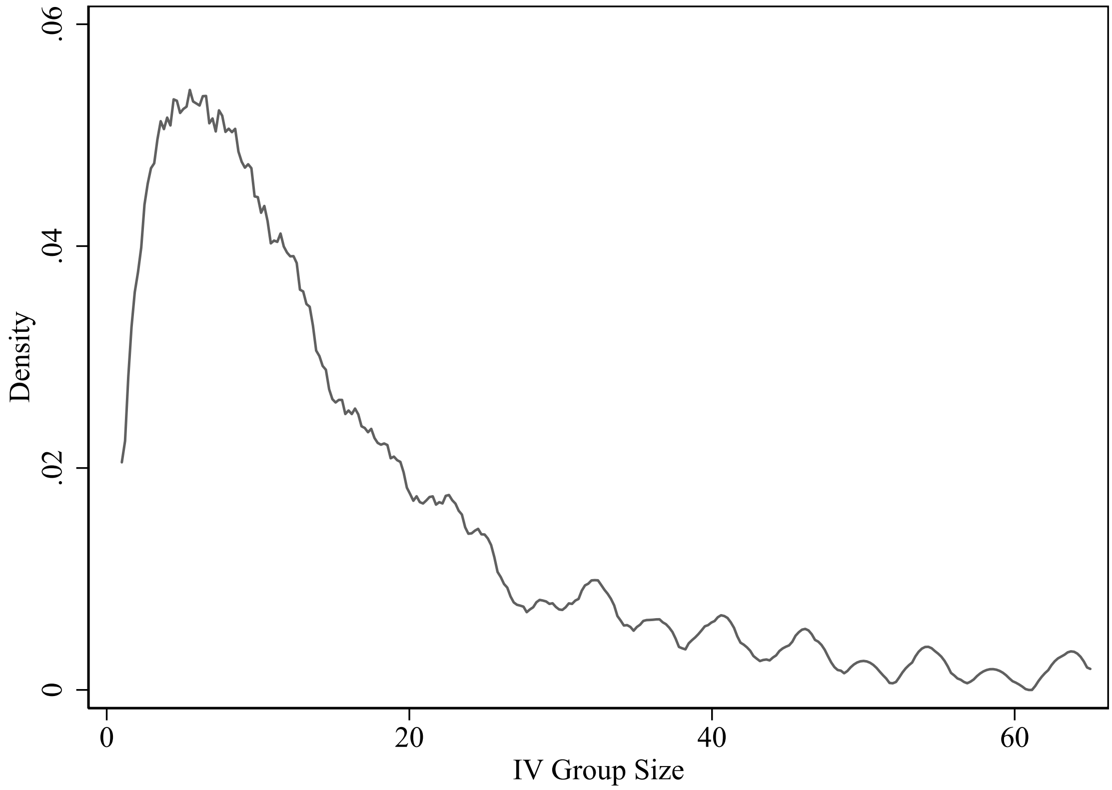
4.9.1 Descriptive Statistics of Full - and Working Sample

Table 4.7. Descriptives for the SOEP

Variable	Sample	Full
Years of education	13.1	11.8
	(0.97)	(0.20)
Female in %	47.5	52.4
	(0.17)	(0.04)
Number of children	0.7	0.7
	(0.32)	(0.07)
Age	45.9	46.1
	(0.04)	(0.01)
Married in %	69.7	62.5
	(0.16)	(0.04)
Migration backgr. in %	13.4	22.9
	(0.12)	(0.03)
Living in East in %	23.2	21.4
	(0.14)	(0.03)
Gross Labor income	2,997	2,224
	(8.75)	(1.86)
Net Wealth	141,282	102,668
	(2,279)	(883)
Financial wealth	18,050	14,033
	(315)	(154)
Home-ownership in %	59,2	47,2
	(0.17)	(0.04)
Financial Share without Housing	32,8	41,2
	(0.12)	(0.09)
Financial Share with Housing	19,2	17,4
	(0.10)	(0.05)

Note. Compiled by the authors using SOEP v35. The table provides the unweighted mean values for the Sample and Full dataset *Sample* describes our working sample with the working population aged 18 to 65 in the years 2002, 2007, 2012, and 2017. *Full* describes the complete dataset in these 4 waves. Bootstrapped standard errors, calculated by using 500 replica weights, are provided in parentheses.

4.9.2 Descriptive Statistics of the Instrumental Variable Groups



Note. Compiled by the authors using SOEP v35. The kernel density was calculated with a bandwidth of 0.8.

Figure 4.6. Kernel Density of the IV Group Size

Table 4.8. IV Group Sizes

Man	SD	p25	p50	p75	p90	N
15.44	13.06	6	11	20	34	3,427

Note. Compiled by the authors using SOEP v35. Table provides descriptive statistics on the group size.

4.9.3 Auxiliary Descriptive Statistics

Table 4.9. Sample Comparison: Owners and Non-Owners of Financial Assets

	owners	non-owners	p-value
Residual Wage Variance	0.19	0.23	0
Years of Educ.	13.29	11.78	0
Female	0.51	0.52	0.01
Number of Children	0.70	0.83	0
Age	44.55	40.80	0
Married	0.66	0.63	0
East	0.21	0.21	0.14
Risky Assets	26,296	0	0
N	34,328	44,506	

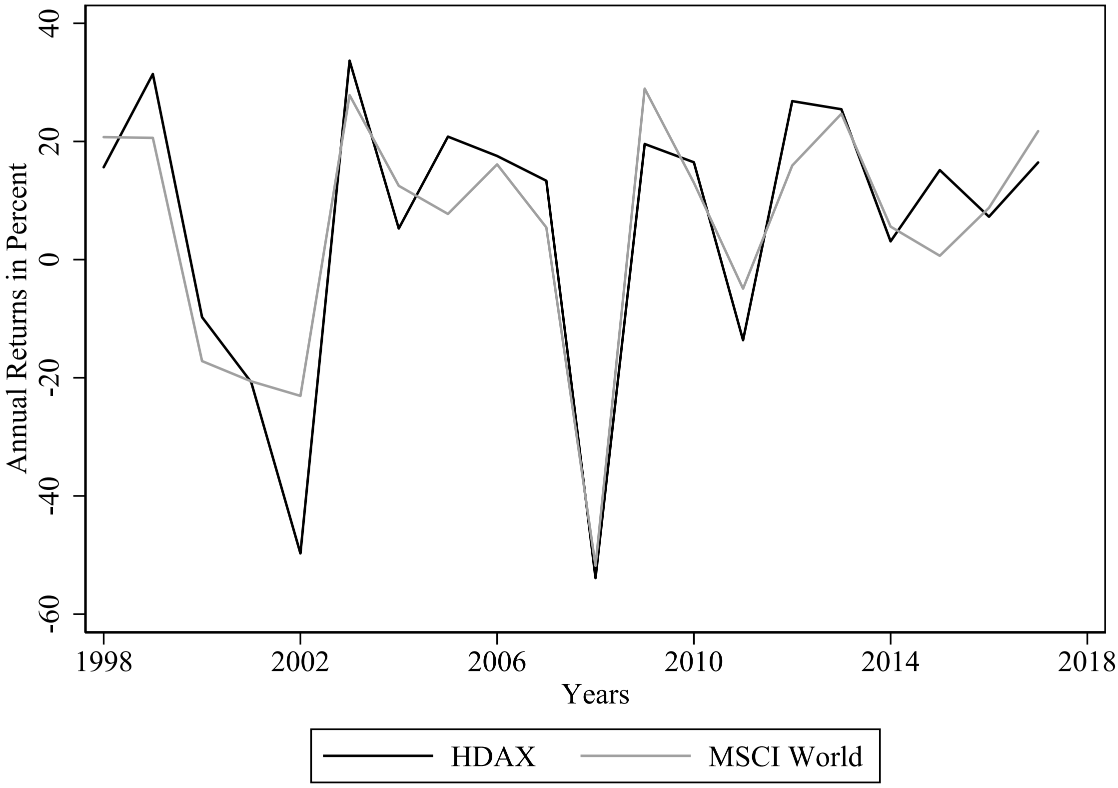
Note. Compiled by the authors using SOEP v35. Means of variables and p-values of t-tests. Sample is not restricted to working sample.

Table 4.10. Descriptives of Focal Variables for the Full Working Sample

	fs_{it}	V_{it}	C_{it}
mean	0.33	0.20	-0.01
sd	0.36	0.36	0.56
min	0	0.00	-1
max	1	5.64	1
N		24,080	

Note. Compiled by the authors using SOEP v35.

4.9.4 Excess Returns Over Time



Note. Compiled by the authors using Thomson Reuters Eikon Database. The annual mean of the excess return is calculated by subtracting the interest rates of the 10-years German sovereign bond.

Figure 4.7. Excess Returns

4.9.5 Robustness Checks

Table 4.11. Influence of Wage Risk and its Correlation with Asset Returns on the Financial Portfolio Share: No Control Variables.

	(1) FE	(2) IVFE	(3) FE	(4) IVFE
V_{it}	-0.0190*	-0.0735**	-0.0190*	-0.0734**
	(0.0101)	(0.0330)	(0.0101)	(0.0330)
C_{it}			0.0021	-0.0026
			(0.0050)	(0.0128)
constant	0.3297***	0.3406***	0.3297***	0.3405***
	(0.0020)	(0.0066)	(0.0020)	(0.0066)
Obs.	24,080	24,080	24,080	24,080
Fstat		510.59		510.59

Note. Compiled by the authors using SOEP v35. Robust standard errors in parentheses. IVFE represents the instrumental variables specification. Log wealth refers to the log of net wealth. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4.12. Influence of Wage Risk and its Correlation with Asset Returns on the Financial Portfolio Share: Only Age Controls

	(1) FE	(2) IVFE	(3) FE	(4) IVFE
V_{it}	-0.0301***	-0.0874***	-0.0301***	-0.0878***
	(0.0099)	(0.0338)	(0.0099)	(0.0338)
C_{it}			0.0011	0.0039
			(0.0049)	(0.0132)
age	-0.0347***	-0.0363***	-0.0347***	-0.0363***
	(0.0036)	(0.0037)	(0.0036)	(0.0038)
age^2	0.0004***	0.0004***	0.0004***	0.0004***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
constant	1.1200***	1.1695***	1.1198***	1.1694***
	(0.0816)	(0.0869)	(0.0816)	(0.0869)
Obs.	24,080	24,080	24,080	24,080
Fstat		92.33		92.33

Note. Compiled by the authors using SOEP v35. Robust standard errors in parentheses. IV is an instrumental variables specification. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4.13. Influence of Wage Risk and its Correlation with Asset Returns on the Financial Portfolio Share: Only Wealth controls

	(1) FE	(2) IVFE	(3) FE	(4) IVFE
V_{it}	-0.0218**	-0.0796**	-0.0218**	-0.0797**
	(0.0099)	(0.0329)	(0.0099)	(0.0329)
C_{it}			0.0025	0.0021
			(0.0050)	(0.0128)
home ownership	-0.0597***	-0.0591***	-0.0596***	-0.0590***
	(0.0119)	(0.0119)	(0.0119)	(0.0119)
log wealth	-0.0179***	-0.0187***	-0.0180***	-0.0187***
	(0.0040)	(0.0040)	(0.0040)	(0.0040)
constant	0.5599***	0.5793***	0.5602***	0.5796***
	(0.0410)	(0.0427)	(0.0410)	(0.0427)
Obs.	24,080	24,080	24,080	24,080
Fstat		259.25		259.25

Note. Compiled by the authors using SOEP v35. Robust standard errors in parentheses. IV is an instrumental variables specification. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4.14. Influence of Wage Risk and its Correlation with Asset Returns on the Financial Portfolio Share: 5 Years Rolling Average

	(1) FE	(2) IVFE	(3) FE	(4) IVFE
V_{it}	-0.0302***	-0.0909**	-0.0302***	-0.0908**
	(0.0109)	(0.0417)	(0.0109)	(0.0417)
C_{it}			-0.0021	-0.0014
			(0.0058)	(0.0160)
age	-0.0300***	-0.0318***	-0.0300***	-0.0318***
	(0.0037)	(0.0039)	(0.0037)	(0.0039)
age^2	0.0003***	0.0004***	0.0003***	0.0004***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
home ownership	-0.0494***	-0.0484***	-0.0495***	-0.0484***
	(0.0121)	(0.0121)	(0.0121)	(0.0121)
log wealth	-0.0167***	-0.0170***	-0.0167***	-0.0170***
	(0.0044)	(0.0044)	(0.0044)	(0.0044)
constant	1.1824***	1.2391***	1.1819***	1.2387***
	(0.0852)	(0.0941)	(0.0852)	(0.0943)
Obs.	24,080	24,080	24,080	24,080
Fstat		66.95		66.95

Note. Compiled by the authors using SOEP v35. Robust standard errors in parentheses. IVFE represents the instrumental variables specification. Log wealth refers to the log of net wealth. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4.15. Influence of Wage Risk and its Correlation with Asset Returns on the Financial Portfolio Share: Excluding Self-Employed Individuals

	(1) FE	(2) IVFE	(3) FE	(4) IVFE
V_{it}	-0.0266**	-0.0817*	-0.0266**	-0.0817*
	(0.0115)	(0.0457)	(0.0115)	(0.0458)
C_{it}			-0.0004	-0.0001
			(0.0052)	(0.0139)
age	-0.0297***	-0.0313***	-0.0297***	-0.0313***
	(0.0039)	(0.0041)	(0.0039)	(0.0041)
age^2	0.0003***	0.0004***	0.0003***	0.0004***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
home ownership	-0.0545***	-0.0536***	-0.0545***	-0.0536***
	(0.0130)	(0.0130)	(0.0130)	(0.0130)
log wealth	-0.0150***	-0.0154***	-0.0150***	-0.0154***
	(0.0047)	(0.0047)	(0.0047)	(0.0047)
constant	1.1631***	1.2139***	1.1632***	1.2139***
	(0.0888)	(0.0987)	(0.0888)	(0.0987)
Obs .	22,741	22,741	22,741	22,741
Fstat		56.14		56.14

Note. Compiled by the authors using SOEP v35. Robust standard errors in parentheses. IVFE represents the instrumental variables specification. Log wealth refers to the log of net wealth.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4.16. Influence of Wage Risk and its Correlation with Asset Returns on the Financial Portfolio Share: Including Housing Wealth in the Financial Share Definition.

	(1) FE	(2) IVFE	(3) FE	(4) IVFE
V_{it}	-0.0141*	-0.0211	-0.0141*	-0.0213
	(0.0080)	(0.0241)	(0.0080)	(0.0241)
C_{it}			0.0046	0.0021
			(0.0037)	(0.0098)
age	-0.0208***	-0.0210***	-0.0207***	-0.0209***
	(0.0031)	(0.0032)	(0.0031)	(0.0032)
age^2	0.0002***	0.0002***	0.0002***	0.0002***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
home ownership	-0.1659***	-0.1658***	-0.1658***	-0.1657***
	(0.0115)	(0.0115)	(0.0115)	(0.0115)
log wealth	-0.0331***	-0.0331***	-0.0331***	-0.0331***
	(0.0040)	(0.0040)	(0.0040)	(0.0040)
constant	1.1096***	1.1158***	1.1092***	1.1158***
	(0.0743)	(0.0778)	(0.0743)	(0.0778)
Obs.	24,080	24,080	24,080	24,080
Fstat		79.45		79.45

Note. Compiled by the authors using SOEP v35. Robust standard errors in parentheses. IVFE represents the instrumental variables specification. Log wealth refers to the log of net wealth.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4.17. Influence of Wage Risk and its Correlation with Asset Returns on the Financial Portfolio Share: Excluding Civil Servants

	(1) FE	(2) IVFE	(3) FE	(4) IVFE
V_{it}	-0.0255**	-0.1120***	-0.0256**	-0.1119***
	(0.0104)	(0.0378)	(0.0104)	(0.0381)
C_{it}			0.0014	-0.0005
			(0.0060)	(0.0160)
age	-0.0268***	-0.0289***	-0.0268***	-0.0289***
	(0.0045)	(0.0047)	(0.0045)	(0.0047)
age^2	0.0003***	0.0003***	0.0003***	0.0003***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
home ownership	-0.0638***	-0.0626***	-0.0638***	-0.0626***
	(0.0143)	(0.0143)	(0.0143)	(0.0143)
log wealth	-0.0179***	-0.0179***	-0.0179***	-0.0179***
	(0.0051)	(0.0051)	(0.0051)	(0.0051)
constant	1.1175***	1.1844***	1.1173***	1.1844***
	(0.1030)	(0.1078)	(0.1030)	(0.1078)
Obs .	17,611	17,611	17,611	17,611
Fstat		56.27		56.27

Note. Compiled by the authors using SOEP v35. Robust standard errors in parentheses. IVFE represents the instrumental variables specification. Log wealth refers to the log of net wealth.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4.18. Jackknifed IV Groups: Regression Estimates

Variable	mean	sd	p25	p50	p75	p90
V_{it}	-0.0846	0.0007	-0.0848	-0.0846	-0.0845	-0.0841
C_{it}	0.0029	0.0003	0.0028	0.0029	0.0030	0.0032

Note. Compiled by the authors using SOEP v35. Table provides the beta coefficients of the instrumented wage risk and correlation from the baseline IVFE regressions after jackknifing the IV groups successively.

List of Tables

List of Figures

Bibliography

Acemoglu, D. and Autor, D. (2011). Skills, Tasks and Technologies: Implications for Employment and Earnings. In *Handbook of Labor Economics*, volume 4, pages 1043–1171. Elsevier.

Acemoglu, D. and Restrepo, P. (2017). Robots and Jobs: Evidence from US Labor Markets. *NBER Working Paper*, (w23285).

Adler, M. and Dumas, B. (1983). International Portfolio Choice and Corporation Finance: A Synthesis. *The Journal of Finance*, 38(3):925–984.

Aedo, C., Hentschel, J., Luque, J., and Moreno, M. (2013). From Occupations to Embedded Skills: A Cross-Country Comparison. *World Bank Policy Research Working Paper*.

Andreoni, J. (1989). Giving with Impure Altruism: Applications to Charity and Ricardian Equivalence. *Journal of Political Economy*, 97(6):1447–1458.

ASFA (2021). Superannuation Statistics. `https://www.superannuation.asn.au/resources/superannuation-statistics`. Accessed: 2021-05-10.

Atkinson, A. B. (1971). The Distribution of Wealth and the Individual Life-Cycle. *Oxford Economic Papers*, 23(2):239–254.

Attanasio, O. P. and Brugiavini, A. (2003). Social Security and Households' Saving. *the Quarterly Journal of Economics*, 118(3):1075–1119.

Australian Bureau of Statistics (2021a). Household Income and Wealth, Australia. `https://www.abs.gov.au/statistics/economy/finance/household-income-and-wealth-australia`. Accessed: 2021-05-10.

Australian Bureau of Statistics (2021b). Life Tables 2002-2018. `https://www.abs.gov.au/statistics/people/population/life-tables/latest-release`. Accessed: 2021-01-27.

Australian Bureau of Statistics (2021c). Marriages and Divorces, Australia, 2019 . `https://www.abs.gov.au/statistics/people/people-and-communities/marriages-and-divorces-australia/latest-release`. Accessed: 2021-04-20.

Australian Bureau of Statistics (2021d). Retirement and Retirement Intentions, Australia, 2018-19 . `https://www.abs.gov.au/ausstats/abs@.nsf/mf/6238.0`. Accessed: 2021-05-10.

Australian Government (1992). Superannuation Guarantee (Administration) Act 1992. https://www.legislation.gov.au/Details/C2017C00230. Accessed: 2021-05-10.

Australian Government (2009). Social Security and Other Legislation Amendment (Pension Reform and Other 2009 Budget Measures) Bill 2009. https://www.legislation.gov.au/Details/C2009A00060. Accessed: 2021-05-10.

Australian Government (2018). A Guide to Australian Government Payments 2002-2018. Accessed: 2021-05-10.

Australian Government (2021a). Deeming. https://www.servicesaustralia.gov.au/individuals/topics/deeming/29656. Accessed: 2021-05-10.

Australian Government (2021b). Social Security Guide: Overview of Deeming. https://guides.dss.gov.au/guide-social-security-law/4/4/1/10/. Accessed: 2021-05-10.

Australian Prudential Regulation Authority (APRA) (2021). Annual Superannuation Bulletin 2019. https://www.apra.gov.au/annual-superannuation-bulletin. Accessed: 2021-05-10.

Australian Tax Office (2021). Individual Income Tax Rates for Prior Years. https://www.ato.gov.au/Rates. Accessed: 2021-05-10.

Autor, D. H. (2019). Work of the Past, Work of the Future. *AEA Papers and Proceedings*, 109:1–32.

Autor, D. H. and Dorn, D. (2013). The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market. *American Economic Review*, 103(5):1553–97.

Autor, D. H., Levy, F., and Murnane, R. J. (2003). The Skill Content of Recent Technological Change: An Empirical Exploration. *The Quarterly Journal of Economics*, 118(4):1279–1333.

Bach, L., Calvet, L. E., and Sodini, P. (2018). Rich pickings? Risk, Return, and Skill in the Portfolios of the Wealthy. *HEC Paris Research Paper No. FIN-2016-1126*, pages 16–03.

Bagchi, S. and Svejnar, J. (2015). Does Wealth Inequality Matter for Growth? The Effect of Billionaire Wealth, Income Distribution, and Poverty. *Journal of Comparative Economics*, 43(3):505–530.

Benartzi, S. and Thaler, R. (2001). Naive Diversification Strategies in Defined Contribution Saving Plans. *American Economic Review*, 91(1):79–98.

Benhabib, J. and Bisin, A. (2018). Skewed Wealth Distributions: Theory and Empirics. *Journal of Economic Literature*, 56(4):1261–91.

Benhabib, J., Bisin, A., and Luo, M. (2017). Earnings Inequality and Other Determinants of Wealth Inequality. *American Economic Review*, 107(5):593–97.

Benhabib, J., Bisin, A., and Zhu, S. (2011). The Distribution of Wealth and Fiscal Policy in Economies with Finitely Lived Agents. *Econometrica*, 79(1):123–157.

Berg, A., Ostry, J. D., Tsangarides, C. G., and Yakhshilikov, Y. (2018). Redistribution, Inequality, and Growth: New Evidence. *Journal of Economic Growth*, 23(3):259–305.

Bernheim, D. D. et al. (1998). Financial Illiteracy, Education, and Retirement Saving. Technical report, Wharton School Pension Research Council, University of Pennsylvania.

Betermier, S., Jansson, T., Parlour, C., and Walden, J. (2012). Hedging Labor Income Risk. *Journal of Financial Economics*, 105(3):622–639.

Bianchi, M. (2018). Financial Literacy and Portfolio Dynamics. *The Journal of Finance*, 73(2):831–859.

Bilias, Y., Georgarakos, D., and Haliassos, M. (2005). Equity Culture and the Distribution of Wealth. Technical report, CFS Working Paper.

Blau, F. D. and Kahn, L. M. (2006). The US Gender Pay Gap in the 1990s: Slowing Convergence. *ILR Review*, 60(1):45–66.

Blundell, R., Duncan, A., and Meghir, C. (1998). Estimating Labor Supply Responses Using Tax Reforms. *Econometrica*, 66(4):827–861.

Blundell, R., Pistaferri, L., and Saporta-Eksten, I. (2016). Consumption Inequality and Family Labor Supply. *American Economic Review*, 106(2):387–435.

Bönke, T., Grabka, M. M., Schröder, C., and Wolff, E. N. (2019). A Head-to-Head Comparison of Augmented Wealth in Germany and the United States. *The Scandinavian Journal of Economics*, 122(3):1140–1180.

Böhm, M. J. (2020). The Price of Polarization: Estimating Task Prices Under Routine-Biased Technical Change. *Quantitative Economics*, 11(2):761–799.

Böhm, M. J., Gaudecker, H.-M. v., and Schran, F. (2019). Occupation, Growth, Skill Prices, and Wage Inequalityy. *CESifo Working Paper*.

Bonin, H., Dohmen, T., Falk, A., Huffman, D., and Sunde, U. (2007). Cross-Sectional Earnings Risk and Occupational Sorting: The Role of Risk Attitudes. *Labour Economics*, 14(6):926–937.

Bottazzi, L., Pesenti, P., and van Wincoop, E. (1996). Wages, Profits and the International Portfolio Puzzle. *European Economic Review*, 40(2):219 – 254.

Bourguignon, F. (1979). Decomposable Income Inequality Measures. *Econometrica*, pages 901–920.

Brandolini, A. (2007). Measurement of Income Distribution in Supranational Entities: The Case of the European Union. *Bank of Italy Temi di Discussione Working Paper*, (623).

Breunig, C., Huck, S., Schmidt, T., and Weizsäcker, G. (2019). The Standard Portfolio Choice Problem in Germany. Discussion Paper 171, München und Berlin.

Brown, S., Ortiz, A., and Taylor, K. (2006). Educational Attainment and Risk Preference. *Working Paper*.

Buraschi, A., Porchia, P., and Trojani, F. (2010). Correlation Risk and Optimal Portfolio Choice. *The Journal of Finance*, 65(1):393–420.

Burns, S. K. and Ziliak, J. P. (2017). Identifying the Elasticity of Taxable Income. *The Economic Journal*, 127(600):297–329.

Cagetti, M. (2003). Wealth Accumulation over the Life Cycle and Precautionary Savings. *Journal of Business & Economic Statistics*, 21(3):339–353.

Caliendo, M., Fedorets, A., Preuss, M., Schroeder, C., and Wittbrodt, L. (2018). The Short-Term Distributional Effects of the German Minimum Wage Reform. *IZA Discussion Paper*, 11246.

Calvet, L. E., Campbell, J. Y., and Sodini, P. (2007a). Down or out: Assessing the Welfare Costs of Household Investment Mistakes. *Journal of Political Economy*, 115(5):707–747.

Calvet, L. E., Campbell, J. Y., and Sodini, P. (2007b). Down or Out: Assessing the Welfare Costs of Household Investment Mistakes - Appendix.

Cardak, B. A. and Wilkins, R. (2009). The Determinants of Household Risky Asset Holdings: Australian Evidence on Background Risk and Other Factors. *Journal of Banking & Finance*, 33(5):850–860.

Chakravarty, S. R. et al. (2015). Inequality, Polarization and Conflict. *Economic Studies in Inequality, Social Exclusion and Well-Being*, 12.

Chetty, R., Sándor, L., and Szeidl, A. (2017). The Effect of Housing on Portfolio Choice. *The Journal of Finance*, 72(3):1171–1212.

Cocco, J. F., Gomes, F. J., and Maenhout, P. J. (2005). Consumption and Portfolio Choice over the Life Cycle. *The Review of Financial Studies*, 18(2):491–533.

Cochrane, J. H. (2009). *Asset Pricing*. Princeton University Press.

Coelli, M. and Borland, J. (2016). Job Polarisation and Earnings Inequality in Australia. *Economic Record*, 92(296):1–27.

Davies, J. B., Fortin, N. M., and Lemieux, T. (2017). Wealth Inequality: Theory, Measurement and Decomposition. *Canadian Journal of Economics/Revue canadienne d'économique*, 50(5):1224–1261.

Davies, J. B., Sandström, S., Shorrocks, A., and Wolff, E. N. (2010). The Level and Distribution of Global Household Wealth. *The Economic Journal*, 121(551):223–254.

Davies, J. B. and Shorrocks, A. F. (2000). The Distribution of Wealth. *Handbook of income distribution*, 1:605–675.

Davis, S. J. and Willen, P. (2014). Occupation-Level Income Shocks and Asset Returns: Their Covariance and Implications for Portfolio Choice. *The Quarterly Journal of Finance*, 3(03n04):1350011.

De La Rica, S. and Gortazar, L. (2016). Differences in Job De-Routinization in OECD Countries: Evidence from PIAAC. *IZA Discussion Papers*, (9736).

De Nardi, M. (2004). Wealth Inequality and Intergenerational Links. *The Review of Economic Studies*, 71(3):743–768.

De Nardi, M. and Fella, G. (2017). Saving and Wealth Inequality. *Review of Economic Dynamics*, 26:280–300.

DiNardo, J., Fortin, N. M., and Lemieux, T. (1996). Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach. *Econometrica*, (64):1001–44.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., and Wagner, G. G. (2011). Individual Risk Attitudes: Measurement, Determinants, and Behavioral Consequences. *Journal of the European Economic Association*, 9(3):522–550.

Dustmann, C., Ludsteck, J., and Schönberg, U. (2009). Revisiting the German Wage Structure. *The Quarterly Journal of Economics*, 124(2):843–881.

Eckerstorfer, P., Halak, J., Kapeller, J., Schütz, B., Springholz, F., and Wildauer, R. (2016). Correcting for the Missing Rich: An Application to Wealth Survey Data. *Review of Income and Wealth*, 62(4):605–627.

Eeckhoudt, L., Gollier, C., and Schlesinger, H. (1996). Changes in Background Risk and Risk Taking Behavior. *Econometrica*, pages 683–689.

Efron, B. (1979). Bootstrap Methods: Another Look at the Jackknife. *Annals of Statistics*, 7(1):1–26.

Esteban, J.-M. and Ray, D. (1994). On the Measurement of Polarization. *Econometrica*, pages 819–851.

Fagereng, A., Gottlieb, C., and Guiso, L. (2017a). Asset Market Participation and Portfolio Choice over the Life-Cycle. *The Journal of Finance*, 72(2):705–750.

Fagereng, A., Guiso, L., Malacrino, D., and Pistaferri, L. (2016). Heterogeneity in Returns to Wealth and the Measurement of Wealth Inequality. *American Economic Review*, 106(5):651–55.

Fagereng, A., Guiso, L., and Pistaferri, L. (2017b). Firm-Related Risk and Precautionary Saving Response. *American Economic Review*, 107(5):393–97.

Fagereng, A., Guiso, L., and Pistaferri, L. (2018). Portfolio Choices, Firm Shocks, and Uninsurable Wage Risk. *The Review of Economic Studies*, 85(1):437–474.

Firpo, S., Fortin, N. M., and Lemieux, T. (2009). Unconditional Quantile Regressions. *Econometrica*, 77(3):953–973.

Firpo, S., Fortin, N. M., and Lemieux, T. (2011). Occupational Tasks and Changes in the Wage Structure. *IZA Discussion Papers*, (5542).

Firpo, S. P., Fortin, N. M., and Lemieux, T. (2018). Decomposing Wage Distributions Using Recentered Influence Function Regressions. *Econometrics*, 6(2):28.

Fortin, N., Lemieux, T., and Firpo, S. (2011). Decomposition Methods in Economics. In *Handbook of Labor Economics*, volume 4, pages 1–102. Elsevier.

Fossen, F. M. (2011). The Private Equity Premium Puzzle Revisited — New Evidence on the Role of Heterogeneous Risk Attitudes. *Economica*, 78(312):656–675.

Fossen, F. M., König, J., and Schröder, C. (2020). Risk Preference and Entrepreneurial Investment at the Top of the Wealth Distribution. *Working Paper - Institute of Labor Economics (IZA)*.

Frick, J. R. and Headey, B. W. (2009). Living Standards in Retirement: Accepted International Comparisons are Misleading. *Journal of Contextual Economics*, 129(2):309.

Gelbach, J. B. (2016). When Do Covariates Matter? And Which Ones, and How Much? *Journal of Labor Economics*, 34(2):509–543.

Goebel, J., Grabka, M. M., Liebig, S., Kroh, M., Richter, D., Schröder, C., and Schupp, J. (2019). The German Socio-Economic Panel (SOEP). *Jahrbücher für Nationalökonomie und Statistik*, 239(2):345–360.

Goetzmann, W. N. and Kumar, A. (2008). Equity Portfolio Diversification. *Review of Finance*, 12(3):433–463.

Goos, M. and Manning, A. (2007). Lousy and Lovely Jobs: The Rising Polarization of Work in Britain. *The Review of Economics and Statistics*, 89(1):118–133.

Goos, M., Manning, A., and Salomons, A. (2014). Explaining Job Polarization: Routine-Biased Technological Change and Offshoring. *American Economic Review*, 104(8):2509–26.

Graham, J. R., Harvey, C. R., and Huang, H. (2009). Investor Competence, Trading Frequency, and Home Bias. *Management Science*, 55(7):1094–1106.

Green, D. A. and Sand, B. M. (2015). Has the Canadian Labour Market Polarized? *Canadian Journal of Economics/Revue Canadienne d'Économique*, 48(2):612–646.

Grinblatt, M., Keloharju, M., and Linnainmaa, J. (2011). IQ and Stock Market Participation. *The Journal of Finance*, 66(6):2121–2164.

Guiso, L., Haliassos, M., and Jappelli, T. (2003). Household Stockholding in Europe: Where Do We Stand and Where Do We Go? *Economic Policy*, 18(36):123–170.

Guiso, L. and Paiella, M. (2008). Risk Aversion, Wealth, and Background Risk. *Journal of the European Economic Association*, 6(6):1109–1150.

Guiso, L., Sapienza, P., and Zingales, L. (2008). Trusting the Stock Market. *The Journal of Finance*, 63(6):2557–2600.

Guiso, L. and Sodini, P. (2013). Household Finance: An Emerging Field. In *Handbook of the Economics of Finance*, volume 2, pages 1397–1532. Elsevier.

Guvenen, F. (2007). Learning Your Earning: Are Labor Income Shocks Really Very Persistent? *American Economic Review*, 97(3):687–712.

Hagenaars, A. J., De Vos, K., Asghar Zaidi, M., et al. (1994). Poverty Statistics in the Late 1980s: Research Based on Micro-data.

Harding, A. (2002). Trends in Income and Wealth Inequality in Australia. In *Presentation to the Economic and Social Outlook Conference: Towards Opportunity and Prosperity*, pages 4–5.

Hardy, W., Keister, R., and Lewandowski, P. (2018). Educational Upgrading, Structural Change and the Task Composition of Jobs in Europe. *Economics of Transition*, 26(2):201–231.

Harmer, J. (2008). *Pension Review Background Paper*. Department of Families, Housing, Community Services and Indigenous Affairs.

Headey, B., Marks, G., and Wooden, M. (2005). The Structure and Distribution of Household Wealth in Australia. *Australian Economic Review*, 38(2):159–175.

Heaton, J. and Lucas, D. (1999). Stock Prices and Fundamentals. *NBER Macroeconomics Annual*, 14:213–242.

Heaton, J. and Lucas, D. (2000). Portfolio Choice and Asset Prices: The Importance of Entrepreneurial Risk. *The Journal of Finance*, 55(3):1163–1198.

Hilgert, M. A., Hogarth, J. M., and Beverly, S. G. (2003). Household Financial Management: The Connection Between Knowledge and Behavior. *Fed. Res. Bull.*, 89:309.

Hunt, J. and Nunn, R. (2019). Is Employment Polarization Informative About Wage Inequality and Is Employment Really Polarizing? *IZA Discussion Papers*, (12472).

Hurst, E., Lusardi, A., Kennickell, A., and Torralba, F. (2010). The Importance of Business Owners in Assessing the Size of Precautionary Savings. *The Review of Economics and Statistics*, 92(1):61–69.

Islam, M. R. and McGillivray, M. (2019). Wealth Inequality, Governance and Economic Growth. *Economic Modelling*.

Jordà, Ò., Knoll, K., Kuvshinov, D., Schularick, M., and Taylor, A. M. (2019). The Rate of Return on Everything, 1870–2015. *The Quarterly Journal of Economics*, 134(3):1225–1298.

Kaplan, G. (2012). Inequality and the Life Cycle. *Quantitative Economics*, 3(3):471–525.

Kaplan, G., La Cava, G., and Stone, T. (2018a). Household Economic Inequality in Australia. *Economic Record*, 94(305):117–134.

Kaplan, G., Moll, B., and Violante, G. L. (2018b). Monetary policy according to HANK. *American Economic Review*, 108(3):697–743.

Kennickell, A. B. (2008). The Role of Over-sampling of the Wealthy in the Survey of Consumer Finances. *Irving Fisher Committee Bulletin*, 28:403–408.

Kennickell, A. B. and McManus, D. A. (1993). Sampling for Household Financial Characteristics Using Frame Information on Past Income. In *Proceedings of the Section on Survey Research Methods*.

Kimball, M. S. and Shumway, T. (2010). Investor Sophistication and the Home Bias, Diversification, and Employer Stock Puzzles. *Working Paper*.

Knoll, K., Schularick, M., and Steger, T. (2017). No Price Like Home: Global House Prices, 1870-2012. *American Economic Review*, 107(2):331–53.

König, J., Schröder, C., and Wolff, E. N. (2020). *Wealth Inequalities*, pages 1–38. Springer International Publishing.

Kuhn, M., Schularick, M., and Steins, U. I. (2017). Income and Wealth Inequality in America. *Journal of Political Economy*.

Kuhn, U. (2020). Augmented Wealth in Switzerland: The Influence of Pension Wealth on Wealth Inequality. *Swiss Journal of Economics and Statistics*, 156(1):1–16.

Lerman, R. I. and Yitzhaki, S. (1985). Income Inequality Effects by Income Source: A New Approach and Applications to the United States. *The Review of Economics and Statistics*, pages 151–156.

Levy, H. and Levy, M. (2014). The Home Bias is Here to Stay. *Journal of Banking & Finance*, 47:29–40.

Lintner, J. (1965). Security Prices, Risk, and Maximal Gains From Diversification. *The Journal of Finance*, 20(4):587–615.

Longmuir, M. and König, J. (2019). Investment losses and inequality. *Available at SSRN 3446783*.

Lusardi, A., Michaud, P.-C., and Mitchell, O. S. (2017). Optimal Financial Knowledge and Wealth Inequality. *Journal of Political Economy*, 125(2):431–477.

Lusardi, A. and Mitchell, O. S. (2007). Baby Boomer Retirement Security: The Roles of Planning, Financial Literacy, and Housing Wealth. *Journal of Monetary Economics*, 54(1):205–224.

Lusardi, A. and Mitchell, O. S. (2011a). Financial Literacy and Planning: Implications for Retirement Wellbeing. Technical report, National Bureau of Economic Research.

Lusardi, A. and Mitchell, O. S. (2011b). Financial Literacy Around the World: An Overview. *Journal of Pension Economics & Finance*, 10(4):497–508.

Lusardi, A. and Mitchell, O. S. (2014). The Economic Importance of Financial Literacy: Theory and Evidence. *Journal of Economic Literature*, 52(1):5–44.

Machin, S. (2010). Changes in UK Wage Inequality over the Last Forty Years. *The Labour Market in Winter: State of Working Britain*.

Mahutga, M. C., Curran, M., and Roberts, A. (2018). Job Tasks and the Comparative Structure of Income and Employment: Routine Task Intensity and Offshorability for the LIS. *International Journal of Comparative Sociology*, 59(2):81–109.

Markowitz, H. (1952). Portfolio Selection. *The Journal of Finance*, 7(1):77–91.

Massari, R., Naticchioni, P., and Ragusa, G. (2014). Unconditional and Conditional Wage Polarization in Europe. *IZA Discussion Papers*, (8465).

Merton, R. (1969). Lifetime Portfolio Selection Under Uncertainty: The Continuous-Time Case. *The Review of Economics and Statistics*, pages 247–257.

Merton, R. (1975). Optimum Consumption and Portfolio Rules in a Continuous-Time Model. In *Stochastic Optimization Models in Finance*, pages 621–661. Elsevier.

Morningstar (2015). Currency Hedge Index Methodology. Accessed: 2021-04-30.

Mossin, J. (1966). Equilibrium in a Capital Asset Market. *Econometrica*, pages 768–783.

Necker, S. and Ziegelmeyer, M. (2016). Household Risk Taking after the Financial Crisis. *The Quarterly Review of Economics and Finance*, 59:141–160.

Oaxaca, R. L. and Ransom, M. R. (1999). Identification in Detailed Wage Decompositions. *Review of Economics and Statistics*, 81(1):154–157.

OECD (2021). Purchasing Power Parities (PPP) (Indicator). doi: 10.1787/1290ee5a-en. Accessed: 2021-01-27.

Papke, L. E. and Wooldridge, J. M. (1996). Econometric Methods for Fractional Response Variables with an Application to 401 (k) Plan Participation Rates. *Journal of Applied Econometrics*, 11(6):619–632.

Piketty, T. (2018). Brahmin Left vs Merchant Right: Rising Inequality and the Changing Structure of Political Conflict. *mimeo, Paris School of Economics*.

Rios-Avila, F. (2020). Recentered Influence Functions (RIFs) in Stata: RIF Regression and RIF Decomposition. *The Stata Journal*, 20(1):51–94.

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

Schröder, C., Bartels, C., Grabka, M. M., König, J., Kroh, M., and Siegers, R. (2020). A Novel Sampling Strategy for Surveying High Net-Worth Individuals — A Pretest Application Using the Socio-Economic Panel. *Review of Income and Wealth*, online first.

Schröder, C., Bartels, C., Göbler, K., Grabka, M. M., König, J., Siegers, R., Zinn, S., et al. (2020a). Improving the Coverage of the Top-Wealth Population in the Socio-Economic Panel (SOEP). Technical report, DIW Berlin, The German Socio-Economic Panel (SOEP).

Schröder, C., Golan, Y., and Yitzhaki, S. (2014). Inequality and the Time Structure of Earnings: Evidence from Germany. *The Journal of Economic Inequality*, 12(3):393–409.

Schröder, C., König, J., Fedorets, A., Goebel, J., Grabka, M. M., Lüthen, H., Metzing, M., Schikora, F., and Liebig, S. (2020b). The Economic Research Potentials of the German Socio-Economic Panel Study. *German Economic Review*, 21(3):335–371.

Sercu, P. (1980). A Generalisation of the International Asset Pricing Model. *Revue de l'Association Frangaise de Finance*, (1):91–135.

Sharpe, W. F. (1964). Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk. *The Journal of Finance*, 19(3):425–442.

Sila, U. and Dugain, V. (2019). Income, Wealth and Earnings Inequality in Australia: Evidence from the HILDA Survey. *OECD iLibrary*.

Solnik, B. H. (1974). An Equilibrium Model of the International Capital Market. *Journal of Economic Theory*, 8(4):500–524.

Stiglitz, J. E. (2012). *The Price of Inequality: How Today's Divided Society Endangers our Future*. WW Norton & Company.

Superannuation (Objective) Bill (2016). `https://www.legislation.gov.au/Details/C2016B00182`. Accessed: 2021-05-10.

Taber, C. R. and Roys, N. A. (2019). Skill Prices, Occupations, and Changes in the Wage Structure for Low Skilled Men. *NBER Working Paper*, (26453).

The World Bank (2021). Consumer Prices. `https://data.worldbank.org/indicator/FP.CPI.TOTL.ZG?locations=AU`. Accessed: 2021-01-27.

Tobin, J. (1958). Liquidity Preference as Behavior towards Risk. *The Review of Economic Studies*, 25(2):65–86.

Treynor, J. L. (1961). Market Value, Time, and Risk. *Unpublished Script: 'Rough Draft'*.

Tversky, A. and Kahneman, D. (1971). Belief in the Law of Small Numbers. *Psychological bulletin*, 76(2):105.

Vissing-Jorgensen, A. (2002). Towards an Explanation of Household Portfolio Choice Heterogeneity: Nonfinancial Income and Participation Cost Structures. Technical report, National Bureau of Economic Research.

Vivarelli, M. (2014). Innovation, Employment and Skills in Advanced and Developing Countries: A Survey of Economic Literature. *Journal of Economic Issues*, 48(1):123–154.

Von Gaudecker, H. M. (2015). How Does Household Portfolio Diversification Vary with Financial Literacy and Financial Advice? *The Journal of Finance*, 70(2):489–507.

Watson, N. and Wooden, M. (2002). The Household, Income, Labour Dynamics in Australia (HILDA) Survey: Wave 1 Survey Methodology.

Weidner, J. (2014). The Wealthy Hand-to-Mouth. *Brookings Papers on Economic Activity*, 1:77–153.

Wilkins, R. (2016). The 11th Annual Statistical Report of the HILDA Survey.

Wilkins, R., Botha, F., Vera-Toscano, E., and Wooden, M. (2020). The Household, Income and Labour Dynamics in Australia Survey: Selected Findings from Waves 1 to 18.

Wolff, E. N. (1996). International Comparisons of Wealth Inequality. *Review of Income and Wealth*, 42(4):433–451.

Wolff, E. N. (2015a). *Inheriting Wealth in America: Future Boom or Bust?* Oxford University Press, USA.

Wolff, E. N. (2015b). US Pensions in the 2000s: The Lost Decade? *Review of Income and Wealth*, 61(4):599–629.

Yun, M.-S. (2008). Identification Problem and Detailed Oaxaca Decomposition: A General Solution and Inference. *Journal of Economic and Social Measurement*, 33(1):27–38.

English Summary

This dissertation consists of four empirical chapters which contribute to the fields of inequality, household finances, and labor economics.

The first chapter analysis how investment behavior, especially investment inefficiencies, contribute to wealth inequality in five European countries. Systematic differences along the wealth distribution in investment performance will potentially have large consequences for the level and persistence of wealth inequality. These differences in performance are hard to measure except in a few, select countries with detailed information on household portfolios. In this first chapter, we use a modified version of the Global Capital Asset Pricing Model (GCAPM), which relies on classed household portfolio data to measure investment performance in five European countries, where previously no measure of investment performance could be computed. We verify the accuracy of the modified GCAPM using Dutch survey data, which contains unclassed portfolio data enabling direct comparison with the regular GCAPM. In all countries households with less wealth exhibit lower investment performance, even after risk-adjustment. Further, we show that raising investment performance creates large efficiency gains, however, households below the median do not benefit from them.

The second chapter empirically investigates the interaction between job de-routinization and earnings inequality in a broad, cross-country analysis. Routine-Biased Technological Change hypothesis (RBTC) by Acemoglu and Autor (2011) suggests that automation processes have substituted workers operating middle-skilled routine tasks. Consequently, the relative demand of complementary non-routine occupations, i.e., low-skilled service and high-skilled abstract jobs, has increased. These changes in the composition of the labor force imply a polarization of jobs along skills distribution. An aspect of high socio-economical and political relevance is its distributional implications. In this chapter, we quantify polarization of jobs and its implication for earnings distributions using a novel dataset of 35 countries around the globe. We find strong evidence for job polarization in most countries but no clear-cut distributional consequences. We show that this weak link stems from variation *within*, rather than *between*, occupational classes, and from heterogeneous de-routinization effects along the earnings distribution.

The third chapter scrutinizes augmented wealth, i.e. the sum of net worth and pension wealth, in Australia between 2002 and 2018. The omission of pension wealth potentially distorts the international comparison of wealth distributions. Private pension wealth is often included in households' wealth portfolios, while public pension claims are not. Augmented wealth resolves this limitation by including the present value of social security pension wealth. This chapter provides a detailed analysis of augmented wealth in Australia between 2002 and 2018, capturing the

establishment of the compulsory private pension scheme, Superannuation, which was introduced in 1992. Moreover, I depict the interaction of Superannuation with the public scheme Age Pension and how that affects the overall wealth distribution. Augmented wealth in Australia is found to be less equally distributed than wealth in Germany or Switzerland, but more equally than in the United States.

The fourth chapter investigates the causal effect of wage risk on individual portfolio choice in Germany. From standard portfolio-choice theory it is well-understood that background risk, overwhelmingly due to wage risk, is one of the central determinants of individuals' portfolio composition: higher background risk reduces risky investments. However, if background risk is negatively correlated with financial market risk, higher background risk implies more risky investment. We quantify the influence of wage risk on German investors' financial portfolio shares and find that an increase of the residual variance of wages by one standard deviation implies a reduction of the financial portfolio share by 3 percentage points. We do not find that the correlation of wage risk with financial market risk has a significant impact on portfolio choice and provide evidence that this may be due to a lack of salience.

Deutsche Zusammenfassung

Diese Dissertation besteht aus vier empirischen Kapiteln, die zu den Bereichen Ungleichheit, Haushaltsfinanzen und Arbeitsökonomie beitragen.

Das erste Kapitel analysiert, wie das Investitionsverhalten, insbesondere die Ineffizienz von Investitionen, zur Vermögensungleichheit in fünf europäischen Ländern beiträgt. Systematische Unterschiede der Anlageeffizienz entlang der Vermögensverteilung haben möglicherweise große Konsequenzen für das Ausmaß und die Persistenz der Vermögensungleichheit. Diese Unterschiede sind schwer zu messen und bisher in einigen wenigen Ländern mit detaillierten Informationen zu Haushaltsportfolios umsetzbar. In diesem Kapitel verwenden wir eine modifizierte Version des Global Capital Asset Pricing Model (GCAPM), das sich auf klassifizierte Daten von Haushaltsportfolios stützt, um das Anlageverhalten in fünf europäischen Ländern zu messen, in denen dies bisher nicht durchgeführt werden konnte. Wir überprüfen die Genauigkeit des modifizierten GCAPM anhand niederländischer Umfragedaten, welche detaillierte Portfoliodaten enthalten und somit einen direkten Vergleich mit dem regulären GCAPM ermöglichen. In allen Ländern weisen Haushalte mit geringerem Vermögen auch nach Risikoanpassung eine geringere Anlageeffizienz auf. Darüber hinaus zeigen wir, dass eine Steigerung der Anlageeffizienz zu Vermögenszuwächsen führt, Haushalte unter dem Median jedoch nicht davon profitieren.

Das zweite Kapitel untersucht empirisch die Wechselwirkung zwischen der De-routinisierung von Berufsfeldern und der Einkommensungleichheit in einer umfassenden, länderübergreifenden Analyse. Die Routine-Biased Technological Change-Hypothese (RBTC) von Acemoglu and Autor (2011) legt nahe, dass Arbeiter:innen, welche typischerweise Routineaufgaben ausführen, durch technische Automatisierungsprozesse ersetzt wurden. Infolgedessen hat die relative Nachfrage nach komplementären nicht-routinemäßigen Berufen, d.h. geringqualifizierte Berufe im Dienstleistungssektor und hochqualifizierte Berufe mit komplexen Aufgaben, zugenommen. Diese Veränderungen in der Zusammensetzung der Erwerbsbevölkerung implizieren eine Polarisierung der Arbeitsplätze entlang der Verteilung der (beruflichen) Fähigkeiten. Eine Debatte von hoher sozioökonomischer und politischer Relevanz sind die Verteilungswirkung dieser Prozesse. In diesem Kapitel quantifizieren wir die Polarisierung von Berufsfeldern und ihre Auswirkungen auf die Einkommensverteilung anhand eines neuartigen Datensatzes für 35 Länder auf der ganzen Welt. Wir finden in den meisten Ländern starke Anzeichen für eine Polarisierung, aber keine eindeutigen Verteilungsfolgen. Wir zeigen, dass diese unklare Verteilungswirkung auf Variationen *innerhalb*, anstelle von Variation *zwischen* Berufsklassen zurückzuführen ist. Darüber hinaus finden wir heterogene De-routinisierungseffekte entlang der Einkommensverteilung.

Das dritte Kapitel untersucht die *erweiterte Vermögensverteilung*, d.h die Summe aus Vermögen und Rentenvermögen, in Australien zwischen 2002 und 2018. Das Weglassen von Rentenvermögen verzerrt möglicherweise den internationalen Vergleich der Vermögensverteilungen. Privates Rentenvermögen wird häufig in die Vermögensportfolios der privaten Haushalte aufgenommen, öffentliche Rentenansprüche dagegen nicht. Die *erweiterte Vermögensverteilung* behebt diese Einschränkung, indem der Barwert der sozialen Rentenversicherung berücksichtigt wird. Dieses Kapitel enthält eine detaillierte Analyse der Vermögenszuwächse in Australien zwischen 2002 und 2018, wobei diese Zeitspanne die Etablierung des 1992 eingeführten, privaten Rentensystems "Superannuation" erfasst. Außerdem zeige ich die Wechselwirkung von "Superannuation" mit der sozialen Rentenversicherung "Age Pension" und deren Implikation für die gesamte Vermögensverteilung. Die *erweiterte Vermögensverteilung* in Australien ist weniger gleichmäßig verteilt als in Deutschland oder in der Schweiz, aber gleichmäßiger als in den USA.

Das vierte Kapitel untersucht die kausalen Auswirkungen des Lohnrisikos auf die Finanzinvestitionen deutscher Haushalte. Aus der Standardtheorie geht hervor, dass Lohnrisiko eine der zentralen Determinanten für die private Portfoliozusammensetzung ist: Ein höheres Lohnrisiko reduziert riskante Investitionen. Wenn das Lohnrisiko jedoch negativ mit dem Risiko am Finanzmarkt korreliert, impliziert ein höheres Lohnrisiko eine riskantere Investition. Wir quantifizieren den Einfluss des Lohnrisikos auf die Wahl der Finanzportfolios deutscher Anleger:innen und stellen fest, dass eine Erhöhung der Residualvarianz der Löhne um eine Standardabweichung eine Reduzierung des Finanzportfolios um 3 Prozentpunkte bedeutet. Wir stellen jedoch nicht fest, dass die Korrelation des Lohnrisikos mit dem Risiko am Finanzmarkt einen signifikanten Einfluss auf die Portfolioauswahl hat, und liefern Hinweise darauf, dass dies möglicherweise auf Salienz zurückzuführen ist.

Erklärung

Erklärung gemäß §4 Abs. 2

Hiermit erkläre ich, dass ich mich noch keinem Promotionsverfahren unterzogen oder um Zulassung zu einem solchen beworben habe, und die Dissertation in der gleichen oder einer anderen Fassung bzw. Überarbeitung einer anderen Fakultät, einem Prüfungsausschuss oder einem Fachvertreter an einer anderen Hochschule nicht bereits zur Überprüfung vorgelegen hat.

(Unterschrift, Ort, Datum)

Erklärung gemäß §10 Abs. 3

Ich habe meine Dissertation soweit im Folgenden nicht anders vermerkt selbständig verfasst habe.

Folgende Hilfsmittel wurde benutzt

- Statistik und Mathematik: Stata, Mathematica
- Schriftsatz und Formatierung: LaTeX, TeXstudio

(Unterschrift, Ort, Datum)