Inequality and Income Dynamics in Germany*

Moritz Drechsel-Grau Andreas Peichl Johannes Schmieder Kai D. Schmid

Hannes Walz Stefanie Wolter

February 26, 2022

Abstract

We provide a comprehensive analysis of income inequality and income dynamics for Germany over the last two decades. Combining personal income tax and social security data allows us – for the first time – to offer a complete picture of the distribution of annual earnings in Germany. We find that cross-sectional inequality rose until 2009 for men and women. After the Great Recession inequality continued to rise at a slower rate for men and fell slightly for women due to compression at the lower tail. We further document substantial gender differences in average earnings and inequality over the life-cycle. While for men earnings rise and inequality falls as they grow older, many women reduce working hours when starting a family such that average earnings fall and inequality increases. Men's earnings changes are on average smaller than women's but are substantially more affected by the business cycle. During the Great Recession, men's earnings losses become magnified and gains are attenuated. Apart from recession years, earnings changes are significantly right-skewed reflecting the good overall state of the German labor market and increasing labor supply. In the second part of the paper, we study the distribution of total income including incomes of self-employed, business owners, and landlords. We find that total inequality increased significantly more than earnings inequality. Regarding income dynamics, entrepreneurs' income changes are more dispersed, less skewed, less leptokurtic and less dependent on average past income than workers' income changes. Finally, we find that top income earners have become less likely to fall out of the top 1 and 0.1 percent.

JEL Classification: D31, E24, E31, J31

Keywords: Inequality; Income Dynamics; Mobility; Non-Labor Income

***Drechsel-Grau**: University of Zurich; **Peichl**: ifo Institute and LMU Munich; **Schmieder** (corresponding author): Boston University, johannes@bu.edu; **Schmid**: Heilbronn University of Applied Sciences; **Walz**: Friedrich-Alexander University Erlangen-Nürnberg (FAU) and Institute for Employment Research (IAB), Nuremberg; **Wolter**: Institute for Employment Research (IAB), Nuremberg. We thank the Research Data Lab of the German Federal Statistical Agency, especially Maurice Brand and Hannes Fauser, for steady support in accessing the data. Sophie Friebe provided excellent research assistance. We also thank Andreas Ganzer and Stefan Seth for providing imputed wages. Peichl gratefully acknowledges financial support from the German Science Foundation (DFG) via CRC TRR 190 (project number 280092119) and Walz from the joint doctoral program of the Institute of Employment Research (IAB) and Friedrich-Alexander University Erlangen-Nürnberg, GradAB.

1 Introduction

Since the beginning of the 21st century we have seen a renewed interest among economists to study and understand the structure and evolution of income inequality as well as the forces that shape it. Focusing solely on cross-sectional inequality leaves out many aspects of the distribution of welfare. A young worker entering the labor market might be perfectly content with a low starting salary at the bottom of the wage distribution if she may expect rapid wage growth and upward mobility over the coming years. Similarly, a worker with middle class income might experience significant uncertainty about his income in subsequent years, for example, due to uncertainty about bonus pay and hours or due to the risk of job displacement and subsequent earnings losses. This has motivated recent work in studying income inequality jointly with income dynamics to provide a more complete picture of the distribution of economic well-being in an economy (e.g. Guvenen et al., 2021b).

This paper provides a comprehensive analysis of income inequality and income dynamics for Germany over the last two decades. We combine two high quality administrative data sources for this analysis: personal income tax records from the Taxpayer Panel (TPP) as well as social security data from the Institute for Employment Research (IAB). Since each of these datasets has distinct advantages and weaknesses, combining the two allows us, for the first time, to offer a complete picture of the German income structure ranging from workers in marginal employment at the very bottom all the way to the highest paid CEOs at the top of the distribution. We begin by focusing on labor earnings, which is the main source of income for the vast majority and most easily compared across datasets and countries. We first describe the evolution of cross-sectional earnings inequality. We then show patterns in the distribution of earnings changes for workers at different stages of their lives and different parts of the earnings distribution. Finally, we relate to the recent literature on capitalists (see, e.g., Smith et al., 2019; Autor et al., 2020) and add self-employed, business owners and landlords to analyze in detail how this affects the picture of inequality and income dynamics.

Germany is a particularly interesting setting with a number of considerable changes in its labor market throughout the last two decades. After a period of high unemployment rates throughout the 1990s and early 2000s, the German labor market experienced a remarkable turnaround starting around 2005 with unemployment rates falling steadily from over 13% to less than 6% in 2018.[1] With this considerable decline in unemployment, the attention of policymakers and the public increasingly focused on income inequality which had risen dramatically during the 1990s and in the early 2000s (see, e.g., Fuchs-Schündeln et al., 2010). In the following years, pushes for higher wages, for example through more aggressive collective bargaining negotiations, could be observed.[2] And finally, for the first time in Germany, in 2015 a nation-wide minimum wage was introduced.

[1] There are several reasons for this improvement: Some economists credit the 2003-2006 labor market reforms ("Hartz reforms") and the short-time work program (Gehrke et al., 2019), while others claim that the combination of flexible collective bargaining institutions and restructuring of supply chains were the driving forces (e.g. Dustmann et al. 2014, Hoffmann and Lemieux 2016). (Macro)economic analyses of the Hartz reforms (e.g., Krebs and Scheffel, 2013; Launov and Wälde, 2013; Hartung et al., 2018; Bradley and Kügler, 2019; Hochmuth et al., 2021) show that the reforms did play an important role for the decline in unemployment, but other factors were also important.

[2] For example, nominal union wages in manufacturing increased by 2% on average annually from 1998–2006 and by 2.8% from 2007–2018 (3.1% from 2007–2009, i.e. before the Great Recession) – see Statistisches Bundesamt (2021b).

Our paper adds to the literature on the German wage structure and labor income inequality in several important ways. Prior to the 1990s, studies find stable (cross-sectional) wage inequality in West Germany (e.g., Biewen, 2000). Using survey data, Fuchs-Schündeln et al. (2010) show that inequality trended upwards for wages and market incomes since the mid-1990s (until about 2005, where their analysis stops). Using IAB data Antonczyk et al. (2018) show that the increase in earnings inequality was mostly restricted to the right tail of the distribution.[3] More recently papers have looked at the 2010s when labor market conditions were improving (e.g., Lochner et al., 2020). Using survey data, Biewen et al. (2019) show that the general rise in wage inequality became less steep (but did not stop) after 2005, while inequality in annual labor incomes did not increase further after 2005 as employment rates increased. Brüll and Gathmann (2020) show that wage inequality decreased already after the introduction of sectoral minimum wages while Bossler and Schank (2020) focus particularly on the consequences of the general minimum wage introduction in 2015. This literature focuses almost exclusively on wage inequality, in particular daily wages of full-time male workers.[4] Focusing on full-time employees is a natural choice for understanding the supply and demand forces shaping the wage structure as the full-time wage is most easily interpreted as a 'price'. However, focusing on full-time wages is less sensible if one is interested in the distribution of individual welfare and how it is affected by income mobility and risk. This is particularly true for women who have a high propensity for working part-time.[5]

We focus on annual earnings (labor income) as our main outcome variable. In contrast to the existing literature, we analyze both men and women and include individuals in marginal or part-time employment as well as people who work partially during the year. This broad sample is much better suited to analyze and compare the income distributions of women and men. This also allows us to capture the unique phenomenon of marginal employment in so called mini-jobs in Germany (see section 2 for more details).

Another difference is that the vast majority of previous work for Germany has used either survey data – which suffer from all the typical problems of selection bias, measurement error, and small sample sizes – or solely IAB data. While the administrative IAB data is of high quality and provides complete coverage of all employees who are in marginal employment or in jobs liable to social security (that is all jobs excluding civil servants and self-employed), the earnings are top coded at the social security contribution limit (corresponding approximately to the 90th percentile for men and the 96th percentile for women). Hence, these studies cannot address the upper tail

[3] Several further papers have studied the evolution of wage inequality during the 1990s and 2000s (e.g., Dustmann et al. 2009, Card et al. 2013, Goldschmidt and Schmieder 2017 and Bartels 2019). The reasons for the steep increase are manifold. For instance, Bayer and Juessen (2012), Antonczyk et al. (2018) and Biewen et al. (2018) find cohort and composition effects to be especially important while Peichl et al. (2012) document that the increase in inequality is strongly related to changes in household structure and employment behavior. Fuest et al. (2018) find that corporate tax hikes increase wage inequality as low-skilled, young, and female employees bear a larger share of the tax burden.

[4] Biewen et al. (2018) show the importance of part-time and employment interruptions for the increase in income inequality. Bönke et al. (2015) analyze lifetime earnings inequality and mobility of yearly earnings for 35 cohorts of West German men. Moreover, several studies analyze inequality in (equivalized) disposable household incomes – see, e.g., Hufe et al. (2018) for a survey and Stockhausen and Maiworm (2021) or Grabka (2021) for recent studies.

[5] For example, in 2019, 58% of women in Germany worked part-time (Wanger, 2020) and women held 81% of all part-time jobs (Statistisches Bundesamt, 2021a).

of the distribution. In contrast, the TPP data does not have any top-coding and hence includes the very top of the distribution.[6] However, many low income workers (and especially workers in mini-jobs) do not file a tax return and thus are not included in the TPP. Combining IAB and TPP data, allows us, for the first time, to study the top and the bottom of the earnings distribution in Germany simultaneously.

This paper is also the first to combine the study of income inequality with a detailed analysis of income dynamics in Germany based on administrative data.[7] We build on the work by Busch et al. (forthcoming) who study income risk over the business cycle for Germany (and other countries) up to 2010 using survey data, documenting highly pro-cyclical skewness in short-term income growth. We extend their work as our combined tax and social security data allows us to study income dynamics for high earners and to include various forms of income. By extending the time horizon we are also able to study how the prolonged period of a strong labor market has affected inequality, mobility and the distribution of earnings changes.

Our main analysis focuses on the time period from 2001 to 2016. We start in 2001 since this is the first year the TPP data is available and when the IAB data provides high-quality information on mini-jobs. As of 2018 there are about 7.5 million mini-jobs and almost 5 million workers work only in a mini-job, thus including these jobs is crucial for getting a complete picture of the German earnings distribution at the lower end. The downside of focusing on the last two decades is, that this is less comparable to earlier papers for Germany that have focused on time series for wages excluding mini-jobs. For this reason, we provide an analysis in the Appendix based on two alternative samples from the IAB data where we exclude mini-jobs: Germany from 1993 to 2018 (starting right when high quality data for East Germany becomes available) and West Germany from 1985 to 2018.

This paper is part of the Global Income Dynamics (GID) project and as such, we follow the comprehensive guidelines of the project to provide a consistent set of core results presenting the evolution of inequality and dynamics in labor income over time. Apart from focusing on annual earnings, the GID project specifies a number of sample restrictions, such as age range and a minimum annual earnings threshold as well as various key measures for the outcomes of interest.

In the first part of the paper, we present several new key results for the structure and evolution of the earnings distribution in Germany by gender: The first result is that for men, continuing the trend of rising inequality in the 1990s and early 2000s, income inequality kept increasing until the Great Recession both at the top and at the bottom of the distribution. After the Great Recession, incomes throughout the distribution, including the lower half, began to increase, slowing the rise in inequality. For women the picture is more complex: While inequality for most women increased until 2009 as well, the 10th percentile, which is mostly composed of mini-jobs, actually rose throughout the entire sample period. After the Great Recession, women's incomes rose quickly, particularly at the 25th

[6] Following Piketty and Saez (2003), researchers have used tax data to study the upper tail of the distribution. For Germany, Bartels (2019) provides a long-run picture of the evolution of top income shares using tabulated tax data. Moreover, Bach et al. (2009) combine survey data and repeated cross-sections of tax data to document inequality trends from 1992 to 2001. Jenderny (2016) uses a sub-sample of the TPP to study top income mobility from 2001–2006.

[7] Using survey data, Fuchs-Schündeln et al. (2010) estimate stochastic wage and earnings processes for Germany from 1984–2003 and decompose both the levels and the growth rates in transitory and permanent components

percentile, which catches up with higher income levels leading to a substantial decrease in lower tail inequality. In contrast to the lower tail, earnings inequality at the very top increased substantially for both men and women, but the increase is almost twice as large for women. For example, across both genders, the percentile with the highest earnings growth is the 99.99th percentile for women.

A second key result is that earnings inequality is substantially larger for women than for men but converges throughout our sample period due to the different trends by gender. Interestingly, initial conditions (inequality at age 25, i.e. around labor market entry) are virtually identical, but while inequality is falling for men within cohorts, it is rising for women. This is driven by many women opting for part-time work or mini-jobs later in life after having children (Kleven et al., 2019), which leads to large variations in working hours within cohorts over time.

A third key result is that volatility (measured as the dispersion of 1 year log earnings changes) exhibits opposite cyclicality at the bottom and at the top of the distribution with shocks in the lower half of the distribution increasing during downturns, while shocks at the top become more muted. The overall volatility is relatively constant, but the skewness of the shocks becomes markedly more negative during downturns. This holds for men and women. Volatility is also significantly higher for women than for men, especially for younger women and at higher income levels.[8]

In the second part of the paper, we present novel results regarding total income inequality and dynamics. First, we show that non-labor income is a major source of total income especially at the bottom and at the top of the distribution. Taking total income, i.e. the sum of labor, rental, self-employed and business income, as the main outcome measure, we find much higher levels of income inequality. In addition, over time, the top percentiles of the total income distribution increased significantly more than the corresponding percentiles in the earnings distribution.

Second, we compare income dynamics between workers and entrepreneurs (individuals with non-labor income as their main income source) and find that entrepreneurs' income changes are more dispersed, less skewed, and much less leptokurtic. In addition, income changes of entrepreneurs are much less state-dependent in the sense that we do not find significant heterogeneity between low-income and high-income entrepreneurs. Third, we document that income mobility at the top has declined significantly between 2001 and 2016. That is, the probabilities of dropping out of the top 1% and top 0.1% of the income distribution have declined both for 1- and 5-year time intervals.

In the next section, we discuss the institutional and macroeconomic setting in Germany over our analysis period, present our data sources and explain the sample construction. Section 3 presents the core results following the GID framework showing the evolution of inequality and income dynamics for labor earnings. Section 4 expands the analysis to total income. Finally, Section 5 concludes.

[8] In the Appendix we also document results on income mobility. Mobility is fairly high at lower ages and then decreases with worker age. Furthermore, mobility is quite a bit larger for women than for men, perhaps again reflecting the impact of hours reductions after childbirth and increases in hours after children grow older (Kleven et al., 2019).

2 Background and Data

2.1 Institutional and Macroeconomic Background

In this subsection, we give a brief overview of the relevant institutions and the macroeconomic situation in Germany for the period from 1993 to 2018 – see Appendix A for more details. While our main analysis focuses on 2001 to 2016, we provide additional results for this longer time period in Appendix F. Furthermore, starting slightly earlier than our main sample window helps to better understand the economic environment during that period.

Personal Income Tax. Germany applies a comprehensive income tax on income from all sources. Married couples file their tax return jointly. Both features are important when constructing our sample from tax return data – as discussed below.

Marginal Employment ("Mini-Jobs"). Marginal employment contracts, called mini-jobs, are jobs with earnings below a time-varying threshold (see Panel C of Figure A.1). Jobs below this threshold, which currently amounts to 450 Euro per month, are exempted from social security contributions and income tax.[9] Two reforms during our sample period increased the monthly earnings threshold from 325 Euro to 400 Euro (in 2003) and then to 450 Euro (in 2013). Over our sample period in each year around 4.5-5 million workers hold only a mini-job. Another 2.7 million workers have a secondary mini-job next to a regular contract.

Minimum Wage. Germany introduced a statutory national minimum wage of 8.50 Euro in 2015. After that, the minimum wage was gradually increased (see Figure A.1, Panel C). Before 2015 different wage floors existed in 12 industries. Furthermore, some of the larger industries have binding collective agreements that set minimum wages. The impact of the wage floor on wages varied by region and affected about 15% of all employees (Dustmann et al., 2022).

Collective Bargaining. Agreements between unions and employer representatives often have a binding character for most firms (above a certain size) in a specific industry (with some possibilities for firms to opt-out) in Germany. The worker coverage of industry-level collective bargaining agreements varies between former West and East Germany and decreases over time (see Figure A.1, Panel B). Especially start-ups and smaller firms are less likely to be part of a collective agreement.

Macroeconomic Background. The macroeconomic development in Germany from 1993-2018 can be broadly split into two periods: before and after 2005 (see Figure A.2). The first part was characterized by low growth and high unemployment (above 12%) and Germany was often referred to as "the sick man of Europe" (Dustmann et al., 2014). This changed in the mid-2000s after a series of (labor market and tax) reforms[10] The Great Recession did not affect the labor market severely.

[9] A person can hold multiple mini-jobs but then only the first 450 Euro are tax exempt.

[10] The causal effect of these reforms and the exact mechanisms are still discussed in the literature – see Footnote 1.

Moreover, labor force participation rates increased steadily after 2004 and the unemployment rate fell below 6% in 2018. Especially the large increase in labor force participation of women, from around 55% to more than 70%, is notable (see Figure A.2, Panel E). However, unlike in countries such as the US, this increase was almost exclusively driven by women entering the labor market in part-time and marginal employment, so that the full-time share over this period fell from 75% to around 50% for women. For men, labor force participation and the part-time share also increased substantially since 2003, though nowhere near as pronounced as for women.

2.2 Data

For our analysis, we combine two high quality administrative data sources: social security data (IAB) and personal income tax records (TPP). Each of these datasets has distinct advantages but also some weaknesses. The combination of both data sources offers a unique possibility for the analysis of inequality and income dynamics along the whole distribution. We describe the two datasets in the next sub-sections before explaining our sample selection, comparing the income distributions in both datasets and describing how we merge them. Appendices B (IAB data), C (TPP data) and D (combined IAB-TPP data) contain further details.

2.2.1 The IAB Social Security Data

The first source of data, which we refer to as the IAB data, is the Integrated Employment Biographies (IEB, version 13.01) supplied by the Institute for Employment Research (*"Institut für Arbeitsmarkt- und Berufsforschung", IAB*). The data contains information on employment and earnings as well as worker and firm characteristics such as gender, education, year of birth, occupation or industry code.[11] The information is spell based, i.e. accurate to the date and especially with respect to earnings very reliable. However, there are two important limitations to this widely used dataset (see, e.g., Dustmann et al., 2009; Card et al., 2013). First, labor earnings are reported including bonuses and extra pay but only up to the social security contribution limit (see Figure A.1 for its real values over time). This censoring affects men and women in West and East Germany differently with West German men being affected the most as here the top 10% are subject to censoring (see Appendix B and especially Figure B.1 for details).[12] Second, the IAB data does not include self-employed individuals (around 4 million) and civil servants (around 1.9 million individuals).

We use a 10% random sample of the IEB for the years 1993-2018, which gives us 87,012,649 observations. For our main results we focus on the period 2001-2016 (65,900,481 observations on 6,250,877 individuals). We show results for the period 1993-2018 in Appendix F.1 and, restricting to West Germany, for the period 1985-2018 in Appendix F.2.

[11] Note that the education information contains some missing values which we impute using the procedure suggested by the IAB. Moreover, throughout 2011, the reporting procedure for full-time and part-time employment in the data changed leading to a small fraction of workers being falsely classified before 2012. We correct the full-time indicator using a cell-wise reclassification approach. See Appendix B for details on both imputations.

[12] We use the algorithm suggested by Card et al. (2013) to impute daily wages which we then aggregate to annual incomes for our analysis. Note, however, that we do not use these imputed wages for our baseline analysis as we will draw on tax data to complement the IAB data at the top as we explain below.

2.2.2 German Taxpayer Panel (TPP)

The second source of data is the German Taxpayer Panel (TPP), which is an administrative dataset based on the universe of personal income tax returns (Kriete-Dodds and Vorgrimler, 2007).[13] The dataset covers all tax units filing tax returns in Germany in the period 2001-2016. The panel has a total of 58,808,899 unique records for which information is available for at least two years. We use a 25% random sample of these records. The unit of observation is the taxpayer, i.e., either a single individual or a couple filing jointly. In the latter case, incomes are measured on the individual level. Moreover, while we observe both spouses before and after a marriage (or divorce), we can only track the husband over these events. The reason is that the wife is assigned to the husband's tax unit in the event of marriage and withdrawn from it in the event of divorce.[14] As less than 10% of the women in our data could be affected once over the period of analysis, we are confident that our income dynamics results for women are mostly unaffected by this.

The dataset contains all information necessary to calculate a taxpayer's annual income tax. This includes basic socio-demographic characteristics such as year of birth, gender, family status, number of children as well as detailed information on gross income (differentiated by seven different sources) and tax-specific parameters such as deductions. As the data are not top-coded, they are especially suited for the analysis of inequality in the upper tail of the distribution. They are, however, missing the very bottom of the income distribution as incomes below the marginal employment threshold are except from the income tax and hence not included in the data (in the case of the mini-job being the only source of income). Note, however, that information on mini-jobs of secondary earners as well as recipients of income from other sources are included in the data. Furthermore, there is a structural break in the dataset in 2011 for the classification of workers which are subject to social security contributions (which are represented in the IAB data) affecting about 4% of the observations. We describe in Appendix C how we correct the data using reweighting techniques.

2.2.3 Sample and Variable Construction

For comparability with other countries covered in the GID project, we focus our analysis on individuals who are between 25 and 55 years old. Following the GID guidelines, the first part of our analysis (section 3) focuses on labor earnings excluding self-employment. Throughout our analysis, we examine both men and women as well as not only full-time but also part-time and marginally employed workers. The definition of gross annual labor earnings is the same for both IAB and TPP data: annual earnings is broadly defined and include, among others, overtime pay, bonuses, 13^{th} month pay, paid sick leave, severance pay, and vacation allowance. We exclude workers with weak

[13] See Appendix C for more information. The TPP has been, for example, used by Doerrenberg et al. (2017) and Dolls et al. (2018) who also provide additional information on the data.

[14] While we see the wife in all years, we could have 3 independent spells for her: before, during and after marriage. On average about 0.45–0.5% of individuals get married each year while roughly 0.2% file for divorce (https://www.destatis.de/EN/Themes/Society-Environment/Population/Marriages-Divorces-Life-Partnerships/Tables/lrbev06.html). The average duration of a marriage is about 15 years. As we analyze a period of 16 years we would expect that about 8% of all women in the data marry while less than 2% file for divorce.

attachment to the labor force by trimming annual earnings below a threshold $\underline{y}_t$, which corresponds to working part-time for one quarter at the national minimum wage (2,300 Euro in 2018). As we combine IAB and TPP data, and since the IAB data only covers social-security-liable earnings, we restrict our analysis to this group which accounts for more than 93% of all workers. In the second part of our analysis, we investigate the distribution of total income and study total income dynamics for individuals with different main income sources (see Section 4 which also includes some further details on the total income sample). All incomes are deflated using the CPI and Euro figures in the text, tables and figures refer to 2018 Euro.

For both labor earnings and total income, we follow the GID project's conventions and refer to three samples. In the "CS sample" (cross-sectional sample), we only impose the restrictions on age and minimum (labor) income. For the analyses that involve dynamics, we impose additional restrictions on the data and focus on two subsamples. To study trends, we use the "LS sample" (longitudinal sample) which only includes observations with non-missing one-year or five-year income changes. When studying heterogeneity in income dynamics by age and income, we use the "H sample" (heterogeneity sample) where we drop observations for which we cannot compute our measure of permanent (labor) income based on observations of the past three years.

2.2.4 Comparison and Combination of IAB and TPP Data

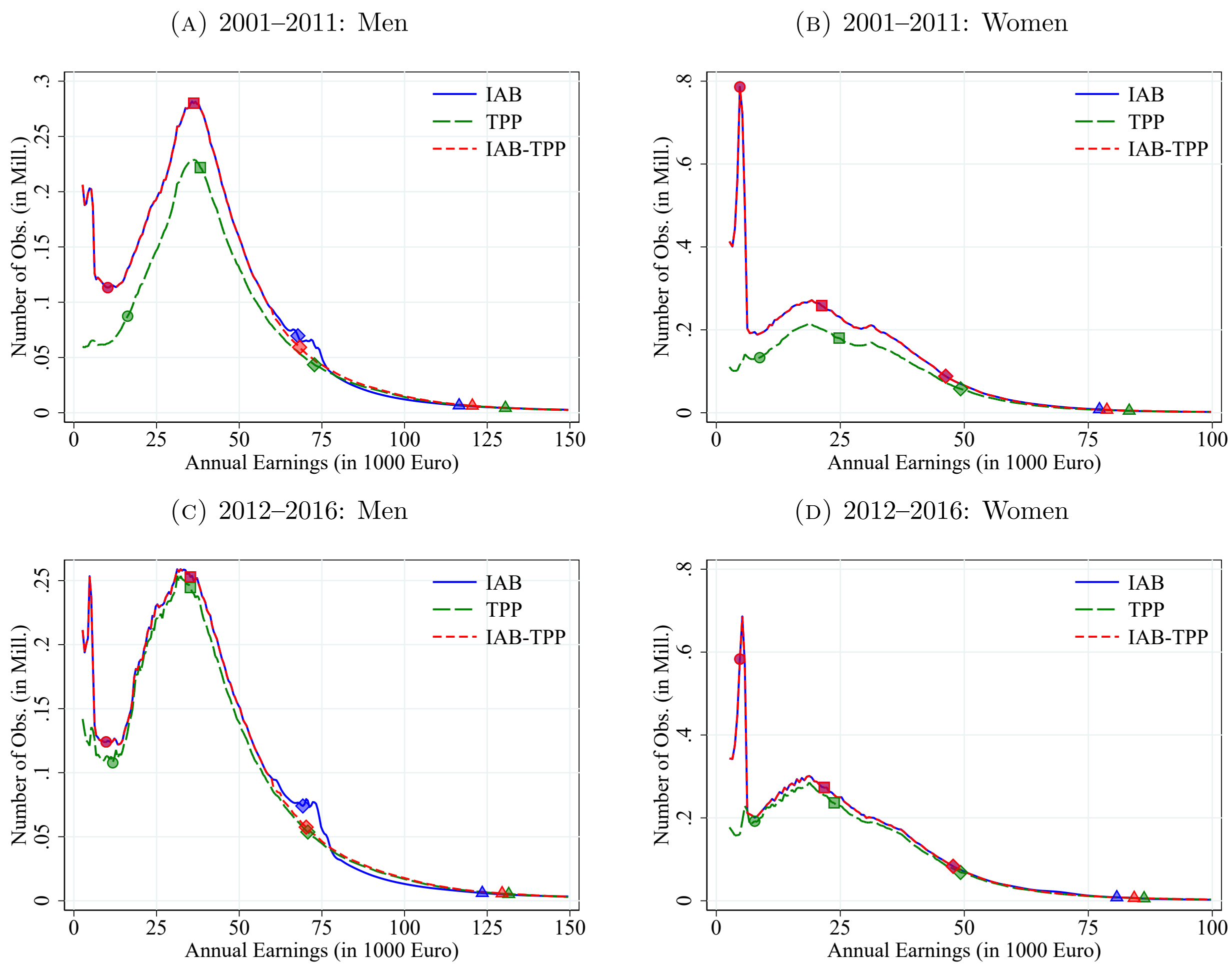
The four key differences between the IAB and TPP with respect to labor income are: several (but not all) missing mini-jobs and the incomplete coverage of regular social security liable jobs due to non-filing (before 2012) in the TPP data as well as the top-coding and the omission of civil servants and self-employed individuals in the IAB data. Table D.1 in the Appendix displays descriptive statistics for the IAB and TPP datasets for the year 2008 separately for men and women who are between 25 and 55 years old. Unsurprisingly, the TPP has fewer observations due to missing non-filers and mini-jobs. As the TPP data contains only very limited demographic information, we can only compare both datasets in terms of age. The TPP population is slightly older which again can be attributed to missing observations representing persons who are more likely to be at the beginning of their career. There are more men than women in both datasets due to their higher labor force participation rate (see Figure A.2, Panel E).

Figure 1 provides a comparison of the earnings distributions in the two raw datasets separately by gender for the periods 2001–2011 and 2012–2016 (as the TPP data includes information on non-filers starting from 2012).[15] For this graph we focus on jobs subject to social security contributions also in the tax data such that all lines refer to the same population. Since only workers who submit a tax return were covered in the TPP until 2012 (Panels A and B) it is not surprising that the green line for the TPP data consistently lies underneath the blue line for the IAB data until about the social security contribution limit at around 70,000 Euro is reached (note that above the contribution limit the IAB data is imputed). From 2012 onward (Panels C and D), both IAB and TPP data

[15] Figure D.1 in the Appendix shows the same information for the full population. Tables D.2 – D.4 show selected earnings percentiles across the different datasets for men, women and in the population respectively for all years.

are much closer together. The biggest difference is still at the bottom of the distribution where the IAB data show a large mass-point stemming from mini-jobs while still about half of the mini-jobs covered in the IAB data are missing in the TPP.

Figure 1: Annual Earnings Distribution in IAB, TPP and Combined Data



Notes: This figure shows the number of observations in real earnings bins for the IAB, the TPP and the combined data (IAB-TPP) by gender (see Figure D.1 in the Appendix for the combined distribution). Panels A and B show averages across the years 2001 to 2011 where non-filing workers (Lohnsteuerfälle) are not included in the TPP and Panels C and D show averages across the years 2012 to 2016 where the TPP data include these workers. We exclude earnings from the TPP that are not subject to social security contributions (e.g. salaries of civil servants) which are not covered in the IAB. The circular, square, diamond and triangle-shaped markers depict the 10^{th}, 50^{th}, 90^{th} and 99^{th} earnings percentile in the respective datasets. We use 500 Euro bins below 80,000 Euro and 1,000 Euro bins above 80,000 Euro but always plot the number of observations per 1,000 Euro bins. The IAB data are imputed above the social security contribution limit. Tables D.2, D.3 and D.4 show selected earnings percentiles across the different datasets for men, women and in the population respectively.

The figure also displays symbols on each line to indicate certain percentiles of each distribution. Interestingly, while both the median (squares) and the 90th percentile (diamonds) values lie relatively close to each other in both datasets, there are larger differences at the 10th percentile (circles), which for women actually lies to the left of the largest mini-job mass point, as well as at the 99th percentile (triangles). Overall, the IAB data is slightly shifted to the left compared to the

TPP data. Note also that the imputation procedure for the IAB data does a fairly good job at approximating the top tail compared to the TPP data but is not fully able to overcome the problem of top-coding and the resulting biases. The earnings distribution of men is more affected by this top coding than the distribution for women, which in turn is more seriously affected by the omission of most mini-jobs at the bottom of the distribution in the TPP data.

For our analysis, we combine both datasets. Due to data protection legislation in Germany, we are not allowed to directly link the individual micro data. Hence, we employ non-parametric matching techniques as described in Appendix D. To do so, we first reweight the TPP data (conditional on gender, age and earnings bin) such that we match the total number of workers liable to social security contributions (from the IAB data). Second, we combine the results from both datasets. For the core analysis in Section 3, we use the (true) earnings distribution from the IAB data below the top-coding threshold. Above the cutoff, we use the conditional earnings distribution from the (reweighted) TPP. For the analysis of total incomes in Section 4, we use the reweighted TPP data.

Figure 1 also shows the combined data which roughly corresponds to the IAB data at the bottom and in the middle of the distribution and to the TPP data at the top. Tables D.2 – D.4 in the Appendix show selected percentiles of the earnings distribution in the combined IAB-TPP data as well as in the IAB and TPP data to confirm this observation.

3 Earnings Inequality and Dynamics in Germany

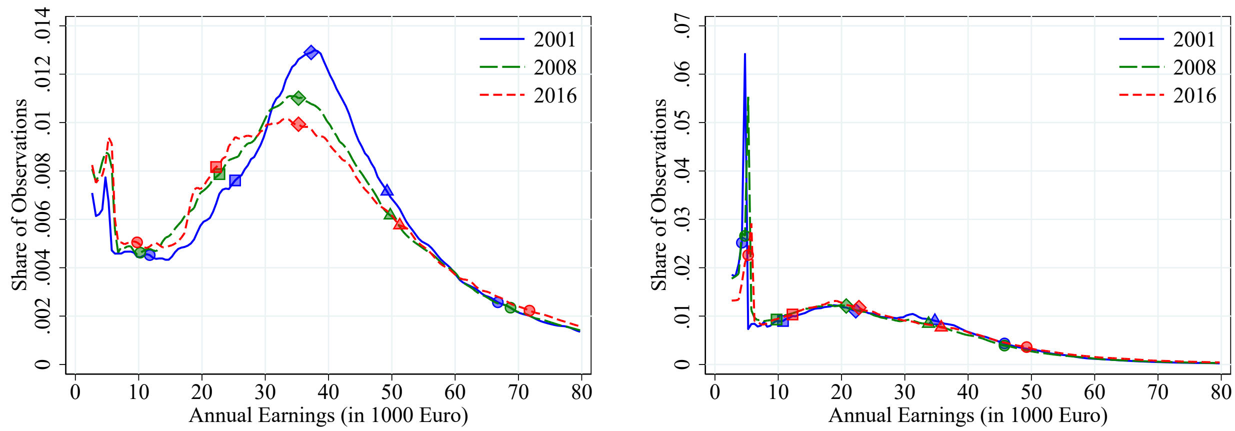
3.1 Earnings Inequality

We start our analysis by documenting the evolution of income inequality over the past 2 decades. Figure 2 shows the density of (real) earnings for men and women for 3 selected years over our sample period. Panel A shows the distribution for men. The mode and median of the distribution in 2001 are just below 40,000 Euro. The distribution also shows a sizable mass (about 6%) at about 5,000 Euro, which corresponds to the earnings of someone working in a mini-job for the full year (325 Euro times 12 adjusted for inflation). Panel B shows the earnings distribution for women: The mini-job spike is much more pronounced for women and the mode of the distribution above the mini-job threshold is close to 20,000 Euro, much lower than for men. The markers on the density correspond to various percentiles of the distribution. Thus, we can see that for men about half of the workers' earnings are between 25,000 and 50,000 Euro per year, while for women the inter-quartile range lies between about 10,000 and 35,000 Euro. Table 1 complements this visual presentation of the percentiles with the exact numbers for selected years. It is striking how much lower the respective percentiles for women are compared to those for men. For example, the median for women is below the 25th percentile for men, while the 75th percentile for women is below the median for men.

Comparing the densities across years in Figure 2 gives a first impression of how the distribution changed over our sample period. For men, there is a clear reduction in mass in the middle of the distribution; e.g. at 40,000 Euro, the density decreased from 0.013 (that is 1.3% of all workers make between 40,000 and 41,000 Euro) in 2001, to 0.01 in 2008, and further to 0.009 in 2016. Instead,

Figure 2: Selected Real Earnings Distributions

(a) Men (b) Women



Notes: This figure shows the distribution of real annual earnings (in 2018 Euro) for selected years in the combined IAB-TPP data (CS sample) by gender. The data is smoothed (by year and gender) using a three-bin moving average for bins above 10,000 Euro. The markers indicate the 10th (circle), 25th (square), 50th (i.e. median; diamond), 75th (triangle) and 90th (circle again) percentiles of the respective distributions.

we have a significant increase in the density in the range of 20,000 to 30,000 Euro, as well as in the mini-job range and at the very top of the distribution. For women, the differences are harder to make out due to the compressed axis which is necessary to show the huge mini-job share (see Figure E.1 for a zoomed-in version). The basic pattern of hollowing out of the middle of the earnings range and increases in the 20,000 to 30,000 Euro range holds for women as well, though is somewhat less pronounced.

Table 1: Percentiles of Real Annual Earnings in Combined IAB-TPP Data

Year	N	Mean	P5	P10	P25	P50	P75	P90	P95	P99	P99.9	P99.99
Men												
2001	13.113	42,479	6,331	11,893	25,563	37,599	50,524	69,852	86,892	141,844	331,421	939,818
2004	12.419	42,455	5,406	10,088	24,398	37,434	51,274	71,866	89,516	145,790	335,445	896,172
2008	12.430	42,459	5,409	10,282	23,013	36,029	50,985	72,925	92,556	158,500	397,202	1,117,735
2012	12.535	42,433	5,121	10,104	22,479	35,528	51,330	74,097	94,212	160,328	387,358	1,025,836
2016	13.096	43,665	5,445	9,991	22,831	36,206	52,817	76,755	97,788	168,152	415,058	1,188,906
Women												
2001	11.476	24,628	3,671	4,673	10,881	22,162	34,723	46,032	54,306	78,226	137,340	285,996
2004	11.101	25,281	3,599	4,686	9,905	21,778	34,760	46,552	55,572	81,110	141,506	281,579
2008	11.228	25,002	3,695	4,736	9,811	20,845	33,761	46,100	55,701	84,335	158,265	335,160
2012	11.510	25,610	3,832	4,861	10,664	21,107	34,085	47,149	57,410	88,128	167,849	363,598
2016	11.799	27,562	4,178	5,366	12,387	22,805	36,241	49,859	61,077	95,403	185,926	414,783

Notes: This table shows the number of observations (in millions) and selected percentiles of real annual earnings (in 2018 Euro) in the combined IAB-TPP data (CS sample) by gender for selected years. Tables D.2 and D.3 in the Appendix show the percentiles in the underlying IAB and TPP data.

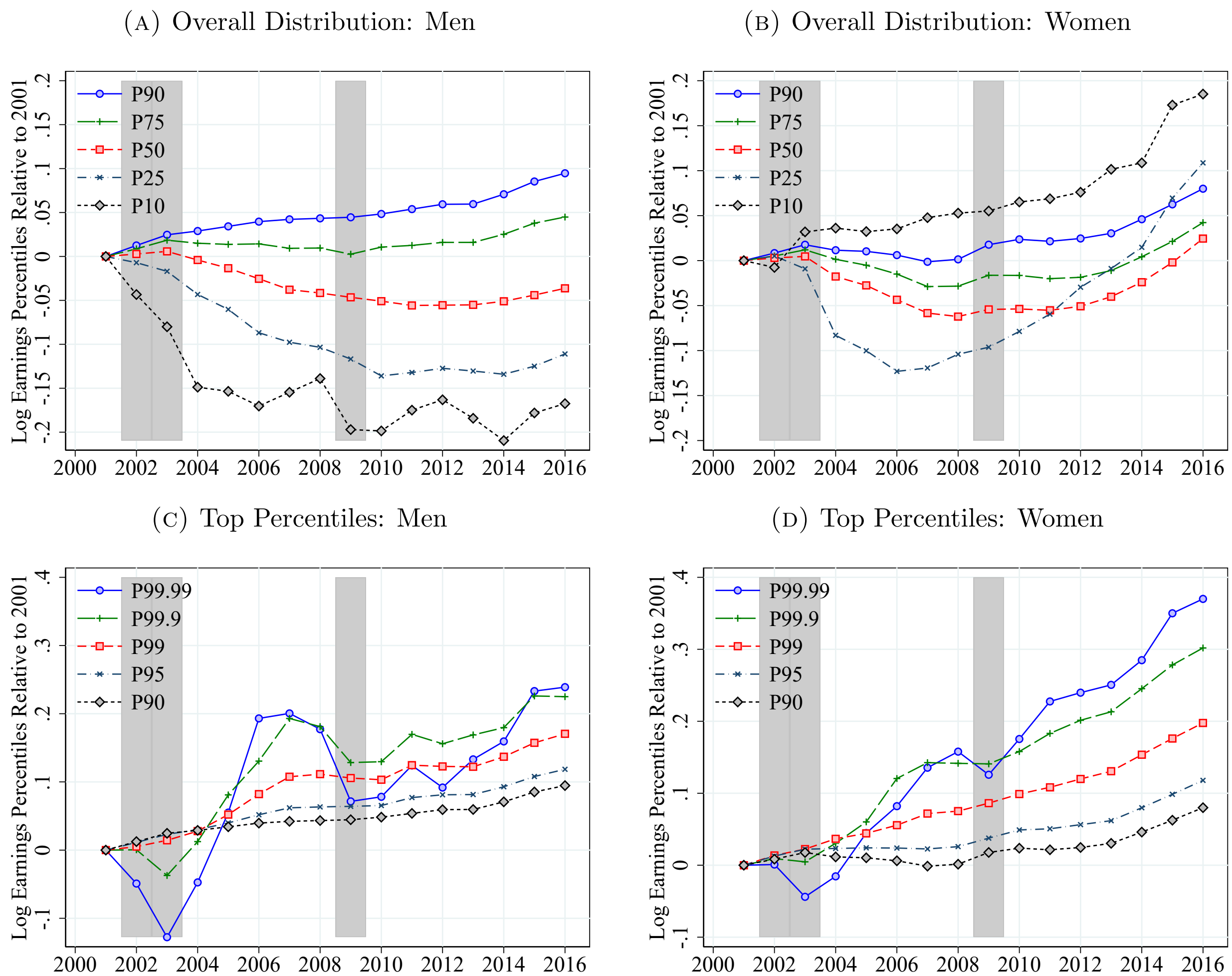
To get a clearer picture of how earnings inequality has evolved, Figure 3 shows a range of percentiles of the log earnings distribution relative to 2001 separately by gender (for the full population - men and women combined - see Figure E.5). Panels A and B illustrate the percentiles up to the 90th for both men and women. Men experienced a pronounced increase in earnings inequality until 2009. At the lower tail of the distribution, the 10th percentile dropped by about 20 log points from 2001 to the Great Recession. Since median earnings also declined, but by only around 5 log points, the log earnings gap between the 50th and 10th percentile increased by around 15 log points. The 90th percentile grew slowly until the Great Recession by around 5 log points, increasing the gap relative to the median. For women inequality also rose in the middle of the distribution as can be seen by the fall of the 25th percentile relative to the 90th percentile. However, the 10th percentile showed a gradual increase after 2001.

After the Great Recession the economy recovered, unemployment rates fell and earnings levels throughout the distribution began to rise slowly for men. Women on the other hand show relatively faster growth in earnings after 2011 throughout the full distribution with the fastest growth at the lower percentiles, in particular the 25th. The faster growth at lower percentiles implies that for women's earnings inequality fell substantially over the second half of our sample period.

Turning to the evolution of incomes at the upper end of the earnings distribution, we can see that for men (Panel C), the top percentiles grow at a faster rate than lower incomes. For example earnings at the 99.9th percentile grew almost twice as much as at the 90th percentile (20 vs. 10 log points). This observation is, for example, consistent with the increase in CEO compensation in Germany during that time (see, e.g., Prinz and Schwalbach, 2020). The fanning out of top incomes is even more pronounced for women (Panel D), where the very top of the earnings distribution (the 99.99th percentile) grew by almost 40 log points or 4 times as much as the 90th percentile. The faster earnings growth for women at the very top also suggests that women at the top end of the income distribution are catching up with the highest-earning men in the economy – although even in 2016 the 99.99th percentile of earnings for women is still only equal to the 99.9th percentile for men (see Table 1). It is also noteworthy that the highest percentiles for both men and women show much more cyclicality than the lower percentiles with marked drops during recessions. The cyclicality is likely due to bonus pay for top managers and highly qualified workers, which in its nature fluctuates over the business cycle. Female top incomes are less cyclical than male top incomes. This might be because female top earners are more often employed in industries that are less prone to business cyclicality. Another potential explanation for the differences between men and women is the increase of the female share in management positions – both at the very top of firms (Kirsch and Wrohlich, 2020) but also at lower management levels (Kohaut and Möller, 2017).

To provide a concise picture of the evolution of inequality, Figure 4 shows different log percentile differentials. These figures support the impression from before: For men, inequality is rising moderately until around the Great Recession, followed by a period of slow growth until 2016. Comparing these differentials with women, we see that the top half of the distribution are very similar between men and women (P90-P50), while the lower two gaps are much larger for women. Regarding

Figure 3: Evolution of Log Earnings Percentiles



Notes: This figure shows the evolution of selected percentiles of log real annual earnings (relative to 2001) in the combined IAB-TPP data (CS sample) by gender. Shaded areas indicate recessions.

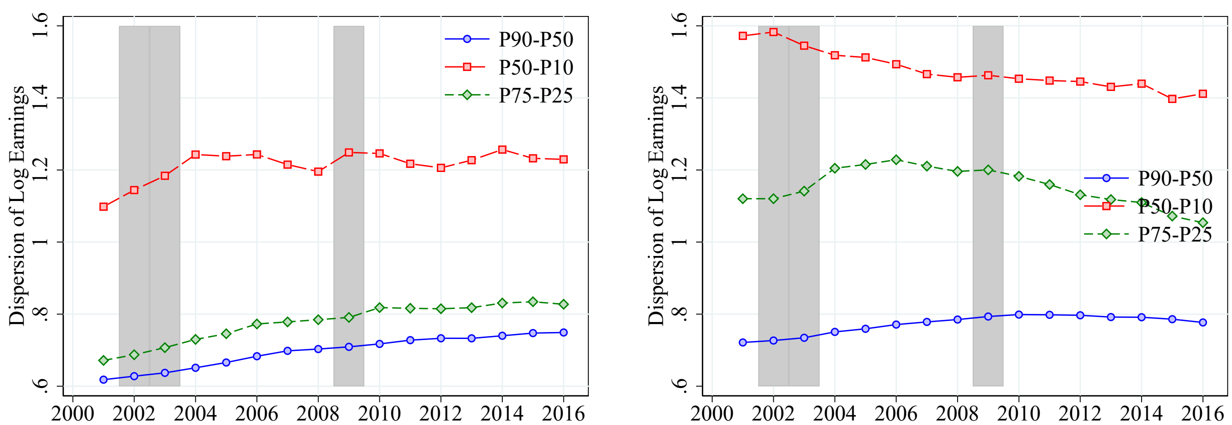
changes over time for women, inequality in the middle and the top half of the distribution (P90-P50 and P75-P25) is increasing slightly from 2001 to 2009, but then decreases until 2016. Inequality at the bottom (measured as P50-P10) by contrast falls throughout the entire sample period, driven by the rising 10th percentile as shown before. As was highlighted by the density figures before, the 10th percentile for women is somewhat special in that it is in the range of mini-job incomes and roughly traces the increases of the mini-job threshold over time. The inverted U-shape of the P75-P25 differential where inequality is rising up to the Great Recession and falling afterward, is probably a better measure for the overall picture of inequality and also in line with measures such as the Gini coefficient or standard deviation (see Figures E.3 and E.14).[16]

What explains the different patterns for men and women? As we saw in Table 1, a key difference between men and women is that the gender-specific percentiles are located at different absolute

[16] In Appendix E, we also report labor income shares of various parts of the income distribution (Figure E.12 and Tables E.1 and E.2) and estimated Pareto coefficients for the top 5% and top 1% (Figure E.13).

Figure 4: Evolution of Earnings Inequality: Log Percentile Differentials

(a) Upper and Lower Inequality: Men (b) Upper and Lower Inequality: Women



Notes: This figure shows the evolution of different log percentile differentials over time in the combined IAB-TPP data (CS sample) by gender. Shaded areas indicate recessions.

income levels. Taking this into account, it appears that at specific income levels men and women are somewhat more similar. For example, the 75th percentile for men is at a similar level as the 90th percentile for women and the two lines look fairly similar in Figure 3. Nevertheless, important differences remain, in particular at the bottom and at the very top of the income distribution.

Turning to the bottom of the distribution, the 10th percentile for women is at about 5,000 Euro, right at the level of someone who works only a mini-job for the full year. Thus, the steady increase of this percentile is essentially earnings growth for mini-job workers. Indeed, the jumps in the 10th percentile line in 2003 and 2013 are driven by the increases in the mini-job threshold in those years (see Figure A.1). As shown by Gudgeon and Trenkle (2020), the adjustment process took several years with many workers initially remaining at the pre-reform levels. This is likely a factor in the rise in the 10th percentile in the years between the reforms. The other marked jump is from 2014 to 2015 when the Federal minimum wage was introduced pushing some mini-jobbers above the earnings threshold (and thus out of the mini-job range). For men, however, the 10th percentile lies substantially above the mini-job threshold (only 2.4% men have only a mini-job in 2008 - see Table D.1) corresponding to individuals working in low-paying part-time jobs. The minimum wage introduction increased the 10th percentile slightly for men, but much less than for women.[17]

There are several potential drivers for these developments. Both men and women are more likely to work part-time in recent years (see Figure A.2). For men, the part-time share increased from around 3% to almost 10%, while for women it increased from around 30% to around 45%. While for men mini-jobs (as a primary job) never played a big role (only around 2-3%), the share for women is substantial (around 12-13% in 2001) and falling to around 6% towards the end of our sample

[17] Dustmann et al. (2022) also report that about two thirds of workers affected by the minimum wage were women and Caliendo and Wittbrodt (2021) find that this reform indeed decreases the gender wage gap.

period. Furthermore, there has been an increase in the college share among workers in our sample over the time period, an increase in non-German immigrants, and average age (see Figure E.15).

Reweighting Analysis. To gauge the importance of these observables for the evolution of the earnings distribution, we reweight our sample such that observable dimensions are held constant at the 2001 level (see Appendix E.2 for details). Figure 5 shows the counterfactual evolution of several percentiles for men and women when we use this reweighting procedure.

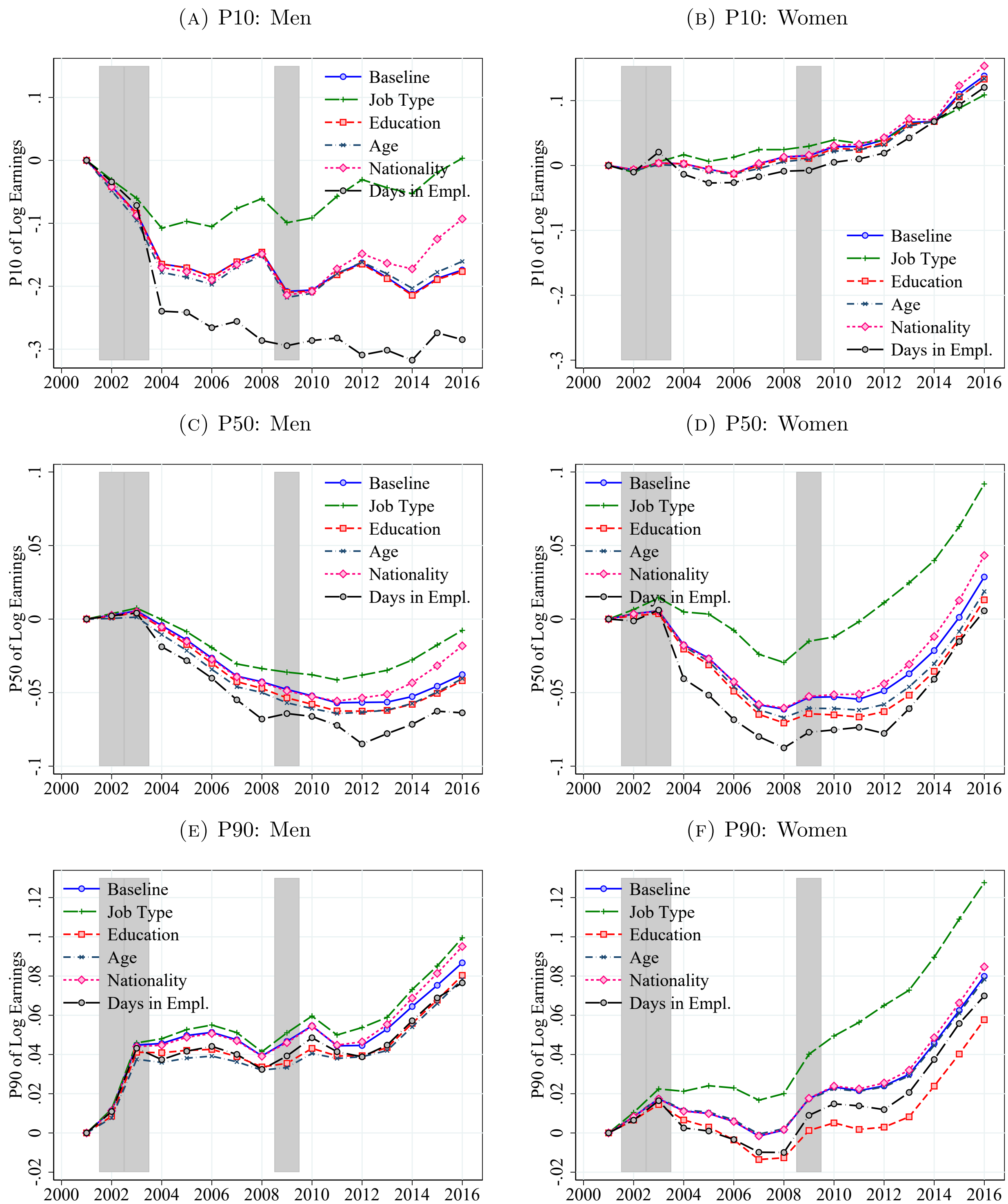
For men, job type (i.e., full-time / part-time / mini-job) plays a key role in explaining the fall in real earnings at the 10th percentile. This suggests that much of the increase in part-time occurred at these very low earnings levels pushing them down. Holding job type constant, the 10th percentile would have fallen less in the early 2000s and fully recovered to the 2001 level by 2016. Job type is also important in explaining the decline of the median up to the Great Recession, and the line is shifted upwards after controlling for this via reweighting. In contrast, for women holding job type constant does not affect the 10th percentile much, since those are always mini-jobbers, but has a huge impact on the median. If the part-time share had not increased, the median would have grown by almost 10 log points compared with the observed growth of around 3 log points. Even at the 90th percentile, the growth in part-time employment led to a significant slowdown in earnings growth, and after controlling for job type composition the P90 grows by an additional 5 log points.

Mean age first increases, peaks between 2010 and 2012 before slightly decreasing again, but staying above initial levels (see Figure E.15). However these modest changes are not large enough to meaningfully affect the earnings structure. Education levels rose gradually over our time period (see Figure E.15). Since education is positively associated with earnings, it is not surprising that after reweighting for this dimension, median earnings and the 90th percentile grow slightly slower.

A different shock to the German labor market was the opening of the market to workers from the new EU member countries in Eastern Europe that occurred on May 1st, 2011. As can be seen in Figure E.15, the share of non-German workers increased sharply from 2010 to 2016 (for men from 9 to 15% and for women from 6 to 10.5%). These workers typically work in low-wage jobs. Indeed, based on the reweighting analysis, they appear to significantly push down the 10th and 50th percentile for men and to a lesser extent the median for women.

Finally, annual earnings are of course in part determined by the number of days worked in a year. For example, if workers move between jobs with intermittent unemployment spells, this will decrease annual earnings for that year, even if wages remain the same. In Figure E.16, we show the average number of days worked for all workers as well as for above and below median earnings individuals. On average, days in employment increased slightly at the beginning of our sample period and decreased again towards the end, dipping during the Great Recession for men. This is consistent with the fall in unemployment after 2005. These changes in days in employment are driven almost solely by workers with below-median earnings. Indeed, the reweighting analysis in Figure 5 shows that this increase did substantially contribute to earnings growth for both genders in the lower half of the distribution.

Figure 5: Counterfactual Evolution of Log Earnings Percentiles (Reweighting)



Notes: This figure shows the evolution of different counterfactual percentiles of the log real annual earnings distribution over time in the IAB data (CS sample) by gender. The P90 for men is imputed in the IAB data as it lies above the social security contribution limit. The counterfactual percentiles are constructed by reweighting the data such that observable dimensions are held constant at the 2001 level (see Appendix E.2 for details). For example, the green line shows how different percentiles would have evolved over time had the job type distribution stayed as it was in 2001. A value of this counterfactual percentile above (below) the baseline value (blue lines) thus means that absent any change in the specific variable, earnings (at the given percentile level) would have been higher (lower) than what was actually observed. Thus, the observed change in the specific variable led to lower (higher) real earnings. Shaded areas indicate recessions. Figures E.17 and E.18 show the evolution of different counterfactual percentiles for each reweighted observable, while Figure E.19 shows counterfactual percentile differentials.

Inequality over the Lifecycle. The previous results showed markedly different developments for men and women. However, women's careers diverge from men's especially after childbirth (Kleven et al., 2019). We, therefore, turn to investigating inequality over the lifecycle. Since we are mostly interested in somewhat younger workers with fewer observations above the top coding threshold (which is always above the 90th percentile in the subsequent analysis), we use only the IAB data and therefore can show results up to 2018 as well as include further demographics.

Figure 6 shows the evolution of labor income over time for the four cohorts who are at the age of 25 in 2001, 2005, 2009, and 2013, respectively. Panel A shows median earnings for men. For all cohorts, earnings grow fast for young workers: around 60 log points in the first 10 years. While earnings at 25 are slightly higher for the 2001 cohort, earnings growth is faster for the later cohorts who make up the initial disadvantage. The decline in initial earnings may be due to the weak labor market from 2001 to 2005, leading to depressed wages, while the later cohorts enter during a much stronger labor market (see Figure A.2). Panel C shows within cohort inequality (P90-P10 gap) over the lifecycle among men for each cohort-by-year cell. It reveals that within cohort inequality falls rather fast over the first 10 years and then flattens out. In Figure E.9 we show that while the share of men working part-time (but more than a mini-job) is fairly constant over the lifecycle (at around 10%); young men are much more likely to work in mini-jobs compared to older men. This decline over the lifecycle is an important factor in reducing inequality. Another factor is that men who attain university degrees enter the labor market relatively late.[18] Moreover, workers at age 25 with college degrees have lower mean earnings than non-college workers but then quickly catch up and overtake the non-college group. This suggests that the decline in within cohort inequality is due to the decline in mini-jobs and entry of college workers in the middle of the income distribution.[19]

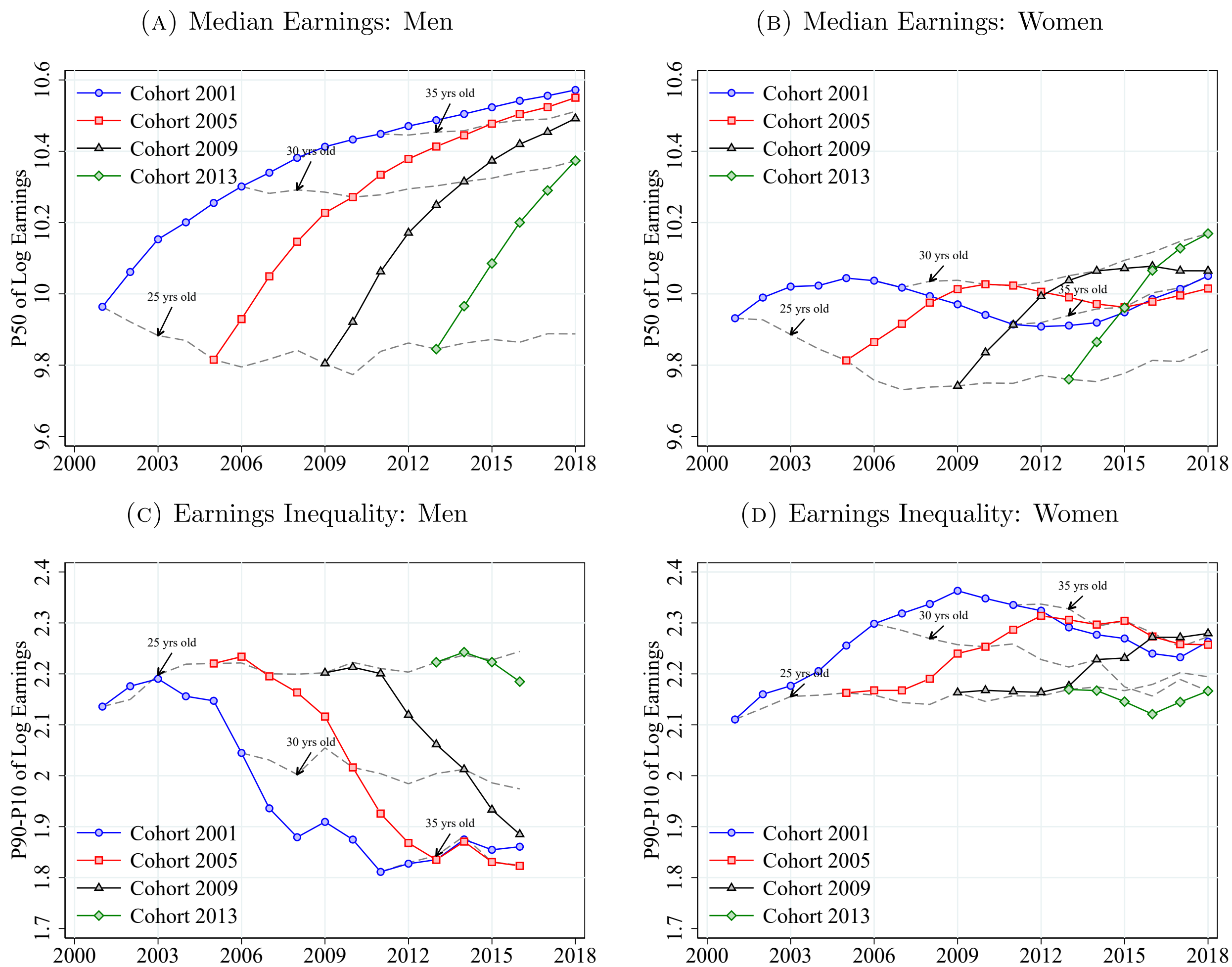
For women, the pattern is markedly different. While median earnings and inequality are almost the same as for men at age 25, median earnings rise much slower. Furthermore, unlike for men, earnings inequality rises continuously from age 25 to 35. This is likely since many women have children in their late 20s and early 30s (Bundesinstitut für Bevölkerungsforschung, 2021) and then transition to working part-time afterward (Schrenker and Zucco, 2020).[20] Figure E.9 shows that while only around 15% of women work part-time at age 25, this increases to over 50% by age 40. In addition, the mini-job share falls much less strongly for women than for men (for the 2001 cohort it even increases somewhat initially). Additional analysis (Figures E.10 and E.11) reveals that this is the combination of a rising mini-job share with age for non-college women and a sharp fall in mini-jobs with age for women with college degrees. Furthermore, while part-time increases with age

[18] During our sample period University-qualifying high school degrees (Abitur) took 13 years and men were regularly spending a year in civil or military service, followed by typically 5-6 years of University. In fact, Figure E.9 shows that the share of workers with college degrees is only around 18% at age 25 and increases to 25% by age 30.

[19] Consistent with this we show in the Appendix, that the decline and flattening of the P90-P10 log earnings differential with age is the result of offsetting trends in lower and upper inequality: Figure E.8 shows that while the P50-P10 declines, the P90-P50 increases as workers pass the age of around 30.

[20] The age for having the first child for women increases over time (Institut Arbeit und Qualifikation, 2021). Kleven et al. (2019) and Bönke et al. (2022) find that child penalties in female earnings are especially pronounced in Germany. Interestingly, the latter paper finds that participation and earnings gaps for women after birth increased until the 1990s but then this trend was reversed. This fits our findings of somewhat larger median earnings profiles and lower/declining inequality profiles of younger female cohorts in Figure 6.

Figure 6: Earnings Profiles and Inequality by Cohort



Notes: This figure shows the evolution of the median as well as the P90-P10 differential of the log real annual earnings distribution over time in the combined IAB-TPP data (CS sample) by gender. As the P90 of men is imputed and the TPP data end in 2016, Panel C also ends in 2016. Grey dashed lines correspond to earnings of 25, 30 and 35 year olds in each year as indicated by arrows. Each colored line corresponds to an individual cohort, where "cohort t" represents the cohort aged 25 in year t.

for all women, the increase is stronger and earlier for women without a college degree (likely because college educated women have children later). Thus, at young ages non-college educated women tend to move to part-time just as more college educated women enter the labor market working full-time, thus pushing up inequality. Then around age 30 college educated women start working part-time (or drop out of the labor force) leading to a mild decline in within cohort inequality.

Long-term Evolution. Our main analysis focused on the period from 2001 to 2016. To put these results into context with the long-term development, we replicate the key figures for two longer samples in the IAB data. The downside of the longer time frame is that prior to 1999 the IAB data does not contain mini-jobs, so we can only capture the earnings distribution above the mini-job threshold. In addition, we have to rely on imputed values for the 90th percentile for men (for women it is still below the top coding threshold).

The first long sample starts right after Germany's reunification covering all workers from 1993 to 2018 (see Appendix F.1). Figures F.1 and F.2 show that the increase in earnings inequality in the 2000s, was preceded by even faster increases in inequality in the 1990s. The overall growth in inequality was also much larger for men than for women, both in the 1990s and 2000s. Despite the data differences, the figure is fairly consistent with the post-2001 sample, except for the 10th percentile for women which is by construction very different. The lifecycle plots for this sample (Figure F.4) show very similar patterns for average earnings growth (much faster for men than for women) for the full period. A difference is that, due to the lack of data on mini-jobs, the fall in within cohort inequality is less pronounced for men, while the increase is even stronger for women.

Focusing on West Germany only, we can show results starting in 1985 (see Appendix F.2).[21] This sample shows that the increase in inequality started in the late 1980s. Even over this very long horizon, there are marked gender differences. Women at the lower percentiles fared much better than men, while the median and 90th percentile are closer to each other. As a result, men experienced a much larger increase in earnings inequality, especially at the bottom of the distribution. The lifecycle plots are similar as in the 1993-2018 sample, showing that the stark differences in lifecycle profiles by gender have been a feature of the German labor market for a long time.

3.2 Earnings Dynamics

From the vantage point of an individual worker, the earnings dynamics over time are arguably just as important as cross-sectional earnings differences as earnings risk affects key economic decisions such as consumption and savings. Therefore, we analyze the distribution of earnings growth, g^1_{it}, over time. It is defined as 1-year changes in residualized log earnings where we take out year- and gender-specific age effects.[22] To limit the effect of extremely large changes, we rely on percentile-based measures and report percentile differentials, Kelley skewness, and excess Crow-Siddiqui kurtosis instead of standard deviation and higher moments of the earnings growth distribution.[23]

Similar to the distribution of earnings, neither IAB nor TPP data alone allow us to estimate the true distribution of earnings growth.[24] To obtain this distribution, we combine IAB and (reweighted)

[21] For a detailed analysis of full-time wage inequality by gender from 1975 to 2004, see Dustmann et al. (2009).

[22] 1-year earnings growth in year t is defined as $g^1_{it} = \epsilon_{i,t+1} - \epsilon_{it}$ where $\epsilon_{it} = y_{it} - x'_{it}\hat{\beta}$ is the residual of a regression of log earnings on gender and year specific age dummies. Figure E.22 shows the density of g^1_{it} for the year 2005. The corresponding results for 5-year earnings growth are reported in Appendix E.3.

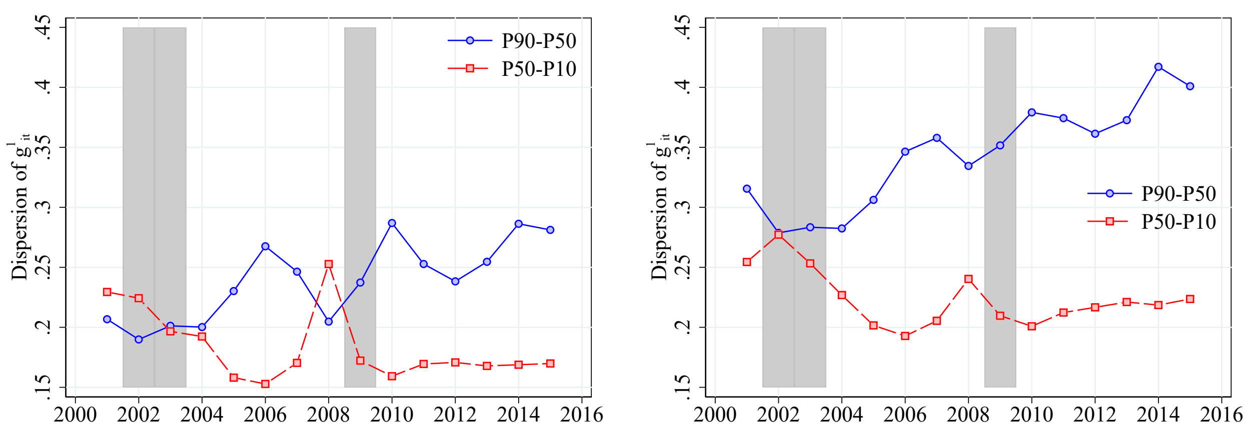
[23] Kelley skewness is defined as $\frac{P90-2P50+P10}{P90-P10}$. A positive (negative) value implies that the right tail of the distribution is longer (shorter) than the left one. Excess Crow-Siddiqui kurtosis is defined as $\frac{P97.5-P2.5}{P75-P25} - 2.91$ (i.e. Crow-Siddiqui kurtosis minus 2.91 – its value for a Gaussian distribution). Larger (positive) values indicate a leptokurtic distribution having a lot of mass around zero and relatively fat tails. In the context of earnings growth, this implies that workers experience fewer medium-sized shocks but more extreme positive or negative shocks.

[24] In the IAB data, top-coding implies that a substantial share of growth rates is calculated from imputed values. As the imputation does not take dynamics into account, this inflates growth rates. In the TPP data, most transitions in and out of mini-jobs are not observed. In addition, the absence of non-filers prior to 2012 distorts the distribution of growth rates. Non-filers are likely to have less volatile earnings, while, for example, workers who switch jobs or receive unemployment insurance are obliged to file a tax return and also experience relatively large earnings changes.

TPP data following a three-step procedure. First, we estimate growth distributions conditional on gender and current earnings separately using IAB and TPP data. Second, we construct a combined conditional distribution of earnings growth. That is, we use the conditional growth distribution from the IAB data for low levels of current earnings where only very few (less than 2%) growth rates are affected by top-coding, and the conditional growth distribution from the TPP data above. This cutoff lies between 45,000 and 50,000 Euro depending on gender and year. Finally, we integrate the conditional distribution with respect to the combined IAB-TPP earnings distribution in the LS sample to obtain the unconditional distribution of earnings growth by year and gender.[25]

Figure 7: Dispersion of 1-Year Log Earnings Changes

(A) Men (B) Women



Notes: This figure shows the P90-P50 and P50-P10 differentials of the distribution of 1-year changes in residualized log real annual earnings (from t to $t+1$) in the combined IAB-TPP data (LS sample) for men and women. See Appendix D.2.2 for details on the construction of the log earnings growth distribution from IAB and TPP data. Shaded areas indicate recessions.

Figure 7 shows the P90-P50 and P50-P10 differentials of 1-year log earnings changes over time. The distribution of men's earnings is more stable in the sense that large negative and (especially) positive changes are less likely.[26] Yet, men's earnings are more strongly affected by the business cycle, in particular by the Great Recession: Large positive (negative) changes become less (more) likely. The fact that the right tail shrinks and the left tail expands in the Great Recession directly explains the drop in the Kelley skewness for men shown in Panel A of Figure 8. Quantitatively, this drop is small compared to other GID countries such as Italy, Spain or the US (Hoffmann et al., 2021; Arellano et al., 2021; McKinney et al., 2021) and of similar magnitude as in the UK and Sweden (Bell et al., 2021; Friedrich et al., 2021). While the qualitative pattern is the same for women, the

[25] Appendix D.2.2 describes this procedure in detail. Figures D.3 and D.4 compare key statistics of the earnings growth distributions conditional on current earnings in IAB and TPP data. Importantly, for intermediate levels of earnings, both datasets deliver highly consistent results. Figure D.5 shows the share of growth rates affected by top-coding in the IAB data. Figure D.2 shows the combined IAB-TPP distribution of earnings in the LS sample that is constructed analogously to the CS sample: Below the top-coding threshold we use the IAB distribution while above we use the TPP distribution rescaled to match the total number of workers above the threshold in the IAB data.

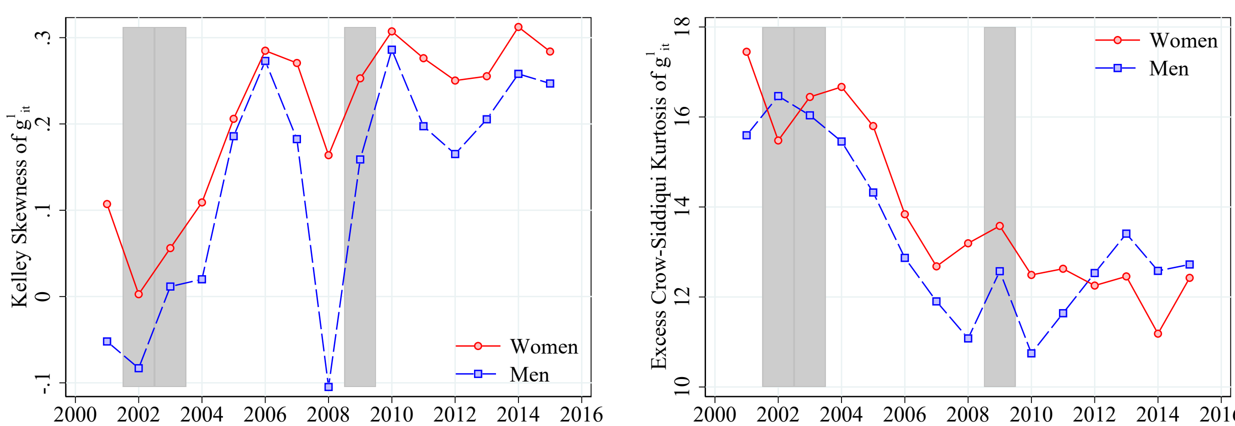
[26] As shown in Figure E.20, the median 1-year log-earnings change is approximately zero for all years and for both men and women. Hence, below-median earnings changes are negative and above-median earnings changes are positive.

cyclicality of earnings changes is much lower. This is consistent with Doepke and Tertilt (2016) and Alon et al. (2021) who show that recessions (prior to COVID-19) tend to affect men more severely.

Besides its pro-cyclicality, which has been documented before (e.g. Guvenen et al., 2014; Busch et al., forthcoming), it is noteworthy that after 2005 the Kelley skewness of earnings growth is more positive than before as well as compared to other countries (e.g. Italy, Spain, Sweden, UK, US – see Hoffmann et al. (2021); Arellano et al. (2021); Friedrich et al. (2021); Bell et al. (2021); McKinney et al. (2021)).[27] This reflects the good overall labor market conditions in Germany over this time period (Figure A.2). The low risk of becoming unemployed limits large earnings losses, compresses the P50-P10 differential, and increases the Kelley skewness. In addition, increasing labor supply may explain the relatively high share of large positive changes.[28] At the extensive margin, a previously unemployed person who starts working has significantly lower annual earnings in the first (incomplete) year than in the second (complete) year of an employment spell. At the intensive margin, workers increasing working hours experience a substantial increase in earnings.

Figure 8: Skewness and Excess Kurtosis of 1-Year Log Earnings Changes

(A) Kelley Skewness (B) Excess Crow-Siddiqui Kurtosis



Notes: This figure shows the evolution of Kelley skewness and excess Crow-Siddiqui kurtosis of 1-year changes in residualized log real annual earnings (from t to $t+1$) in the combined IAB-TPP data (LS sample) by gender. See Footnote 23 for definitions and interpretations of Kelley skewness and excess Crow-Siddiqui kurtosis. See Appendix D.2.2 for details on how we construct the distribution of log earnings growth from IAB and TPP data. Shaded areas indicate recessions.

Figure 8 (B) shows that earnings growth is substantially more leptokurtic than a Gaussian distribution with the same variance. Both the level and evolution of excess Crow-Siddiqui kurtosis are similar for men and women. For both, the kurtosis of earnings growth declined from about 16 to 12 between 2004 and 2007. Hence, medium-sized earnings changes became more important relative to large changes. For women, the decline is mainly driven by the strong increase in the P75-P25 differential of earnings growth that dominates the widening of the tails. For men, the decline is driven both by moderate increases in the P75-P25 and decreases in the P97.5-P2.5 differential.[29]

[27] The Kelley skewness of earnings growth is also positive when looking at 5-year changes (see Figure E.27).

[28] Figure A.2 shows a substantial increase in labor force participation after 2005 (especially for women). Moreover, transition probabilities from mini-job to part- or full-time increase by 5–10 percentage points (see Figure E.24).

[29] See Figure E.21 for a decomposition of the evolution of excess Crow-Siddiqui kurtosis.

3.3 Heterogeneity in Earnings Dynamics by Age and Permanent Earnings

We now study heterogeneity in earnings dynamics by age and workers' position in the permanent earnings distribution using the H sample. Permanent earnings P_{it} are defined as the residual of the log of average earnings between $t-2$ and t.[30] To avoid mechanical mean reversion, we analyze the distribution of (residual) log earnings changes between t and $t+1$ conditional on past permanent earnings $P_{i,t-1}$. As we are now interested in conditional earnings growth distributions, we use those from the IAB below and those from the TPP above the cutoff of 45,000 Euro.[31]

Figure 9 shows how different moments of the 1-year earnings growth distribution depend on age and the permanent earnings distribution. Panels A and B reveal that young workers' earnings are more volatile. However, while for men this is only relevant in the bottom half of the distribution, this holds across the entire distribution for women. The P90-P10 differential for young women is 30–50 log points higher compared to women above 35. Volatility is U-shaped in permanent earnings. Male workers in the bottom half and at the very top experience substantially larger changes than workers between the median and the 95th percentile of the distribution. For women, the P90-P10 decreases steadily until the 80th percentile and spikes back up above the 90th percentile.[32]

The Kelley skewness of earnings growth is decreasing in permanent earnings for both men and women (Panels C and D) albeit this gradient is larger for women. The fact that young women's Kelley skewness is smaller than for the other groups and even negative for the top 70% (Panel D) suggests that earnings volatility is higher for young women because they experience disproportionately many large earnings losses. This mainly reflects reduced female labor supply following childbirth. As children get older and mothers rejoin the labor force or switch from marginal to part-time or from part- to full-time jobs, they experience substantial earnings gains. Hence, the Kelley skewness of earnings changes for women between 35 and 55 is positive (except for the top 30% where it is close to zero). For men, the Kelley skewness measure does not vary as much over the lifecycle (Panel C). Consistent with men's concave lifecycle profile for median earnings (Figure 6), large positive changes become relatively less likely and the Kelley skewness drops as men get older.[33]

Panels E and F show that the relationship between excess Crow-Siddiqui kurtosis and permanent earnings differs between men and women. While earnings growth is more leptokurtic in the bottom half of the male distribution, the opposite is true for women. In terms of lifecycle profiles, older women and younger men differ relative to the other age groups. Especially for older women the excess kurtosis profile is flat across the earnings distribution.

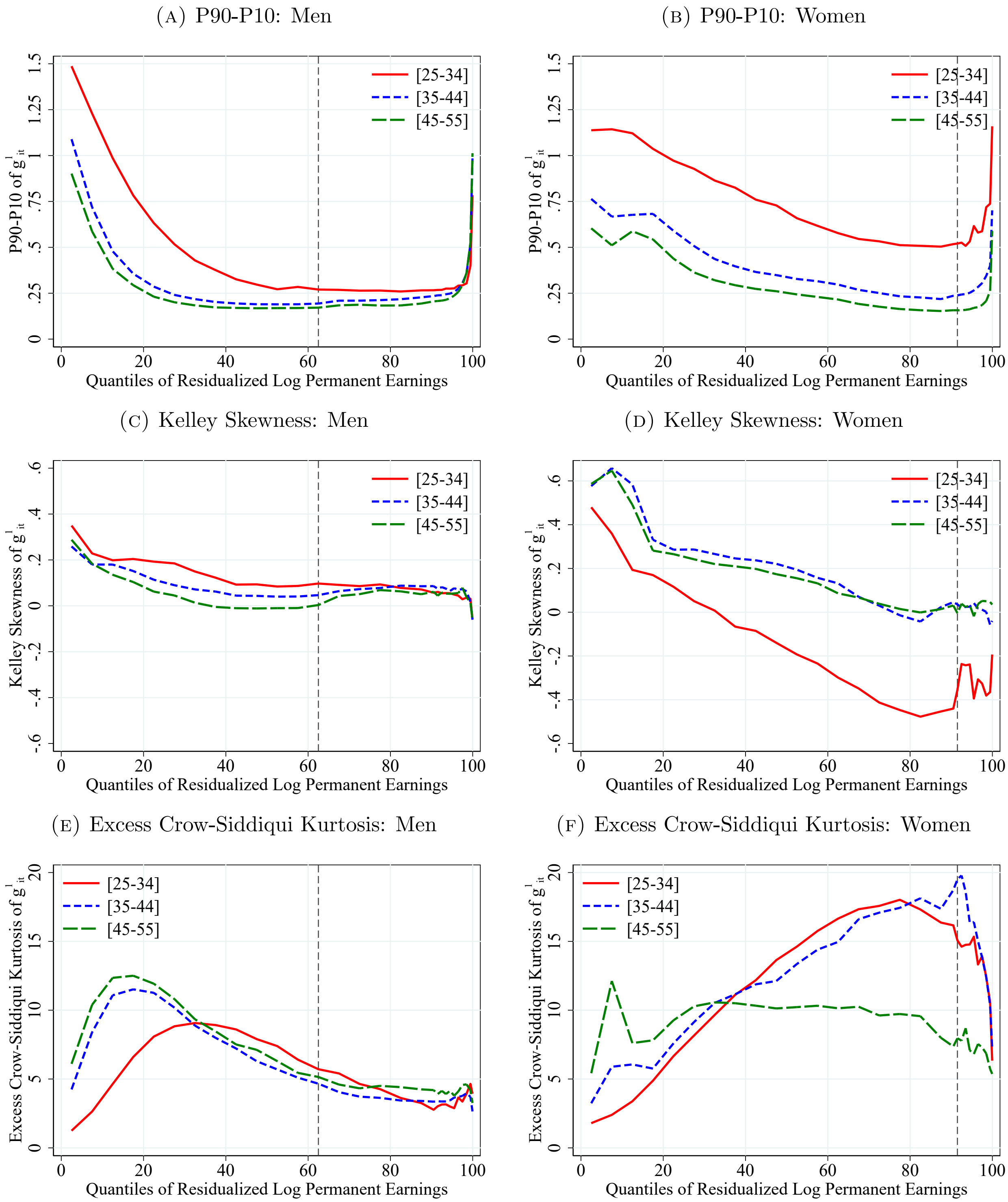
[30] We compute permanent earnings P_{it} by regressing the log of average past earnings on a full set of gender and year-specific age dummies, taking the residual and adding the raw sample mean.

[31] See Appendix D.3 for further details. In particular, we show that the distributions of permanent earnings in the IAB and TPP data are almost identical in the middle of the distribution and differ only at the bottom and very top. We also show the conditional earnings growth distribution by permanent earnings in both datasets separately.

[32] This tilted U-shape is consistent with findings for Sweden, Denmark, Norway, France, and the US (Friedrich et al., 2021; Leth-Petersen and Sæverud, 2021; Halvorsen et al., 2021; Kramarz et al., 2021; McKinney et al., 2021).

[33] Figure E.28 shows that this also holds for 5-year log earnings changes. The main difference is that earnings growth of male workers between 45 and 55 is negatively skewed.

Figure 9: Heterogeneity in Dispersion, Skewness and Kurtosis of 1-Year Log Earnings Changes



Notes: This figure shows the P90-P10 differential, Kelley skewness and excess Crow-Siddiqui kurtosis of 1-year changes in residualized log real annual earnings (from t to $t+1$) by quantiles of residualized permanent earnings and age groups in the combined IAB-TPP data (H sample) as averages from 2004 to 2011 by gender. Permanent earnings $P_{i,t-1}$ are defined as the residual (net of a full set of gender and year specific age dummies) of the log of average earnings between $t-3$ and $t-1$. See Footnote 23 for definitions and interpretations of Kelley skewness and excess Crow-Siddiqui kurtosis. See Figures D.7, D.8 and D.9 for a comparison of the underlying data in both data sources.

4 Inequality and Dynamics of Total Income

In this section, we move beyond labor earnings by adding entrepreneurs and non-labor income to the picture in order to study overall income inequality and income dynamics of all taxpayers.

Total Income and its Components. Total income is defined as the sum of labor and non-labor (i.e. self-employment, business or rental) income. While total income can be negative (less than 2% of the observations in our data), we still impose the same minimum income threshold of 2,300 Euro as for labor income. This has two reasons: First, the results are comparable to those for earnings. Second, we avoid computing growth rates between negative and positive values.[34]

Table 2 shows descriptive statistics for our analysis sample (for the year 2008). We exclude capital income (interest, dividends and capital gains from low-stake investments) from the total income analysis as there are several changes in capital income taxation that make the information on capital income unreliable after 2008.[35] Note, however, that capital income from high-stake investments (ownership share of at least 1%, "wesentliche Beteiligung") are included in the data as they count as business income. Table 2 (B) shows that capital income amounts to only 1.3% of total income for men and even less for women (0.5%) in 2008. As a comparison, other non-labor income accounts for 16.4% of men's and 10.6% of women's total income. Thus, omitting capital income should only have small effects on the total income distribution.[36]

Panels A and B of Figure 10 show the share of each component in total income conditional on total income. As expected, labor income accounts for the lion's share of total income except for very high incomes (above 200,000 Euro). Self-employment income is especially relevant for individuals with income between 100,000 to 1 million Euro. The share of business income rises continuously after 50,000 Euro and it becomes the dominant source after around 1 million Euro. While it is not surprising that the very rich receive mostly business income, the share of non-labor income is not monotonically increasing in total income. At the bottom, non-labor income accounts for up to 20% of total income. This is because, compared to labor income, we observe relatively more individuals with very high but also very low business income.[37] In other words, the distribution of non-labor income is much wider than that of labor income (see also Panels C and D of Figure 10).

Furthermore, compared to Section 3, the sample additionally includes non-social-security workers and individuals without labor but with non-labor income. We refer to individuals who receive most of their income from labor as "workers" and to those who receive at least half of their income

[34] In Figure D.14 we show the share of non-zero and negative values for each income component over time. The fact that there are both trends and breaks in the time series is one reason why we stay away from negative values. Table D.11 shows how our analysis sample compares to the data without the threshold.

[35] After the introduction of a dual income tax in 2009 ("Abgeltungsteuer-Reform") capital income largely disappeared from personal tax records.

[36] While capital income is most important at the top, its share never rises above 9% (see Figure 10, A and B). On average, the share is below 0.9% for men and 0.4% for women (averaged over 2001–2008; the values for 2008 in Table 2 (B) are the highest in this period) – independent of whether we exclude capital income from total income or whether we impose the minimum income threshold.

[37] For women, we observe many low-earnings observations due to the high share of mini-jobs.

Table 2: Summary Statistics for Total Income Data

	Men	Women
Observations (in mill.)	14.667	12.351
A. Income Distribution		
Mean	45,810	26,163
P50	36,620	21,698
P90	78,113	48,393
P99.9	696,521	270,828
P99.99	2,919,253	931,065
B. Share of Total Income		
Labor	0.836	0.894
Non-Labor	0.164	0.106
Self-Empl.	0.054	0.043
Business	0.109	0.053
Rental	0.001	0.011
Capital*	0.013	0.005
C. Main Income Source		
Workers	0.882	0.918
Entrepreneurs	0.118	0.082
Self-Employed	0.026	0.024
Business Owners	0.082	0.036
Landlords	0.010	0.021
D. Non-Zero Income		
Labor	0.895	0.934
Non-Labor	0.300	0.207
Self-Empl.	0.052	0.049
Business	0.165	0.080
Rental	0.144	0.102
Capital*	0.103	0.042

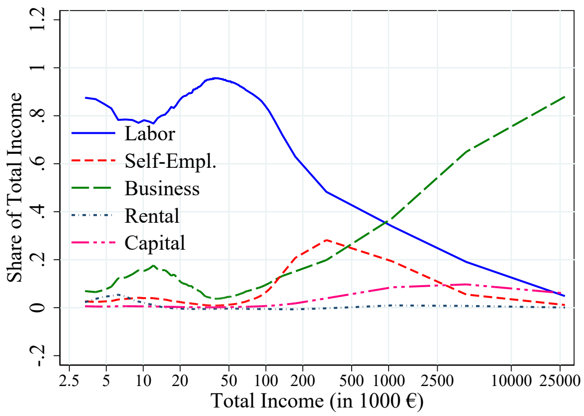
Notes: This table shows descriptive statistics for the combined IAB-TPP data by gender for the year 2008. The data includes all individuals with labor and non-labor income. See Appendix D for details on how we construct this combined dataset. Our analysis sample is restricted to individuals with total income (excluding capital income) above the minimum threshold of 2,300 Euro (2018 prices) and between 25 and 55 years of age. Panel A shows the mean and selected percentiles of the total income distribution (in 2018 Euro). Panel B shows the share of each income source in total income. Panel C reports the share of observations by main income source. Panel D shows the share of observations with non-zero income from different sources.

from non-labor sources as "entrepreneurs". Among the latter, a person is called "self-employed" if the main income source is self-employment, "business owner" if it is business income or "landlord" if it is rental income.[38] Panel C of Table 2 shows that 11.8% of men are entrepreneurs with the majority of them classified as business owners (8.2%) or self-employed (2.6%) and few landlords (1.0%). While the share of self-employed is similar for women (2.4%), fewer women are business owners (3.6%) and more are workers (91.8%) or landlords (2.1%). While this classification is intuitive, we

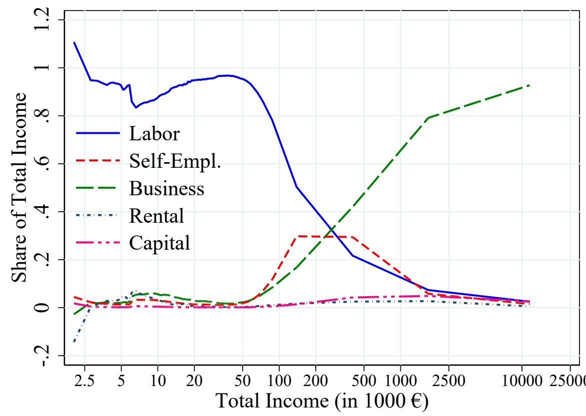
[38]The distinction between self-employed and business owners is somewhat special in the German tax code as both own their business. Self-employed are high-skilled independent professionals with special qualifications ("Freiberufler") such as doctors, lawyers, journalists or tax consultants. Any other self-employed individuals (e.g., delivery drivers) or owners of firms that produce or sell products are referred to as business owners ("Gewerbetreibende").

Figure 10: Income Components

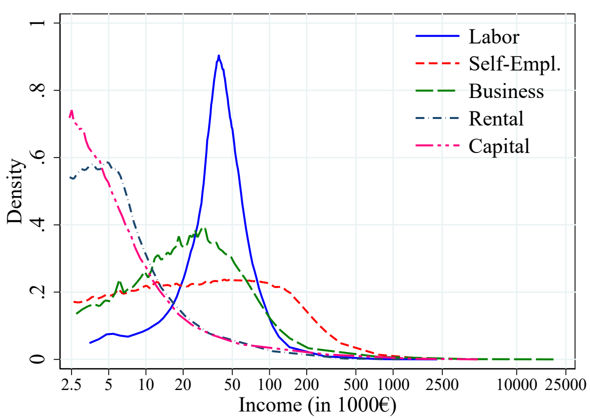
(a) Component Shares by Income: Men



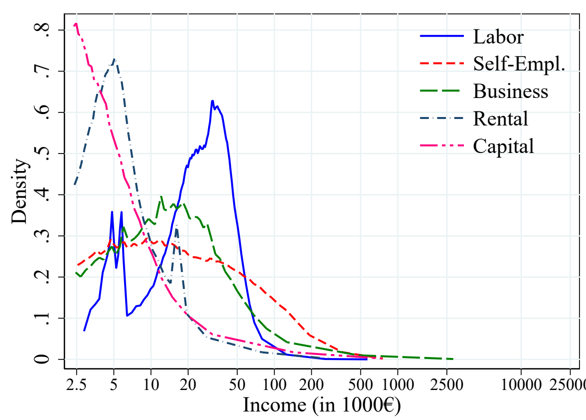
(b) Component Shares by Income: Women



(c) Densities: Men



(d) Densities: Women



Notes: This figure shows different statistics for the components of total income in the combined IAB-TPP data by gender (averages from 2001 to 2008). Panels A and B show how total income (including capital income) is split into labor, self-employment, business, rental and capital income across the total income distribution. Panels C and D show the densities of each income sources for all individuals with income from the respective source above 2,300 Euro (in 2018 Euro).

emphasize that non-labor income is not only relevant for entrepreneurs but also for workers. Indeed, Table 2 (D) shows that 30% of men and 20.7% of women have some non-labor income.

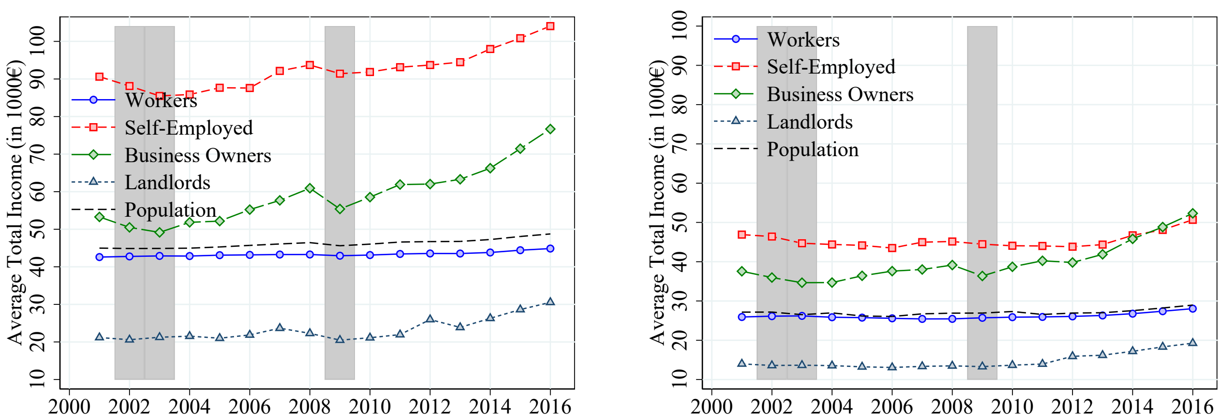
4.1 The Distribution of Total Income

We now turn to the evolution of the total income distribution over time. Total income inequality is driven by income differences among workers (as discussed in Section 3) or entrepreneurs, but also by income differences between those groups.

Figure 11 shows that average incomes differ substantially between workers and entrepreneurs. While male workers (Panel A) receive average annual income of around 43,000 Euro in 2001, business owners and self-employed receive 53,000 and 90,000 Euro respectively. Landlords, on the other hand, earn only 20,000 Euro in 2001. Panel B reveals a similar pattern for women. However, there is a substantial gender gap in average income not just among workers but also among self-employed (53.3%), business owners (34.8%) and landlords (40.3%). Strikingly, the overall gender gap increased from 39.6% in 2001 to 40.6% in 2016 despite the fact that the gender gap among workers decreased from 39.1% to 37.5%.[39] Hence, the gender gap among non-workers has increased significantly.

FIGURE 11: AVERAGE TOTAL INCOME BY MAIN COMPONENT

(A) Men (B) Women



Notes: This figure shows average total income (excluding capital income) for different sub-populations defined by the main income source as well as the full population (gray dashed line) over time in the combined IAB-TPP data by gender. Shaded areas indicate recessions. Figure G.2 shows the evolution in terms of log differences relative to 2001.

Over time, we find that average incomes of workers are remarkably flat compared to entrepreneurs.[40] For men, average incomes of self-employed and business owners decrease in the recession of 2002 and 2003, but then increase significantly. While workers' incomes only rise from around 43,000 to 45,000 Euro between 2003 and 2016, business owners' incomes grow from around 50,000 to 75,000 Euro and incomes of self-employed men go from roughly 85,000 to 105,000 Euro over that period. Average incomes of landlords also increase by much more than workers' incomes but the increase takes place only after the Great Recession (possibly due to a combination of lower mortgage interest payments and rising real estate prices). Panel B shows that female business owners and landlords exhibit a similarly strong increase in average income, while average income of self-employed women is essentially flat between 2003 and 2013 before it starts to increase. Overall, these differences between entrepreneurs and workers underline the importance of looking at the distribution of total income to get a more complete picture of income inequality over time.

Table D.12 in the Appendix shows selected percentiles of total income. While percentiles below the 90th percentile are very similar to those of labor income (Table 1), the right tail of the distribution is much longer for total incomes. In 2016, for example, the 90th percentile of total income for men is about 9% higher than that of labor income (84 vs. 77 Euro). This gap grows substantially as we move to the very top of the distribution. While the 99.99th percentile of labor income is around

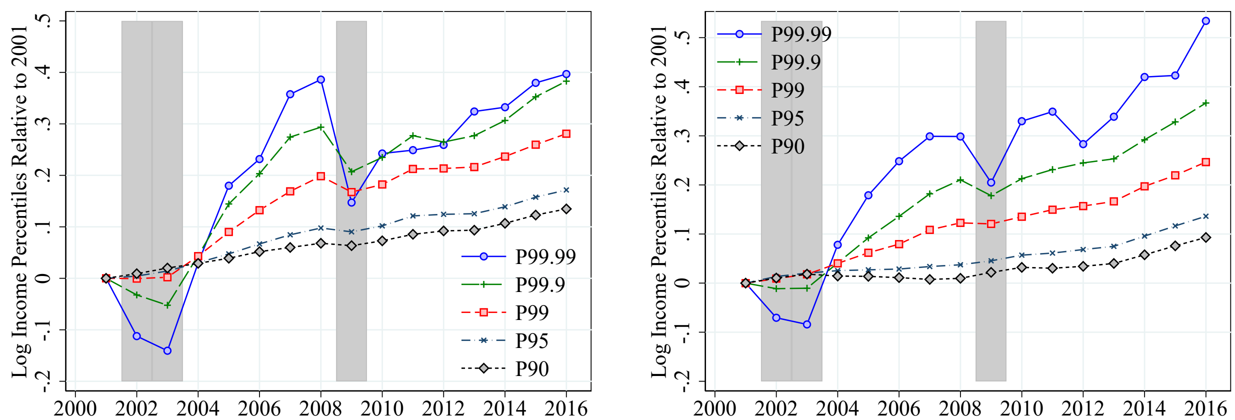
[39] In the data on social security workers in Section 3, the gender gap in average earnings decreased even more strongly from 42% to 37% between 2001 and 2016 (see Table D.1).

[40] Figure G.2 shows the evolution of average income by main income source as log differences relative to 2001.

1.2 million Euro, it is close to 3 million Euro for total income – a factor of 2.5. For women, this factor is even larger (2.8) although the levels are considerably lower (400,000 vs. about 1.2 million Euro). These findings again highlight the large gender differences throughout the distribution and especially at the top: While men earn more than women at all percentiles, at the very top men's incomes are almost three times as large as women's.

Figure 12: Evolution of Top Log Total Income Percentiles

(A) Men (B) Women



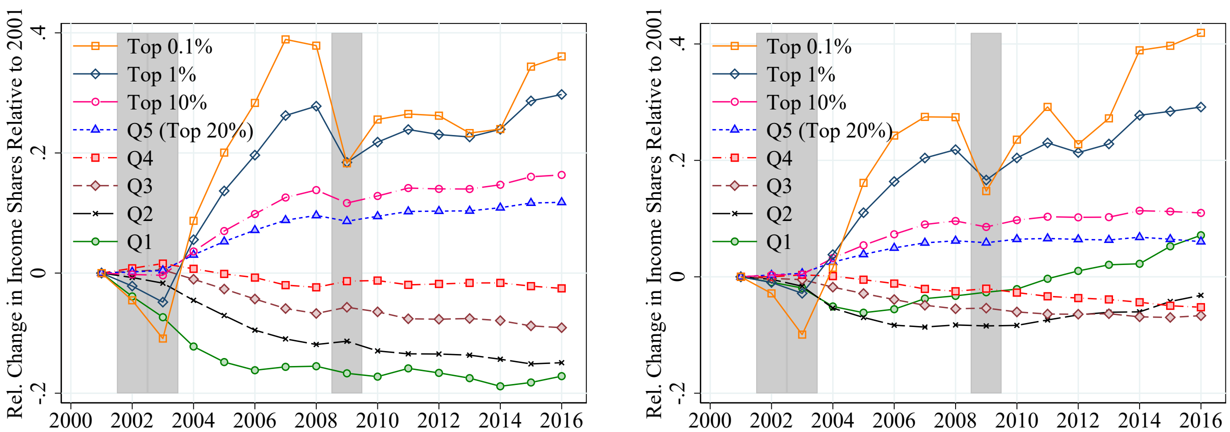
Notes: This figure shows the evolution of selected top percentiles of log real annual total income (relative to 2001) in the combined IAB-TPP data (CS sample) by gender. Shaded areas indicate recessions. See Figure 3 for the same analysis of only labor earnings (albeit for a slightly different sample as discussed in the text) and Figure G.3 for more percentiles of the total income distribution.

To analyze of the changes in total income inequality over time, Figure 12 shows the evolution of selected top percentiles relative to 2001. Not surprisingly, lower percentiles of the overall distribution appear very similar to those for labor income (compare Panels A and B in Figures G.3 and 3). However, at the top of the distribution (Figure 12) total income percentiles grow much faster. For example, while at the 99.99th percentile men's labor incomes grew by 20 log points, total income grew by around 40 log points. For women the growth at the top of total income is also stronger than for labor income but the difference is not as pronounced as for men. Indeed while top earning women seemed to clearly catch up with men based on Figure 3 (C) and (D), in Figure 12 (A) and (B) only the 99.99th percentile grows faster than for men, while all other top percentiles grow less.

Taking a closer look at the very top of the income distribution, it becomes apparent that total incomes show larger fluctuations over the business cycle than labor earnings. Especially the P99.99 (i.e. the top 0.01%) of men are hit by large negative shocks of 15 (in 2003) and 25 (in 2009) log points in the two recessions in our period of analysis. While top income men recover very quickly after the 2003 recession, the recovery is more muted after 2009 and only reaches the 2008 level again in 2016. The top 0.1% also show a similar pattern albeit of lesser magnitude while the top 1% grew relatively steadily over time. Business income at the high end is of course dependent on the profitability of the respective businesses which may fluctuate much more over the cycle than wages. Business income may also be affected by deductions of negative incomes in subsequent years.

Figure 13: Changes in Total Income Shares Relative to 2001 – Analysis Sample

(a) Men (b) Women



Notes: This figure shows the evolution of selected income shares of real annual total income (relative to 2001) in the combined IAB-TPP data (CS sample) by gender. The absolute shares for all groups can be found in Appendix Tables G.2, G.3 and G.4 for men, women and the entire population. Shaded areas indicate recessions.

As an alternative way to highlight changes in the income distribution, Figure 13 shows the relative changes in income shares of various groups of the total income distribution.[41] For men (Panel A), the bottom 4 quintiles show substantial losses, especially in the 2001-2008 window. For example, the bottom quintile share fell from 5.6 to just 4.6% (or by around 10%). By contrast, the income share of the top quintile rose from 42 to 47% (or by around 11% relative to 2001). For women (Panel B) the increase at the top is somewhat more muted from 42 to 45%. The relative increase of the lowest quintile (Q1) is striking, however this is from a low level (4 to 4.2%). Zooming in at the very top shows a similar strong (cyclical) increase as before both for men and women.

4.2 Income Dynamics for Workers and Entrepreneurs

We now study the distribution of total income growth, measured as 1-year changes in residualized log total income using the IAB-reweighted TPP data (as the IAB data itself does not contain information on total income)[42] for workers and entrepreneurs.[43] For both groups we analyze 1-year

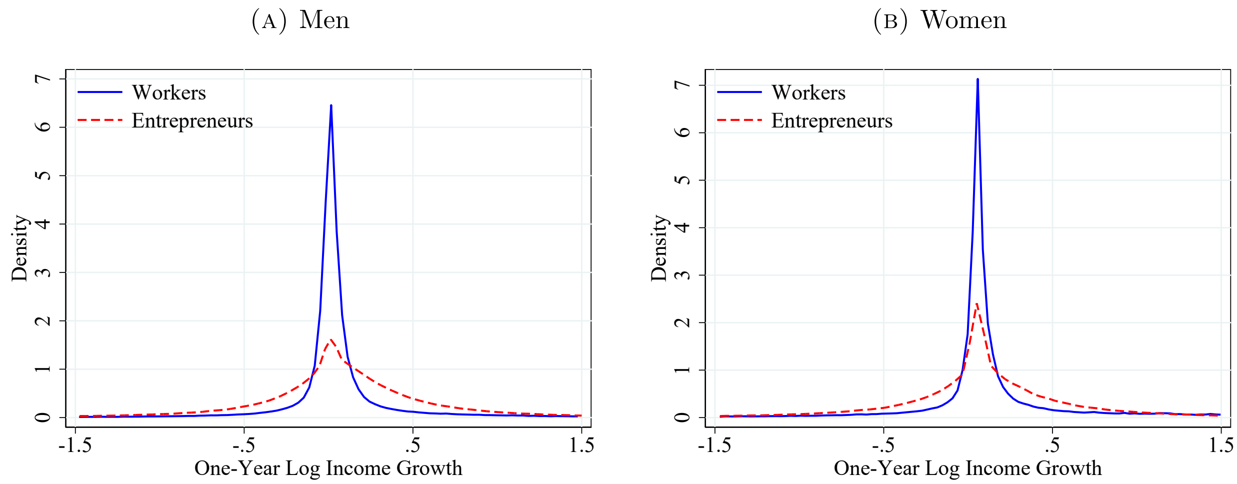
[41] Our results are consistent with those of Bartels (2019) or the World Inequality Database, though our top shares are a bit lower. Differences are a narrower age window (25–55 vs 20–60), analyzing individuals instead of tax units as well as not imputing capital incomes. Jenderny and Bartels (2015) analyze how the absence of capital incomes from the post-2009 tax data affects top income shares.

[42] This means that we cannot use the procedure used in Section 3.2 to correct for non-random attrition in the TPP before 2012. However, this only affects pure labor income earners as observations with any non-labor income always have to file a tax return. In addition, attrition will be more problematic for women than for men as most mini-jobs are missing in the TPP. However, the evolution (and levels) of the statistics of total income growth for workers reported in this section are consistent with the results for earnings growth in Section 3.2. In addition, there are no clear breaks from 2011 to 2012 when most of the attrition problem vanishes with the inclusion of non-filers in the TPP.

[43] The distribution of 1-year income changes in the population is very similar to that of workers as 88% of men and 92% of women are classified as workers. In Appendix G, we show that income changes of self-employed are slightly more similar to those of workers but still much more similar to those of other business owners.

changes in *total* income. The densities in Figure 14 give a first indication that the distribution of 1-year income growth differs substantially between workers and entrepreneurs.

Figure 14: Density of 1-Year Income Growth by Main Income Source (Year 2005)



Notes: This figure shows the density of 1-year changes of residualized log total income separately for workers (labor income as main income source) and entrepreneurs (non-labor income as main income source) for men and women in the combined IAB-TPP data (LS sample) for the year 2005.

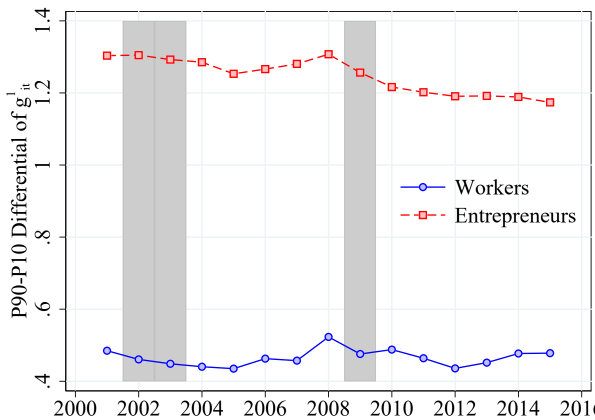
Figure 15 compares the evolution of the P90-P10 differential, Kelley skewness and excess Crow-Siddiqui kurtosis of 1-year income growth between workers and entrepreneurs. Focusing on men, we find that entrepreneurs' income are substantially more volatile (Panel A). That is, the P90-P10 differential is three times as large as that of male workers (see Panel A of Figure 15). This means that about 80% of workers' 1-year income changes are smaller than 25 log points while about 20% of entrepreneurs' 1-year income changes exceed 70 log points (in absolute value). Panel C shows that while workers' log income changes are (mostly) positively skewed, Kelley skewness of entrepreneurs' income changes is essentially zero. Perhaps surprisingly, male entrepreneurs' log income changes are far less cyclical than male workers' income changes. In particular, during the Great Recession Kelley skewness dropped sharply for workers but only mildly for entrepreneurs. Panel E documents that entrepreneurs' income changes are much less leptokurtic than workers' as excess Crow-Siddiqui kurtosis of 1-year income growth is around 3 throughout the sample period.[44] Strikingly and in contrast to workers, we find that gender differences play almost no role for the dispersion, skewness and excess kurtosis of entrepreneurs' income growth. Note that this is true despite the fact that we observe substantial gender gaps in non-labor income (Figure 11).

In Figure 16, we show how our percentile-based measures of dispersion, skewness, and excess kurtosis of total income growth depend on permanent income. As in the core analysis of earnings changes, permanent income is defined as the residual of the log of average past income where we take out gender and year specific age effects. Again, we compare entrepreneurs to workers. The

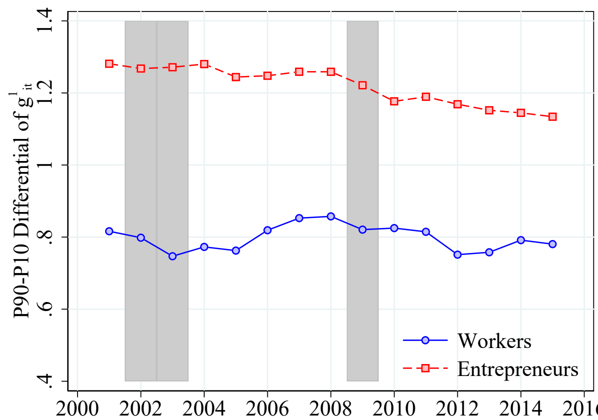
[44] Figure G.8 shows that this difference is mostly driven by the shoulders instead of the tails of the distribution. This means that entrepreneurs experience intermediate income changes more frequently than workers.

Figure 15: Dispersion, Skewness and Kurtosis of 1-Year Log Income Changes

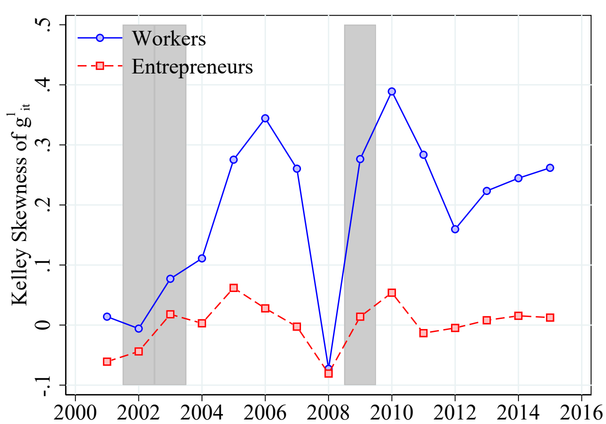
(a) P90–P10: Men



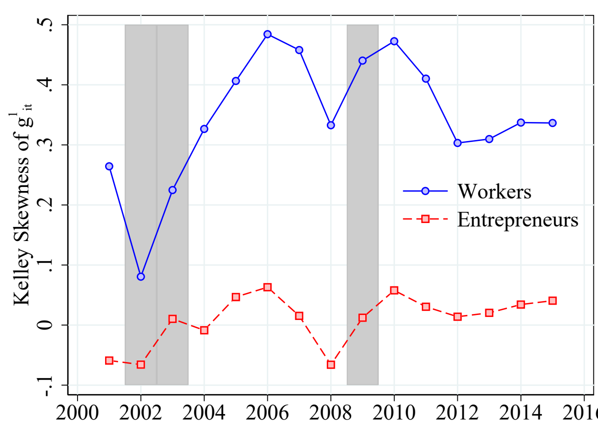
(b) P90–P10: Women



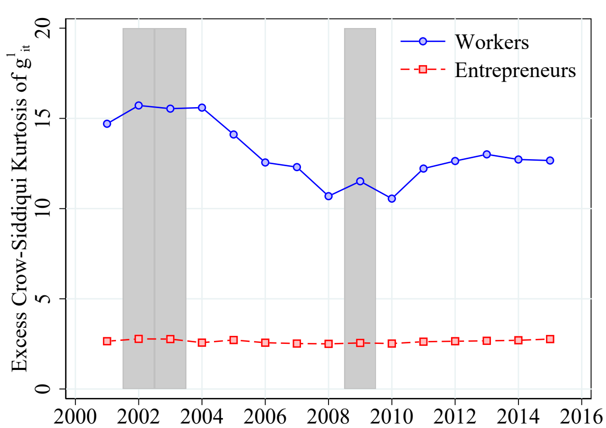
(c) Kelley Skewness: Men



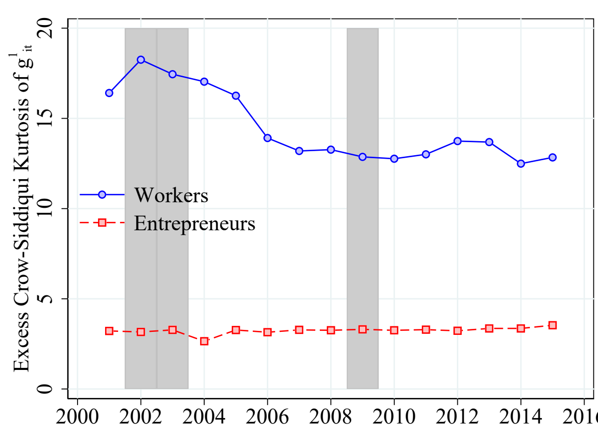
(d) Kelley Skewness: Women



(e) Excess Crow-Siddiqui Kurtosis: Men



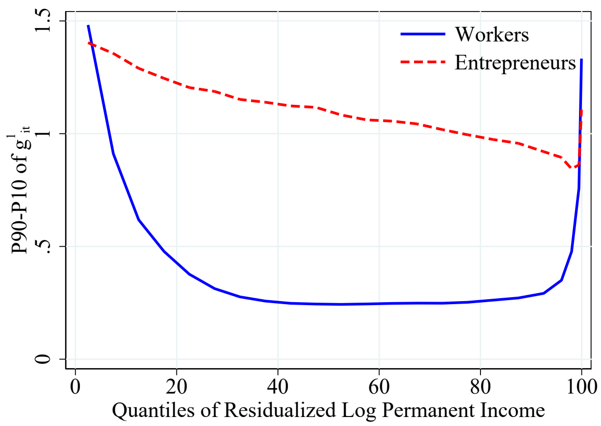
(f) Excess Crow-Siddiqui Kurtosis: Women



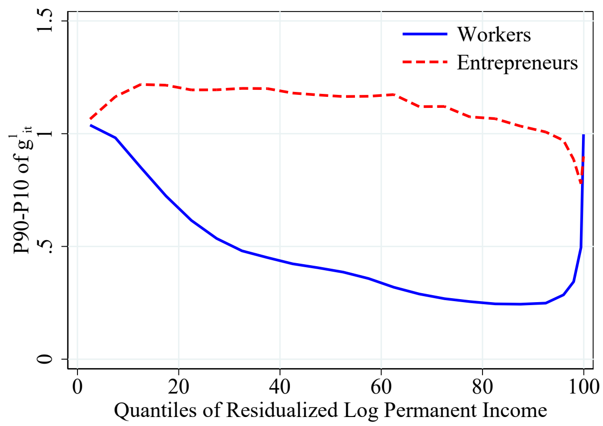
Notes: This figure shows the evolution of P90-P10 differentials, Kelley skewness and excess Crow-Siddiqui kurtosis of 1-year changes in residualized log real annual total income (from t to $t+1$) in the combined IAB-TPP data (LS sample) by main income source (labor, non-labor) and gender. A person's main income source is labor whenever more than 50% of her income comes from (dependent) employment instead of self-employment, corporate business income or rental income. See Footnote 23 for definitions and interpretation of Kelley skewness and excess Crow-Siddiqui kurtosis. Shaded areas indicate recessions.

Figure 16: Heterogeneity in Dispersion, Skewness and Kurtosis of 1-Year Log Income Growth by Main Income Source

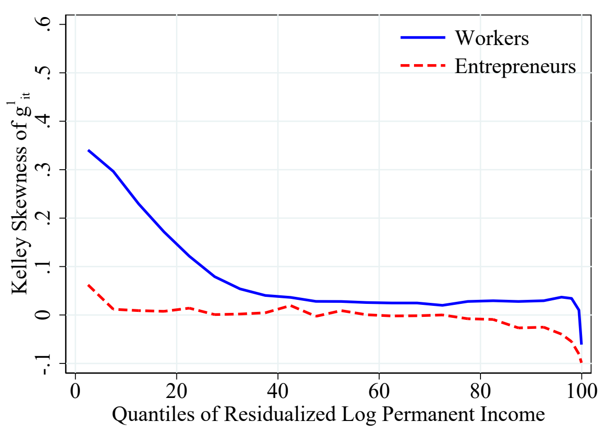
(a) P90–P10: Men



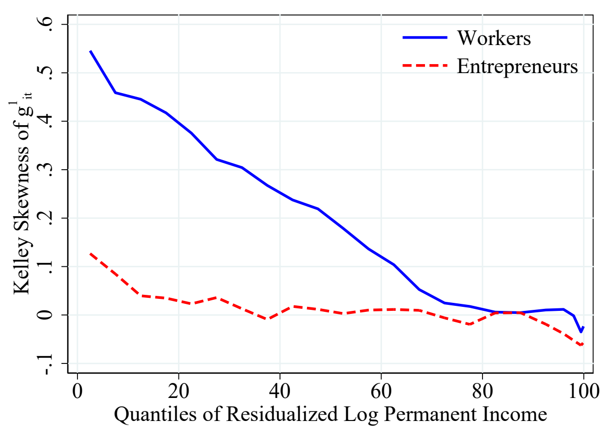
(b) P90–P10: Women



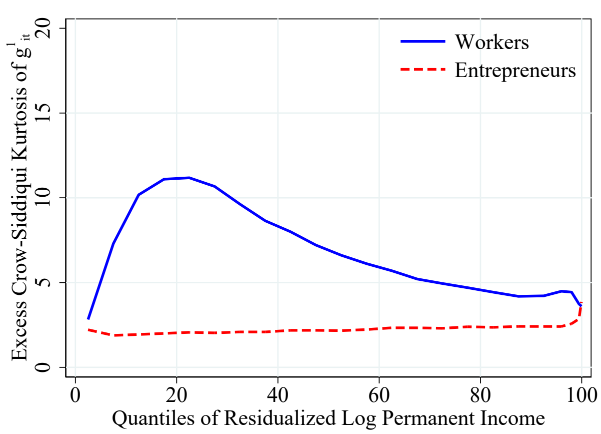
(c) Kelley Skewness: Men



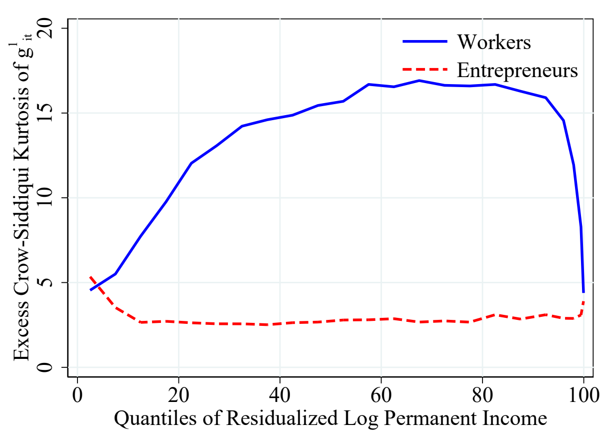
(d) Kelley Skewness: Women



(e) Excess Crow-Siddiqui Kurtosis: Men



(f) Excess Crow-Siddiqui Kurtosis: Women



Notes: This figure shows the P90-P10 differential, Kelley skewness and excess Crow-Siddiqui kurtosis of 1-year changes in residualized log real total income by quantiles of the distribution of permanent total income (from t to $t+1$) in the combined IAB-TPP data (H Sample) as averages from 2004 to 2011 by main income source (labor, non-labor) and gender. The (gender-specific) ranking of permanent income is based on the distribution of total income of both workers and entrepreneurs. A person's main income source is labor (worker) whenever more than 50% of her income comes from (dependent) employment instead of self-employment, corporate business income or rental income. See Footnote 23 for definitions and interpretation of Kelley skewness and excess Crow-Siddiqui kurtosis. Shaded areas indicate recessions.

P90-P10 differential of male entrepreneurs' income growth decreases steadily in permanent income but spikes up for the top 1%. While heterogeneity for workers is more pronounced, the difference in the P90-P10 between entrepreneurs in the bottom 10% and the top 10% (without the top 1%) still differs by about 50 log points. For female entrepreneurs, this relationship is hump-shaped. However, this difference compared to male entrepreneurs is largely driven by differences in the distribution of permanent income (see Figure G.11). In stark contrast to workers, there is almost no dependence on permanent income for Kelley skewness and excess Crow-Siddiqui kurtosis of entrepreneurs' income growth. Overall, the fact that excess Crow-Siddiqui kurtosis is much closer to zero and that Kelley skewness is close to zero implies that entrepreneurs' incomes change are more similar to a Gaussian distribution than those of workers. In addition, the moments of entrepreneurs' income growth exhibit much less state-dependence with respect to (permanent) income than those of workers.

4.3 Top Income Mobility

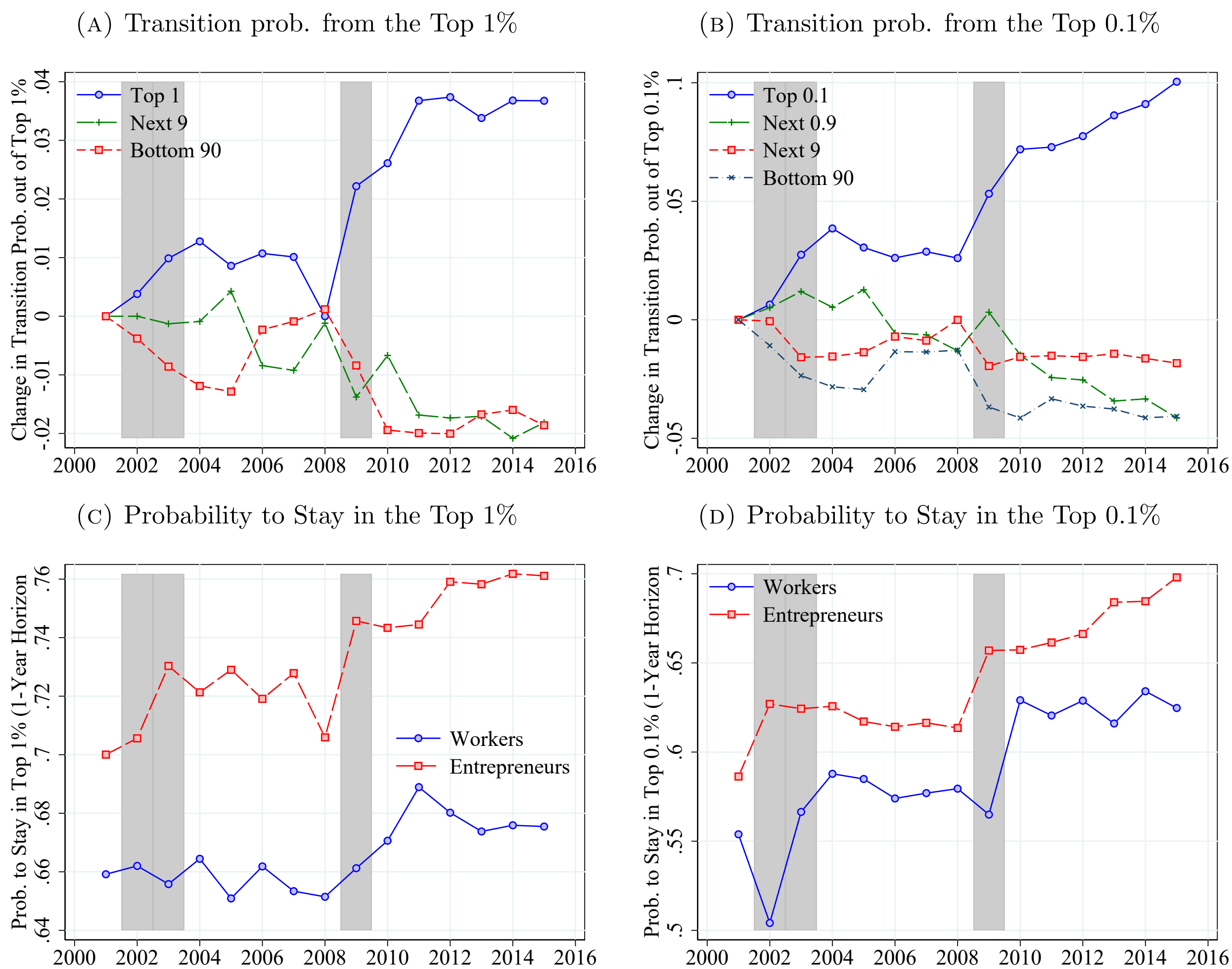
In this final part of our analysis, we ask how the probability of someone someone dropping out the top 1% or 0.1% of the (age-specific) income distribution has evolved over time. Panel A of Figure 17 shows that the probability of staying in the top 1% of the income distribution has increased by 4 percentage points between 2001 and 2016.[45] Both the probability of falling into the next 9%, i.e. between the 90th and 99th percentile of the income distribution, as well as the probability of falling out of the top 10% have decreased. This decrease in top income mobility seems to have occurred mainly around the two recessions in our sample.

This becomes even more evident in Panel B which shows the evolution of transition probabilities of income earners in the top 0.1% whose probability to stay at the very top has increased by 10 percentage points. After the Great Recession, the continued increase in the probability to stay is mirrored by a decline in the probability to fall into the next 9% instead of the bottom 90%.

Panels C and D show the probabilities to stay in the top 1% and top 0.1% separately for workers and entrepreneurs. Workers are more likely to fall out of the top than entrepreneurs. In addition, this gap has widened after 2010 when entrepreneurs' likelihood to stay at the top kept increasing while workers' did not. While we focus on 1-year transition rates here, we show in Figure G.13 that the same patterns (both qualitatively and quantitatively) hold also for 5-year transition probabilities. Both magnitudes and trends reported here are very similar to the findings for US workers studied in Guvenen et al. (2021a). If anything, German top income earners are slightly more likely to stay at the top.

[45] In 2001, the probability of staying in the top 1% is equal to 0.676 and the corresponding probabilities to fall into the next 9% and bottom 90% are 0.257 and 0.067 respectively.

Figure 17: Top Income Mobility – 1-Year Transition Probabilities



Notes: This figure shows statistics on top income mobility in the combined IAB-TPP data (LS sample but including negative incomes). Panels A and B show the evolution of 1-year transition probabilities from the top 1% and top 0.1% of the income distribution into selected parts of the income distribution from one year to the next. The rankings are computed conditional on age. The "Top 1" ("Top 0.1%")is the probability of staying in the Top 1% (0.1%). The "Next 9" is the part of the distribution between the P90 and P99 and the "Next 0.9" is the part between the P99 and the P99.9. The lines sum to zero. Panels C and D show the 1-year probability of staying in the top 1% or top 0.1% for workers and entrepreneurs. The ranking is based on the total income distribution and not conditional on the main income source. Shaded areas indicate recessions. Figure G.13 shows the same statistics for 5-year transition probabilities.

5 Conclusion

This paper provides a comprehensive analysis of inequality and income dynamics for Germany over the last two decades. By combining two high quality administrative data sources – personal income tax and social security records – this is the first paper to offer a complete picture of the German income distribution ranging from the very bottom to the very top.

The first part of the analysis focuses on labor earnings, which is the main source of income for the vast majority of individuals and is most easily compared across datasets and countries. We find that earnings inequality among men has been increasing over the entire sample period from 2001 to 2016 and in particular before the Great Recession – a period where only the top 25% experienced real earnings growth and the bottom half realized real earnings losses of 5 to 20 % and more. After the Great Recession, earnings below the median stabilized while those above the median continued to grow. For women, the evolution of earnings inequality is a tale of two halves. While bottom inequality has been falling due to strong earnings growth at the bottom, top inequality has been rising.

A striking finding is that female top earners (above the 90th percentile) have seen the strongest growth in real earnings. In fact, women's earnings have been catching up with male earnings throughout most of the distribution. This happened even though the share of women working full-time has been declining. While we provide some evidence on how changes in observable job and worker characteristics have contributed to this, a fruitful avenue for future research could be to disentangle the economic forces behind this closing of the gender earnings gap.

Our analysis of individual earnings changes reveals that the earnings growth distribution has been significantly skewed to the right for both men and women since 2005 (apart from the Great Recession). This is in contrast to various other countries (e.g. Italy, Spain, Sweden, UK, US). The fact that large positive gains are more likely than large negative shocks reflects the low risk of job-loss and increasing labor force participation.

In the second part of the paper, we investigate how taking the incomes of self-employed, business owners and landlords into account enriches the overall picture on total income inequality. While workers' incomes are more stable over the business cycle, non-labor incomes have increased substantially relative to labor incomes. Between 2001 and 2016, average incomes of workers grew by around 5% while average incomes of entrepreneurs increased by around 25%. Hence, total income inequality is higher and increased more strongly than labor income inequality.

Our analysis also shows that there exist large gender differences in non-labor incomes. While we find some convergence over time, we document large gaps between men and women at the very top of the total income distribution driven by women being less likely to have high business incomes.

Finally, we contrast income dynamics of workers with those of entrepreneurs and find that the latter face significantly more volatility. From a modeling point of view, non-labor income risk is much better approximated by a normal distribution and exhibits less dependence on age or permanent income.

References

Alon, T., S. Coskun, M. Doepke, D. Koll, and M. Tertilt (2021): "From Mancession to Shecession: Women's Employment in Regular and Pandemic Recessions," NBER Working Papers 28632.

Antonczyk, D., T. DeLeire, and B. Fitzenberger (2018): "Polarization and Rising Wage Inequality: Comparing the U.S. and Germany," *Econometrics*, 6.

Arellano, M., S. Bonhomme, M. D. Vera, L. Hospido, and S. Wei (2021): "Income Risk Inequality: Evidence from Spanish Administrative Record," GID Working Paper.

Autor, D., D. Dorn, L. F. Katz, C. Patterson, and J. Van Reenen (2020): "The Fall of the Labor Share and the Rise of Superstar Firms*," *The Quarterly Journal of Economics*, 135, 645–709.

Bach, S., G. Corneo, and V. Steiner (2009): "From bottom to top: The entire distribution of market income in Germany, 1992–2003," *Review of Income and Wealth*, 55, 303–330.

Bartels, C. (2019): "Top incomes in Germany, 1871–2014," *The Journal of Economic History*, 79, 669–707.

Bayer, C. and F. Juessen (2012): "The life-cycle and the business-cycle of wage risk — Cross-country comparisons," *Economics Letters*, 117, 831–833.

Bell, B., N. Bloom, and J. Blundell (2021): "Income Dynamics in the United Kingdom 1975-2020," GID Working Paper.

Biewen, M. (2000): "Income Inequality in Germany during the 1980s and 1990s," *Review of Income and Wealth*, 46, 1–19.

Biewen, M., B. Fitzenberger, and J. de Lazzer (2018): "The role of employment interruptions and part-time work for the rise in wage inequality," *IZA Journal of Labor Economics*, 7.

Biewen, M., M. Ungerer, and M. Löffler (2019): "Why Did Income Inequality in Germany Not Increase Further After 2005?" *German Economic Review*, 20, 471–504.

Bossler, M. and T. Schank (2020): "Wage inequality in Germany after the minimum wage introduction," LASER Discussion Papers 117.

Bradley, J. and A. Kügler (2019): "Labor market reforms: An evaluation of the Hartz policies in Germany," *European Economic Review*, 113, 108–135.

Brenke, K., U. Rinne, and K. F. Zimmermann (2013): "Short-time work: The German answer to the Great Recession," *International Labour Review*, 152, 287–305.

Brüll, E. and C. Gathmann (2020): "Evolution of the East German Wage Structure," ZEW Discussion Paper 20-081.

Bundesinstitut für Bevölkerungsforschung (2021): "Nach Altersgruppen zusammengefasste Geburtenziffern deutscher Frauen in Deutschland (1991-2019)," https://www.bib.bund.de/Permalink.html?id=10244004, accessed: 2022-01-04.

Busch, C., D. Domeij, F. Guvenen, R. Madera, et al. (forthcoming): "Skewed Idiosyncratic Income Risk over the Business Cycle: Sources and Insurance," *American Economic Journal: Macroeconomics*.

Bönke, T., G. Corneo, and H. Lüthen (2015): "Lifetime Earnings Inequality in Germany," *Journal of Labor Economics*, 33, 171–208.

Bönke, T., U. Glogowosky, E. Hansen, H. Lüthen, and D. Sachs (2022): "Family Policies and Child-Related Earnings Gaps in Germany," Mimeo.

Caliendo, M. and L. Wittbrodt (2021): "Did the Minimum Wage Reduce the Gender Wage Gap in Germany?" workingpaper 40, Center for Economic Policy Analysis (CEPA).

Card, D., J. Heining, and P. Kline (2013): "Workplace Heterogeneity and the Rise of West German Wage Inequality," *The Quarterly Journal of Economics*, 128, 967–1015.

Dauth, W. and J. Eppelsheimer (2020): "Preparing the sample of integrated labour market biographies (SIAB) for scientific analysis: a guide," *Journal for Labour Market Research*, 54.

DiNardo, J., N. Fortin, and T. Lemieux (1996): "Labor Market Institution and the Distribution of Wages," *Econometrica*, 1001–1044.

Doepke, M. and M. Tertilt (2016): "Families in Macroeconomics," Elsevier, vol. 2, chap. Chapter 23, 1789–1891.

Doerrenberg, P., A. Peichl, and S. Siegloch (2017): "The elasticity of taxable income in the presence of deduction possibilities," *Journal of Public Economics*, 151, 41–55.

Dolls, M., P. Doerrenberg, A. Peichl, and H. Stichnoth (2018): "Do retirement savings increase in response to information about retirement and expected pensions?" *Journal of Public Economics*, 158, 168–179.

Dustmann, C., B. Fitzenberger, U. Schönberg, and A. Spitz-Oener (2014): "From Sick Man of Europe to Economic Superstar: Germany's Resurgent Economy," *Journal of Economic Perspectives*, 28, 167–88.

Dustmann, C., A. Lindner, U. Schönberg, M. Umkehrer, and P. vom Berge (2022): "Reallocation Effects of the Minimum Wage," *The Quarterly Journal of Economics*, 137, 267–328.

Dustmann, C., J. Ludsteck, and U. Schönberg (2009): "Revisiting the German wage structure," *The Quarterly Journal of Economics*, 124, 843–881.

Ellguth, P. and S. Kohaut (2019): "Tarifbindung und betriebliche Interessenvertretung. Aktuelle Ergebnisse aus dem IAB-Betriebspanel 2018," *WSI-Mitteilungen*, 72, 290–297.

——— (2020): "Tarifbindung und betriebliche Interessenvertretung: Aktuelle Ergebnisse aus dem IAB-Betriebspanel 2019," *WSI-Mitteilungen*, 73.

Fitzenberger, B. and A. Seidlitz (2020): "The 2011 break in the part-time indicator and the evolution of wage inequality in Germany," *Journal for Labour Market Research*, 54.

Friedrich, B., L. Laun, and C. Meghir (2021): "Income Dynamics in Sweden 1985-2016," GID Working Paper.

Frodermann, C., A. Schmucker, S. Seth, and P. vom Berge (2021): "Sample of Integrated Labour Market Biographies (SIAB) 1975-2019," FDZ-Datenreport 01, Nuremberg.

Fuchs-Schündeln, N., D. Krueger, and M. Sommer (2010): "Inequality trends for Germany in the last two decades: A tale of two countries," *Review of Economic Dynamics*, 13, 103–132, special issue: Cross-Sectional Facts for Macroeconomists.

Fuest, C., A. Peichl, and S. Siegloch (2018): "Do Higher Corporate Taxes Reduce Wages? Micro Evidence from Germany," *American Economic Review*, 108, 393–418.

Gehrke, B., W. Lechthaler, and C. Merkl (2019): "The German Labor Market during the Great Recession: Shocks and Institutions," IZA Discussion Papers 11858.

Goldschmidt, D. and J. F. Schmieder (2017): "The rise of domestic outsourcing and the evolution of the German wage structure," *The Quarterly Journal of Economics*, 132, 1165–1217.

Grabka, M. M. (2021): "Income Inequality in Germany Stagnating over the Long Term, But Decreasing Slightly during the Coronavirus Pandemic," DIW Weekly Report 17/18, IW.

Gudgeon, M. and S. Trenkle (2020): "The speed of earnings responses to taxation and the role of firm labor demand," IZA Discussion Paper 13931.

Guvenen, F., G. Kaplan, and J. Song (2021a): "The Glass Ceiling and the Paper Floor: Changing Gender Composition of Top Earners since the 1980s," *NBER Macroeconomics Annual*, 35, 309–373.

Guvenen, F., F. Karahan, S. Ozkan, and J. Song (2021b): "What do data on millions of US workers reveal about life-cycle earnings risk?" *Econometrica*, 89, 2303–2339.

Guvenen, F., S. Ozkan, and J. Song (2014): "The Nature of Countercyclical Income Risk," *Journal of Political Economy*, 122, 621–660.

Halvorsen, E., S. Ozkan, and S. Salgado (2021): "Earnings Dynamics and Its Intergenerational Transmission: Evidence from Norway," GID Working Paper.

Hartung, B., P. Jung, and M. Kuhn (2018): "What Hides Behind the German Labor Market Miracle? Unemployment Insurance Reforms and Labor Market Dynamics," CESifo Working Paper Series 7379, CESifo.

Hochmuth, B., B. Kohlbrecher, C. Merkl, and H. Gartner (2021): "Hartz IV and the decline of German unemployment: A macroeconomic evaluation," *Journal of Economic Dynamics and Control*, 127, 104114.

Hoffmann, E. B., D. Malacrino, and L. Pistaferri (2021): "Labor Reforms and Earnings Dynamics: The Italian Case," GID Working Paper.

Hoffmann, F. and T. Lemieux (2016): "Unemployment in the Great Recession: A Comparison of Germany, Canada, and the United States," *Journal of Labor Economics*, 34, S95–S139.

Hohendanner, C. and J. Stegmaier (2012): "Geringfügig Beschäftigte in deutschen Betrieben: Umstrittene Minijobs," *IAB-Kurzbericht*.

Hufe, P., A. Peichl, and M. Stöckli (2018): "Ökonomische Ungleichheit in Deutschland – ein Überblick," *Perspektiven der Wirtschaftspolitik*, 19, 185–199.

Institut Arbeit und Qualifikation (2021): "Durchschnittsalter der Frauen bei der Geburt des ersten Kindes, 1965 – 2020," https://www.sozialpolitik-aktuell.de/files/sozialpolitik-aktuell/_Politikfelder/Bevoelkerung/Datensammlung/PDF-Dateien/abbVII4.pdf, accessed: 2022-01-04.

Jenderny, K. (2016): "Mobility of Top Incomes in Germany," *Review of Income and Wealth*, 62, 245–265.

Jenderny, K. and C. Bartels (2015): "The Role of Capital Income for Top Income Shares in Germany," World Top Incomes Database (WTID) Working Paper.

Kirsch, A. and K. Wrohlich (2020): "Frauenanteile in Spitzengremien großer Unternehmen steigen - abgesehen von Aufsichtsräten im Finanzsektor," *DIW Wochenbericht*, 87, 38–49.

Kleven, H., C. Landais, J. Posch, A. Steinhauer, and J. Zweimüller (2019): "Child Penalties across Countries: Evidence and Explanations," *AEA Papers and Proceedings*, 109, 122–26.

Kohaut, S. and I. Möller (2017): "Führungspositionen in Betrieben und Verwaltungen: Oberste Chefetage bleibt Männerdomäne," IAB-Kurzbericht 24/2017, Nürnberg.

Kramarz, F., E. Nimier-David, and T. Delemotte (2021): "Inequality and Earnings Dynamics in France: National Policies and Local Consequences," GID Working Paper.

Krebs, T. and M. Scheffel (2013): "Macroeconomic Evaluation of Labor Market Reform in Germany," *IMF Economic Review*, 61, 664–701.

——— (2017): "Labor Market Institutions and the Cost of Recessions," IMF Working Papers 2017/087, International Monetary Fund.

Kriete-Dodds, S. and D. Vorgrimler (2007): "The German Taxpayer-Panel," *Schmollers Jahrbuch*, 127, 497–509.

Krolage, C., A. Peichl, and D. Waldenström (forthcoming): "Long Run Trends in Top Income Shares: the Role of Income and Population Growth," *Journal of Economic Inequality*.

Launov, A. and K. Wälde (2013): "Estimating Incentive and Welfare Effects of Nonstationary Unemployment Benefits," *International Economic Review*, 54, 1159–1198.

Leth-Petersen, S. and J. Sæverud (2021): "Trends in Income Risk in Denmark 1987-2016," GID Working Paper.

Lochner, B., S. Seth, and S. Wolter (2020): "Decomposing the large firm wage premium in Germany," *Economics Letters*, 194, 109368.

McKinney, K. L., J. M. Abowd, and H. P. Janicki (2021): "U.S. Long-Term Earnings Outcomes by Sex, Race, Ethnicity, and Place of Birth," GID Working Paper.

Peichl, A., N. Pestel, and H. Schneider (2012): "Does Size Matter? The Impact of Changes in Household Structure on Income Distribution in Germany," *Review of Income and Wealth*, 58, 118–141.

Piketty, T. and E. Saez (2003): "Income Inequality in the United States, 1913–1998," *The Quarterly Journal of Economics*, 118, 1–41.

Prinz, E. and J. Schwalbach (2020): "Vergütung der DAX-Vorstände in den 30-Jahren seit 1987 verzehnfacht," *Forschungsbericht, Humboldt-Univerität zu Berlin.*

Riebesell, P. (1922): *Steuer-Mathematik. Die Fehler in den Reichssteuertarifen.*, Hamburg: Henri Grand.

Schrenker, A. and A. Zucco (2020): "Gender Pay Gap steigt ab dem Alter von 30 Jahren stark an," *DIW Wochenbericht*, 87, 137–145.

Smith, M., D. Yagan, O. Zidar, and E. Zwick (2019): "Capitalists in the Twenty-First Century*," *The Quarterly Journal of Economics*, 134, 1675–1745.

Statistisches Bundesamt (2021a): "Abhängig Erwerbstätige: Deutschland, Jahre, Beschäftigungsumfang, Geschlecht," Tabellen.

——— (2021b): "Verdienste und Arbeitskosten," Index der Tarifverdienste und Arbeitszeiten – Lange Reihe, Destatis.

Stockhausen, M. and K. Maiworm (2021): "IW-Verteilungsreport 2021. Wohlstand für alle? Inklusives Einkommenswachstum vor Corona," IW-Report 36, IW.

Thomsen, U., J. Ludsteck, and A. Schmucker (2018): "Skilled or unskilled – Improving the information on qualification for employee data in the IAB Employee Biography," FDZ-Methodenreport 09, Nuremberg.

Wanger, S. (2020): "Entwicklung von Erwerbstätigkeit, Arbeitszeit und Arbeitsvolumen nach Geschlecht," *IAB-Forschungsbericht*, 16.