Earnings Dynamics and Its Intergenerational Transmission: Evidence from Norway*

Elin Halvorsen† Serdar Ozkan‡ Sergio Salgado§

February 19, 2021

Abstract

As part of a cross-country research consortium, in the first part we use administrative data from Norway between 1993 and 2017 to present stylized facts about individual earnings dynamics. Some of our key findings are as follows. (i) Norway has not been immune to the recent increase in top income inequality observed in other countries. (ii) Earnings dispersion below the 90th percentile declines sharply over the life cycle but increases significantly for those in the top 10%. (iii) The earnings growth distribution is left-skewed and leptokurtic, the extent of which varies with age and past earnings. (iv) Finally, earnings in the top 1% are highly persistent even relative to those in the top 2% or 5%.

In the second part, we exploit a longer panel dating back to 1967 to investigate the intergenerational transmission of income dynamics. First, we find that children of high-income and high-wealth fathers enjoy steeper income growth over the life cycle and face more volatile but more positively skewed income changes, suggesting that they are more likely to pursue high-return and high-risk careers. Second, the income dynamics of fathers and children are strongly correlated. In particular, the children of fathers with steeper life-cycle income growth, more volatile incomes, or higher downside risk also have income streams of similar properties. These findings shed light on the determinants of intergenerational mobility.

Keywords: earnings dynamics, income inequality, Pareto tails, heterogeneity, intergenerational mobility
JEL codes: E24, J24, J31

*We thank Fatih Guvenen, Luigi Pistaferri, and Gianluca Violante as well as the participants of the Global Income Dynamics conferences at Stanford University and the University of Minnesota for helpful comments and suggestions. All comments welcome. Click here for the most recent version.

†Statistics Norway
‡University of Toronto
§The Wharton School, University of Pennsylvania

1 Introduction

Income inequality and stability have been at the forefront of recent economic and policy debates. In the first part of the paper, we document secular and cyclical trends in individual earnings dynamics as well as its life-cycle evolution between 1993 and 2017 using administrative data from Norway. In contrast to earlier literature, we focus more on the top earners and the non-Gaussian features of income dynamics. This contribution is part of the Global Income Dynamics (GID) Project, a research consortium that aims to produce cross-country statistics on earnings dynamics from administrative data sources for 13 economies (see Guvenen, Pistaferri, and Violante (2021)).[1]

In the second part of the paper, we use administrative data dating back to 1967 to investigate the intergenerational transmission of earnings *dynamics*. Unlike earlier literature, which studied the relationship between the income *levels* of parents and children (see Black and Devereux (2011) for a review), we focus on fathers' role in workers' income *dynamics*. In particular, we document the relationship between fathers' and children's income dynamics (i.e., the mean, variance, and skewness of income *changes* over the life cycle) as well as the variation in idiosyncratic risk by fathers' lifetime income and wealth.

Earnings Dynamics in Norway. We start with the distribution of earnings in Section 3. While earnings inequality in Norway is low relative to other developed economies (e.g., Atkinson and Søgaard (2016), Domeij and Floden (2010)), Norway has not been immune to the recent increase in top income inequality seen mainly in developed countries (see Atkinson *et al.* (2011)). In particular, the share of income accrued to the top 1% earners has increased by 25% among men and 28% among women since 1993.

Unlike most other developed economies (see Lagakos *et al.* (2018) for a cross-country comparison), in Norway within-cohort inequality *below* the 90th percentile of the earnings distribution declines significantly over the life cycle (between ages 25 and 55). In contrast, the dispersion *above* the 90th percentile increases significantly. We also find that recent cohorts enter the labor market more unequal at age 25, which is mainly a result of the higher dispersion above the median (Guvenen *et al.* (2017) document a similar trend for the United States).

[1] To ensure the harmonization of the statistics in the database, the authors of this paper have also provided a set of Stata programs to all the teams (and to future researchers) that require only minor adaptations to fit the specifics of each country's administrative dataset. This set of programs will be available on GitHub.

Turning to the properties of the distribution of *earnings growth* (Section 4), we exploit our large administrative dataset to focus on its non-Gaussian features. The distribution of earnings changes exhibits negative skewness and excess kurtosis. The extent of these non-normalities varies with age and permanent earnings, which are qualitatively and quantitatively similar to those documented for the United States and other countries (Halvorsen *et al.* (2019); De Nardi *et al.* (2019)). We do not find strong secular trends in earnings risk during our sample period, except that income volatility rose significantly for men until 2007. Also, the cyclical variation in income risk is not economically large except during the Great Recession, which features a large decline in the skewness of earnings growth (Guvenen *et al.* (2014b); Busch *et al.* (2021) find similar procyclical skewness for the United States and other countries).

Finally, to shed some light on the persistence of incomes, in Section 5, we document long-term mobility patterns. We find that income mobility declines significantly after age 35 and is higher for women. Furthermore, mobility declines sharply in the top 5% of the earnings distribution, with top incomes being the most persistent. Interestingly, there are very few transitions from the lower end of the income distribution to the top end. For example, more than 99.5% of workers in the top 0.1% of the distribution in year $t + 10$ were already in the top 5% in year t. Similarly, less than 25% of the top 0.1% earners fell below the 95th percentile in year $t + 10$. Finally, we find that a quarter (one third) of the top 1% male (female) earners never leave that income group over a 10-year period.

Intergenerational Transmission of Income Dynamics. In Section 6, we switch gears and investigate fathers' role in heterogeneity in income *dynamics* using a longer panel dataset that covers the entire Norwegian population from 1967 to 2017. First, we study the variation in the properties of the distributions of workers' earnings changes by their fathers' economic resources (measured by lifetime income and wealth). This variation can arise, for example, because of parental differences in human capital investment, *dynastic* precautionary savings motive of parents (as in Boar (2020)), and/or because workers may pursue different careers depending on available parental insurance. We also investigate whether the children of fathers with steeper life-cycle income growth, more volatile incomes, or higher downside risk also have income streams of similar properties. Such correlation can emerge, for example, because of similar risk attitudes or similar jobs and occupations.

For both completeness and consistency with the earlier work, we start our analysis by documenting the relation between parents' and children's *lifetime* incomes. Our results confirm the findings of the earlier literature of strong intergenerational persistence in income: 10% increase in fathers' lifetime income is associated with a 2.4% (1.8%) increase in sons' (daughters') income. Furthermore, we find that the intergenerational income elasticity is remarkably uniform across the fathers' lifetime income distribution, indicating a roughly linear relationship between incomes of two generations.

Turning to the features of income growth, we first investigate how workers' *average income growth* varies by their family resources. We find that workers from more affluent families enjoy higher annual income growth with very sizable differences. For example, the sons and daughters of fathers at the 90th percentile of the wealth (or lifetime income) distribution experience annual income growth that is approximately a 2 log points higher relative to those with parents at the 10th percentile.[2]

We also document a strong positive correlation between fathers' and children's life-cycle income growth. For this purpose, for each individual with at least 20 years of income observations, we compute the median income growth over the life cycle. An increase in the father's median income growth from 0 to 5 log points results in roughly 1 log point increase in the son's median income growth over the life cycle. These findings emphasize the importance of using lifetime incomes when measuring intergenerational income elasticity (e.g., Haider and Solon (2006)) and suggest developing models of multi-dimensional intergenerational skill transmission (e.g., Lochner and Park (2020)).[3]

As for the *second moment*, we find that children's income volatility follows a U-shaped pattern by fathers' lifetime income and net worth, with workers from middle-class and upper middle-class families experiencing the most stable incomes. Children of very affluent fathers face particularly more volatile incomes over the life cycle. For example, the standard deviation of income growth for workers with fathers at the top 1% of the lifetime income or wealth distribution is around 10 to 12 log points higher compared to those of fathers at the median. This higher income volatility, combined with the higher average income growth, suggests that children of affluent parents can pursue careers with

[2] This figure roughly corresponds to one-third of the differential between the 90th and 10th percentiles of heterogeneous income growth rates estimated from the US data. For example, Guvenen *et al.* (2018) estimate its standard deviation at around 2 log points from the administrative data.

[3] Recently, Lochner and Park (2020) use data from Canada to estimate a two-factor model of intergenerational skill transmission. In contrast to the results presented in this paper, they find no correlation between the earnings growth of fathers and sons in their data. However, they estimate significant, positive correlations between the initial skills of one generation and the skill growth rates of the other.

higher growth potential but also higher risk.

We also provide strong evidence for the intergenerational transmission of income risk. In particular, we find that fathers with more volatile incomes have children with riskier income streams too. In particular, income growth dispersion—measured by the log 90–10 differential of an individual's income stream—increases from 35 to 45 (45 to 55) log points for sons (daughters) as fathers' dispersion increases from 10 to 50 log points.[4]

Next, we investigate how the tail risk (i.e., *skewness and kurtosis* of income growth) of workers is affected by fathers' characteristics. We find that the skewness of the distribution of children's income growth increases as we move from children of poorer to richer fathers. The increase in skewness—which is accompanied by a decrease in dispersion—is mainly a result of a reduction in the likelihood of a sharp fall in incomes of children up to around the 85th percentile of the fathers' lifetime income or wealth. As for children of the top 10% fathers, and in particular, of the top 1%, both tails of the income growth distribution stretch, but the right tail expands more than the left tail, thereby generating an increase in skewness and dispersion. These findings are consistent with the conjecture that workers with more parental insurance can pursue higher-return and higher-risk careers. Furthermore, we find a positive and significant correlation between fathers' and children's skewness of income changes over the life cycle. In other words, fathers with higher left tail risk also tend to have children with higher downside risk. Along with the intergenerational transmission of volatility, this finding is consistent with fathers and children having similar risk attitudes and/or working in occupations, sectors, and jobs with similar risk profiles.

Finally, we examine whether fathers' strong role in workers' income *dynamics* is simply spurious because of omitted variables such as workers' own permanent income. For this purpose, we run "horse race" regressions with all four factors investigated above—father's lifetime income, net wealth, and moments of income changes as well as worker's own permanent income—included as explanatory variables in the same model. We find that all four regressors are statistically significant at the 1% level and economically important.

[4] Shore (2011) also shows that parents with more volatile incomes have children with risker income streams in the Panel Study of Income Dynamics (PSID). He argues that this positive correlation is partly explained by the intergenerational transmission of risk tolerance and by the propensity for self-employment.

2 Institutional Background and Data

2.1 Institutional Background

Norway is a welfare state with comprehensive social security benefits and redistributive economic policy as a fundament. Welfare provision is financed primarily through high tax rates on income (from labor services and capital) and, to a smaller extent, by returns from a sovereign wealth fund.[5] Two other institutional features are important for this study. First, there is centralized wage bargaining, and second, a substantial part of employment is within the public sector, especially for women. Both of these features contribute to a high labor force participation rate (slightly above 80% during most of our sample period), a low unemployment rate (hovering around 3.5%), and a compressed earnings distribution (which we will discuss in detail in the next section).[6]

The collective bargaining system—called "tripartism"—fosters a broad agreement among unions, employers, and government to maintain a high level of coordination in wages and employment. As a consequence, there is little downward real wage flexibility, which alleviates the effects of negative shocks to the economy. A potential drawback of this tripartism is, however, the difficulty in reducing employee absences and disability. The number of people of working age in Norway who receive sickness leave compensation or disability benefits is among the highest in the OECD (Hemmings and Prinz (2020)). The level of employment protection is high, and firms may find it convenient to pass on the costs of temporary redundant labor to the social security administration. After only one year of sick leave, continued absence becomes a legitimate cause for dismissal. Furthermore, incentives in the sickness compensation system cause both workers and employers to prefer long-term sick leaves to temporary unemployment (Fevang *et al.* (2014)). Studies also show that a large proportion of disability insurance claims in Norway can be directly attributed to job displacement and other adverse shocks to employment opportunities (Bratsberg *et al.* (2013)).

The welfare state has, in many respects, bolstered women's labor participation, but women's own work effort has been a precondition for the development of the welfare state. A higher labor supply means higher tax revenues, which in part can be utilized to offer more comprehensive public services. At the same time, a higher labor supply is

[5]This sovereign fund is currently more than three times the GDP (around 1.1 trillion US$).

[6]Barth *et al.* (2014) develop a model with wage bargaining, creative job destruction, and welfare spending to explain these features of small open economies in Scandinavia.

also needed to perform these services. Many of these services, in turn, allow for a higher female labor supply—for instance, by moving the care of children and the elderly out of the family and under public responsibility instead. In Norway, about 30% to 34% of total employment is in the public sector, of which two-thirds are employed in local government jobs. The public sector in Norway is female dominated with 70% of jobs being held by women, whereas women's share in private sector employment is only 37%. One possible explanation for this larger participation of women in public sector jobs is that women with children are more likely to work in the public sector than those without children, since the public sector is considered more family friendly than the private sector—for example, through more available part-time jobs and more standardized work hours.

2.2 Data Description

In this paper, we use two high-quality datasets covering different time spans. Both datasets are collected for administrative purposes, which reduces the concerns about measurement error or sample attrition that plague survey data. Here we describe the short-panel dataset used in the first part of our analysis. We postpone the description of the long-panel dataset to Section 6, where we use it in our analysis.

The dataset used in the first part of our analysis is derived from a combination of administrative registers covering the whole Norwegian population from 1993 to 2017. Data are assembled based on annual tax records as well as other registers, such as the one administered by the Norwegian Labor and Welfare Administration, the National Education Register, and the Population Register. The tax records are of high quality; most information is third-party reported to the tax authorities as employers are obliged to send earnings information to the individual and to the tax authority.

Earnings include wages and salaries from all employments, including bonuses and other remunerations. The earnings measure reported on tax forms also includes certain earnings-related benefits such as sick and parental leave benefits because these benefits are reimbursed by the social security administration to employers rather than paid out directly to workers. Using additional administrative data, we have deducted these benefits to ensure that our earnings measure consists only of wages and salaries. Therefore, compared to other studies using Norwegian data on earnings from tax returns (e.g., Blundell *et al.* (2014)), we find somewhat higher variance and more volatility of earnings at the bottom of the distribution (similar to Halvorsen *et al.* (2019)).

2.3 Sample Selection and Descriptive Statistics

Our base sample includes all prime-age residents in Norway between ages 25 and 55 who have a personal identification number. The number of individual-year observations between 1993 and 2017 is 51.3 million in total. All nominal incomes are deflated to their 2018 real values using the Consumer Price Index in Norway. Furthermore, to make our results comparable across countries, we convert Norwegian kroner (NOK) values to US dollars using the average exchange rate in 2018.

As a standard practice in the literature, we trim observations below a certain time-varying annual minimum earnings threshold (Y_t^{min}), both to focus on workers with a meaningful labor force attachment and also to avoid few observations of very low earnings affecting our results. Norway does not have a national minimum wage defined by law.[7] Thus, we follow Guvenen *et al.* (2014b) and define Y_t^{min} as the annual earnings of an individual who works 40 hours a week for a full quarter at half the US minimum wage. The minimum income threshold roughly equals NOK 12,000 in 2017. After this trimming, we lose roughly 16% of the total sample available in the administrative records. However, most of these individuals are out of the labor force and receive zero labor earnings during a given year, whereas only 2% of individuals with positive labor earnings are below Y_t^{min}.

Table I shows selected summary statistics for the cross-sectional sample (CS) after the selection explained above, which we use for most of our analysis. Three aspects are worth noting. First, our sample is almost evenly split between men and women (52% men versus 48% women). Second, similarly to other developed economies, the Norwegian working population has become older and more educated: Over our sample period, the share of college graduates or those with more advanced degrees and the share of 46- to 55-year-old workers increased by around 13 and 5 percentage points (pp.), respectively. Finally, between 1995 and 2015 (roughly corresponding to the beginning and end of our sample period), average incomes have grown for both men and women by around 50%, and this increase has been shared by individuals across all income levels.

2.4 Stata Programs for the Global Income Dynamics Database

One of the main goals of the GID project is to provide a set of harmonized statistics on individual earnings dynamics across different countries. To this end, we provide a unified set of Stata programs that can easily be implemented by other researchers with

[7] Although Norway has no general minimum wage, a minimum wage level has been agreed upon in certain sectors.

TABLE I – DESCRIPTIVE STATISTICS FOR CROSS SECTIONAL SAMPLE

Panel A: Sample Statistics										
	Obs. (1000s)		Mean Earnings		Age Shares %			Education Shares %		
Year	Men	Women	Men	Women	$[25, 35]$	$[36, 45]$	$[46, 55]$	$< HS$	HS	$CD+$
1995	1,046	954	44,467	27,805	42.1	31.8	26.1	41.9	27.4	30.8
2015	1.017	943	65,509	46,094	36.4	32.3	31.3	23.8	32.7	43.5

Panel B: Percentiles of the Income Distribution (2018 US$)										
Year	P1	P5	P10	P25	P50	P75	P90	P95	P99	P99.9
1995	369	2,829	8,062	25,061	43,871	57,716	74,860	90,833	134,531	540,776
2015	888	5,691	13,269	35,014	56,122	75,409	102,645	126,169	196,130	1,278,346

Notes: Table I shows summary statistics for the CS sample. All nominal values are deflated to their 2018 real values using the Consumer Price Index in Norway and converted to US dollars using the average exchange rate in 2018. In the right columns of Panel A, we separate workers into three groups. $< HS$ are workers with less than a high school diploma, HS are workers with a high school degree, and $CD+$ are workers with a college degree or more advanced degrees.

access to longitudinal data on individual earnings. The main statistics generated by these programs form the core set of results presented by papers in this issue of the journal.

The codes are structured in eight interconnected files requiring minimum input from the user (e.g., specifying the names of the variables in the dataset, the time period dataset covers, providing aggregate price indices). These programs produce a set of baseline results such as descriptive statistics (Section 2.3), measures of earnings inequality (Section 3) and volatility (Section 4) as well as facts on individual income mobility (Section 5). The programs also generate a harmonized set of core figures that can easily be compared over time and across countries as they are based on similar measures of income and a common set of sample selection criteria. Additional details on the codes and their implementation can be found in Appendix E and in the codes themselves, which will be available on Github.

3 Earnings Inequality

3.1 Trends in Earnings Inequality

We begin our analysis by discussing the salient features of the earnings distribution and its evolution between 1993 and 2017 in the cross-sectional sample. Throughout the paper, we present results for men and women separately (see Appendix A.1 for findings from a combined sample of men and women). The results in this section are from the distribution of raw earnings without controlling for changes in educational attainment or the age composition of the population. We find qualitatively similar patterns if we control for these observable characteristics of workers (see Appendix A.2), which suggests that compositional changes are not likely to drive the evolution of income inequality.

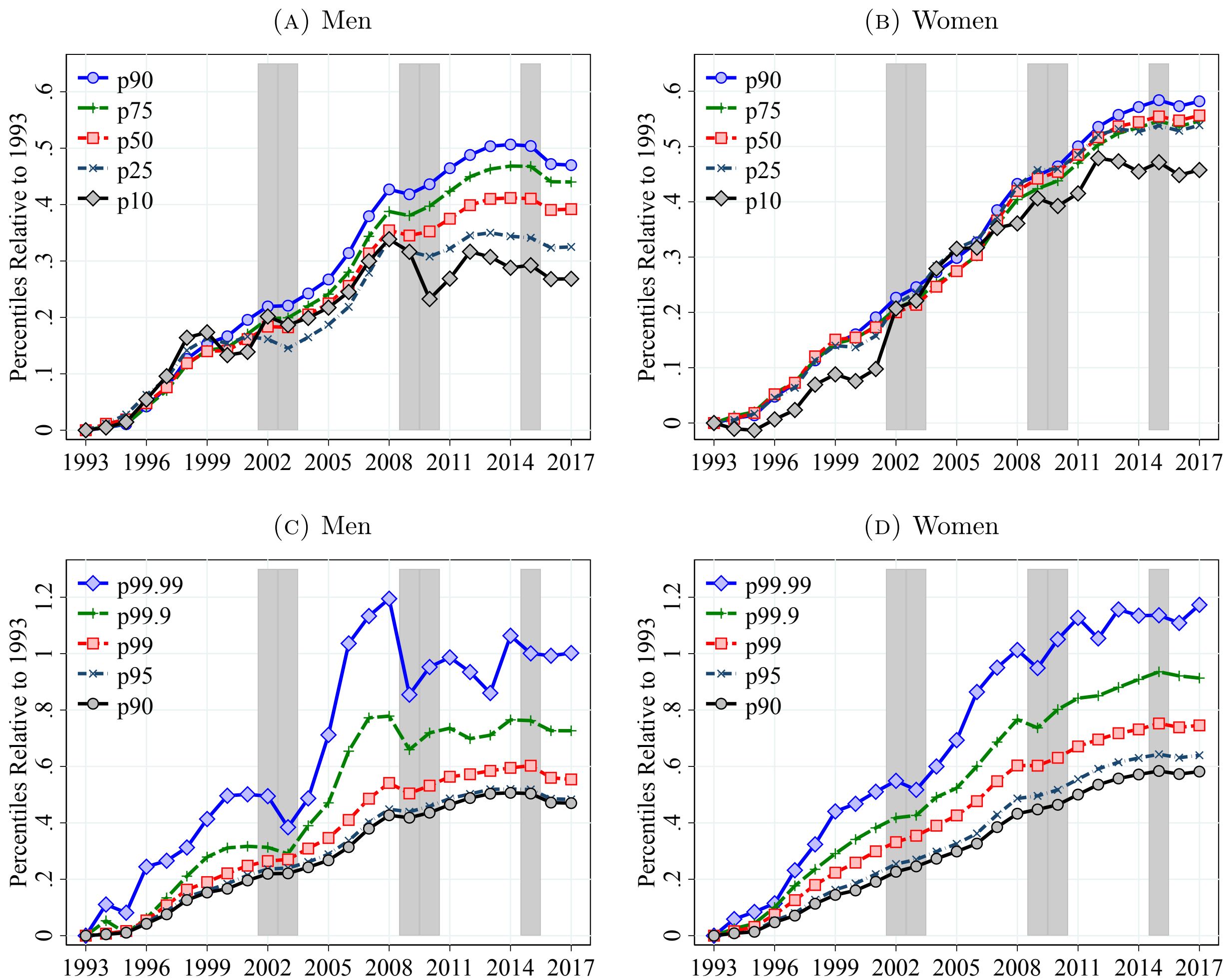
Similar to other Scandinavian countries, Norway has a relatively compressed earnings distribution compared to other developed economies (e.g., Roine and Waldenström (2008), Atkinson and Søgaard (2016), Domeij and Floden (2010), Guvenen *et al.* (2014a), and other studies in this issue). For example, the differential between the 90th and 50th percentiles (hereafter P90-P50) and between the 50th and 10th percentiles (P50-P10) of log labor earnings for men is on average 55 and 115 log points over our sample period, respectively, compared to 100 and 150 log points for a similar sample from the United States in 2010 (see Guvenen *et al.*, 2014b). The earnings distribution for women is relatively more dispersed than for men, mainly because of the higher inequality below the median (e.g., P50-P10 for women is around 140 log points versus 115 log points for men).

We next turn to the evolution of the earnings distribution during our sample period. The Norwegian economy experienced strong growth in the 1990s and 2000s, in the aftermath of a banking crisis that occurred in Norway, Sweden, and Finland between 1991 and 1993. This strong growth is partly a result of the expanding oil industry and the many economic reforms introduced during or after the crisis.[8] Figure 1 shows the changes in selected percentiles of the log earnings distribution for men and women relative to their corresponding values in 1993. Until the Great Recession, earnings of men around and below the 75th percentile grew steadily and at a similar pace by roughly 35 log points (Figure 1a), whereas high-income workers (those around and above the 90th percentile) enjoyed relatively steeper growth. Thus, overall inequality, measured as the differential between the 90th and 10th percentiles (P90-P10) or as the standard deviation of the log earnings distribution, increased during this period mainly because of the 8 log points rise in the upper-tail inequality (Figures 2a and 2c).

During and in the aftermath of the Great Recession, however, this steady income growth almost halted for most workers, and the earnings inequality below the 90th percentile grew significantly. For example, between 2008 and 2017, incomes between the 90th and 50th percentiles grew by a mere 5 log points, whereas the 10th percentile

[8] These reforms included a tax reform in 1992, a closer cooperation between the government and the unions to reduce unemployment and increase labor market competitiveness, a series of product market reforms to increase competition and efficiency, fiscal policy rules to limit discretion in using the sovereign wealth fund, and fiscal plans for age-related welfare spending as well as central bank independence.

Figure 1 – Percentiles of the Log Real Earnings Distribution



Notes: Figure 1 shows the evolution of the following variables: (a) men: P10, P25, P50, P75, P90 (b) women: P10, P25, P50, P75, P90, (c) men: P90, P95, P99, P99.9, P99.99, (d) women: P90, P95, P99, P99.9, P99.99. All percentiles are normalized to 0 in 1993. Shaded areas represent recession years, defined as years with an unemployment growth rate of 0.4 pp. or more and an output gap of -0.5 or less. See Section 2 for sample selection and definitions.

declined during this 10-year period. As a result, P50-P10 widened by 10 log points after 2008 (Figure 2c). These concerning developments are partly because of a sharp drop in oil prices from 2014 to 2015, which hit the oil-dependent Norwegian economy hard and caused another recession in 2015. Although the oil sector was the most affected (with a loss of about 50,000 jobs), the shock had ripple effects throughout the economy. Thus, in the aftermath of the oil price shock, income growth slowed down across the entire distribution, but more so for low-income workers.[9]

[9]Other factors that contributed to the slower income growth after the Great Recession are lower productivity growth and a decline in labor force participation (partly linked to a long-standing problem

Similar to men, women with earnings below the 90th percentile of the income distribution experienced a steady income growth by around 50 log points until 2008 (Figure 1b). Thus, P90-P10 was stationary for them until the Great Recession.After 2008 income growth slowed down in all percentiles but more so in the bottom of the distribution, thereby increasing P50-P10 (Figure 2b). However, women experienced a smaller increase in the lower tail inequality compared with the men (Figure 2a). These heterogeneous effects of recessions on the earnings of men and women are in line with the fact that Norwegian men work predominantly in the private sector, which is more exposed to business cycles, whereas women are mainly employed in the public sector.[10]

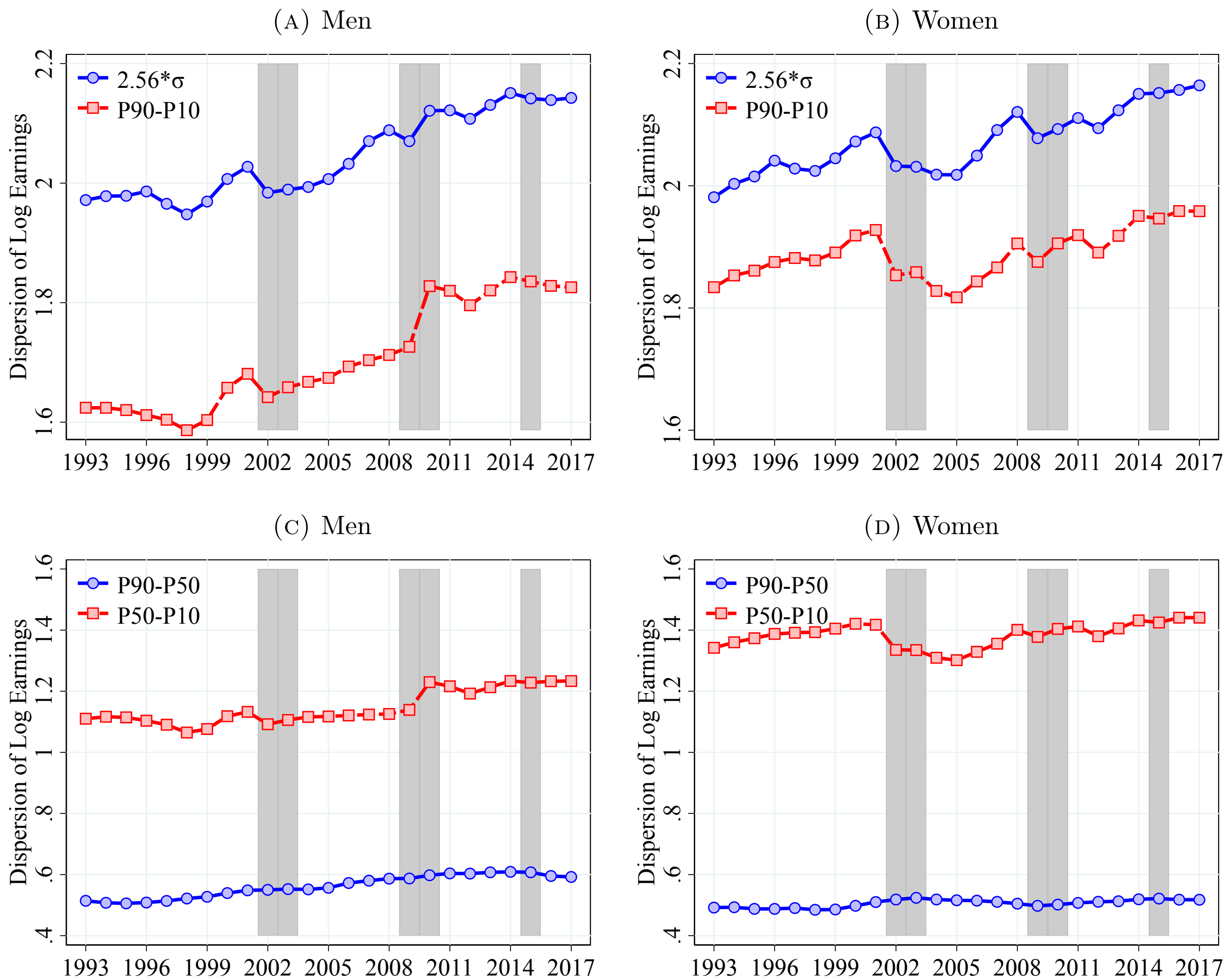
Incomes in the top 10% have evolved differently from the rest of the earnings distribution. In particular, we find that the top end of the distribution fans out for both men and women until 2008. For example, male workers at the 90th percentile in 2008 earn 53% (43 log points) more relative to 1993, whereas the earnings of those at the 99th, 99.9th, and 99.99th percentiles grew by around 75%, 120%, and 230%, respectively (Figure 1c). However, after 2008, like those in the rest of the distribution, high-earnings men experienced stagnant incomes with an average growth of merely a few log points over a 10-year period. As a result, P90-P50 grew significantly before 2008 but remained roughly constant afterward (Figure 2c). Interestingly, inequality in the top 10% has shrunk during this period for men as the 99.99th percentile of the distribution declined by 20 log points between 2008 and 2017.[11] For women, however, growth in top incomes has not slowed as much as it did for men. For example, income in the top 1% increased by around 20 log points between 2008 and 2017. As a result, inequality in the top 10% has continued to increase even after 2008, albeit at a slower pace.

of early retirement via disability benefits).

[10]Income inequality in the private and public sectors has evolved quite differently over the last 30 years (see Appendix A.3). In fact, P90-P10 for workers employed in the private sector increased by 30 log points between 1993 and 2014. Among men, most of the increase in inequality is accounted for by an expansion of the left tail of the distribution. For women, the increase in inequality is due to the stretching of both tails of the distribution. In contrast, inequality in the public sector has remained the same among men and declined quite significantly among women, especially after 2001. These results highlight the prominent role that the public sector plays in reducing income inequality in Norway.

[11]We take a closer look at the right tail of the earnings distribution by examining its empirical log density (Figure A.1 in Appendix A). First, in line with previous literature (e.g., Atkinson *et al.*, 2011), the right tail is much thicker compared with a normal distribution and declines almost linearly, implying a Pareto distribution. Second, consistent with our previous discussion on increasing top income inequality, the right tail has become flatter during our sample period for both men and women. Third, the right tail of the earnings distribution is declining more slowly for men than for women (with a lower tail index for men), which indicates a more pronounced top income inequality for men.

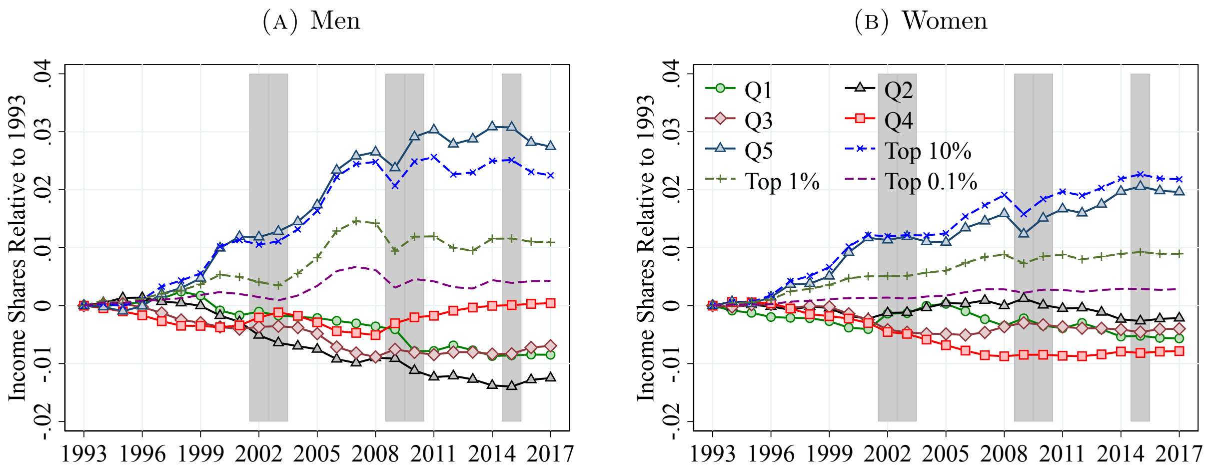
FIGURE 2 – INCOME INEQUALITY



Notes: Figure 2 plots the following variables against time: (a) men: P90-P10 and 2.56*SD of log income (b) women: P90-P10 and 2.56*SD of log income, (c) men: P90-P50 and P50-P10, (d) Women: P90-P50 and P50-P10. Shaded areas are recessions. The value of 2.56*SD corresponds to the differential between the 10th and the 90th percentiles in a Normal distribution. Shaded areas represent recession years, defined as years with an unemployment growth rate of 0.4 pp. or more, and an output gap of -0.5 or less. Results based on the CS sample. See Section 2 for sample selection and definitions.

As in other developed countries, earnings in Norway are concentrated at the top of the distribution, albeit to a smaller extent. For instance, workers in the top 1% of the distribution earned 4% of the aggregate income in 1993 compared to a 10% income share of top earners in the United States in the same year (Kopczuk *et al.*, 2010). In line with our previous discussion, the share of income accrued to top earners—for both men and women—has risen significantly (Figure 3). Especially after the late 1990s, the income shares of top percentile and decile workers increased by around 1 pp. and 2 pp., respectively, at a cost of a decline in the income shares of workers at the bottom four

Figure 3 – Income Shares Relative to 1993



Notes: Figure 3 shows the share of income accrued to five income quintiles and top income groups for men and women. All shares are normalized to 0 in 1993. The line labeled Q1 represents the share of income accrued to the first quintile of the income distribution, Q2 is the income share of the second quintile and so on. The shaded areas represent recession years defined as years with: i)a growth in the unemployment rate of 0.4 pp. or more and ii) an output gap of -0.5 or less. See Section 2 for sample selection and definitions.

quintiles of the distribution.[12] Growth in top income inequality is also reflected by a decrease in the Pareto tail index of the earnings distribution between 1995 and 2015 (Figure A.1 in Appendix A). Therefore, we conclude that Norway is not immune to the recent increase in top income inequality seen mainly in English-speaking countries (see Atkinson *et al.* (2011)).

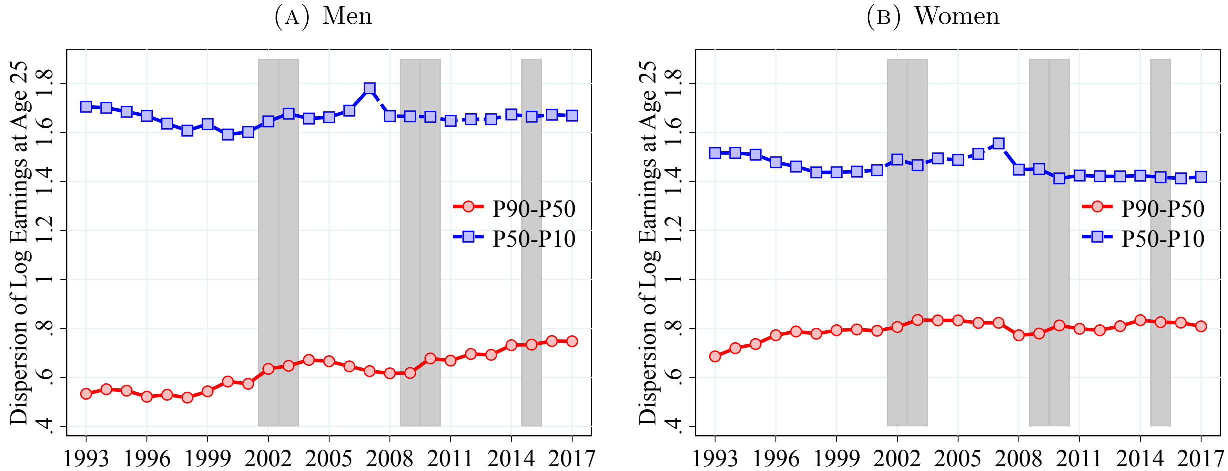
3.2 Evolution of Life-Cycle Earnings Inequality

The increase in top income inequality documented in the previous section could, in principle, be attributed to higher inequality at young ages for newer cohorts, a higher increase in the earnings dispersion over the life cycle, or both. To dissect these changes, in this section we explore how within-cohort inequality evolves over the life cycle and how it varies across different cohorts.

Figure 4 shows the dispersion of earnings above and below the median at age 25 for cohorts entering the labor market from 1993 to 2017. The recent cohorts enter the labor market with a higher initial inequality above the median compared to older cohorts, especially so for men. In particular, the P90-P50 of the log earnings distribution for 25-year-old men is 75 log points in 2017 versus 45 log points in 1993. For women, the increase is only around 10 log points, from 70 to 80 log points. As for the dispersion

[12] The World Inequality Database reports that the income share of the top 1% increased from 8% to 10% during the same period. This measure include all income sources.

Figure 4 – Earnings Inequality at age 25



Notes: Figure 4 shows (a) men: P90-50 and P50-10 at age 25, (b) women: P90-50 and P50-10 at age 25. The shaded areas represent recession years defined as years with: i) growth in the unemployment rate of 0.4 pp. or more and ii) an output gap of -0.5 or less. See Section 2 for sample selection and definitions.

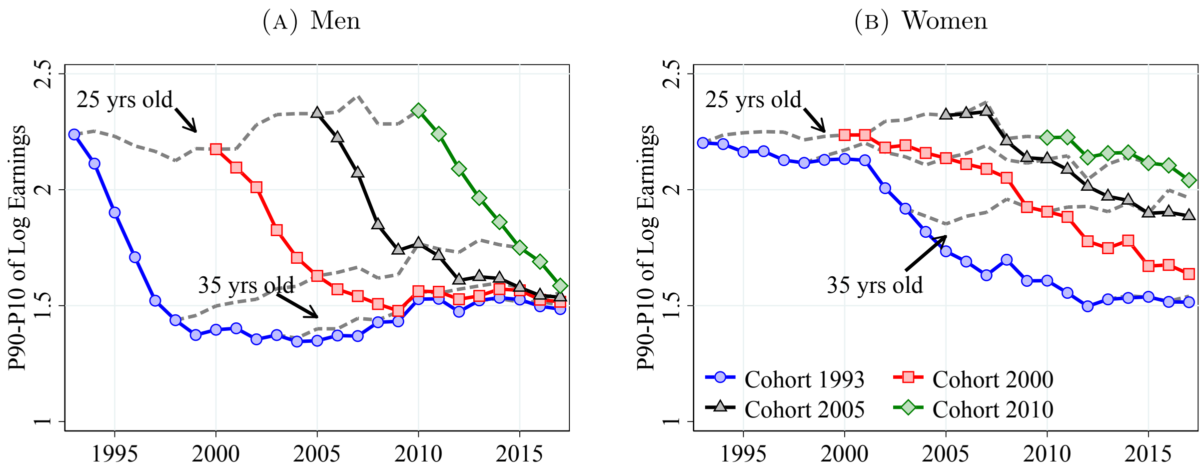
below the median (P50-P10), we see no significant change for men during our sample period and a decline of 10 log points for women (from 150 log points in 1993 to 140 log points in 2017).[13] The higher initial inequality for newer cohorts means that life-cycle inequality is also higher in recent years (which we will discuss in detail shortly). These findings suggest that the rise in inequality between 1993 and 2017 is partly a result of the higher initial inequality of newer cohorts.

To obtain a more complete picture of the evolution of inequality over time and over the life cycle, in Figure 5 we plot the overall dispersion (P90-P10) of log earnings in each age for 18 cohorts of workers entering the labor market at age 25 between 1993 to 2010. The colored markers connect ages of the same cohort, thus showing how within-cohort inequality evolves over the life cycle. Dashed lines correspond to particular ages of different cohorts and show how the inequality in a particular age evolves over time.

The life-cycle profile of the P90-P10 of earnings is roughly similar for all cohorts. For men between ages 25 and 35, within-cohort inequality declines by around 75 log points and remains relatively constant afterward, whereas women experience a rapid decline in inequality only after the cohort turns 35. These findings are in sharp contrast to the age profile of earnings inequality documented for the United States, where within-cohort dispersion fans out steeply over the life cycle (Storesletten *et al.* (2004a); Karahan

[13]Similarly, Guvenen *et al.* (2017) find that the initial dispersion is dramatically higher for younger cohorts in the United States, which is mainly a result of the higher inequality above the median.

Figure 5 – Evolution of Labor Earnings Inequality by Cohorts



Notes: Figure 5 uses the log earnings from the CS sample and shows: (a) men: P90-P10 over the life cycle for selected cohorts and (b) women: P90-P10 over the life cycle for selected cohorts. A cohort is defined by the year in which the cohort turns 25. Dashed lines connect individuals of the same age. The plot consider cohorts born between 1969 and 1986 and turn 25 from 1993 to 2010, respectively. See Section 2 for sample selection and definitions.

et al. (2019); Lagakos *et al.* (2018)). For example, Guvenen *et al.* (2018) document that the variance of log earnings increases by 55 log points between ages 25 and 55. The decline in earnings inequality over the life cycle also explains why Norway has a relatively compressed earnings distribution compared to other developed economies (see Guvenen *et al.* (2014a)).

To dig a bit further, we decompose the overall earnings inequality into its above- and below-median components. Among men, the decline in P90-P10 over the life cycle is mainly a result of the closing of the gap between the bottom and median workers. To see this, notice that the within-cohort P50-P10 declines by around 70 log points between ages 25 and 35 and remains roughly constant afterward (Figure 6a).[14] In contrast, the P90-P50 declines only between ages 25 and 28 and then shows a steady increase over the life cycle (Figure 6c). This picture is remarkably different for women, for whom life-cycle inequality declines at both ends of the earnings distribution (Figures 6b and 6d).

In contrast to the marked decline in earnings inequality below the 90th percentile of the distribution, top income inequality increases significantly over the life cycle. As we show in Figure A.6 in Appendix A.2, the difference between the 99th and 90th percentiles of log earnings (P99-P90) increases over the life cycle by around 40 log points for men

[14]This result also holds when we exclude from our sample those who are still students between ages 25 and 35. Therefore, the decline in P50-P10 is not driven by lower rates of school enrollment.

and 30 log points for women. Furthermore, P99-P90 is higher for newer cohorts at all ages compared to older cohorts, which explains the rise in top income inequality over our sample period. These facts suggest that different economic mechanisms may be at play in determining the inequality at different parts of the earnings distribution. For instance, Karahan *et al.* (2019) show that in the United States, earnings differences between the bottom and median earners are mainly a result of the differences in unemployment risk, whereas the right-skewed distribution of returns to experience explains the within-cohort inequality between the top and median earners.

4 Distribution of Earnings Growth

In this section we characterize the distribution of individual earnings growth in Norway. Similar to other papers (e.g., Dynan *et al.* (2012); Guvenen *et al.* (2014b)), we follow a descriptive approach and document the properties of the cross-sectional distribution of individual earnings changes. The measure of income change, we use is the familiar log growth rate of residual earnings between years t to $t+k$, $g_{it}^{k} = \tilde{\varepsilon}_{it+k} - \tilde{\varepsilon}_{it}$ for $k = \{1, 5\}$. We obtain residual earnings, $\tilde{\varepsilon}_{it}$, by regressing log earnings in each year on a set of age dummies for men and women separately.

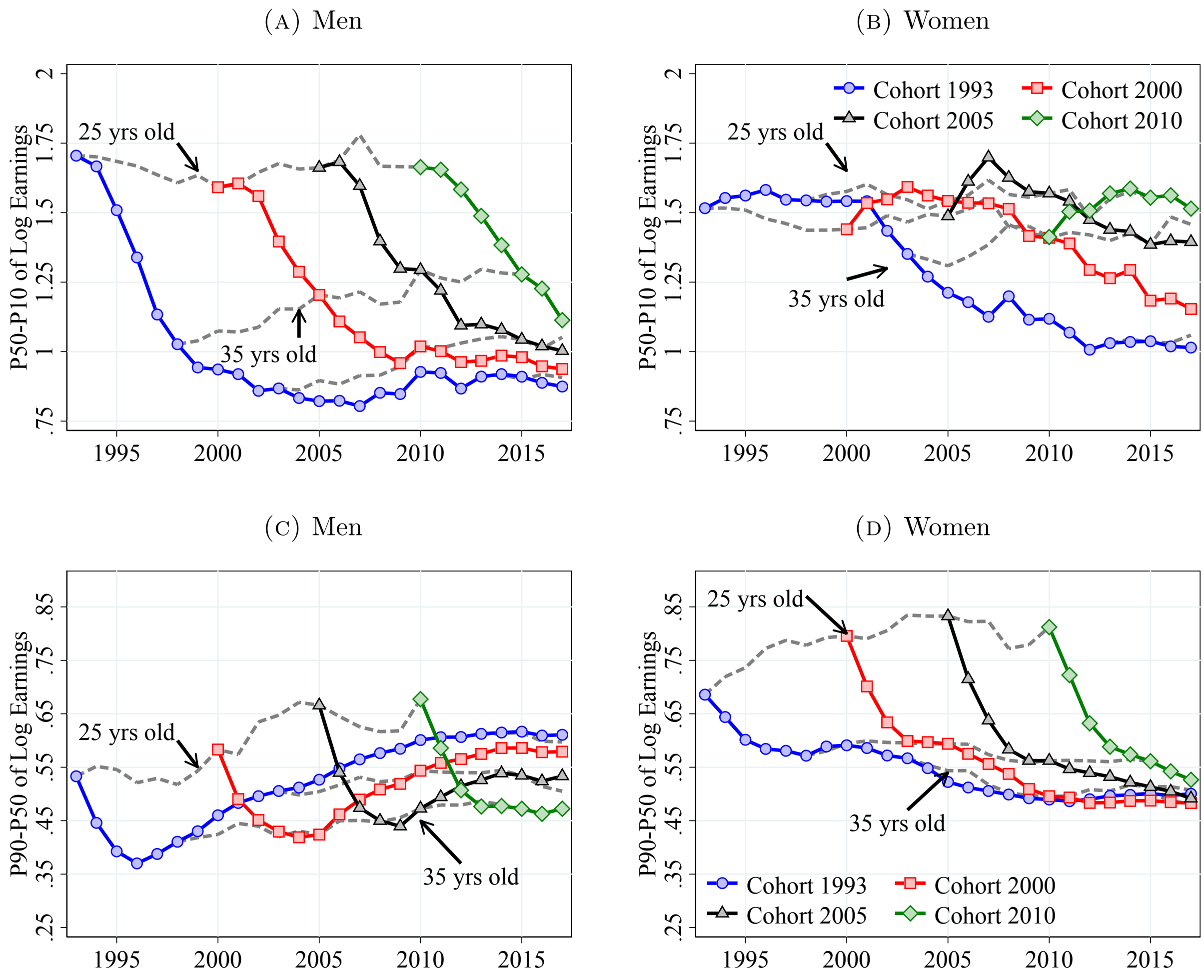
Notice that the log growth measure is applicable only for individuals with earnings above the minimum threshold, Y_t^{min}, in periods t and $t+k$ and therefore does not capture the earnings changes of individuals that are nonemployed during most of the year and thus have little or zero earnings.[15] Because of this key drawback of the log growth measure, we replicate the main results in this section using the arc-percent growth measure, which is more robust to this caveat. In particular, the arc-growth measure is defined as $arc_{it}^{k} = \frac{Y_{t+k}^{i} - Y_t^{i}}{0.5 \times \left(Y_{t+k}^{i} + Y_t^{i}\right)}$, where Y_{it} is the level of earnings normalized by average earnings in each year and age. The key results for the arc-percent change measure are reported in Appendix B.3 and show qualitatively similar patterns.

4.1 Higher-Order Moments of Individual Earnings Growth

Recent literature has shown that the distribution of idiosyncratic earnings changes displays strong non-Gaussian features—such as left skewness and excess kurtosis—and that the extent of these non-normalities varies significantly with age and earnings level

[15] To be more precise, we construct log income growth between t and $t+k$ for those who have earnings Y_{it} above the minimum threshold, Y_t^{min} in t and above one-third of Y_t^{min} in $t+k$ (i.e., $Y_{it+k} > \frac{Y_t^{min}}{3}$) so that we can better capture large declines in annual earnings. This slight addition to the sample selection does not lead to any material difference in our results.

Figure 6 – Evolution of Below- and Above-Median Inequality by Cohorts

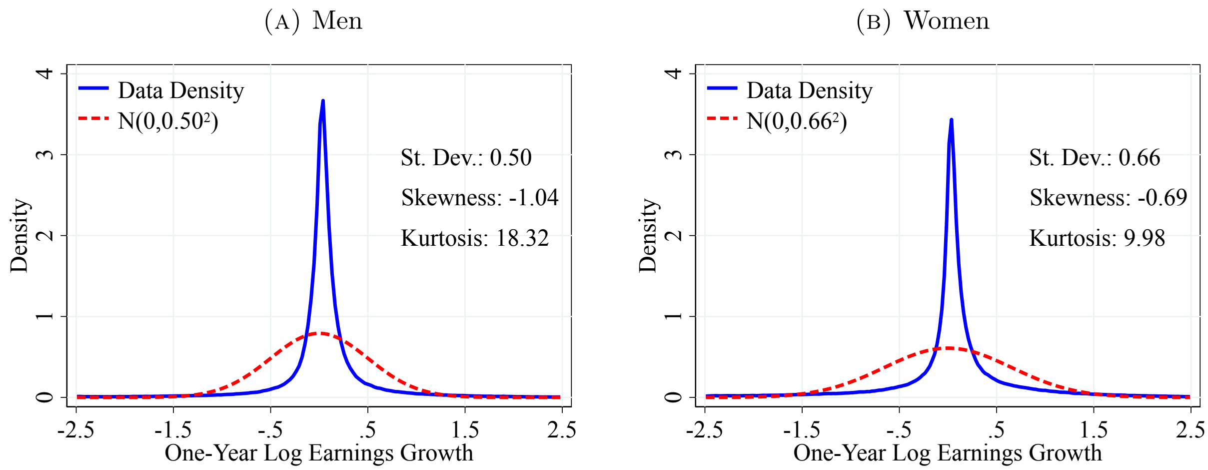


Notes: Figure 6 uses the CS sample to show the life cycle inequality of log earnings for selected cohorts: (a) men, P50-P10 dispersion, (b) women, P50-P10 dispersion (c) men, P90-P50 dispersion, and (d) women, P90-P50 dispersion. A cohort is defined by the year in which the cohort turns 25. Dashed lines connect individuals of the same age. See Section 2 for sample selection and definitions.

(e.g., Arellano *et al.* (2017); Guvenen *et al.* (2018)). Furthermore, several papers have already shown that these features have important implications for the consumption-savings behavior of households and individuals (e.g., Kaplan *et al.* (2018); De Nardi *et al.* (2020)). Exploiting our dataset's sheer size and high quality, we also focus on these moments, in addition to the usual second moment, in our analysis.

As shown in Figure 7, the one-year earnings growth distribution displays left (negative) skewness and excess kurtosis relative to a normal density, which is chosen to have the same variance as in the data. These features of the data are qualitatively and quan-

Figure 7 – Empirical Density of One-Year Log Earnings Change



Notes: Figure 7 shows the empirical density and corresponding cross-sectional moments of the distribution of one-year log earnings growth for men and women in 2005. See Section 2 for sample selection and definitions.

titatively very similar to those documented for the United States (see Guvenen *et al.* (2018)). We also find similar patterns in the distribution of five-year changes (Figure B.3 in Appendix B.1). The left skewness implies that the left tail of the distribution is longer than the right tail, with workers being more likely to experience very large declines in earnings (disaster shocks) than very large increases. For example, almost twice as many workers experience a decline of more than three standard deviations as those experiencing increases of the same size (Table B.1 in Appendix B.2).

In addition, the excess kurtosis captures the fact that, in a given year, far more people have very small and extreme changes and fewer people have middling ones relative to a normal distribution—notice the peaky center and narrow shoulders of densities (Figure 7). For instance, the fraction of workers who experience a change in their labor earnings of less than 5% is 32.2% in the data versus 7.1% implied by a normal distribution with the same standard deviation (Table B.1 in Appendix B.2). Similarly, in the data, 3.1% of workers see their incomes change by more than three standard deviations or more in any given year, whereas only 0.3% of individuals would experience such changes if income changes were normally distributed.

4.2 Trends and Business-Cycle Variation

Next, we investigate how the earnings growth distribution evolves over time and over the business cycle. Below we present cross-sectional moments of one-year earnings growth to better capture the high-frequency business-cycle variation. The corresponding figures for long-term (five-year) changes—which capture persistent innovations—are reported in Appendix B.1 and show similar qualitative patterns. Also, in the main text, we report percentile-based moments (e.g., Kelley's skewness and Crow-Siddiqui kurtosis), which are robust to outliers but ignore valuable information in the tails of the distribution. Therefore, we also document variations in standardized moments in Appendix B.2 and discuss any substantial differences in our findings from these measures.

Dispersion. Figure 8 shows the evolution of the dispersion of earnings growth below and above the median between 1993 and 2016. Several remarks are in order. First, overall earnings volatility for women—as measured by the P90-P10 of log income changes—is almost twice as large as the earnings volatility for men, hovering around 115 log points for women versus 60 log points for men. This is likely a result of the generous maternity leave benefits provided by the Norwegian government (up to nine months of full pay), the fact that women are more likely to work in part-time jobs with more flexible hours, and an overall weaker labor market attachment of women.[16] Second, the income volatility of men has risen significantly over our sample period, with P90-P10 increasing from 54 log points in 1993 to 65 log points in 2016, whereas for women, income volatility has remained relatively stable over the same period. These findings are in contrast with the evidence from the United States—also based on administrative data—that shows that earnings volatility is roughly similar for men and women and has been trending down since the 1980s for both (Bloom *et al.*, 2017).

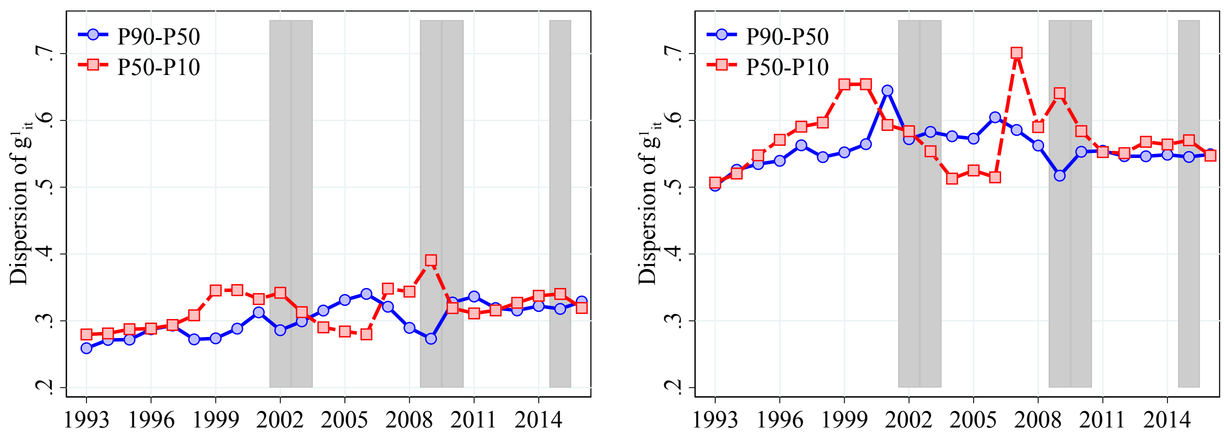
Skewness. We now turn to the asymmetry in the distribution of earnings growth as measured by the Kelley skewness (Kelley, 1947), which is defined as $\mathcal{S}_{\mathcal{K}} = \frac{\text{(P90-P50)}-\text{(P50-P10)}}{\text{P90-P10}}$. A positive (negative) value of this measure indicates that the distribution is skewed to the right (left). In Norway, the Kelley skewness of log earnings growth is, on average, slightly negative hovering around -0.06 during our sample period (Figure 9a), indicating a slightly negative skewed distribution with 53% of the total dispersion being accounted for by the left tail versus 47% by the right tail. However, when measured with the third centralized moment (Figure B.4), earnings growth displays strong left skewness in all years, indicating a stronger asymmetry between the bottom and the top deciles of the earnings growth distribution.

Kurtosis. Figure 9b shows the excess kurtosis of one-year earnings changes for men and women as measured by the Crow-Siddiqui kurtosis (Crow and Siddiqui, 1967), which

[16] The gap between men's and women's income volatility shrinks significantly once public benefits are taken into account (see Figure D.1 in Appendix D.1).

FIGURE 8 – DISPERSION OF EARNINGS CHANGES

(A) Men (B) Women



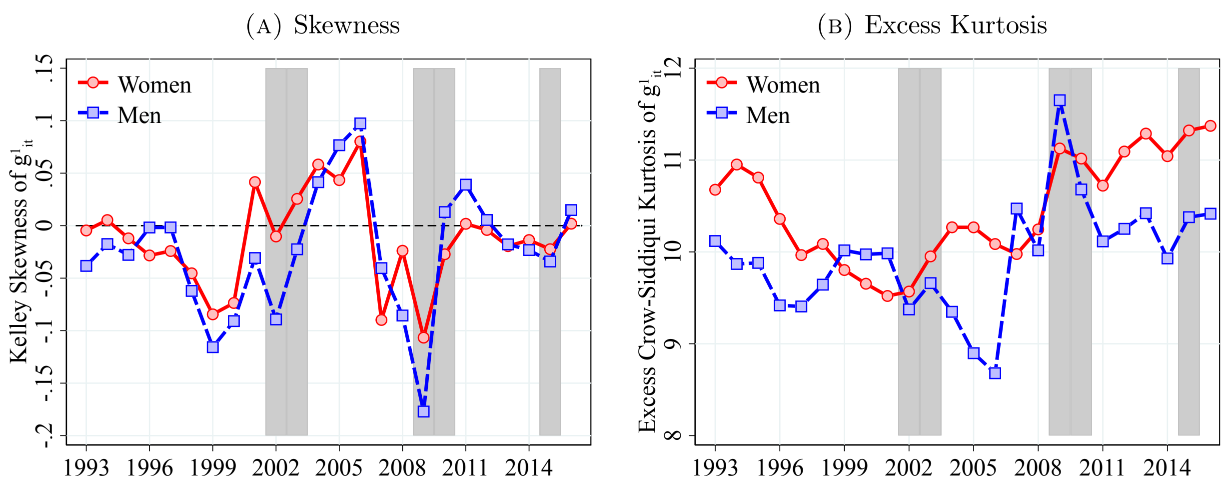
Notes: Figure 8 shows the 90th-to-50th and 50th-to-10th percentiles differential of earnings growth for men and women. The shaded areas represent recession years, defined as years with: i) growth in the unemployment rate of 0.4 pp. or more and ii) an output gap of -0.5 or less. See Section 2 for sample selection and definitions.

is defined as $\mathcal{C}_{\mathcal{K}} = \frac{\text{(P97.5-P2.5)}}{\text{P75-P25}} - 2.91$, where 2.91 corresponds to $\frac{\text{(P97.5-P2.5)}}{\text{P75-P25}}$ of a normal distribution. According to this measure, the kurtosis of earnings growth has remained relatively stable around 10 over our sample period showing only a slight increase after the Great Recession for women.[17]

Cyclical Nature of Individual Earnings Growth. To measure the business-cycle variation in income risk, we estimate simple time-series regressions of different moments of the earnings growth distribution on the unemployment rate and GDP growth (both standardized). We find mixed evidence for the cyclical nature of income risk (see Table II). (i) Dispersion is countercyclical for both men and women (as in Storesletten *et al.*, 2004b). (ii) For men, the Kelley skewness is procyclical (as in Guvenen *et al.*, 2014b) but not the third centralized moment. For women, Kelley skewness is procyclical only when regressed on unemployment growth. The decline in the Kelley skewness of earnings growth is more noticeable during the Great Recession, reaching a value of -0.18 for men and -0.10 for women in 2009 (Figure 9a). In other words, during the Great Recession, almost 60% of the total dispersion for men was accounted for by the left tail, whereas 40% was accounted for by the right tail of the distribution. (iii) Kurtosis is acyclical for

[17] In contrast, the forth standardized moment of income changes for men is more than twice as big as that for women—around 15 versus 7 (Figure B.4 in Appendix B.2). Similar to the skewness, the percentile-based and standardized moments of kurtosis show substantial differences, underscoring the importance of extreme observations for higher- order moments.

Figure 9 – Skewness and Kurtosis of Earnings Changes



Notes: Figure 9 shows the Kelley skewness and excess Crow-Siddiqui kurtosis of earnings growth for men and women. The shaded areas represent recession years, defined as years with: i) growth in the unemployment rate of 0.4 pp. or more, and ii) an output gap of -0.5 or less. The excess Crow-Siddiqui kurtosis is defined as the annual Crow-Siddiqui measure minus 2.91, which is the corresponding value of Crow-Siddiqui for a Normal distribution. See Section 2 for sample selection and definitions.

men but seems to be procyclical for women.[18]

Even when it is statistically significant, however, the cyclical variation in income risk is not economically large except during the Great Recession. For example, for men, when the unemployment rate increases by 1%, P90-P10 increases by 1 pp. and the Kelley skewness declines by 0.04. These mixed results also explain why the gray recession bars in Figures 8 and 9—defined as the years during which the unemployment rate increases by more than 0.4 pp. and GDP is below trend by more than 0.5 pp.—do not seem to display any significant pattern for the income risk over the business cycle.

4.3 Heterogeneity in Idiosyncratic Earnings Changes

Several papers have shown that there is significant unobserved (e.g., Alvarez *et al.* (2010)) and observed (e.g., Guvenen *et al.* (2018) and Karahan and Ozkan (2013)) heterogeneity in the nature of idiosyncratic earnings risk. Following this literature, in this section, we investigate how the properties of the earnings growth distribution in Norway vary by age and permanent income, which we define next.

Our measure of permanent earnings in $t-1$ is defined as the average earnings between $t-1$ and $t-3$ net of age and year effects. In particular, the average earnings of a worker

[18] As we show in Table B.2 in Appendix B.2, these results do not change much if we consider moments of the arc-percent change, indicating that the extensive margin of annual employment is not likely to change the cyclical properties of different moments of the earnings growth distribution.

TABLE II – CYCLICALITY OF CROSS-SECTIONAL MOMENTS OF EARNINGS CHANGES

	(1)	(2)	(3)	(4)	(5)	(6)
	Dispersion		Skewness		Kurtosis	
	P90-P10	Std. Dev.	Kelley	Third.	Crow-Siddiqui	Kurtosis
			Men			
ΔGDP_t	-0.01***	-0.01**	0.03**	0.09	-0.16	0.06
	(0.00)	(0.01)	(0.01)	(0.10)	(0.11)	(0.32)
			Women			
ΔGDP_t	-0.04***	-0.02**	0.02	-0.01	0.27***	0.21**
	(0.01)	(0.01)	(0.01)	(0.03)	(0.09)	(0.08)
			Men			
ΔUnemp_t	0.01**	0.01*	-0.04***	-0.03	0.22	-0.29
	(0.00)	(0.00)	(0.01)	(0.05)	(0.16)	(0.21)
			Women			
ΔUnemp_t	0.02**	0.00	-0.02*	0.01	-0.12	-0.18**
	(0.01)	(0.00)	(0.01)	(0.03)	(0.12)	(0.07)
N	24	24	24	24	24	24

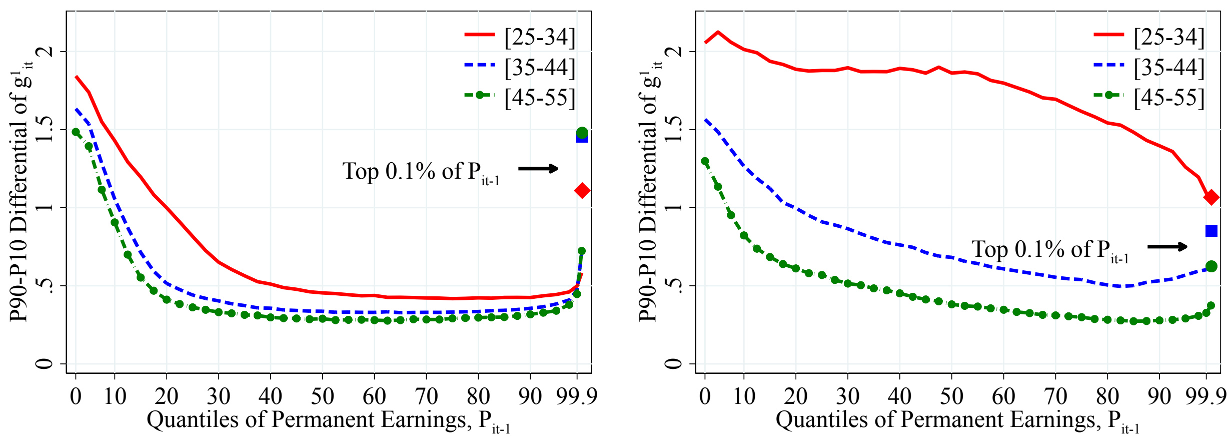
Notes: Table II shows the coefficients from regressions of different moments of log earnings growth on either GDP or unemployment growth for men and women. The growth rate of unemployment (real GDP) is calculated as the (log) difference of the average unemployment rate (real GDP) between years t and $t+1$. Notice each regression is run separately. The unemployment rate is obtained from Statistics Norway and real GDP is obtained from the Federal Reserve Economic Data, FRED. Newey-West standard errors are in parentheses, estimated using one lag. In each regression, we standardize the right-hand-side variable so that the coefficient can be directly interpreted as the impact of a one-standard-deviation change on the dependent variable. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

i between years $t-1$ and $t-3$ is given by $\overline{P}_{t-1}^{i} = \frac{1}{3}\sum_{j=1}^{3} Y_{it-j}$, where Y_{it} denotes real labor earnings in year t. We construct $\overline{P}_{t-1}^{i}$ only for workers who have earnings above the minimum income threshold, Y_t^{min}, in year $t-1$ and at least one more year. We then control for age and year effects by regressing the log of $\overline{P}_{t-1}^{i}$ on a set of age dummies separately in each year and define the residuals as permanent earnings (PE), P_{t-1}^{i}.

Our focus is on how different moments of the earnings growth distribution vary across the joint distribution of PE and age for men and women separately. For this purpose, in each year t, we first divide workers into three age groups: 25–34, 35–44, and 45–55. Then, within each gender-age group, we rank individuals into 40 quantiles with respect to their level of P_{t-1}^{i}. Furthermore, to capture the earnings risk at the top of the distribution, we place the top 0.1% workers into a separate group. Finally, within each quantile, we compute the cross-sectional moments of earnings growth between periods t and $t+k$. This conditional distribution of residual earnings growth can be thought of as the income uncertainty that workers of the same gender with similar age and permanent income face looking ahead. In our figures, we plot the average of these moments over years between

Figure 10 – Dispersion of Earnings Growth by Permanent Income and Age

(a) Men (b) Women



Notes: Figure 10 shows the P90-P10 of the log growth rate of residual earnings for men and women within quantiles of the permanent income distribution, P_{it-1}. In each plot, the solid markers represent P90-P10 for those workers at the top 0.1% of the permanent income distribution for different age groups (diamond for 25 to 34 years old, square for 35 to 44 years old, and circle for 45 to 55 years old). See Section 2 for sample selection and definitions.

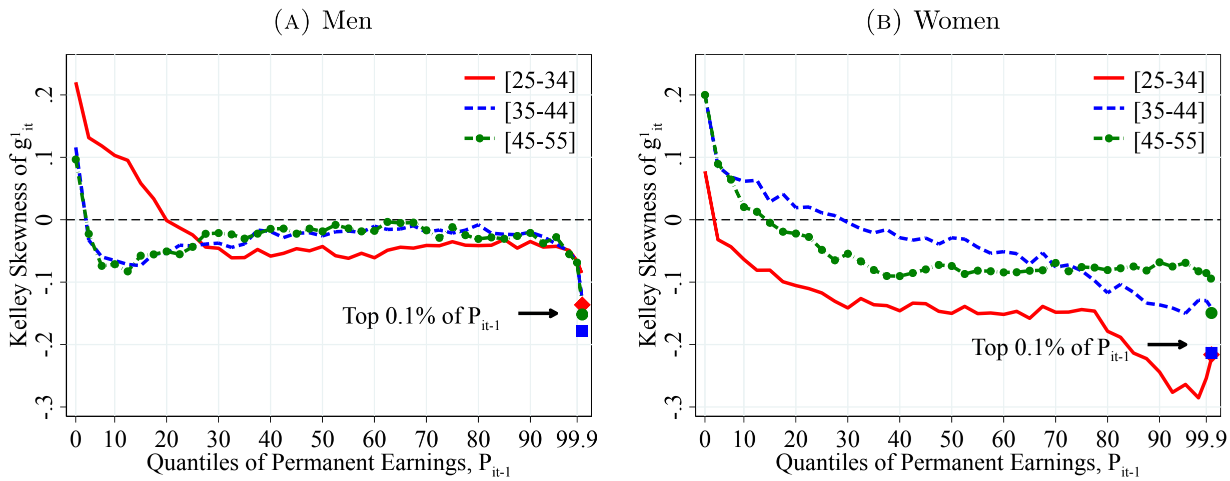
1997 and 2017-k.[19] In this section, we present the results for one-year log earnings growth ($k = 1$). Appendix B.4 show similar results for five-year log earnings growth ($k = 5$) and for the arc-percent growth measure.

Heterogeneity in Dispersion. We start with heterogeneity in the dispersion of earnings growth, shown in Figure 10. For men (Panel A), we find that the dispersion of earnings changes follows a pronounced U-shaped pattern over the PE distribution for every age group. For example, for 35- to 44-year-old men, P90-P10 declines from 1.5 for individuals at the 10th percentile of the permanent earnings distribution to 0.40 for those at the 90th percentile, and then rises rapidly to 0.7 for those at the 99th percentile. Furthermore, within the top 0.1%, earnings growth—identified by the solid markers in Figure 10—dispersion increases further, reaching almost the same level of those individuals at the bottom of the permanent earnings distribution. This finding is consistent with top earners being more likely to have performance-based compensation and therefore risker incomes (e.g., Parker and Vissing-Jørgensen, 2010). For women (Panel B), P90-P10 declines by permanent income except at the very top end of the distribution and except for the oldest group.[20] Notice also that the age variation is smaller relative to the variation over the PE distribution. Still, we find significant differences, with earnings

[19]The first year we can compute P^i_{t-1} is 1996 using data between 1993 and 1995.

[20]We rank workers within their gender groups; therefore, it is possible that individuals of different gender in the same quantiles have different incomes.

Figure 11 – Skewness of Earnings Growth by Permanent Income and Age



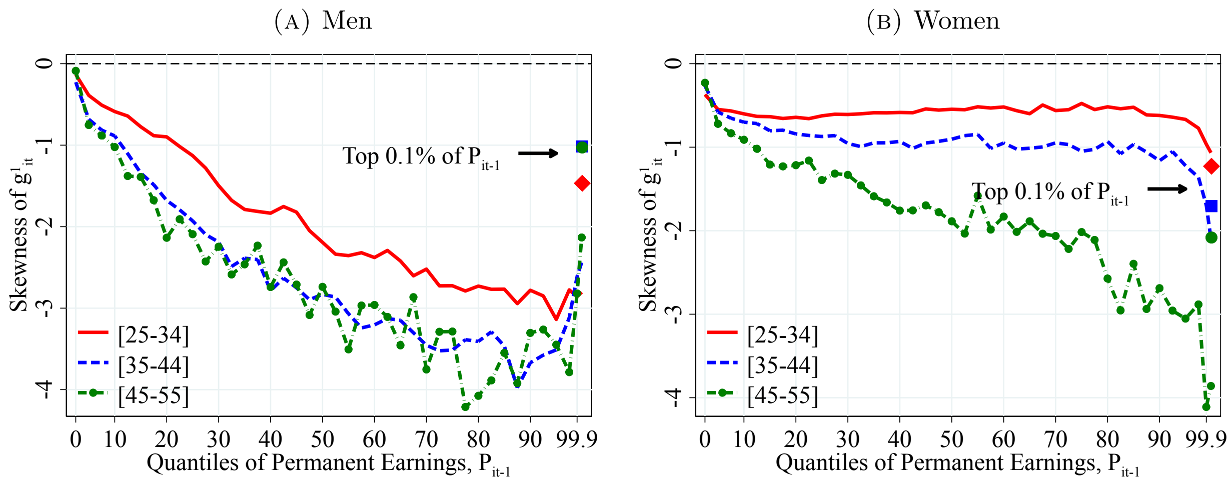
Notes: Figure 11 shows the Kelley skewness of the log growth rate of residual earnings for men and women within quantiles of the permanent income distribution, P_{it-1}. Kelley skewness is defined as $\mathcal{S}_{\mathcal{K}}$ = ((P90-P50) − (P50-P10)) / (P90-P10). In each plot, the solid markers represent the Kelley skewness for those workers at the top 0.1% of the earnings distribution for different age groups (diamond for 25 to 34 years old, square for 35 to 44 years old, and circle for 45 to 55 years old).

volatility being the largest among young women. These patterns are qualitatively similar to those found for the United States, albeit that the overall dispersion is lower in Norway (see Karahan and Ozkan, 2013).

Heterogeneity in Skewness. Figure 11 displays the Kelley skewness of one-year log earnings growth for men and women. Because the skewness patterns shown by the third standardized moment, and our conclusions drawn from them, are significantly different from those of the Kelley skewness, we complement our analysis by reporting the corresponding standardized moment in Figure 12.

Several aspects of Figures 11 and 12 are worth noticing. First, except for those at the bottom of the permanent earnings distribution, both male and female workers face a strongly left- (negatively) skewed earnings growth in all age groups. Second, the distribution of income changes become increasingly left skewed as we move from low- to high-earnings workers. Thus, as workers climb the earnings ladder, the more room there is to fall and the less room there is to move up. However, for men the decline in the Kelley skewness measure by permanent income is neither as strong nor as monotonic as that from the standardized moment. For example, in Panel A of Figure 12, the standardized moment ranges from 0 for bottom earners to -4 for 45- to 55-year-old workers in the 80th percentile of the PE distribution, whereas Kelley skewness ranges non-monotonically between 0.2 for workers at the bottom of the income distribution to a value -0.2 for

Figure 12 – Skewness of Earnings Growth by Permanent Income and Age



Notes: Figure 12 shows the third standardized moment of the log growth rate of residual earnings for men and women within quantiles of the permanent income distribution, P_{it-1}. In each plot, the solid markers represent the corresponding measure of skenwess for those workers at the top 0.1% of the permanent income distribution for different age groups (diamond for 25 to 34 years old, square for 35 to 44 years old, and circle for 45 to 55 years old).

workers at the top, but stays close to zero for middle-income workers. Furthermore, when measured with the third standardized moment, we find that for workers at the top 0.1% of the PE distribution (except for women in the youngest age bin), earnings growth is less left skewed relative to their other high-income peers, suggesting that a significant fraction of workers at the top of the PE distribution experience large, positive earnings changes. Recall that the Kelley skewness ignores income changes in the top and bottom 10 percentiles of the distribution.

Third, the distribution of earnings growth becomes more left skewed over the life cycle, with the largest differences being between the youngest and the oldest age groups. In other words, like high-income workers, for the oldest workers there is more room to fall and less room to move up on the earnings ladder. Again, the Kelley and standardized measures of skewness show different patterns for men in different age groups: According to the Kelley measure, the age variation is less pronounced overall and highest for low-income workers, whereas the third standardized measure exhibits large age differences across the permanent earnings distribution except for the bottom and top ends. Overall, the variation in skewness by PE and age is remarkably similar to the variation found for the United States (see Guvenen *et al.*, 2018). The marked differences between the Kelley and third standardized moment highlight the importance of considering the entire earnings growth distribution when analyzing the higher-order moments of earnings risk.

Heterogeneity in Kurtosis. Lastly, we investigate the (excess) kurtosis of one-year earnings growth in Figure 13 for men (Panel A) and women (Panel B) using the Crow-Siddiqui measure. Several remarks are in order. First, we find that the distribution of earnings changes displays excess kurtosis across all workers, regardless of their permanent earnings level, age groups, or gender. Second, the extent of kurtosis varies significantly across the income distribution and over the life cycle. In particular, kurtosis exhibits a hump-shaped profile over the PE distribution and is usually higher for older workers. For example, excess kurtosis is highest at roughly 11 for the oldest men around the 20th to 30th percentiles, compared to a mere 2 for the bottom earners in the youngest age group.[21] Among women, the kurtosis of earnings growth peaks at a higher level of 14 for middle-aged women around the 80th to 90th percentiles, versus 2, again, for young low income -workers.

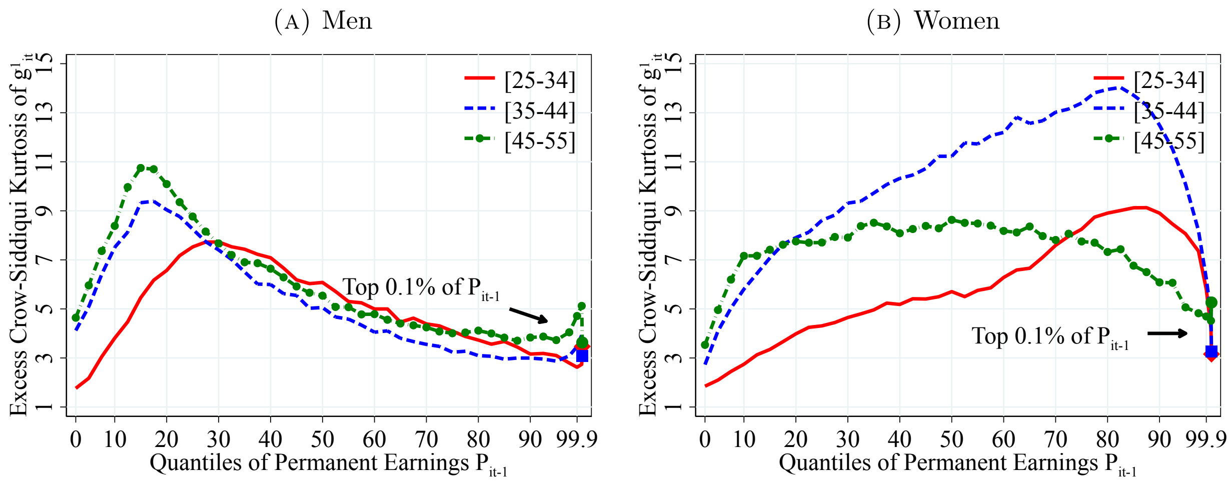
As in the case for skewness, the fourth standardized moment depicts relatively different patterns compared to the Crow-Siddiqui measure (shown in Figure B.7 Appendix B.4). In particular, for men, kurtosis increases up to the 80th and 90th percentiles of the PE distribution and only starts to decline at the 95th percentile. And, age variation is also larger when the standardized kurtosis measure is used. For women, the standardized measure of kurtosis is increasing over the entire PE distribution except for top earners who are 35 years and older. These results again suggest that extremely large changes in earnings (which are ignored by the Crow-Siddiqui measure) have an important effect on the higher-order moments of earnings changes.

5 Earnings Mobility

Having studied the properties of the distribution of earnings levels and growth, we now take a more longitudinal perspective and turn to the (intragenerational) persistence in earnings dynamics. Instead of identifying the persistence parameter from an estimated income process (e.g., Meghir and Pistaferri, 2004; Abowd and Card, 1989), we follow a descriptive approach and investigate how individuals' rank in the income distribution changes over time. For this purpose, we estimate two related measures of mobility. First, we calculate a measure of average rank-rank mobility that shows the expected position of an individual in the income distribution in year $t + k$ conditional on the

[21] The higher kurtosis for older workers who are between the 20th and 30th percentiles of the PE distribution indicates that typical annual earnings changes are small. However, a small but still significant share of workers in this age range experience extreme positive or negative changes in earnings.

FIGURE 13 – KURTOSIS OF EARNINGS GROWTH BY PERMANENT INCOME AND AGE



Notes: Figure 13 shows the excess Crow-Siddiqui kurtosis of the log growth rate of residual earnings for men and women with quantiles of the permanent income distribution, P_{it-1}. Excess Crow-Siddiqui kurtosis is defined as $\mathcal{C}_{\mathcal{K}}$ = (P97.5-P2.5) / (P75-P25) − 2.91 where 2.91 is the value of the Crow-Siddiqui measure for a Normal distribution. In each plot, the solid markers represent the corresponding measure of kurtosis for those workers at the top 0.1% of the earnings distribution for different age groups (diamond for 25 to 34 years old, square for 35 to 44 years old, and circle for 45 to 55 years old).

individual's position in year t. Second, we construct income transition matrices that show the probability that an individual in income quantile i in period t transitions into quantile j after k years.

Similarly to previous sections, here we use a measure of "permanent income" to isolate the persistent component of earnings. This measure, however, is slightly different from the permanent income used in Section 4.3 (P_{it-1}). In particular, the new permanent income is estimated by averaging levels of earnings of a worker i between years t and $t-2$ to obtain $P^*_{it} = \frac{1}{3}\sum_{j=0}^{2} Y_{it-j}$. We compute this measure for workers who have at least one year of labor earnings above the minimum income threshold, Y_t^{min}. Unlike the permanent income measure in Section 4.3, we do not residualize P^*_{it} out of year and age effects. Instead, we rank workers within each year and age, which controls for age and time effects not only in means but also in other moments.

The top row of Table III shows the average permanent earnings in selected percentiles of P^*_{it} in 2015. We find substantial heterogeneity across the distribution. For example, for the middle 40% group, average permanent earnings are \$84,157 and \$60,381 per year for men and women, respectively. In the bottom decile of the P^*_{it} distribution, the average annual permanent earnings are less than \$12,000 (or less than \$1,000 per month). This sizable fraction of prime-age men with very little labor earnings raises the ques-

Table III – Permanent Earnings Distribution in 2015

	Average Income (2018 US$) by Percentiles of P_{it}^*									
	Men					Women				
$P_{it}^* \rightarrow$	1-10	11-30	31-70	71-90	91-100	1-10	11-30	31-70	71-90	91-100
Earnings	11,149	48,636	84,157	124,065	205,345	7,657	32,054	60,381	87,184	134,858
SE Inc	14,257	3,322	648	401	526	4,198	1,552	460	260	312
Benefits	22,348	9,675	3,367	2,212	1,915	28,742	18,930	10,726	6,032	5,231

Notes: Table III shows the average permanent earnings, self-employment income (SE Inc), and benefits for individuals in selected quintiles of the permanent earnings (P_{it}^*) distribution in 2015. All nominal values are deflated to their 2018 real values using the Consumer Price Index in Norway. To make our results comparable across countries, we convert NOK values to US dollars using the 2018 exchange rate.

tion of whether they have other sources of income such as self-employment income or social safety benefits. Our data from administrative sources allow us to investigate this question: The next two rows of Table III document average permanent self-employment income and permanent benefits in the same percentiles of the permanent income distribution.[22]

Indeed, workers at the lower end of the P_{it}^* distribution have substantial income from self-employment and from public benefits. For example, the average self-employment income of men in the bottom decile of P_{it}^* is higher than their permanent earnings (14,250 US$ versus 11,149 US$). However, self-employment income declines sharply to less than $1,000 for workers above the 30th percentile. Public benefits are an even more important source of income throughout the P_{it}^* distribution, especially for women, ranging from almost $29,000 in the bottom P_{it}^* decile to more than $5,000 in the top decile.

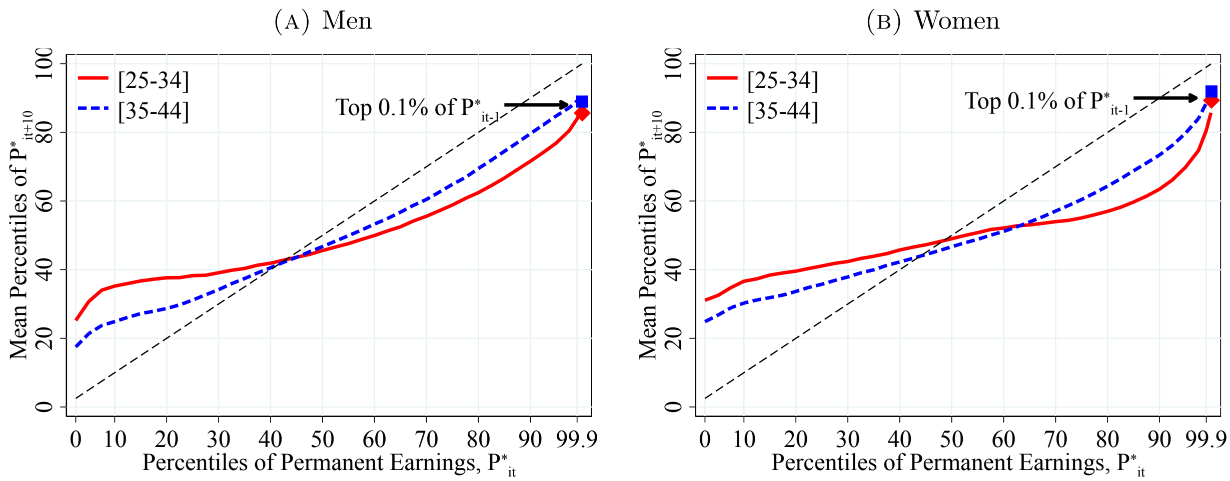
5.1 Average Rank-Rank Mobility

We rank workers into 40 quantiles in period t within each gender and age with respect to their permanent income, P_{it}^*. To provide a more granular view at the top of the income distribution, we put the top 0.1% earners in a separate group. Then, for each income quantile, age, and gender, we calculate individuals' average rank (out of 100) in the future permanent income distribution in $t+k$.[23] In this section, we present this average rank-rank mobility measure between t and $t+10$ (10-year mobility). The results for

[22] Benefits include unemployment benefits, sickness benefits, paid parental leave, remuneration for participation in various government activity programs, disability benefits, public pensions, and other social welfare payments. Self-employment income includes business income. We construct permanent self-employment income and permanent benefits in the same way we compute P_{it}^* (i.e., by averaging them between $t-2$ and t).

[23] In the analysis of mobility between t and $t+k$, our sample includes individuals who have non-missing observations of permanent income in both t and $t+k$.

Figure 14 – Income Mobility: Rank-Rank Measures by Age



Notes: Figure 14 shows the average rank obtained by individuals in period $t+10$ in the distribution of (alternative) permanent earnings, P^*_{it+10}, within different percentiles of the distribution of (alternative) permanent earnings in period t, P^*_{it}. To construct this figure, we calculate the average rank in $t+10$ for each year in our sample between 1993 and 2007 (the last years in which a 10-year change can be calculated) for each age group. We then average across all years in our sample.

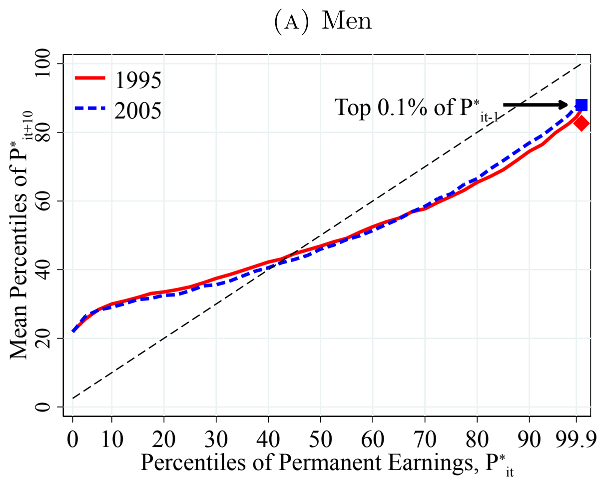
5-year ($k = 5$) mobility are presented in Appendix C.

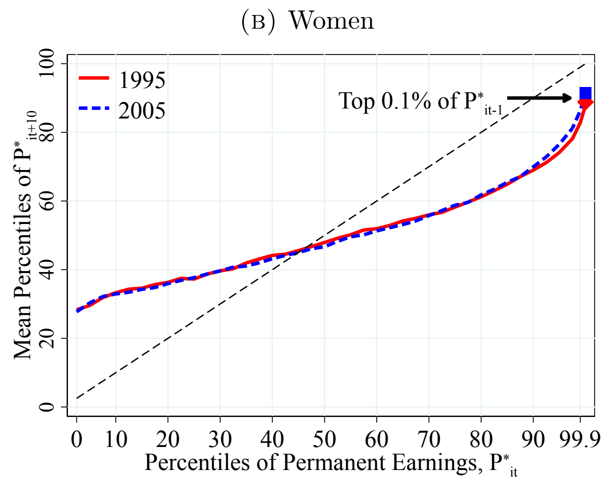
Income Mobility over the Life Cycle. Figure 14 presents average rank-rank mobility for two age groups, 25-34 and 35-44, separately. In particular, this figure shows the expected rank in year $t+10$ on the y-axis conditional on permanent income rank in year t on the x-axis. Notice that the 45-degree (dashed black) line represents the perfect immobility case (on average, individuals remain in the same percentile after $t+k$ years). Therefore, the closer the expected rank is to the 45-degree line, the more persistent the incomes are.

Several remarks are in order. First, income mobility declines significantly after the first decade of working life. For instance, male workers in the 10th percentile of the permanent income distribution in year t find themselves, on average, in the 35th percentile after ten years when they are between 25 and 34 years old, versus the 25th percentile when they are 35 years and older. This finding is consistent with the previous findings (see Karahan and Ozkan, 2013; Guvenen *et al.*, 2018) in that income growth becomes more persistent as individuals advance in their career until around the ages of 45 to 55.

Second, income mobility is the highest around the 10th and 90th percentiles of the income distribution, especially for young workers. Furthermore, upward mobility in the bottom half of the income distribution is higher relative to the downward mobility in

Figure 15 – Income Mobility: Rank-Rank Measures for Selected Years





Notes: Figure 15 shows the average rank obtained by individuals in period $t+10$ in the distribution of (alternative) permanent earnings, P^*_{it+10}, within different percentiles of the distribution of (alternative) permanent earnings in period t, P^*_{it}. To construct this figure, we calculate the average rank in $t+10$ for each year in our sample between 1993 and 2007 (the last years in which a 10-year change can be calculated) for each age group. We then average across all years in our sample.

the upper half of the distribution. In fact, higher upward mobility in the bottom half (compared to the downward mobility in the upper half) leads the rank-rank measure to cross the 45-degree line below the 50th percentile. Interestingly, incomes are increasingly more persistent above the 95th and below the 5th percentiles. We will discuss more about income mobility in the tails of the income distribution shortly.

Third, relative to men, income mobility is higher for women, and even more so for young women in the upper half of the distribution (Panel B of Figure 14). For example, young women (those between 25 and 34 years old) who are in the 90th percentile of the income distribution in year t find themselves, on average, in the 65th percentile after 10 years, whereas the similarly ranked young men roughly fall to the 80th percentile. For earners at the top 0.1% of the distribution, however, the persistence is slightly higher for women than for men.

Trends in Income Mobility. We then investigate whether income mobility has changed in Norway over the last 20 years. For this purpose, Figure 15 shows the average rank-rank mobility for two selected years pooling workers of all age groups. Considering that the income distribution in Norway does not seem to have gone through major changes during our sample period, it is not surprising to find that income mobility, measured by rank-rank correlations, has remained almost unaltered over the last 20 years for both men and women.[24] If anything, mobility seem to have declined for men, especially at higher ranks of the earnings distribution, as the average rank for 2005 is closer to the 45-degree line relative to 1995.

5.2 Income Transition Matrices

So far, our analysis has focused on the *average* rank-rank mobility of permanent earnings. To capture a more complete picture of workers' income transition dynamics, here we investigate where exactly individuals end up in the income distribution in year $t+k$, conditional on their rank in year t, by constructing first-order Markov transition matrices. In our analysis, we again use P^*_{it+k}, as our measure of income and rank workers within age and gender groups. We then define the following states in our transition matrices: the first four quintiles of the P^*_{it+k} distribution, the next 15 percentiles (81st-95th percentiles), the next 4 percentiles (96th-99th percentiles), the top 1% excluding the top 0.1%, and finally, the top 0.1% of the distribution. Furthermore, instead of dropping individuals who have no significant labor income three years in a row in $t+k$ (i.e., missing P^*_{it+k} observations) from our transition matrices, we explicitly investigate whether individuals have other sources of income. In particular, we add three more states that describe the status of individuals with missing P^*_{it+k} observations: self-employed workers (who have permanent self-employment income above the minimum income threshold Y_t^{min}), individuals with permanent public benefits greater than Y_t^{min}, and individuals who do not have any significant income (i.e., total permanent income less than Y_t^{min}).[25]

To save space, in Figure 16 we present 10-year transition matrices for men and women between 35 and 44 years old. To understand this figure, notice that the color intensity of each cell reflects the transition probability between the corresponding row and column shown in the cell. So, the darker the cell, the more likely the transition between two quantiles. For both men and women, the diagonal cells and their close neighbors are darker than the rest, indicating that most individuals remain in their original states even after 10 years, and if they move, they do not move far. This is especially true at the top and bottom of the distribution. For instance, among men, the probabilities in the diagonal cells (i.e., probability of staying the same state) decrease from 44% for the bottom quintile to 35% for the third quintile and then increase to 49% between the 81st

[24] This finding is in line with Aaberge *et al.* (2013) who found no significant change in market income mobility from 1969 to 1992, followed by a period of growth and decline that the authors attribute primarily to capital income.

[25] In every year, 1.8% of individuals have missing income observations because of emigration or death.

Figure 16 – Permanent Earnings Mobility: Transition Matrix

(A) Men (B) Women

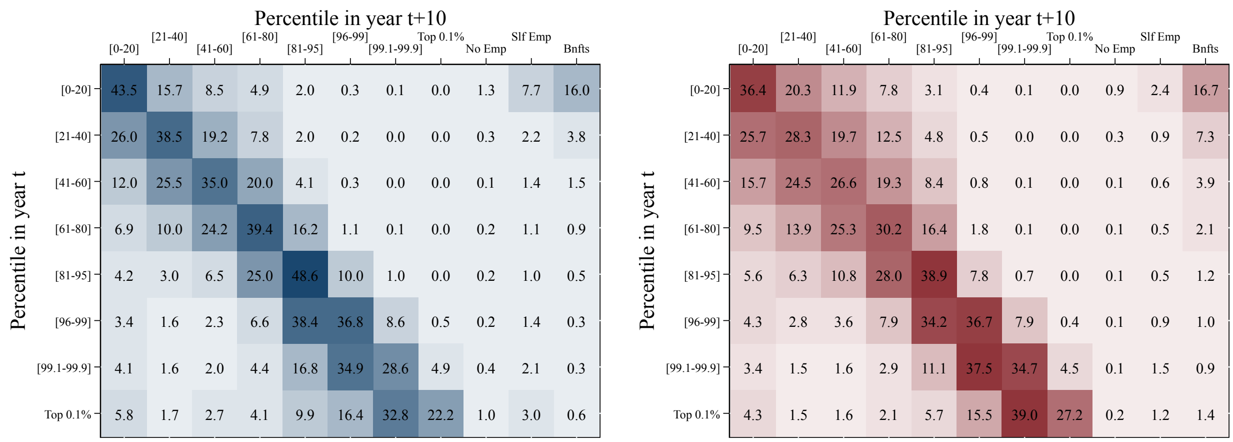


Figure 16 shows a first-order transition matrix of individuals' permanent earnings between periods t and $t+10$ for a sample of workers between 35 and 44 years old. To construct this figure, we calculate permanent earnings for workers between the years 1995 and 2007 (the first and last years for which we can calculate permanent earnings and 10-year changes). No Emp. corresponds to individuals whose permanent earnings is below the minimum income threshold and those who do not have significant self-employment income or social security benefits in period $t+10$. Slf Emp (Bnfts) corresponds to individuals whose permanent earnings are below the minimum income threshold but the average level of self-employment income (benefits) over the last three years is above the minimum income threshold. We then calculate the share of individuals transitioning between the predefined states for each year. Finally, we average the shares across all possible years.

and 95th percentiles. More broadly, remaining in the same state or transitioning into one of the neighboring states constitutes more than 60% of the cases. These findings suggest that individual rankings in the income distribution are quite persistent. These results hold for different age groups and different transition periods (see Appendix C).

Zooming into the top 1% of the distribution, we find that persistence is even higher at the top of the income distribution. For example, 35.6% of male workers who are in the top 1% in year t appear again in the same income bracket after 10 years. More interestingly, there are very few transitions between the lower and top ends of the distribution and vice versa. For example, most (more than 99.5% of) workers in the top 0.1% of the distribution in year $t+10$ were already in the top 5% in year t. Similarly, very few workers who are in the top 0.1% of the income distribution in year t end up outside of the top 5% in year $t+10$. Specifically, less than 25% of the top 0.1% earners fell below the 95th percentile in year $t+10$. This finding is inconsistent with calibrations of earnings processes with shocks that increase earnings to very high levels (e.g., the top 0.1%) but only temporarily (see Castaneda *et al.*, 2003). For women, top incomes are even more persistent, with a 42% probability of staying in the top 1% after 10 years.

When we say that 35.6% of workers appear again in the top 1% after 10 years we do

FIGURE 17 – NUMBER OF YEARS STAYING AT THE TOP 1% OVER 10 YEARS

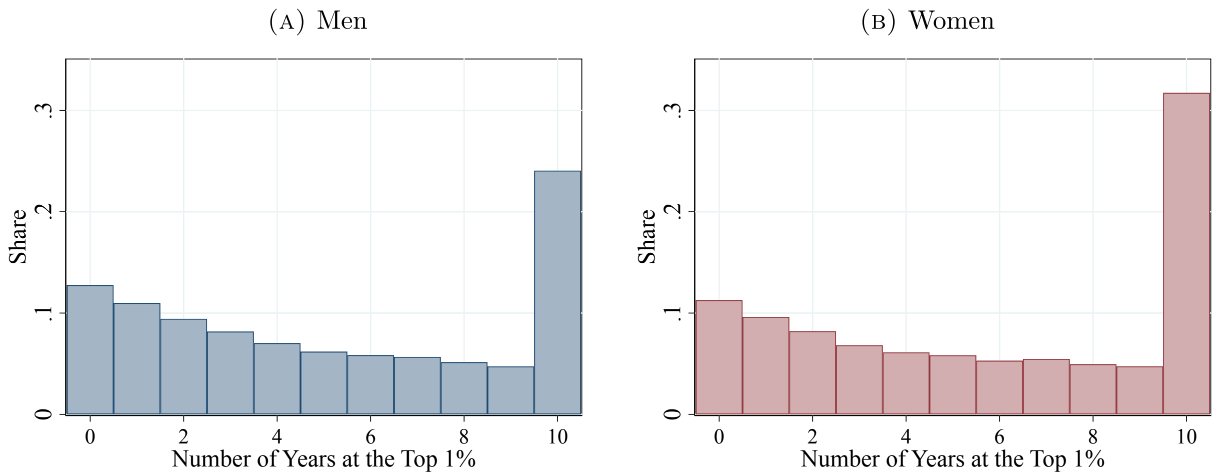


Figure 17 shows the fraction of top 1% workers in year t that appear in the same income group between $t+1$ and $t+10$ for 0 years, for 1 year, for 2 years, and so on. To construct this figure, we pool all observations between the years 1995 and 2007 (the first and last years for which we can calculate permanent earnings and 10-year changes).

not know whether this transition probability is the same for all workers just by looking at the results shown in Figure 16. For example, it may be that 35.6% of workers are always in the top income group and the rest temporarily appear in the top 1% only in year t, or that all top earners have the same probability of staying in the top 1%. These two different income dynamics have very different implications for consumption and saving decisions and portfolio allocation. To investigate the possible heterogeneity in the persistence of top incomes, we calculate the number of years a top earner in year t reappears in the top 1% over the next 10 years. In other words, we follow the top earners for the next 10 years and document the numbers of years they stay in the top 1%.

Our results, displayed in Figure 17, show that 12% of men at the top 1% of the permanent earnings distribution (Panel A) in year t do not appear at the top again over the next 10 years, whereas 11% will appear only one more time during the same period, and so on. Interestingly, around a quarter of the top 1% earners never leave that group over the next 10 years. The results are even more striking for women (Panel B): Almost one-third of the female top earners stay in the top 1% for 10 years in a row. This finding is consistent with our results from the transition matrices for women that show a higher probability of staying in the top 1%. Whether these findings are consistent with a simple first-order Markov process or whether there are ex ante differences in income dynamics among the top earners is an open question and beyond the scope of this paper.

Finally, we turn to workers who have no significant earnings (i.e., less than Y_t^{min}) in

$t+10$, which is shown in the three last columns of the matrices in Figure 16. Transitioning into self-employment follows a hockey-stick-shaped pattern over the P_{it}^* distribution: The probability of switching to self-employment declines from 7.7% for the bottom quintile wage earners to 1% between the 81st and 95th percentiles and then increases to 3% for the top 0.1% wage earners.[26] Finally, the probability of transitioning into relying on public benefits declines sharply from 16% in the bottom quintile of permanent earnings to less than 1% for workers above the median.

6 Intergenerational Transmission of Income Dynamics

The relationship between parents' and children's incomes has been a long-standing question of great importance in economics and public policy (see Piketty (2000), Corak (2013), and Jäntti and Jenkins (2015) for recent reviews of the literature). Following Solon's seminal work (Solon, 1992), several papers have documented a positive correlation between parents' and children's income in the United States, the United Kingdom (Long and Ferrie, 2013), and several other countries, including Norway (Bratberg *et al.* (2005), Pekkarinen *et al.* (2017), and Markussen and Røed (2019)).

Unlike most of the literature, however, which has focused on the relationship between parents' and children's income *levels*, in this section we investigate the intergenerational transmission of income *dynamics*.[27] In particular, we investigate whether the children of fathers with steeper life-cycle income growth, more volatile incomes, or higher downside risk also have income streams of similar properties. Such correlations can arise, for example, if fathers and children genetically share similar risk attitudes or work in similar jobs and sectors. Another question is how the income dynamics of workers vary with parents' financial resources. The answer to this question can shed light, for instance, on the roles of (i) the *dynastic* precautionary savings motive of parents in wealth accumulation (as in Boar (2020)), (ii) the importance of family resources for children's human capital accumulation (as in Holter, 2015), and (iii) the importance of parental insurance for the career choices of young adults (i.e., whether children of families with more resources can pursue high-risk, high-return careers, as in Fawcett, 2020).

[26] Herreno and Ocampo (2020) also show that the self-employment rate follows a U-shaped pattern over the earnings distribution in the United States as well as in developing countries.

[27] Two other papers study some aspect of the intergenerational transmission of income risk. Shore (2011) shows that the children of parents with more volatile incomes have risker income streams in the PSID. Similar to our approach, Jäntti and Lindahl (2012) find a U-shaped relationship between parents' income and the dispersion of children's log earnings for Sweden.

We start our analysis in Section 6.1 by documenting the intergenerational *lifetime* income mobility for both completeness and consistency with the earlier work. Next, we investigate how workers' income risk varies according to their fathers' financial resources. For this purpose, we follow our graphical approach in Section 4.3 and document the variation in moments of children's idiosyncratic income changes by their fathers' lifetime income and wealth (Section 6.2). Finally, in Section 6.3, we study the intergenerational transmission of income *dynamics* by documenting the relation between (the first three) moments of parents' and children's income changes over the life cycle.

Intergenerational Data. The analysis in this section is based on a dataset with a longer panel, dating back to 1967 and ending with the most recent data from 2017. The main purpose of this administrative register is to calculate social security pensions benefits, which was first institutionalized in 1967. The income measure in this dataset reflects this goal and therefore is different from the baseline measure of earnings we used in previous sections. In particular, it includes all pension-point-generating labor earnings and government transfers such as wages, self-employed income, unemployment benefits, paid sick leave, and parental leave.[28] All nominal values are deflated to their 2018 real values using the Consumer Price Index in Norway.

All administrative registers include personal identifiers that allow us to link them together, and in particular to link children to their parents. The information on family links is collected from the Norwegian population register, which was established in the early 1960s using information from the 1960 census. All individuals born after 1950 can be linked to their parents. For earlier cohorts, we identify most parents, but not all. We focus on father-children pairs to prevent our results being affected by the increase in female labor force participation in Norway over our sample period.

The very long panel dimension of the data—specifically, 51 years long—is crucial

[28]Certain measurement issues warrant specific attention. First, some observations are top-coded as a result of the maximum earnings limit for pension rights. However, inspection of the data shows that this top coding has not been applied systematically and occurs mainly in the years prior to 1979, was applied more spuriously in the period between 1979 and 1986, and has not been applied since 1987. Second, there are changes to how benefits are included in the data throughout the sample period: (i) Paid sick leave and parental leave was included in 1978 (paid sick leave was included even before 1978 if more than 90% of annual income); (ii) unemployment benefits were included in 1980; and (iii) various work activity benefits were included in 2002, 2004, and 2010. There are no separate long time series that allow us to systematically remove unemployment, sickness or parental benefits. Since work activity benefits were included after 1993, we were able to check whether removing these benefits would affect the results. They did not, but then again, these are small benefits compared to sickness and parental leave benefits.

for our purposes for at least two reasons. First, it allows us to precisely measure each individual's income risk over the life cycle. This is also why we are working with a sample of workers from a wider age range, specifically, between 23 (typical college graduation age) and 60 (instead of the 25-55 age range used in previous sections) to maximize the number of observations for each individual. For example, we have at least 20 years of data for cohorts born before 1998, which allows us to compute individual-specific income risk measures (e.g., percentile-based dispersion, skewness, and kurtosis measures of income growth) using 20 or more observations of annual incomes. Second, it is crucial to use a dataset with a long panel to study the link between parents' and children's income dynamics because a short-panel dataset can end up having only a few observations for one of the individuals in a father-child pair because of the short span of overlapping working lives.

Trends in Inequality and Volatility since 1967: A Brief Digression. Our long administrative database provides a long-term overview of the evolution of income inequality and instability in Norway, and several interesting facts emerge from this longer dataset (see Appendix D). For instance, we find that the small increase in earnings inequality for men between 1993 and 2017 (see Section 3.1) is part of a longer trend that started in the early 1970s (see Figure D.1). However, inequality among women seems to have remained relatively stable since the 1980s. We also find that income volatility has been declining in Norway over the last 50 years, especially among women (see Figure D.2), whereas skewness and kurtosis show the same patterns as those found in our baseline results (see Figure D.3). The differences between our baseline results and those derived from our long administrative sample are mainly a result of government transfers that reduce income volatility. These differences are, however, not a cause for concern for us because the analysis in this section pertains to the comparison of earnings risk within cohorts and not its variation over time.

Wealth Data. The wealth measure used in this section includes financial and non-financial assets derived from Norwegian administrative records and available from 1993 on. This high-quality dataset is mostly third-party reported to the tax authorities, and very little is self-reported. Employers, banks, brokers, insurance companies, and other financial intermediaries are obliged to send information on payment of earnings, the value of the asset owned, and information on the income earned on these assets. In our analysis, we use household net wealth that accounts for all financial wealth (e.g., stocks and bonds) and non-financial wealth (e.g., imputed value of houses, value of vehicles) net of the value of short- and long-term liabilities (e.g., credit card debt and mortgages).[29]

6.1 Intergenerational Income Mobility

We start by documenting the relation between fathers' and children's *income levels.* It is crucial to use a permanent income measure when estimating intergenerational mobility because transitory changes in fathers' and children's incomes attenuate the estimates (see Black and Devereux (2011); Solon (1992)). Therefore, in our analysis we use individuals' lifetime incomes, measured as the average income from ages 23 to 60. Notice, however, that our sample consists of cohorts who were born between 1908 and 1995, covering a 51-year time span. Furthermore, we observe some cohorts for only a short period of time, either when they are young or when they are old. Thus, to make the lifetime incomes comparable across different cohorts, we normalize incomes in each year and age by their corresponding year-age cell average. Then, we define the lifetime income of an individual i born in cohort c (birth year) as

$$LI_{i,c} = \frac{\left[\sum_{t=\max\{25+c,1967\}}^{\min\{60+c,2017\}} \frac{I_{it}}{d_{t,h(i,t)}}\right]}{\min\{60+c,2017\} - \max\{25+c,1967\} + 1}, \tag{1}$$

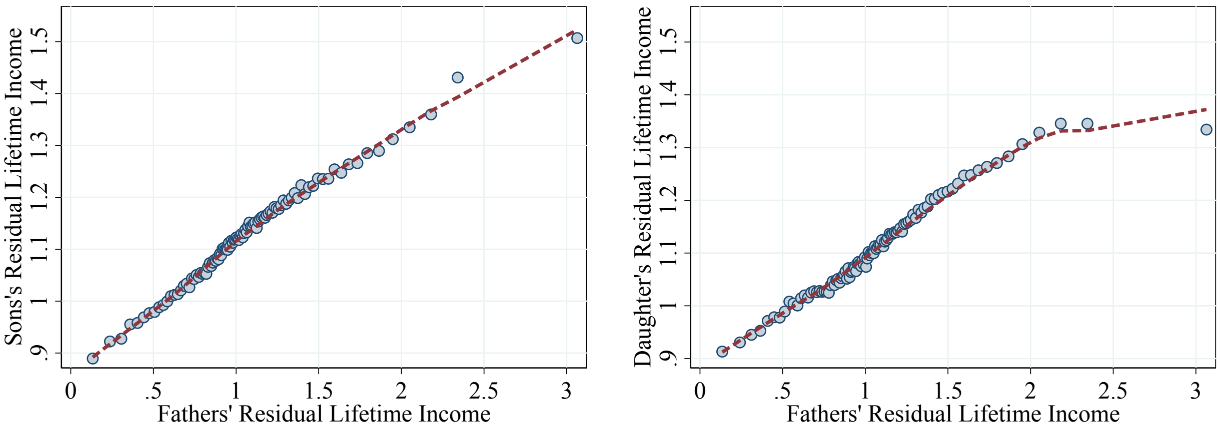
where I_{it} and $h(i,t)$ denote the real income and age of individual i in year t, respectively. We denote by $d_{t,h(i,t)}$ the average income of workers who are $h(i,t)$ years old in year t. Furthermore, to ensure that we have a reliable measure of lifetime income, we only include individuals who have at least eight years of income, I_{it}, above the minimum threshold, Y_t^{min}. The final sample includes 2.16 million father-child pairs, of which 1.04 million are father-daughter pairs.

Figure 18 shows a binned scatter plot of lifetime incomes of fathers and sons and daughters. In particular, on the x-axis we rank fathers into 100 bins with respect to their lifetime incomes and plot the average lifetime incomes of children in each bin on the y-axis. Our results confirm the findings of the earlier literature of a strong intergenerational persistence in income. We find that the relation between fathers' and children's lifetime income is fairly linear with an elasticity of intergenerational income of 0.24. This implies that an increase of 50 log-points in fathers' lifetime income is associated with an average

[29] See Fagereng *et al.* (2016) and Fagereng *et al.* (2019) for additional details of the measurement of wealth in Norwegian administrative resources, who use it to study return rate heterogeneity and saving behavior across the wealth distribution, respectively.

Figure 18 – Fathers and Children Income Correlation

(a) Sons (b) Daughters



Notes: Figure 18 shows a binned scatter plot of fathers' and children's residual lifetime income. All nominal values are deflated to their 2018 real values using the Consumer Price Index in Norway. The dashed line is a non-linear trend calculated using a LOWESS smoothing estimator. The plot is based on a sample of fathers and children with 20 years of data or more between 1967 and 2017. The final sample considers 2.16 million father-child pairs.

increase in sons' earnings of 18 log points. The results are quite similar for the father-daughter pairs, with a slightly weaker elasticity of 0.20 (Panel B of Figure 18).[30]

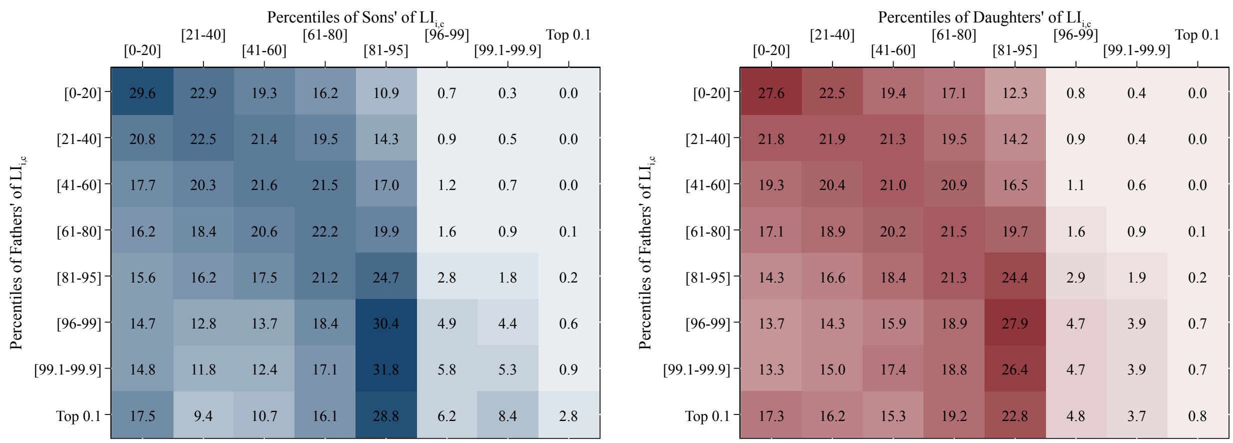
To obtain a more granular view of the relation between fathers' and children's income, we construct intergenerational transition matrices of lifetime income (Figure 19), similar to those presented in Section 5 (Figure 16). In this case, these matrices show the probability that a father in the ith percentile of lifetime income distribution (in the rows of the matrix) has a child at the jth percentile of the lifetime income distribution (in the columns of the matrix).[31] We find a significant degree of intergenerational persistence, with around 30% of the father-son (28% of the father-daughter) pairs sharing the bottom quintile of the income distribution. Furthermore, only 11% (1%) of the sons (13% (1.2%) of the daughters) of low-income fathers reach the top quintile (top 5%). Persistence is even stronger in the higher end of the distribution, with at least 32% (31%) of top-quintile fathers having sons (daughters) in the same quintile. Zooming into the top

[30] Bratsberg *et al.* (2007) find an intergenerational elasticity of 0.16 for Norway using the same Norwegian administrative data. However, they only include all males from the 1958 birth cohort matched with biological fathers' earnings records from 1971 and 1976, whereas our sample includes individuals from all cohorts that have at least eight years of income above a minimum income threshold. In fact, using a less restricted sample, we find an elasticity of 0.16, which falls within a range of elasticities for a variety of developed countries (between 0.15 and 0.55 as reported by Corak (2013)). Studies on the elasticity of fathers' and daughters' income are less common in the literature.

[31] In particular, we rank fathers, sons, and daughters separately among their peers with respect to their lifetime incomes, $LI_{i,c}$.

Figure 19 – Lifetime Income Mobility: Measures for Fathers and Children

(a) Sons (b) Daughters



Notes: Figure 19 uses fathers' and children's income data for a pooled sample of individuals between 1967 and 2012. The matrix shows the transition between selected quantiles of fathers' lifetime incomes (rows) and children's lifetime incomes (columns) for men and women. Lifetime income is the average residual income for individuals with at least three observations. Each row sums to 100%.

0.1% of the fathers' distribution, we find that 2.8% of their sons end up in the top 0.1% of the income distribution; that is, their sons are 28 times more heavily represented in the top 0.1% income group. Still, a significant fraction of children (17.5%) of top-earner (top 0.1%) fathers fall to the bottom quintile of the lifetime income distribution.

6.2 Fathers' Resources and Children's Income Dynamics

Arguably, parents' resources also have a significant effect on children's income dynamics. For example, high-income or high-wealth parents tend to spend more resources on their children's education, which allows them to accumulate more human capital and enjoy high incomes from more stable jobs. Alternatively, the children from rich families might be able to pursue high-risk, high-return careers or can simply afford to live off their inheritances without maintaining stable employment. Furthermore, parents may accumulate wealth to insure their children against income risk (i.e., dynastic precautionary saving, as in Boar (2020)).

Motivated by these ideas, we now investigate the relation between family resources—measured as fathers' lifetime income and net wealth—and children's income dynamics. For this purpose, we employ a methodology similar to Section 4.3 and analyze how the mean, dispersion, skewness, and kurtosis of one-year log income changes of children vary across the distribution of fathers' lifetime income and wealth. In particular, we rank fathers into 40 quantiles with respect to either their lifetime income, $LI_{i,c}$, or their wealth, denoted by $W_{i,c}$. Furthermore, to provide a more granular view of the children of very rich parents, we group fathers at the top 1% of the income or wealth distribution in a separate bin. We then calculate the first four moments of residual income growth, $g_{it}^{I,k}$, of children within each quantile.[32] The results for the distribution of five-year changes, which capture more persistent changes in earnings, show qualitatively similar patterns and are shown in Appendix D.2.

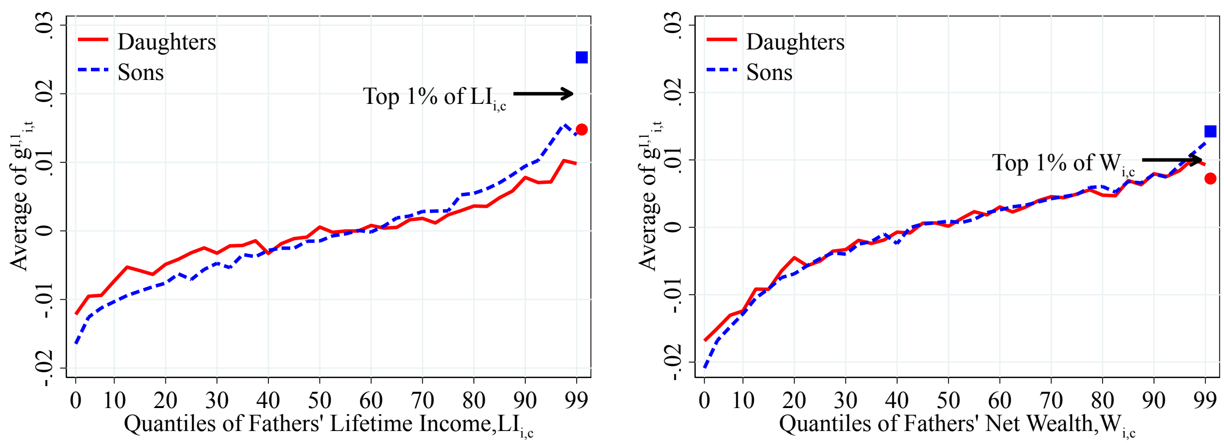
To make the measures of wealth comparable between different cohorts, similar to how we construct the lifetime income measure, we normalize each individual's net wealth by a year-age cell average. The data on wealth only span from 1993 to 2014; therefore, we cannot observe the wealth of most fathers when they were young (before 1993). However, this is a minor issue since wealth is a stock variable, unlike income (a flow variable). Therefore, in our analysis we consider the average residual net worth of a household calculated between ages 45 and 54. Some cohorts either are not in our sample or have only a few observations in this age interval. For them, we use observations in years closest to these ages. We require fathers to have observations of wealth for at least two years to be included in the sample. Our final sample consists of 1.69 million father-child pairs.

Average Income Growth. We start by discussing how sons' and daughters' average income growth varies by family resources, specifically, by fathers' lifetime income and wealth (Figure 20). We find that life time income growth is significantly higher for workers born into richer families. For example, children of parents at the 90th percentile of the lifetime income distribution enjoy annual income growth that is around 2 log points higher relative to those children with parents at the 10th percentile. This difference is economically significant too, considering that the typical estimates of the standard deviation of heterogeneous income profiles are around 2% for the United States (see Guvenen *et al.* (2018)). Children of top earners enjoy an exceptionally steeper average income growth over the life cycle—specifically, 1 log point higher compared to those from the next 2.5%. The variation over the father's wealth distribution is qualitatively and quantitatively similar.

[32]The residual income growth is defined similarly to the residual earnings growth, $g_{i,t}^{k}$ of Section 4.3 In particular, the residual income growth between t to $t+k$ is given by $g_{it}^{I,k} = \tilde{\varepsilon}_{it+k}^{I} - \tilde{\varepsilon}_{it}^{I}$ for $k = 1, 5$, where residual income $\tilde{\varepsilon}_{it}^{I}$ is obtained by regressing log income in each year and for each gender on a set of age dummies for those above minimum income threshold Y_t^{min}.

Figure 20 – Mean Log Earnings Growth by Fathers' Resources

(A) By Lifetime Income (B) By Net Wealth



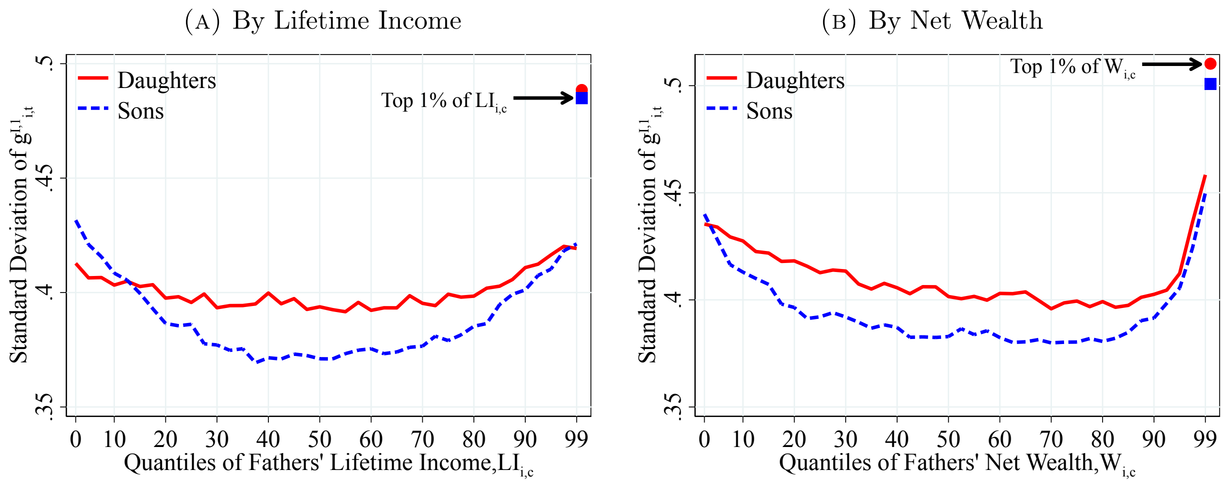
Notes: Figure 20 shows the average of one-year residual earnings growth for men and women within quantiles of fathers' lifetime income distribution (Panel A) and fathers' household net wealth distribution (Panel B) in 40 quantiles. Each line was been normalized to have a mean of 0. The top 2.5% of the distribution is further separated in two groups (97.5th to 99th and 99th percentile and above) for a total of 41 quantiles. We show the average across annual moments between 1990 and 2017. Darker markers correspond to the top 1% group.

Volatility of Income Growth. Figure 21 shows that the volatility of children's income growth follows a U-shaped pattern by their fathers' lifetime incomes (left panel) and wealth (right panel). This pattern is more pronounced for sons than for daughters. The variation by father's financial resources is relatively small, however, except at the top end of the lifetime income and wealth distributions.

The children very affluent fathers face significantly more volatile incomes over the life cycle. For example, the standard deviation of income growth for workers with fathers from the top 1% of the lifetime income or wealth distribution is around 10 to 12 log points higher compared to those of fathers at the median. This higher income volatility for children with rich fathers, combined with their exceptionally higher average income growth (shown in Figure 20), suggests that they can pursue high-risk and high-return careers that children from modest backgrounds cannot.

Recall that we also find a U-shaped pattern in the dispersion of earnings growth over workers' own permanent earnings in Section 4.3 (Figure 10). However, the U-shaped pattern is tilted toward the right over the fathers' lifetime income and wealth distribution compared to the variation by workers' own permanent earnings. That is, children of high-income and high-wealth fathers experience the most volatile incomes, whereas the volatility of earnings shocks is the highest for workers with the lowest permanent earnings. These results suggest that the relation between workers' income volatility and their

Figure 21 – Dispersion of Log Earnings Growth by Fathers' Resources



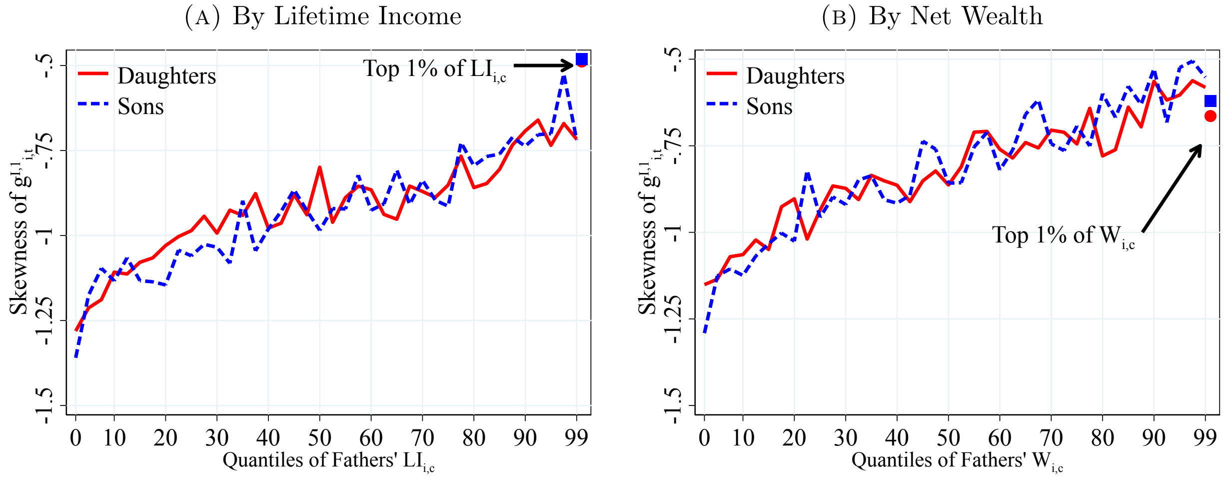
Notes: Figure 21 shows the standard of the one-year residual earnings growth for men and women within quantiles of fathers' lifetime income distribution (Panel A) and fathers' household net wealth distribution (Panel B) in 40 quantiles. The top 2.5% of the distribution is further separated in two groups (97.5th to 99th and 99th percentile and above) for a total of 41 quantiles. We show the average across annual moments between 1990 and 2017. Markers show the average for children whose parents were at the top 1% of the corresponding distribution.

fathers' financial resources is not due to the omitted variable of workers' permanent earnings, which is correlated with fathers' resources. We investigate this conjecture in the next section when we control for other worker and father characteristics.

Skewness of Income Growth. We next turn to the variation in skewness of children's income growth across the fathers' lifetime income and wealth distributions. Consistent with the previous results, Figure 22 shows that the distribution of income growth is left skewed regardless of fathers' income and wealth. More importantly, however, we find that the distribution of income growth becomes increasingly less left skewed for both sons and daughters as we move from poorer to richer families. This finding suggests that children from more affluent families experience higher upside income potential or lower left-tail risk or both. Differences between high- and low-resource fathers are substantial and economically significant.

Another important question is whether skewness becomes less negative over fathers' resources because of a compression of the lower tail (less risk of large declines) or because of an expansion in the upper tail (more opportunities for large gains). To answer this question, we investigate how the left and right tails of the children's income growth distribution change between poor and rich parents. In particular, Figures 23 and 24 show the differential between the 50th and 5th percentiles (P50-P5) and the 95th and 50th percentiles (P95–P50) of children's income growth. First, up to around the 85th

Figure 22 – Skewness of Log Earnings Growth by Fathers' Resources



Notes: Figure 22 shows the skewness (the third standardized moment) of the one-year residual earnings growth for men and women within quantiles of fathers' lifetime income distribution (Panel A) and fathers' household net wealth distribution (Panel B) in 40 quantiles. The top 2.5% of the distribution is further separated in two groups (97.5th to 99th and 99th percentile and above) for a total of 41 quantiles. We show the average across annual moments between 1990 and 2017. Markers show the average for children whose parents were at the top 1% of the corresponding distribution.

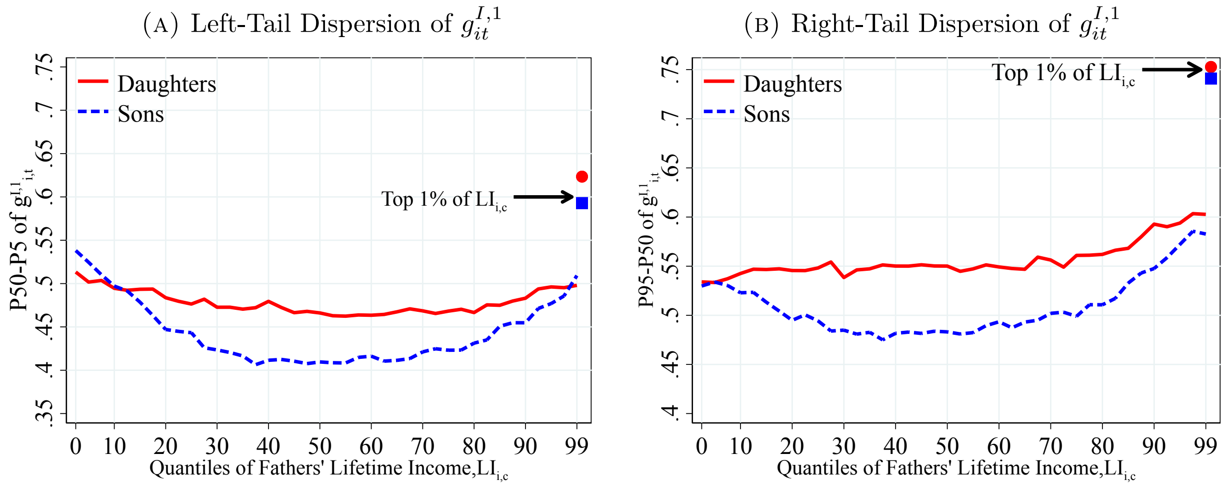
percentile of the fathers' income and wealth distributions, the P95-P50 is more or less constant, whereas the lower tail dispersion compresses substantially as we move from poorer to richer parents. Therefore, the decline in income volatility in this range reflects mostly a reduction in the left-tail risk of workers (see also Kelley's skewness in Figure 22 and Figure D.7 in Section D.2). However, beyond the 85th percentile we see that both tails open up sharply, the upper tail even more so, thereby resulting in both an increase in volatility and a reduction in negative skewness. This finding is consistent with our conjecture that children from affluent families pursue high-risk high-return careers.

Kurtosis of Income Growth. Finally, in Figure 25 we look at the kurtosis of income growth conditional on fathers' lifetime income and wealth. Similar to the results shown in Section 4.3, here we find a hump-shaped profile with low kurtosis of earnings growth among children whose fathers were at the top or bottom of the lifetime income distribution relative to those whose fathers were in the middle of the lifetime income distribution. This hump-shaped pattern is more pronounced for men than for women. Children of the richest fathers face the least leptokurtic distribution of income changes.

6.3 Fathers' and Children's Income Dynamics

In this section, we study whether the income dynamics of fathers and children have similar features, which we measure from their individual income fluctuations over the

Figure 23 – Dispersion of Log Earnings Growth by Fathers' Lifetime Income



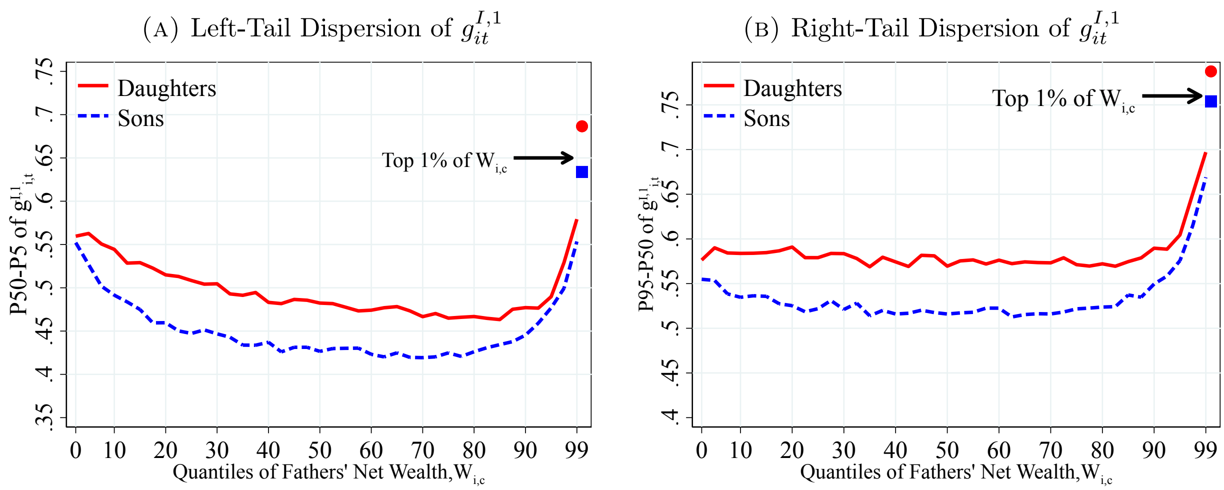
Notes: Figure 23 shows P50-P5 and P95-P50 of the one-year residual earnings growth for men and women within quantiles of fathers' lifetime income distribution (Panel A) and fathers' household net wealth distribution (Panel B) in 40 quantiles. The top 2.5% of the distribution is further separated in two groups (97.5th to 99th and 99th percentile and above) for a total of 41 quantiles. We show the average across annual moments between 1990 and 2017. Markers show the average for children whose parents were at the top 1% of the corresponding distribution.

life cycle. In particular, for each individual i, we construct a time series of permanent income $\tilde{P}_{it}$, which is defined as the average income between periods t and $t-2$ winsorized at the minimum income threshold (Y_t^{min}), $\tilde{P}_{it} = \max\left\{Y_t^{min}, \frac{1}{3}\sum_{j=0}^{2} Y_{it-j}\right\}$.[33] Then, for each individual we compute the first three moments from a log permanent income growth stream, $\Delta\tilde{P}_{it} = \log\tilde{P}_{it} - \log\tilde{P}_{it-1}$. To ensure a robust measure of moments, we restrict our sample to individuals with at least 20 years of income observations, half of which are above Y_t^{min}. This leaves us with 155,300 father-child pairs, of which 75,600 are father-daughter pairs. Furthermore, given the relatively short length of the individual time series—between 20 and 40 years—we use percentile-based moments (the median, P90-P10, and Kelley skewness) to avoid having our results being driven by outliers. The results for the standardized moments display similar qualitative patterns (Appendix D.2).

Median Income Growth. We begin our discussion by comparing fathers' and children's life-cycle income growth. Figure 26 shows a binned scatter plot of the median log permanent income growth of fathers and sons (left panel) and fathers and daughters (right panel). We find a marked non-linear relation between fathers' and children's lifetime income growth. Fathers' and children's income growth does not seem to be strongly correlated for around 10% of our sample when fathers have negative life-cycle income

[33]Different from permanent income measures used in previous sections, we compute this variable even for workers with zero earnings three years in a row.

Figure 24 – Dispersion of Log Earnings Growth by Fathers' Wealth



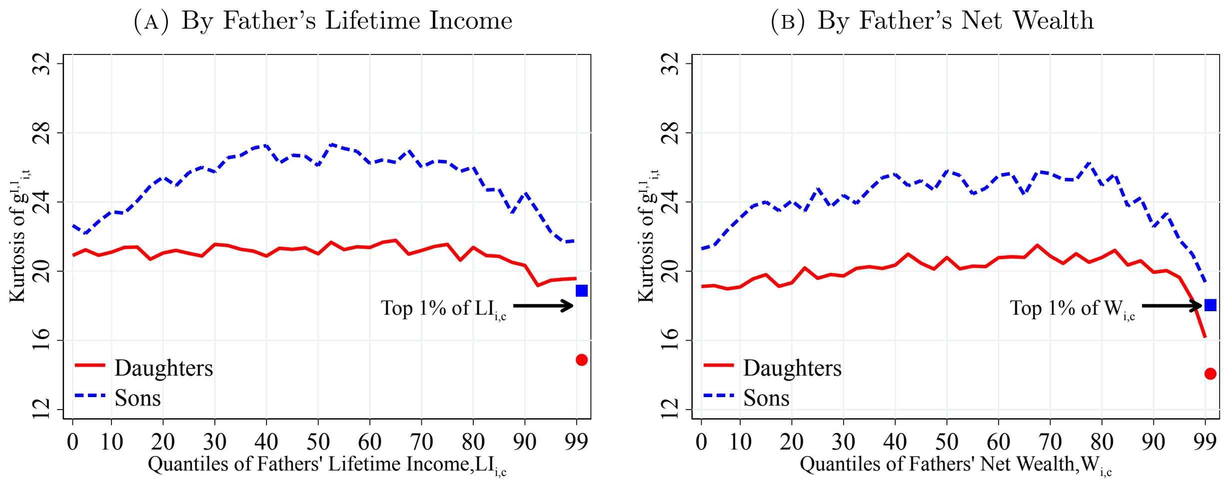
Notes: Figure 24 shows P95-P50 and P50-P5 of one-year residual earnings growth for men and women within quantiles of fathers' lifetime income distribution (Panel A) and fathers' household net wealth distribution (Panel B) in 40 quantiles. The top 2.5% of the distribution is further separated in two groups (97.5th to 99th and 99th percentile and above) for a total of 41 quantiles. We show the average across annual moments between 1990 and 2017. Markers show the average for children whose parents were at the top 1% of the corresponding distribution.

growth. In the rest of the sample, however, the children of fathers that experienced steeper income growth during their lifetime are more likely to similarly experience high income growth. This correlation is also economically significant: An increase in a father's median income growth from 0 to 5 log points results in an increase of roughly 1 log point in the son's median income growth over the life cycle. For daughters, this number is slightly lower (around 0.8 log points) but still significant. The overall intergenerational elasticity of life-cycle income growth is 0.16 for men and 0.09 for women.[34]

Income Growth Volatility. We now investigate whether the children of fathers with volatile incomes also have riskier incomes. Figure 27 shows a binned scatter plot of the P90-P10 of fathers' and children's permanent income growth stream. We find a strong and economically significant correlation between fathers' and children's volatility of income. For example, when the father's dispersion of income changes increases from 10 to 50 log points, where the bulk of the sample is, the son's (daughter's) P90-P10 of income growth increases from 35 to 45 (45 to 55) log points, which implies an intergenerational elasticity of dispersion for sons (daughters) of 0.25. For more volatile incomes of fathers,

[34]Using Canadian administrative data, Lochner and Park (2020) find no significant correlation between fathers' and children's income growth. They show that the covariance between children's earnings growth at a given age (27 years old in their baseline case) and fathers' earnings growth is not statistically significant. Our analysis differs from theirs in that we calculate the correlation between the *lifetime* earnings growth of fathers and children rather than at a particular age.

Figure 25 – Kurtosis of Log Earnings Growth by Fathers' Resources



Notes: Figure 25 shows the excess kurtosis (the fourth standardized moment minus 3) of the one-year residual earnings growth for men and women within quantiles of fathers' lifetime income distribution (Panel A) and fathers' household net wealth distribution (Panel B) in 40 quantiles. The top 2.5% of the distribution is further separated in two groups (97.5th to 99th and 99th percentile and above) for a total of 41 quantiles. We show the average across annual moments between 1990 and 2017. Markers show the average for children whose parents were at the top 1% of the corresponding distribution.

the association becomes flatter, though still significant: An increase in the P90-P10 of fathers' income from 50 to 150 is associated with an increase in children's dispersion of only around 10 log points.

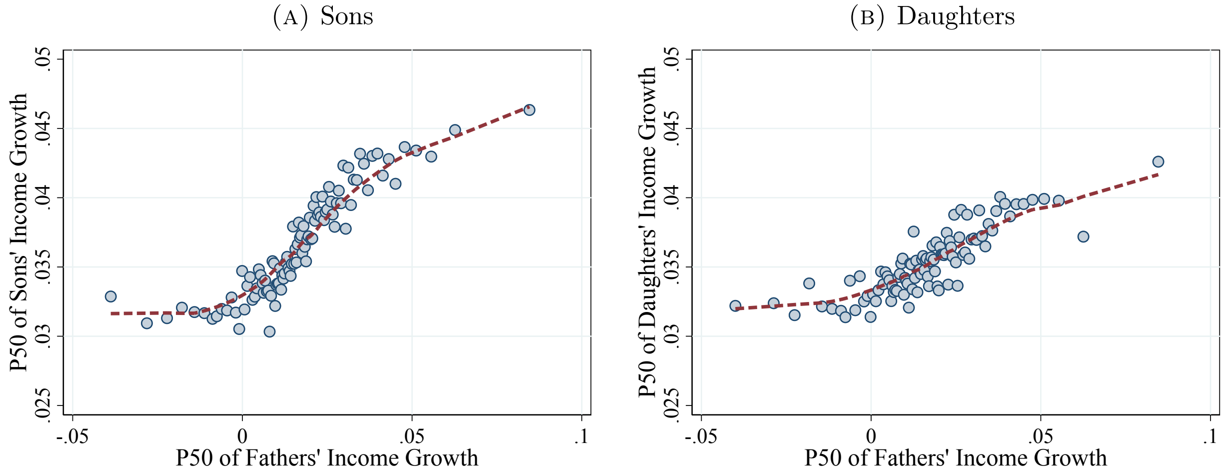
Skewness of Income Growth. We end this section by looking at the relation between fathers' and children's skewness of income growth. As we have seen for the first two moments, Figure 28 shows a positive and strong correlation for Kelley skewness. Quantitatively, we find that an increase in fathers' Kelley skewness from -0.25 to 0.25—where most of the distribution is found—is associated with an average increase of 0.05 in the skewness of children's income growth.[35]

6.4 Determinants of the Transmission of Income Risk

The results presented in this section so far indicate that children's income dynamics are strongly connected to their fathers' economic resources and income dynamics. However, there are two main potential concerns regarding our descriptive analysis. First, simple bivariate correlations cannot quantify the relative importance of different factors (i.e., fathers' lifetime income, net wealth, and moments of income growth) on children's income dynamics as these factors are also correlated with each other. Second, some of the

[35] We do not find a significant relation between the levels of kurtosis of fathers' and children's earnings growth (Figure D.15 in Appendix D.2).

Figure 26 – Median Income Growth of Fathers and Children



Notes: Figure 26 shows a binned scatter plot of fathers' and children's median lifetime income growth. The dashed line is the non-linear correlation estimated from a LOWESS estimator. The scatter plot is based on a sample of 155,300 fathers-child pairs. The sample is divided into 100 bins.

strong relations documented above may simply be spurious as a result of omitted variables. For example, as we have shown in Section 4.3, workers' differences in permanent income are an important determinant of their income risk, and fathers' and children's lifetime incomes are strongly correlated (Figure 18).

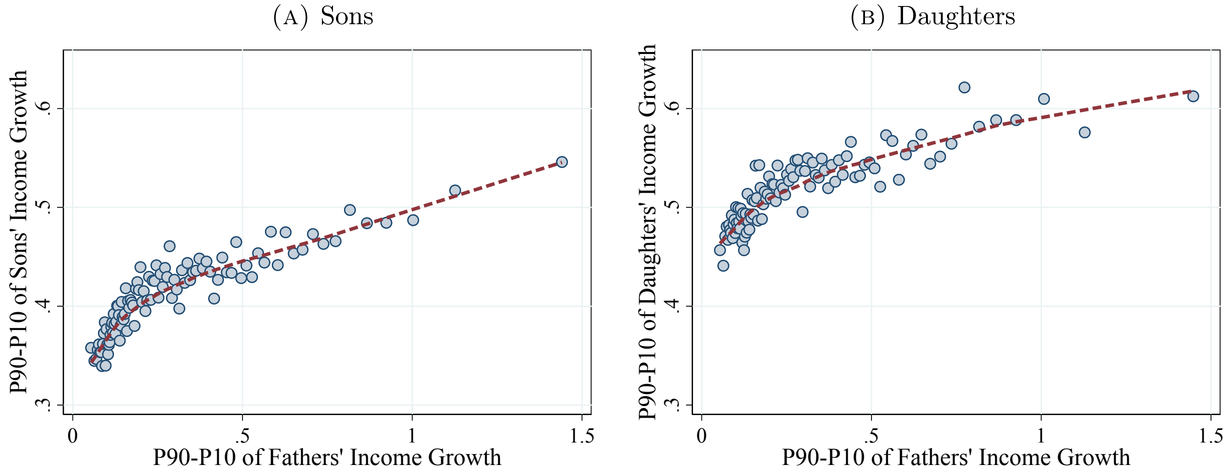
To address these concerns and examine the importance of different factors for workers' income dynamics, we run a series of individual-level regressions of the form

$$x_i^c \;=\; \beta_0 + \beta_1 x_i^f + X_i \Gamma + \varepsilon_i, \tag{2}$$

where x_i^c is a moment from child i's permanent income growth stream over the life cycle (i.e., median, P90-P10, and Kelley skewness) and x_i^f is the same moment but for the child's father. The matrix X_i contains a set of controls for child i that includes the log of the father's lifetime income, LI_i^f, the log of the father's wealth, W_i^f, and the log of the child's lifetime income, LI_i^c. As before, we focus on father-child pairs with at least 20 years of income data and with a non-missing observation for father's wealth variable.

Table IV shows estimation results for equation (2) for different moments of the income growth distribution for sons and daughters. We also report the dispersion of the independent variables, measured as the standard deviation and P90-P10 differential in columns (1) and (2). In all cases, we find that all four regressors are statistically significant at the 1% level—except for fathers' wealth on daughters' skewness. These results

Figure 27 – Dispersion of Income Growth of Fathers and Children



Notes: Figure 27 shows a binned scatter plot of fathers' and children's income growth dispersion measured by the individual-level 90th-to-10th percentiles differential. The dashed line is the non-linear correlation estimated from a LOWESS estimator. The scatter plot is based on a sample of 155,300 fathers-child pairs. The sample is divided into 100 bins.

are quite robust and remain almost unaltered if we include additional controls (see Table D.1 in Appendix D.2) or if we consider centralized moments instead of percentile-based measures as dependent variable (Table D.2).

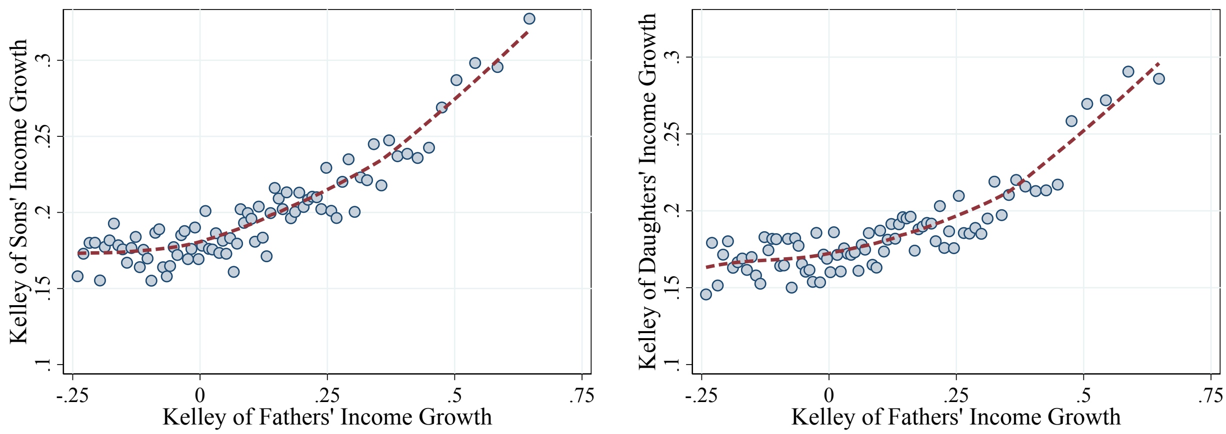
Median Income Growth. Columns (3) and (4) of Table IV show the results for the median income growth over the life cycle. Two points are worth noticing. First, similar to the results presented in Figure 26, we find that fathers' and children's median income growth is positively correlated both for sons and daughters, with an intergenerational elasticity of median income growth of 0.08 and 0.06, respectively.[36] This means that sons (daughters) of fathers at the 90th percentile of the median income growth distribution experience a 0.31% (0.23%) higher annual income growth relative to children of fathers at the 10th percentile. Second, consistent with earlier evidence (Figure 20), the coefficients of fathers' lifetime income and wealth are positive as children of more affluent fathers tend to experience higher income growth over their lifetime. For instance, the median growth for sons (daughters) of fathers at the 90th percentile of the lifetime income distribution is approximately 0.33% (0.16%) higher compared to sons of fathers at the 10th percentile. These magnitudes are economically quite substantial, with 0.3% steeper annual income growth, implying 13% higher income over a 40-year working life.

Volatility of Income Growth. We then turn to the income growth dispersion (columns (5) and (6) of Table IV). We find that the P90-P10 of income growth for sons (daughters)

[36] The intergenerational elasticity of lifetime income for this sample is 0.18 (0.15) for sons (daughters).

Figure 28 – Skewness of Income Growth of Fathers and Children

(a) Sons (b) Daughters



Notes: Figure 28 shows a binned scatter plot of fathers' and children's income growth skewness measured by the individual-level Kelley skewness. The dashed line is the non-linear correlation estimated from a LOWESS estimator. The scatter plot is based on a sample of 155,300 fathers-child pairs. The sample is divided into 100 bins.

of fathers at the 90th percentile of the income volatility distribution is 11 (8) log points higher compared to those with fathers at the 10th percentile. Quantifying the importance of family resources, we find an elasticity of 0.16 with respect to fathers' lifetime income, which implies that sons of fathers at the 90th percentile of the lifetime income distribution face an income volatility that is 13 log points higher than the workers with fathers at the 10th percentile. For daughters, the corresponding figure is about 9.7 log points. These differences are quite significant, considering that the average P90-P10 of income growth among sons (daughters) is 0.41 (0.51). Hence, our regression results suggest that children of more affluent families experience a more volatile income stream during their lifetime. However, recall that we find a U-shaped pattern for the volatility of income over fathers' economic resources in Section 6.2; therefore, our linear regression results are probably a lower bound for the true effect.

Skewness of Income Growth. Finally, columns (7) and (8) in Table IV present the regression results for the skewness of income growth. We find that a one standard deviation of increase in the father's skewness of income growth implies a 2.6 pp. (2.0 pp.) increase in the Kelley skewness of the son's (daughter's) income growth. This magnitude is relatively small considering that the average skewness in the sample is 0.19 for both sons and daughters. The elasticity of skewness with respect to the father's income is also statistically significant but of a smaller magnitude.

In summary, the results in this section show that fathers have a significant effect on

Table IV – Determinants of Children's Income Dynamics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	σ	P90-P10	$P50_i^c$		P90-P10$_i^c$		${\mathcal{S}_\mathcal{K}}_i^c$	
			Sons	Daughters	Sons	Daughters	Sons	Daughters
$P50_i^f$	0.020	0.04	0.077***	0.058***				
			(0.006)	(0.006)				
P90-P10$_i^f$	0.26	0.57			0.195***	0.133***		
					(0.005)	(0.005)		
${\mathcal{S}_\mathcal{K}}_i^f$	0.28	0.72					0.094***	0.072***
							(0.004)	(0.004)
$\log LI_i^c$	0.42	0.88	0.022***	0.010***	-0.317***	-0.438***	0.087***	0.107***
			(0.000)	(0.000)	(0.003)	(0.003)	(0.003)	(0.003)
$\log LI_i^f$	0.36	0.83	0.004***	0.002***	0.157***	0.115***	0.046***	0.027***
			(0.000)	(0.000)	(0.004)	(0.004)	(0.004)	(0.004)
$\log W_i^f$	1.60	3.80	0.001***	0.000***	0.005***	0.005***	0.009***	0.001
			(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)
R^2			0.119	0.028	0.179	0.275	0.032	0.030
N (000s)	155.3	155.3	79.7	75.6	79.7	75.6	79.7	75.6

Notes: Columns (1) and (2) of Table IV show the standard deviation and P90-P10 differential of the cross-sectional distribution of different moments of the fathers' and children's lifetime income, fathers' wealth, and fathers' income growth. For children's lifetime income, we report the moments for men. The corresponding moments for women are $\sigma^c = 0.44$ and $P90 - P10^c = 0.99$. Columns (1) to (6) show the coefficients of a series of cross-sectional regressions of worker-level measures of median lifetime growth, P90-P10, and Kelley Skewness ($\mathcal{S}_\mathcal{K}$), with the superscript c denoting children and f denoting fathers. Income growth is measured as the one-year log change of a measure of permanent income, calculated as the average income of an individual between years t and $t-2$. In the sample, we consider fathers and children with more than 20 years of data. The lifetime income of fathers and children is calculated as in equation (1). The measure of lifetime wealth is calculated as the fathers' average wealth between ages 45 and 55 (or the nearest age to this age range for individuals that are observed when they are too young (below 45) or too old (above 55).

children's income dynamics beyond the strong intergenerational transmission of income. Our empirical findings in this section can be explained, for instance, by fathers and children pursuing similar careers, having similar attitudes toward risk, or children making career choices by relying on parents' resources. Quantifying these different economic forces is beyond the scope of this paper, and further research is needed.

7 Conclusions

Using administrative data from Norway between 1993 and 2017, we documented several stylized facts on individual earnings dynamics with a special focus on top earners and non-Gaussian features. Our key findings can be summarized as follows. First, even though Norway is a country of relative equality and stability, it has experienced a substantial increase in top income inequality over our sample period. Second, in contrast to most other developed economies, inequality declines sharply over the life cycle below the 90th percentile. However, dispersion in the top 10% fans out, suggesting that different economic forces drive inequality in different parts of the earnings distribution. Third, the earnings growth distribution is left skewed and leptokurtic. Similar to other countries, these features vary substantially over the life cycle and across the earnings distribution. Fourth, earnings in the top 1% are very persistent even relative to those in the top 2% to 5%. These stylized facts—along with many others—are part of the Global Income Dynamics Database Project.

In the second part of the paper, we switched gears and investigated the intergenerational transmission of income *dynamics*. For this purpose, we used a different administrative dataset dating back to 1967, the very long panel of which allows us to precisely measure each individual's income risk over two generations. We find that workers from richer families experience steeper but more volatile income growth over the lifecycle. The higher volatility is mainly driven by a longer right tail (arising from more opportunities for large gains). These findings suggest that children of more affluent families can pursue high-risk, high-return careers, possibly because of the availability of parental insurance.

Finally, we find strong evidence of the transmission of income dynamics across generations. In particular, children of fathers with more volatile incomes or with higher tail risk also have riskier income streams, suggesting either that fathers and children share similar risk attitudes or that they work in similar jobs and sectors with similar risk profiles, or both. Our findings are important to understand the consumption and savings behavior of parents as well as parents' role in human capital accumulation and the determinants of intergenerational income mobility.

References

Aaberge, R., Atkinson, A. B. and Modalsli, J. (2013). The ins and outs of top income mobility. *IZA Discussion Paper*, **7729**.

Abowd, J. M. and Card, D. (1989). On the covariance structure of earnings and hours changes. *Econometrica*, **57** (2), 411–45.

Alvarez, J., Browning, M. and Ejrnï¿œs, M. (2010). Modelling income processes with lots of heterogeneity. *Review of Economic Studies*, **771-4**, 1353–1381.

Arellano, M., Blundell, R. and Bonhomme, S. (2017). Earnings and consumption dynamics: a nonlinear panel data framework. *Econometrica*, **85** (3), 693–734.

Atkinson, A. B., Piketty, T. and Saez, E. (2011). Top incomes in the long run of history. *Journal of Economic Literature*, **49** (1), 3–71.

— and Søgaard, J. E. (2016). The long-run history of income inequality in denmark. *The Scandinavian Journal of Economics*, **118** (2), 264–291.

Barth, E., Moene, K. O. and Willumsen, F. (2014). The scandinavian model—an interpretation. *Journal of Public Economics*, **117**, 60 – 72.

Black, S. E. and Devereux, P. J. (2011). Recent developments in intergenerational mobility. *Handbook of Labor Economics*, **4**, 1487–1541.

Bloom, N., Guvenen, F., Pistaferri, L., Sabelhaus, J., Salgado, S. and Song, J. (2017). *The Great Micro Moderation*. Tech. rep., Working Paper.

Blundell, R., Graber, M. and Mogstad, M. (2014). Labor income dynamics and the insurance from taxes, transfers, and the family. *Journal of Public Economics*.

Boar, C. (2020). *Dynastic precautionary savings*. Tech. rep., National Bureau of Economic Research.

Bratberg, E., Anti Nilsen, Ø. and Vaage, K. (2005). Intergenerational earnings mobility in norway: Levels and trends. *Scandinavian Journal of Economics*, **107** (3), 419–435.

Bratsberg, B., Fevang, E. and Røed, K. (2013). Job loss and disability insurance. *Labour Economics*, **24**, 137 – 150.

—, Røed, K., Raaum, O., Naylor, R., Ja¨ ntti, M., Eriksson, T. and O¨ sterbacka, E. (2007). Nonlinearities in intergenerational earnings mobility: consequences for cross-country comparisons. *The Economic Journal*, **117** (519), C72–C92.

Busch, C., Domeij, D., Guvenen, F. and Madera, R. (2021). Skewed idiosyncratic income risk over the business cycle: Sources and insurance. *American Economic Journal: Macroeconomics*.

Castaneda, A., Diaz-Gimenez, J. and Rios-Rull, J.-V. (2003). Accounting for the us earnings and wealth inequality. *Journal of political economy*, **111** (4), 818–857.

Corak, M. (2013). Income inequality, equality of opportunity, and intergenerational mobility. *Journal of Economic Perspectives*, **27** (3), 79–102.

Crow, E. L. and Siddiqui, M. (1967). Robust estimation of location. *Journal of the American Statistical Association*, **62** (318), 353–389.

De Nardi, M., Fella, G., Knoef, M. G., Paz-Pardo, G. and Van Ooijen, R. (2019). *Family and Government Insurance: Wage, Earnings, and Income Risks in the Netherlands and the US*. Tech. rep., National Bureau of Economic Research.

—, — and Paz-Pardo, G. (2020). Nonlinear household earnings dynamics, self-insurance, and welfare. *Journal of the European Economic Association*, **18** (2), 890–926.

Domeij, D. and Floden, M. (2010). Inequality trends in sweden 1978-2004. *Review of Economic Dynamics*, **13** (1), 179–208.

Dynan, K., Elmendorf, D. and Sichel, D. (2012). The evolution of household income volatility. *The BE Journal of Economic Analysis & Policy*, **12** (2).

Fagereng, A., Guiso, L., Malacrino, D. and Pistaferri, L. (2016). Heterogeneity in returns to wealth and the measurement of wealth inequality. *American Economic Review*, **106** (5), 651–55.

—, Holm, M. B., Moll, B. and Natvik, G. (2019). *Saving behavior across the wealth distribution: The importance of capital gains*. Tech. rep., National Bureau of Economic Research.

Fawcett, K. W. (2020). Parental insurance, directed search and the labor market outcomes of young adults.

Fevang, E., Markussen, S. and Røed, K. (2014). The sick pay trap. *Journal of Labor Economics*, **32** (2), 305–336.

Guvenen, F., Kaplan, G., Song, J. and Weidner, J. (2017). *Lifetime incomes in the United States over six decades*. Tech. rep., National Bureau of Economic Research.

—, Karahan, F., Ozkan, S. and Song, J. (2018). *What do data on millions of U.S. workers reveal about life-cycle earnings risk?* Tech. rep., National Bureau of Economic Research.

—, Kuruscu, B. and Ozkan, S. (2014a). Taxation of human capital and wage inequality: A cross-country analysis. *Review of Economic Studies*, **81**, 818–850.

—, Ozkan, S. and Song, J. (2014b). The nature of countercyclical income risk. *Journal of Political Economy*, **122** (3), 621–660.

Haider, S. J. and Solon, G. (2006). "life-cycle variation in the association between current and lifetime earnings. *American Economic Review*, **96** (4), 1308–1320.

Halvorsen, E., Holter, H., Ozkan, S. and Storesltetten, K. (2019). *Dissecting Idiosyncratic Income Risk: Evidence from Norwegian Registry Data*. Tech. rep., Technical report, mimeo.

Hemmings, P. and Prinz, C. (2020). Sickness and disability systems: comparing outcomes and policies in norway with those in sweden, the netherlands and switzerland. (1601).

Herreno, J. and Ocampo, S. (2020). Self-employment and development.

Holter, H. A. (2015). Accounting for cross-country differences in intergenerational earnings persistence: The impact of taxation and public education expenditure. *Quantitative Economics*, **6** (2), 385–428.

Jäntti, M. and Jenkins, S. P. (2015). Income mobility. In *Handbook of income distribution*, vol. 2, Elsevier, pp. 807–935.

— and Lindahl, L. (2012). On the variability of income within and across generations. *Economics Letters*, **117** (1), 165–167.

Kaplan, G., Moll, B. and Violante, G. L. (2018). Monetary policy according to hank. *American Economic Review*, **108** (3), 697–743.

Karahan, F. and Ozkan, S. (2013). On the persistence of income shocks over the life cycle: Evidence, theory, and implications. *Review of Economic Dynamics*, **16** (3), 452–476.

—, — and Song, J. (2019). Anatomy of lifetime earnings inequality: Heterogeneity in job ladder risk vs. human capital. *FRB of New York Staff Report*, (908).

Kelley, T. L. (1947). *Fundamentals of statistics*. Harvard University Press.

Kopczuk, W., Saez, E. and Song, J. (2010). Earnings inequality and mobility in the united states: Evidence from social security data since 1937. *Quarterly Journal of Economics*, **125** (1).

Lagakos, D., Moll, B., Porzio, T., Qian, N. and Schoellman, T. (2018). Life cycle wage growth across countries. *Journal of Political Economy*, **126** (2), 797–849.

Lochner, L. and Park, Y. (2020). Earnings dynamics and intergenerational transmission of skill. *NBER Working Paper*, (w28141).

Long, J. and Ferrie, J. (2013). Intergenerational occupational mobility in great britain and the united states since 1850. *American Economic Review*, **103** (4), 1109–37.

Markussen, S. and Røed, K. (2019). Economic Mobility Under Pressure. *Journal of the European Economic Association*, **18** (4), 1844–1885.

Meghir, C. and Pistaferri, L. (2004). Income variance dynamics and heterogeneity. *Econometrica*, **72** (1), 1–32.

Parker, J. A. and Vissing-Jørgensen, A. (2010). The increase in income cyclicality of high-income households and its relation to the rise in top income shares. *Brookings Papers on Economic Activity*, **41** (2), 1–70.

Pekkarinen, T., Salvanes, K. G. and Sarvimaki, M. (2017). The evolution of social mobility: Norway during the twentieth century. *The Scandinavian Journal of Economics*, **119** (1), 5–33.

Piketty, T. (2000). Theories of persistent inequality and intergenerational mobility. *Handbook of income distribution*, **1**, 429–476.

Roine, J. and Waldenström, D. (2008). The evolution of top incomes in an egalitarian society: Sweden, 1903–2004. *Journal of public economics*, **92** (1-2), 366–387.

Shore, S. H. (2011). The intergenerational transmission of income volatility: is riskiness inherited? *Journal of Business & Economic Statistics*, **29** (3), 372–381.

Solon, G. (1992). Intergenerational income mobility in the united states. *The American Economic Review*, pp. 393–408.

Storesletten, K., Telmer, C. I. and Yaron, A. (2004a). Consumption and risk sharing over the life cycle. *Journal of Monetary Economics*, **51** (3), 609–633.

—, — and — (2004b). Cyclical dynamics in idiosyncratic labor market risk. *Journal of Political Economy*, **112** (3), 695–717.

Supplemental Online Appendix

NOT FOR PUBLICATION

A Additional Figures on Distribution of Earnings

Figure A.1 – Top Income Inequality: Pareto Tail at top 1%

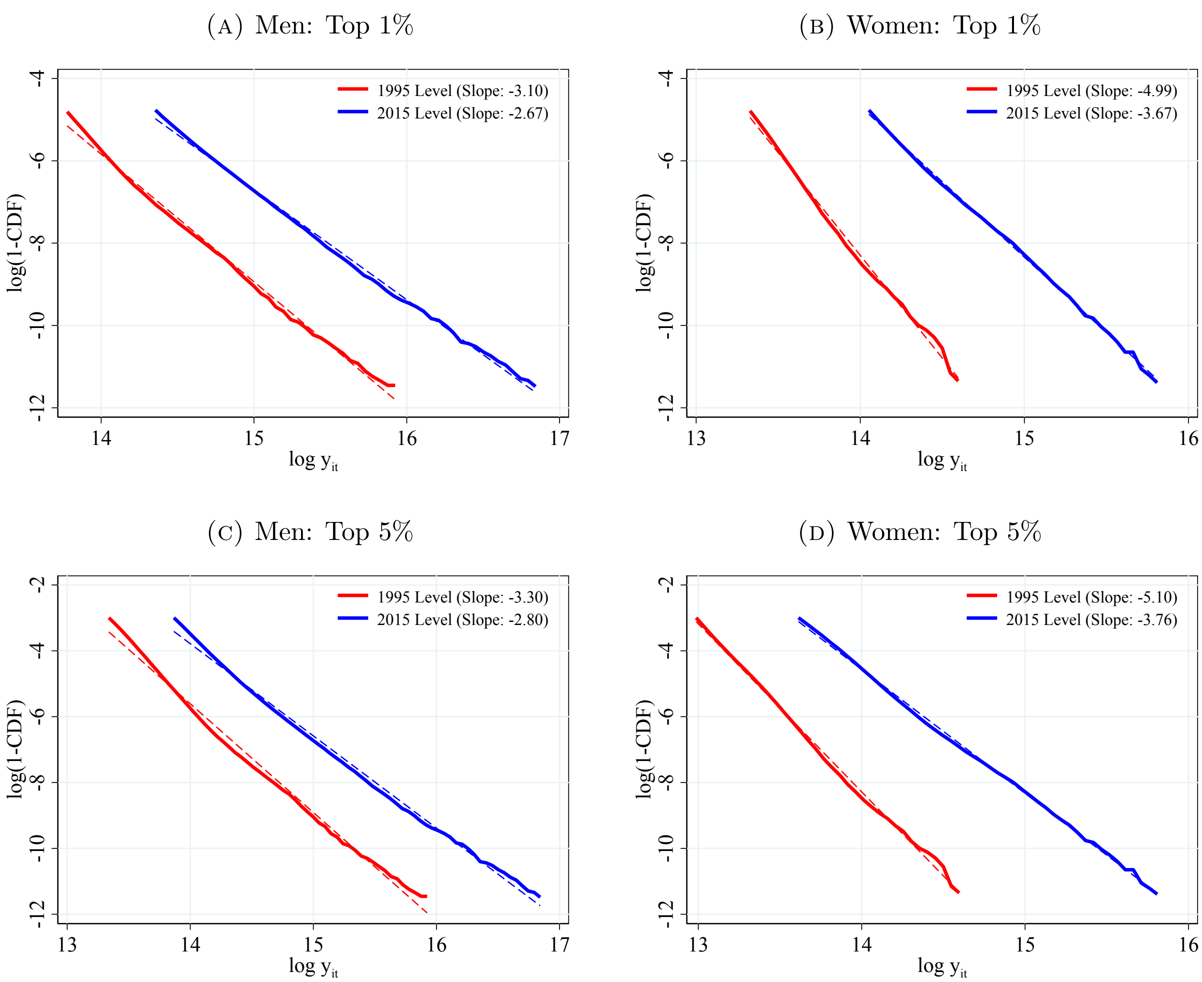
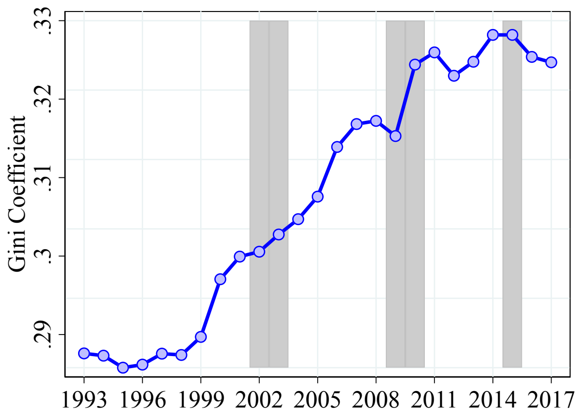


Figure A.1 shows the tail of the distribution of log-earnings above the 99th percentile of the distribution (panels A and B) and above the 95th percentile (panels C and D).

Figure A.2 – Gini Coefficient

(a) Men



(b) Women

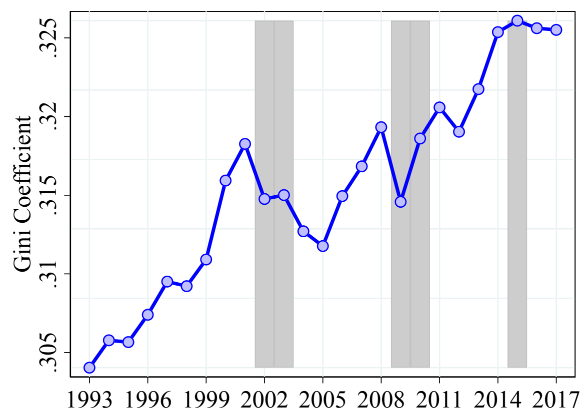


Figure A.2 shows the Gini coefficient of the distribution of log-earnings.

A.1 Figures for the Combined Sample (Men and Women)

Figure A.3 – Distribution of Earnings in the Population

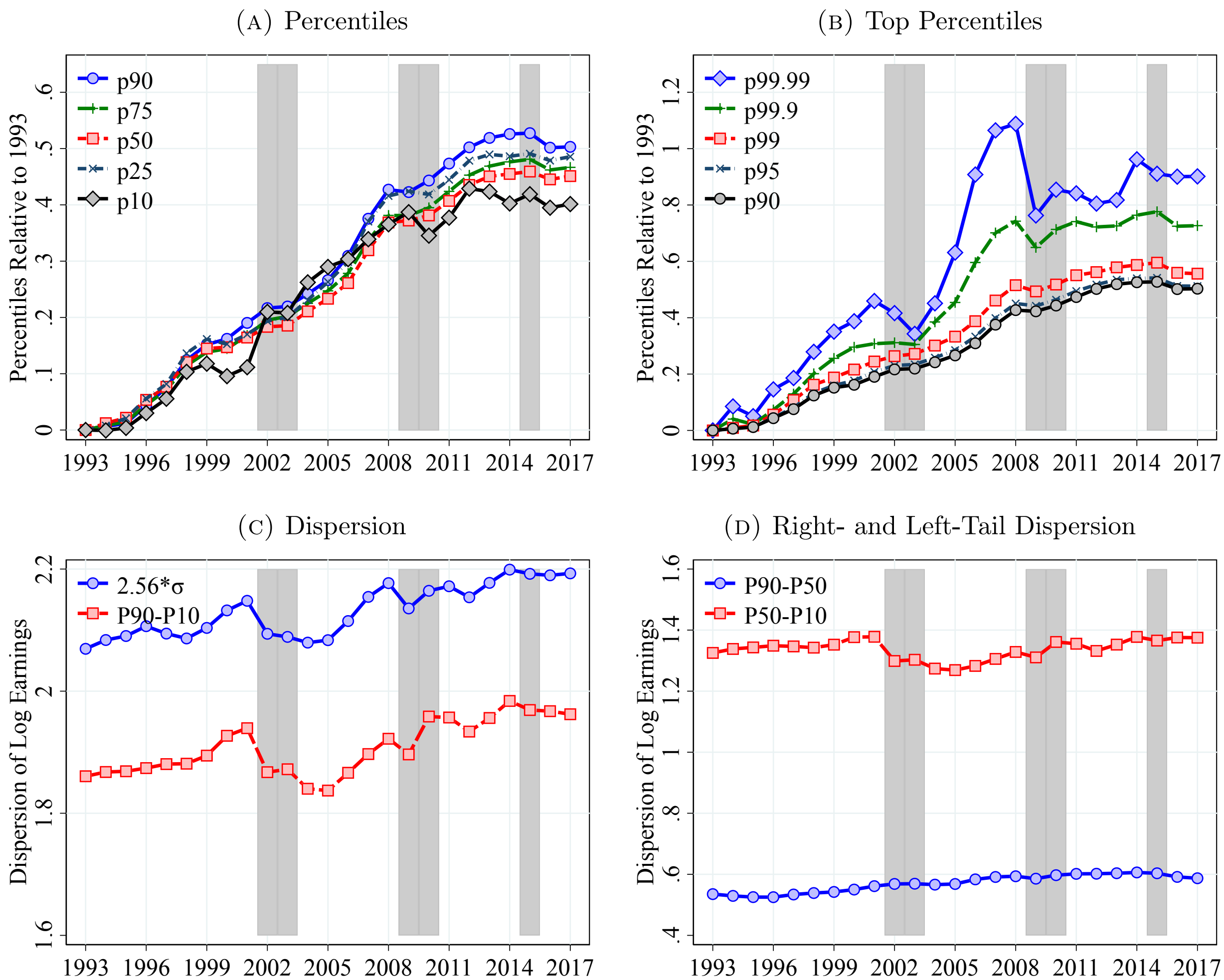


Figure A.3 shows the evolution of the following variables: (a) P10, P25, P50, P75, P90 (b) P90, P95, P99, P99.9, P99.99, (c) P90-P10 and 2.56*SD of log income, (d) P90-P50 and P50-P10. Percentiles in (a) and (b) are normalized to 0 in 1993. Shaded areas represent recession years as defined as years with unemployment rate growth 0.4 pp. or more and an output gap of -0.5 or less. In all figures we consider a joint sample of men and women. See Section 2 for sample selection and definitions.

A.2 Figures for Residual Earnings

Figure A.4 – Residual Earnings Controlling for age

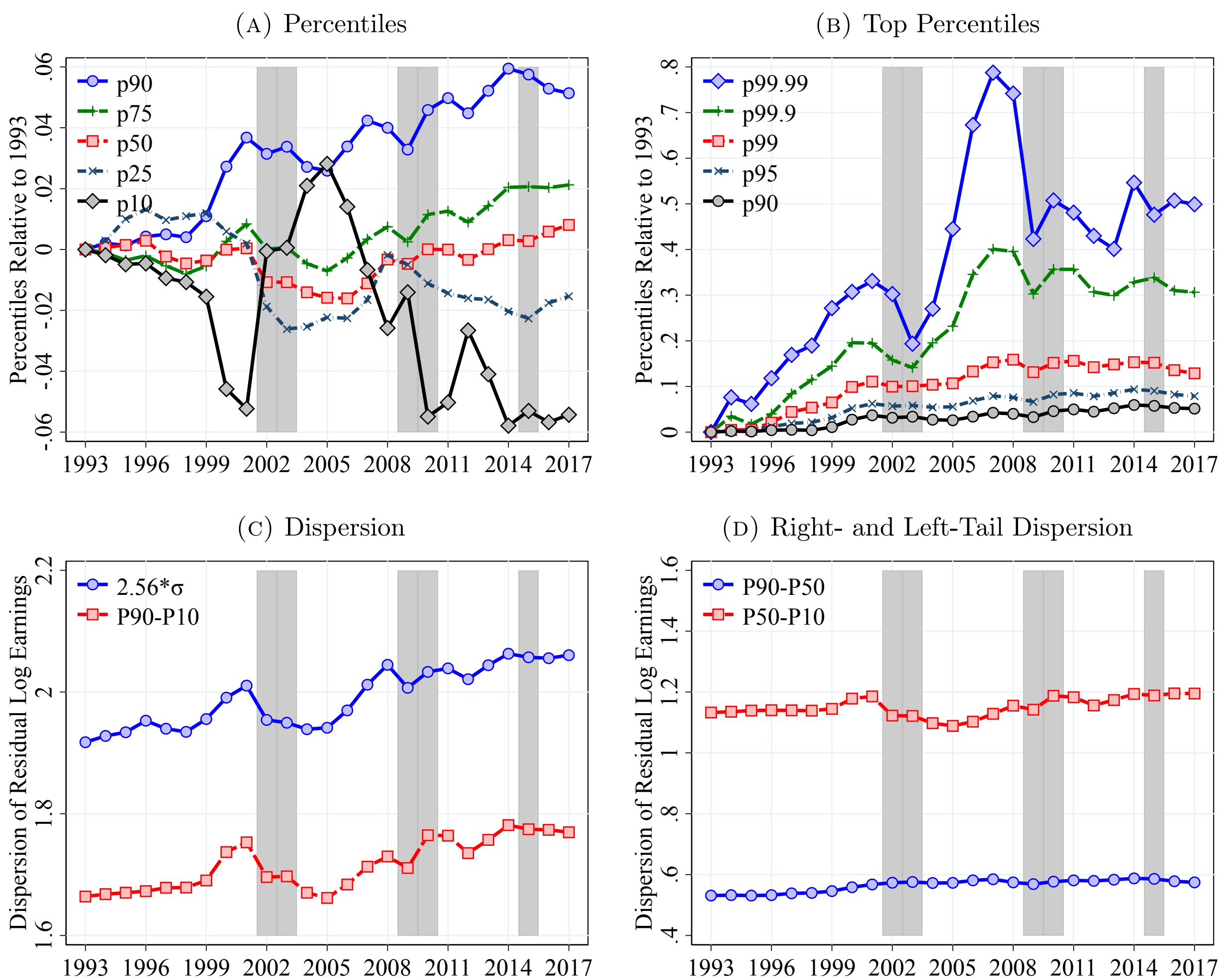


Figure A.4 shows the evolution of the following variables: (a) P10, P25, P50, P75, P90 (b) P90, P95, P99, P99.9, P99.99, (c) P90-P10 and 2.56*SD of log income, (d) P90-P50 and P50-P10. Percentiles in (a) and (b) are normalized to 0 in 1993. Shaded areas represent recession years as defined as years with unemployment rate growth 0.4 pp. or more and an output gap of -0.5 or less. In all figures we consider a joint sample of men and women. We residualize log-income from age fixed effects by year and gender. See Section 2 for sample selection and definitions.

Figure A.5 – Residual Earnings Controlling for Age and Education

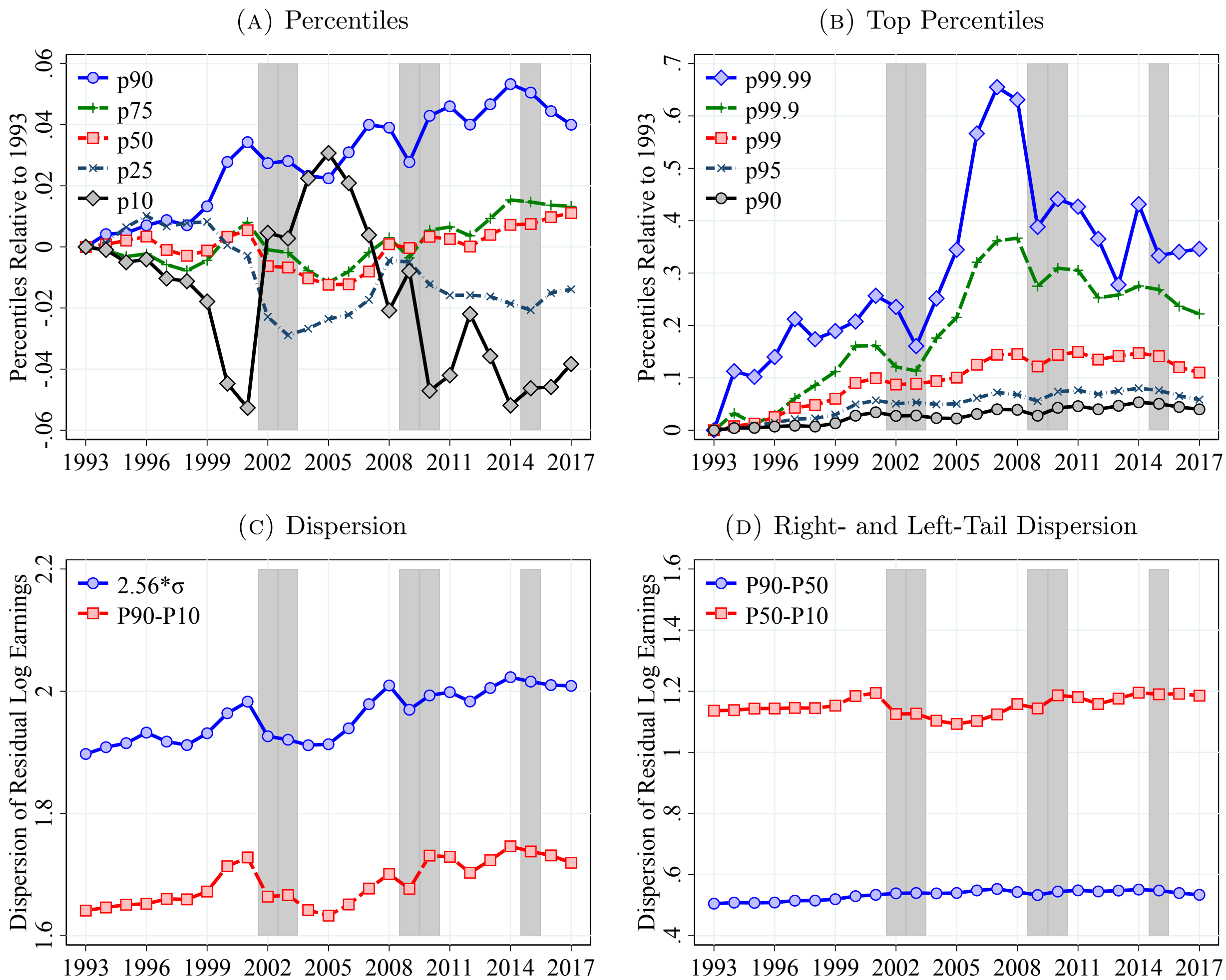
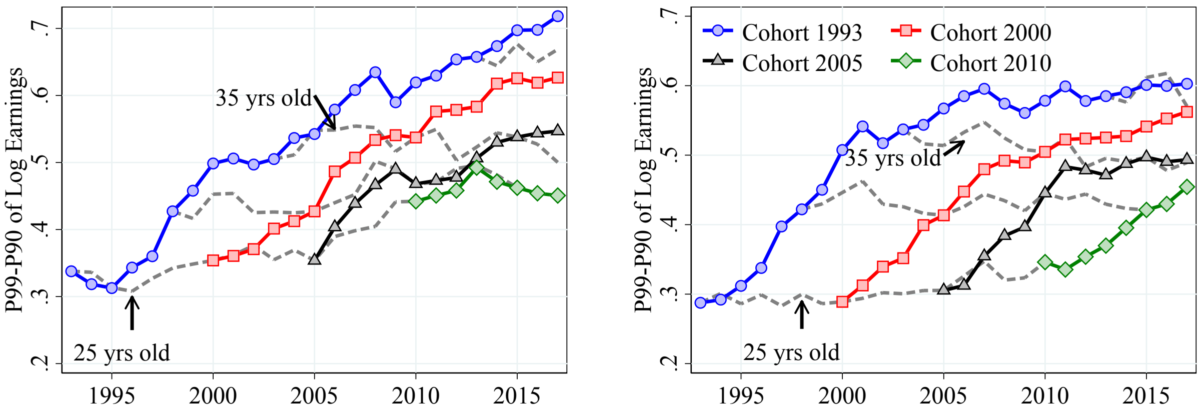


Figure A.5 shows the evolution of the following variables: (a) P10, P25, P50, P75, P90 (b) P90, P95, P99, P99.9, P99.99, (c) P90-P10 and 2.56*SD of log income, (d) P90-P50 and P50-P10. Percentiles in (a) and (b) are normalized to 0 in 1993. Shaded areas represent recession years as defined as years with unemployment rate growth 0.4 pp. or more and an output gap of -0.5 or less. In all figures we consider a joint sample of men and women. We residualize log-income from age and education fixed effects (three groups: less than high-school, high-school graduates, and college graduate or more) by year and gender. See Section 2 for sample selection and definitions.

Figure A.6 – Evolution of Within-Cohort Top-Income Inequality

(a) Men (b) Women

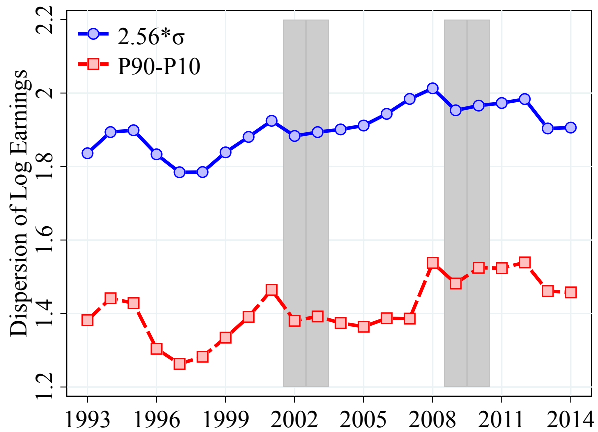


Notes: Figure A.6 uses the log earnings and the CS and shows: (a) Men: P99-P90 over the life cycle for selected cohorts and (b) Women: P99-P90 over the life cycle for selected cohorts. A cohort is defined by the year in which the cohort turns 25 years old. Dashed lines connect individuals of the same age. See Section 2 for sample selection and definitions.

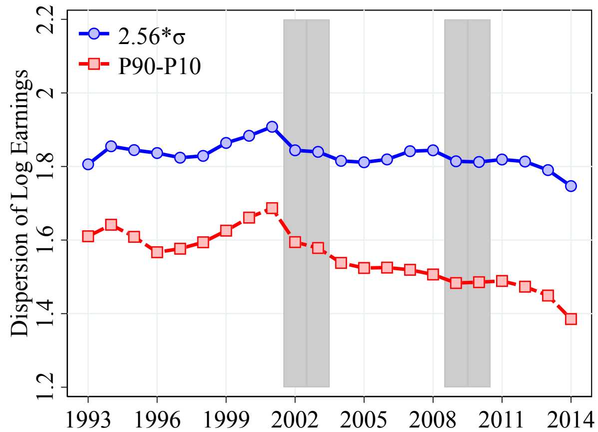
A.3 Figures for Public and Private Sectors

Figure A.7 – Income Inequality in Public and Private Sector

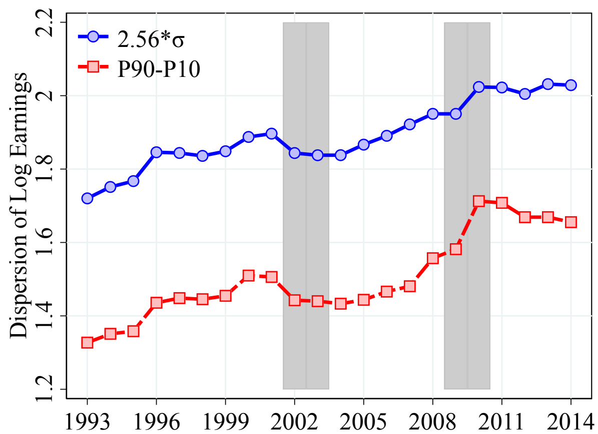
(a) Men Public Sector



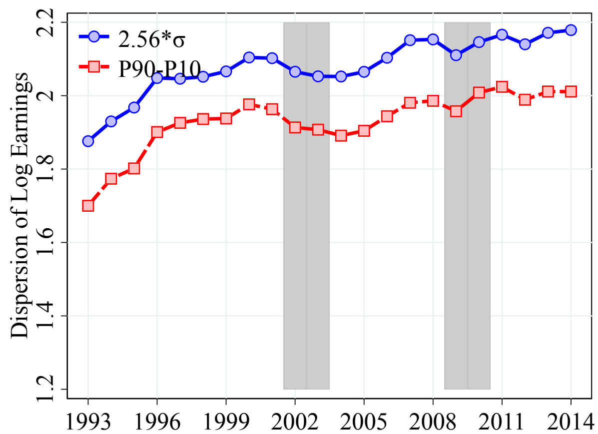
(b) Women: Public Sector



(c) Men: Private Sector

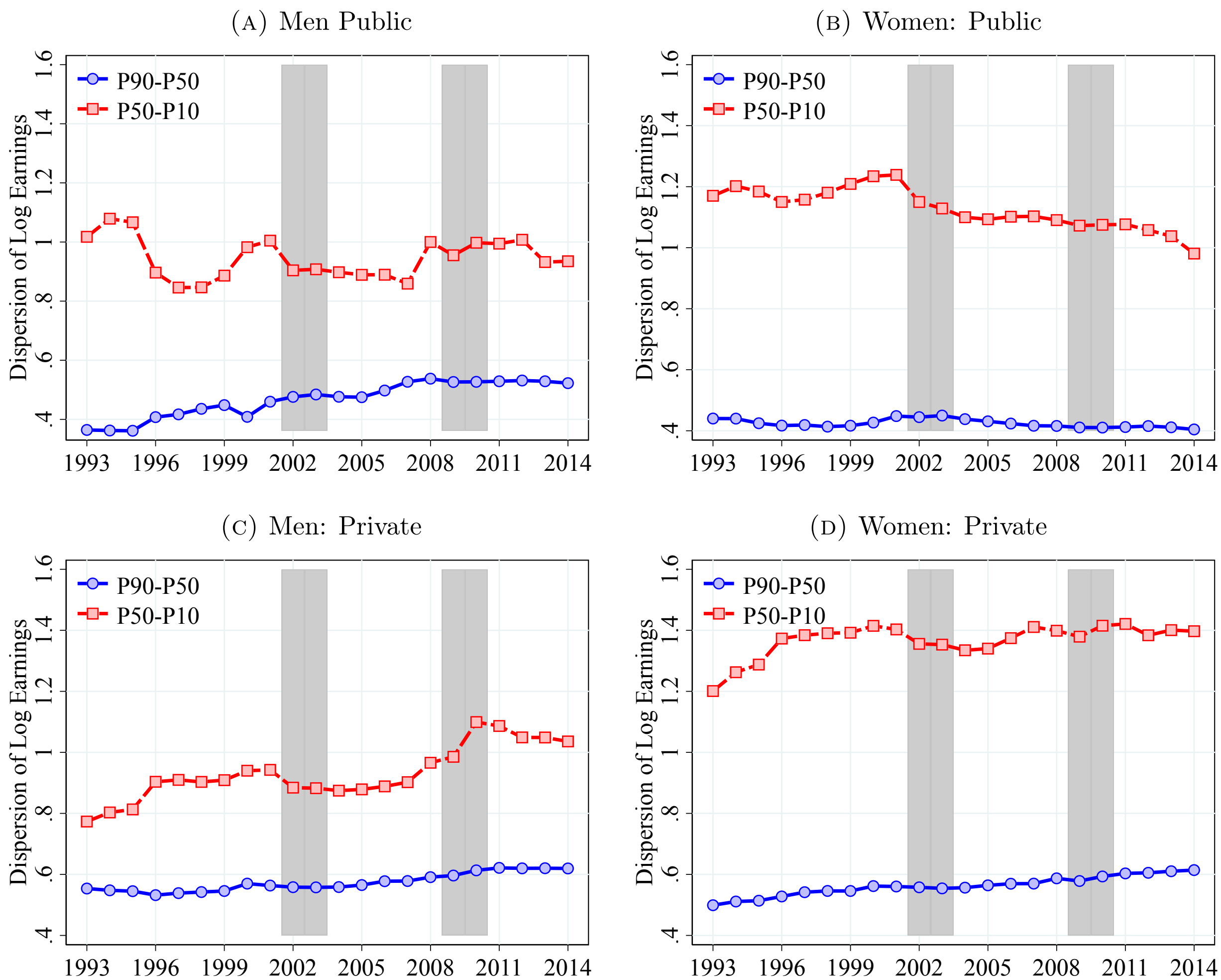


(d) Women: Private Sector



Notes: Figure A.7 plot against time the following variables: (a and c) Men: P90-P10 and 2.56*SD of log income (b and d) Women: P90-P10 and 2.56*SD of log income. The value of 2.56*SD corresponds to the differential between the 10th and the 90th percentiles in a Normal distribution. Shaded areas represent recession years defined as years with unemployment rate growth 0.4 pp. or more, and an output gap of -0.5 or less. Results based on the CS sample. See Section 2 for sample selection and definitions. We have information on worker's sector only until 2014.

Figure A.8 – Right- and Left-Tail Inequality IN for Public and Private Sector



Notes: Figure A.8 plot against time the following variables: (a and c) Men: P90-P50 and P50-P10, (b and d) Women: P90-P50 and P50-P10. Shaded areas are recessions. The value of 2.56*SD corresponds to the differential between the 10th and the 90th percentiles in a Normal distribution. Shaded areas represent recession years defined as years with unemployment rate growth 0.4 pp. or more, and an output gap of -0.5 or less. Results based on the CS sample. See Section 2 for sample selection and definitions. We have information on worker's sector only until 2014.

B Appendix for the Distribution of Earnings Growth

B.1 Moments of Five-Years Earnings Growth

Figure B.2 – Skewness and Kurtosis of Five-Years Earnings Changes

(a) Skewness (b) Excess Kurtosis

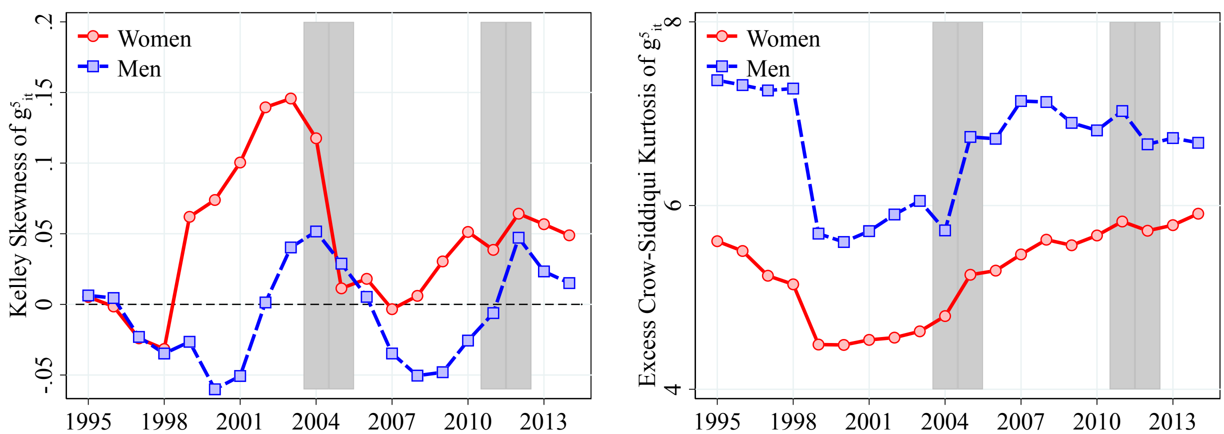


Figure B.2 plot against time the following variables: (a) Men and Women: Kelley skewness, (b) Men and Women: Crow-Siddiqui kurtosis. Shaded areas are recessions.

Figure B.1 – Dispersion of Five-Years Earnings Changes

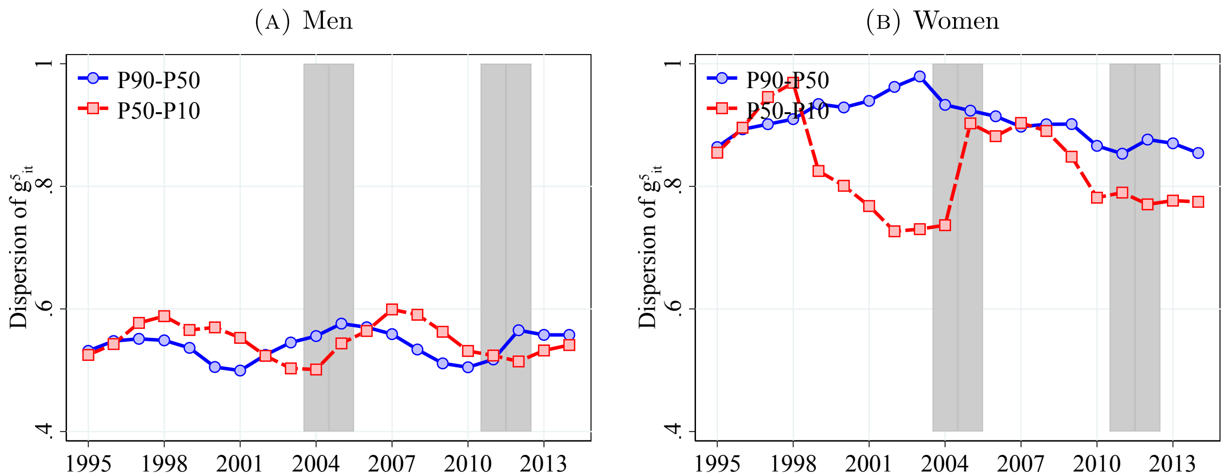


Figure B.1 plot against time the following variables: (a) Men: P90-10 differential, (b) Women: P90-10 differential. Shaded areas are recessions.

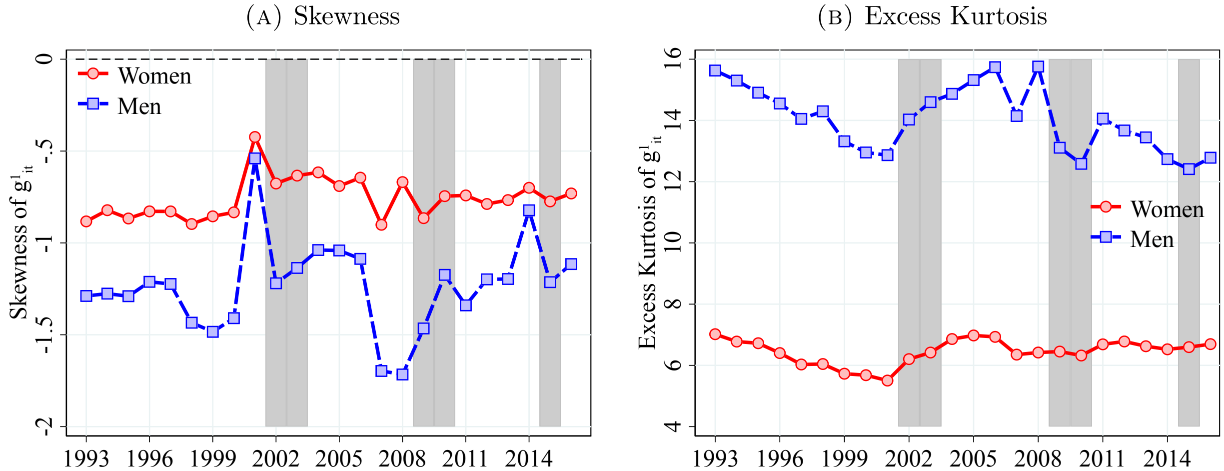
B.2 Centralized Moments

Table B.1 – Fraction of Individuals at Selected Ranges of Log Earnings Changes

	(1)	(2)	(3)	(4)	(5)	(6)
Range	$\Delta\varepsilon_t^1$	$\mathbb{N}(0, 0.58)$	Ratio	$\Delta\varepsilon_t^5$	$\mathbb{N}(0, 0.81)$	Ratio
$(-\infty, -3\sigma]$	2.2	0.1	16.0	1.8	0.1	13.6
$(-3\sigma, -2\sigma]$	1.8	2.1	0.8	1.9	2.1	0.9
$(-2\sigma, -\sigma]$	3.8	13.6	0.3	4.6	13.6	0.3
$(-\sigma, -0.05]$	26.3	30.7	0.9	34.5	31.7	1.1
$(-0, 05, 0.05]$	31.8	6.8	4.7	15.3	4.9	3.1
$(0.05, \sigma]$	27.9	30.7	0.9	34.1	31.7	1.1
$(\sigma, 2\sigma]$	4.6	13.6	0.3	5.6	13.6	0.4
$(2\sigma, 3\sigma]$	1.7	2.1	0.8	2.3	2.1	1.1
$(3\sigma, +\infty]$	1.1	0.1	7.8	0.9	0.1	7.0

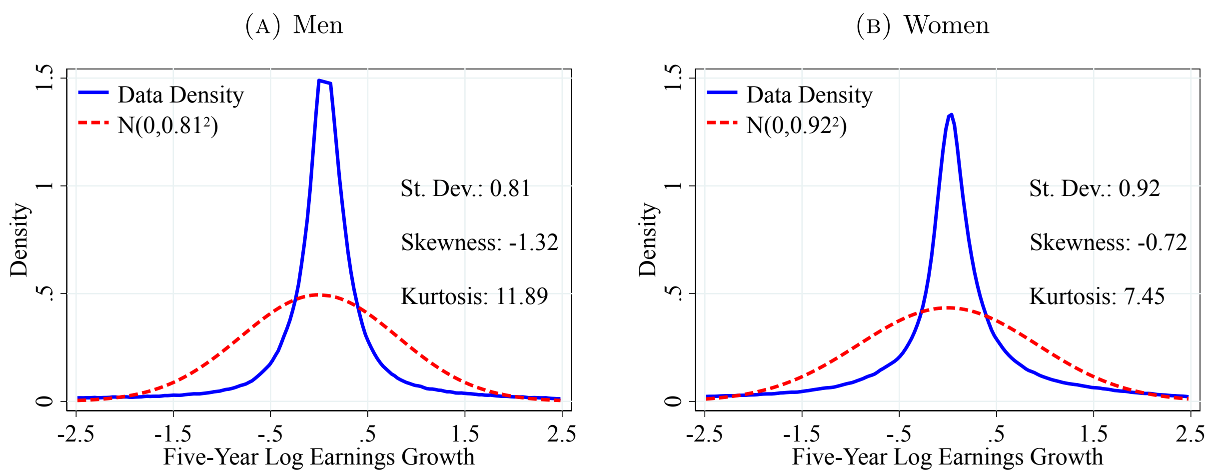
Figure B.1 shows the fraction of individuals between different cuts of the one- and five-year change distribution of log earnings growth for a sample of men in 2005. Columns (2) and (5) shoe the corresponding moments from a normal distribution with the standard deviation.

Figure B.4 – Skewness and Kurtosis of Earnings Changes



Notes: Figure B.4 shows the third and fourth standardized moments earnings growth for men and women. Shaded areas represent recession years as defined as years with unemployment rate growth of 0.4 pp. or more and an output gap of -0.5 or less. See Section 2 for sample selection and definitions.

Figure B.3 – Empirical Log-Densities of Five-year Earnings Growth



Notes: Figure B.3 shows the empirical density and corresponding cross-sectional moments of the distribution of give-year log earnings growth for men and women in 2005. See Section 2 for sample selection and definitions.

Table B.2 – Cyclicality of Cross-Sectional Moments of Arc-Earnings Changes

	(1)	(2)	(3)	(4)	(5)	(6)
	Dispersion		Skewness		Kurtosis	
	P90-P10	Std. Dev.	Kelley	Third	Crow-Siddiqui	Kurtosis
			Men			
ΔGDP_t	-0.01**	-0.01***	0.03**	0.05	-0.06	0.16**
	(0.01)	(0.00)	(0.01)	(0.06)	(0.09)	(0.06)
			Women			
ΔGDP_t	-0.04***	-0.01**	0.01	-0.01	0.29***	0.15***
	(0.01)	(0.00)	(0.01)	(0.02)	(0.09)	(0.04)
			Men			
ΔUnemp_t	0.01**	0.00*	-0.05***	-0.05	0.16	-0.11**
	(0.00)	(0.00)	(0.01)	(0.04)	(0.13)	(0.04)
			Women			
ΔUnemp_t	0.02**	0.00	-0.02**	-0.01	-0.12	-0.08**
	(0.01)	(0.00)	(0.01)	(0.02)	(0.12)	(0.03)
Obs.	24	24	24	24	24	24

Notes: Table B.2 shows the coefficients from regressions of different moments of earnings growth calculated as the arc-percent change on either GDP or unemployment growth for Men (Panel A) and Women (Panel B). The growth rate of unemployment (real GDP) is calculated as the (log) difference of the average unemployment rate (real GDP) between years t and $t+1$. Notice each regression is run separately. The unemployment rate is obtained from Statistics Norway whereas real GDP is obtained from the Federal Reserve Economic Data, FRED Newey-West standard errors in parentheses, estimated using one lag. In each regression, we standardize the right-hand-side variable so the coefficient can be directly interpreted as the impact of a one-standard deviation change on the dependent variable. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Figure B.5 – Dispersion of One-Year Arc-Percent Earnings Changes

(a) Men | (b) Women

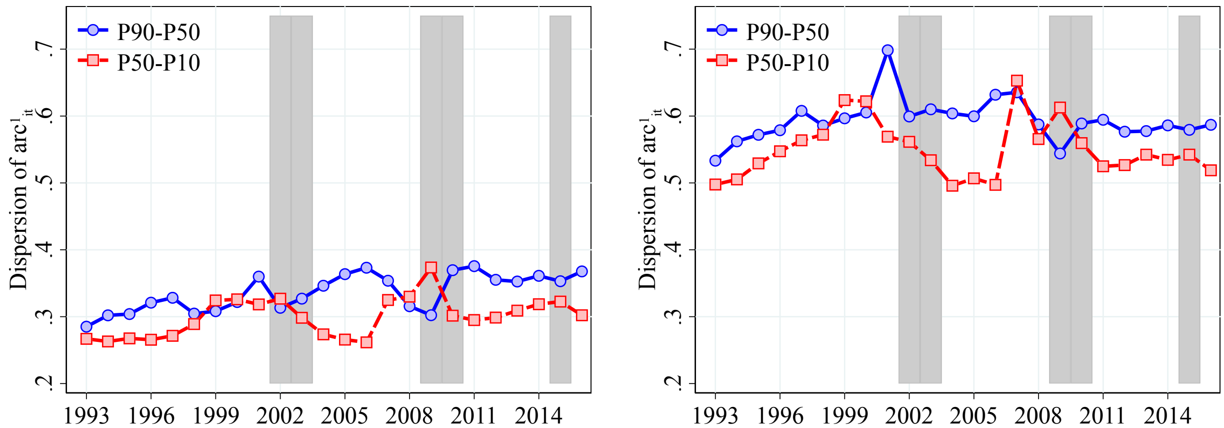


Figure B.5 plot against time the following variables: (a) Men: P90-10 differential, (b) Women: P90-10 differential. Shaded areas are recessions.

B.3 Arc-Percent Earnings Growth Distribution

Figure B.6 – Skewness and Kurtosis of One-Year Arc-Percent Earnings Changes

(a) Skewness | (b) Excess Kurtosis

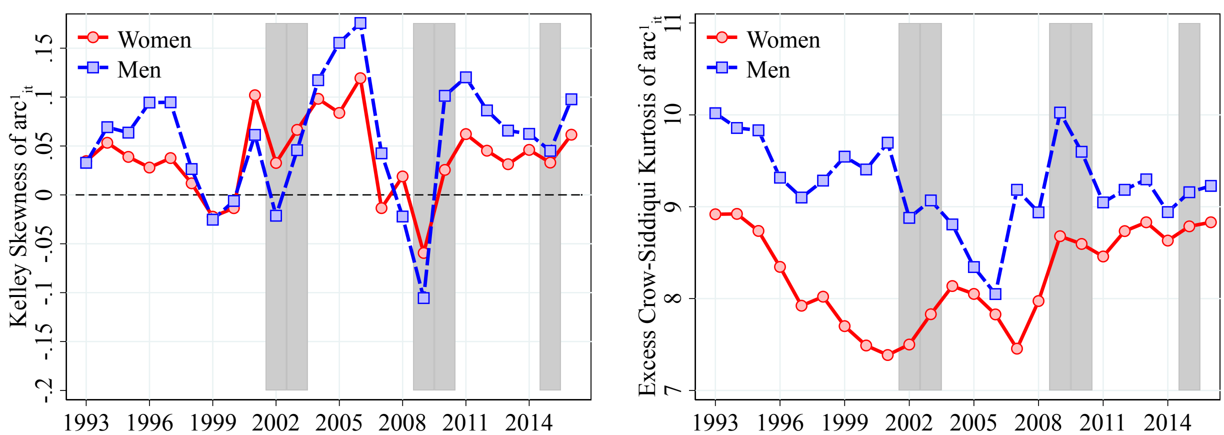


Figure B.6 plot against time the following variables: (a) Men and Women: Kelley skewness, (b) Men and Women: Crow-Siddiqui kurtosis. Shaded areas are recessions.

B.4 Heterogeneity in Idiosyncratic Earnings Changes

Figure B.7 – Kurtosis of Earnings Growth by Earnings Level and Age

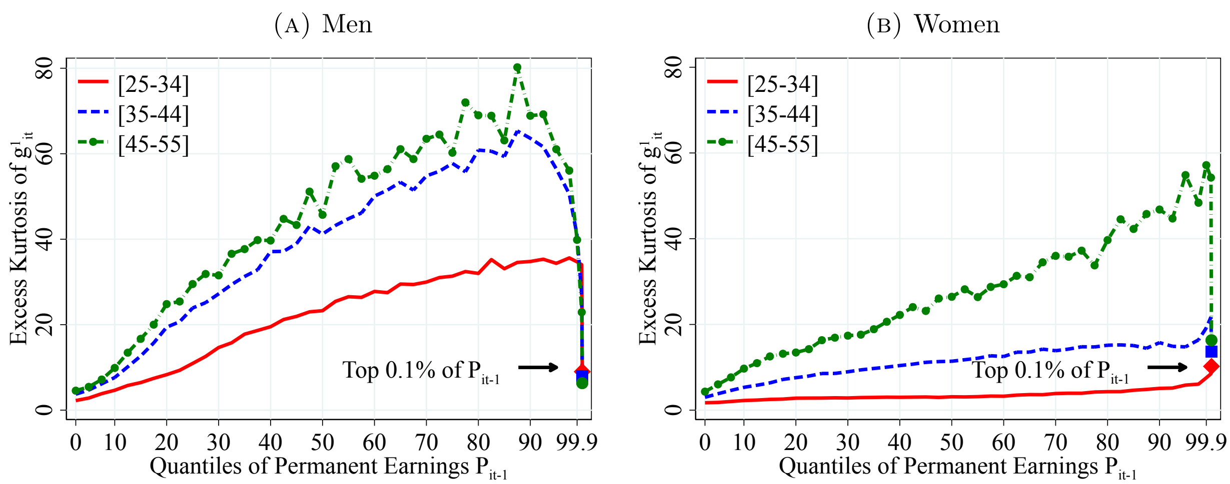
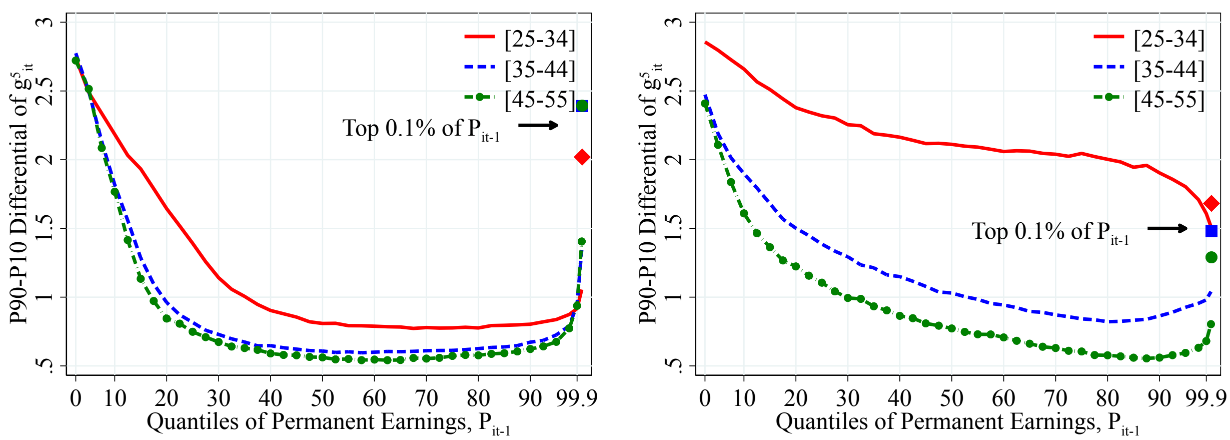


Figure B.7 shows the excess fourth standardized moment of log earnings changes for men and women with quantiles of the permanent income distribution, P_{it-1}. Excess kurtosis is defined as the value of kurtosis minus 3 which is the corresponding value for a Normal distribution. In each plot, the solid markers represent the corresponding measure of kurtosis for those workers at the top 0.1% of the earnings distribution for different age groups (diamond for 25 to 34 years old, square for 35 to 44 years old, and circle for 45 to 55 years old).

B.4.1 Heterogeneity of Idiosyncratic Earnings for Five-Year Changes

Figure B.8 – Dispersion of Earnings Growth by Permanent Income and Age

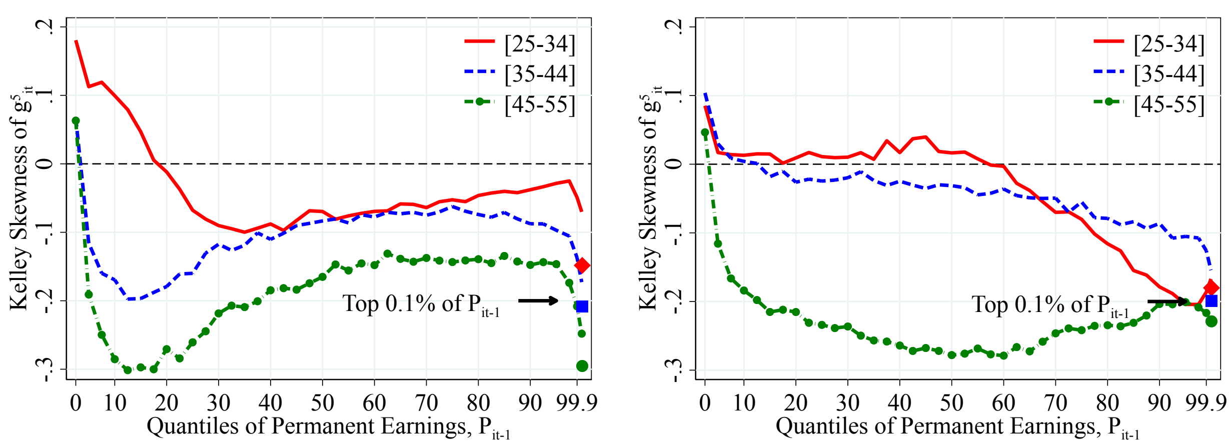
(a) Men (b) Women



Notes: Figure B.8 shows the P90-P10 of log growth rate of residual earnings for men and women within quantiles of the permanent income distribution, P_{it-1}. In each plot, the solid markers represent P90-P10 for those workers at the top 0.1% of the permanent income distribution for different age groups (diamond for 25 to 34 years old, square for 35 to 44 years old, and circle for 45 to 55 years old). See Section 2 for sample selection and definitions.

Figure B.9 – Kelley Skewness of Earnings Growth by Earnings Level and Age

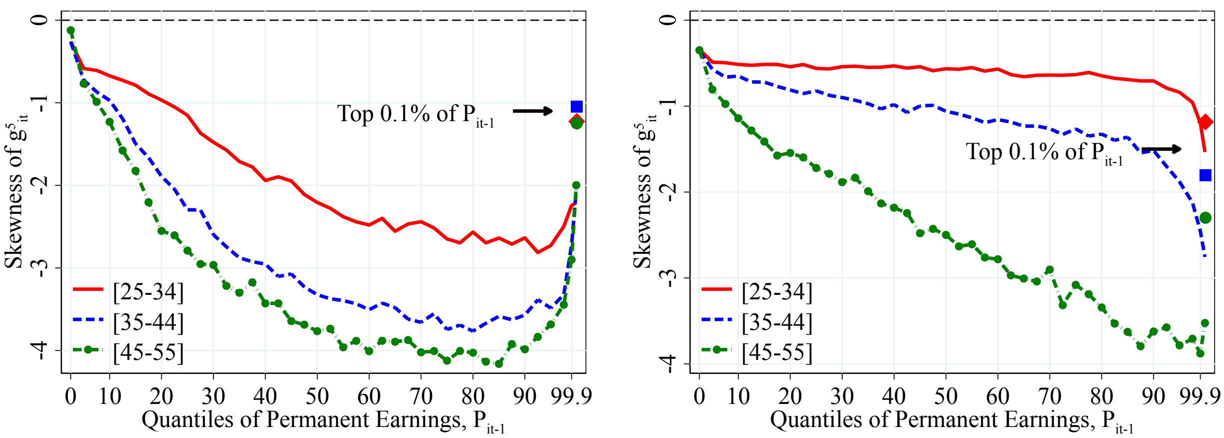
(a) Men (b) Women



Notes: Figure B.9 shows the Kelley skewness of log growth rate of residual earnings for men and women within quantiles of the permanent income distribution, P_{it-1}. Kelley skewness is defined as $\mathcal{S}_{\mathcal{K}}$ = ((P90-P50) − (P50-P10)) / (P90-P10). In each plot, the solid markers represent the Kelley skewness for those workers at the top 0.1% of the earnings distribution for different age groups (diamond for 25 to 34 years old, square for 35 to 44 years old, and circle for 45 to 55 years old). See Section 2 for sample selection and definitions.

Figure B.10 – Skewness of Earnings Growth by Earnings Level and Age

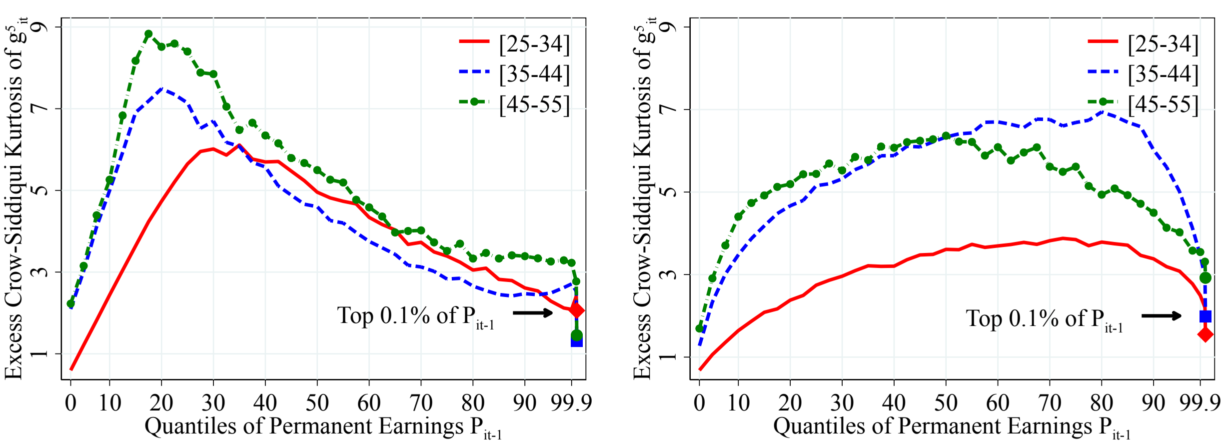
(a) Men (b) Women



Notes: Figure B.10 shows the third standardized moment of log growth rate of residual earnings for men and women with quantiles of the permanent income distribution, P_{it-1}. In each plot, the solid markers represent the corresponding measure of kurtosis for those workers at the top 0.1% of the earnings distribution for different age groups (diamond for 25 to 34 years old, square for 35 to 44 years old, and circle for 45 to 55 years old). See Section 2 for sample selection and definitions.

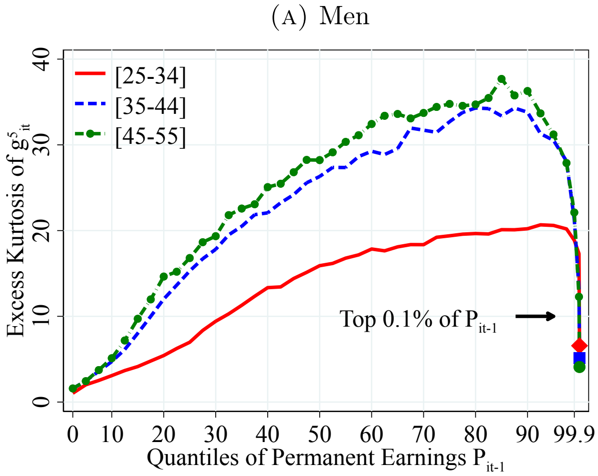
Figure B.11 – Kurtosis of Earnings Growth by Earnings Level and Age

(a) Men (b) Women



Notes: Figure B.11 shows the excess Crow-Sidiqqui kurtosis of log growth rate of residual earnings for men and women with quantiles of the permanent income distribution, P_{it-1}. Excess Crow-Sidiqqui kurtosis is defined as $\mathcal{C}_{\mathcal{K}}$ = (P97.5-P2.5) / (P75-P25) − 2.91 where 2.91 is the value of of the Crow-Siddiqui measure for a Normal distribution. In each plot, the solid markers represent the corresponding measure of kurtosis for those workers at the top 0.1% of the earnings distribution for different age groups (diamond for 25 to 34 years old, square for 35 to 44 years old, and circle for 45 to 55 years old). See Section 2 for sample selection and definitions.

Figure B.12 – Kurtosis of Earnings Growth by Earnings Level and Age



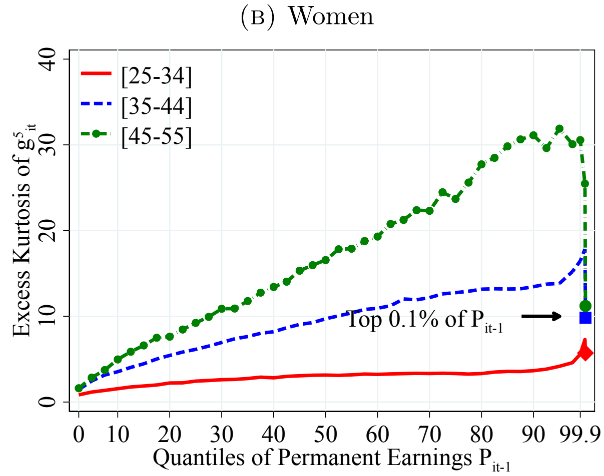
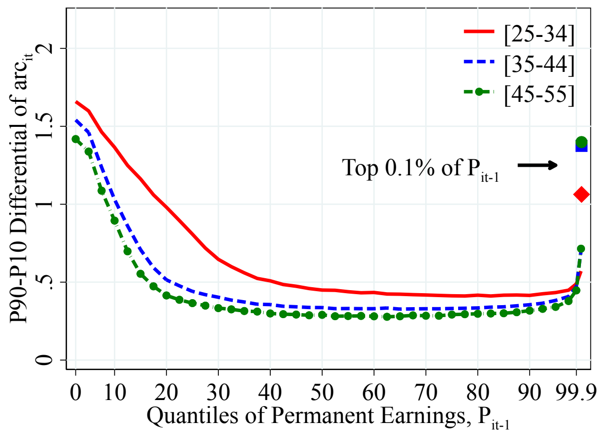


Figure B.12 shows the excess fourth standardized moment of log earnings changes for men and women with quantiles of the permanent income distribution, P_{it-1}. Excess kurtosis is defined as the value of kurtosis minus 3 which is the corresponding value for a Normal distribution. In each plot, the solid markers represent the corresponding measure of kurtosis for those workers at the top 0.1% of the earnings distribution for different age groups (diamond for 25 to 34 years old, square for 35 to 44 years old, and circle for 45 to 55 years old).

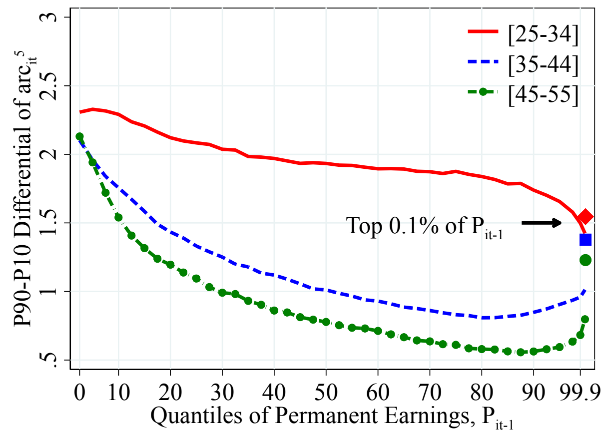
B.4.2 Heterogeneity if Idiosyncratic Earnings for One-Year Arc-Percent Change

Figure B.13 – Dispersion of Earnings Growth by Permanent Income and Age

(a) Men



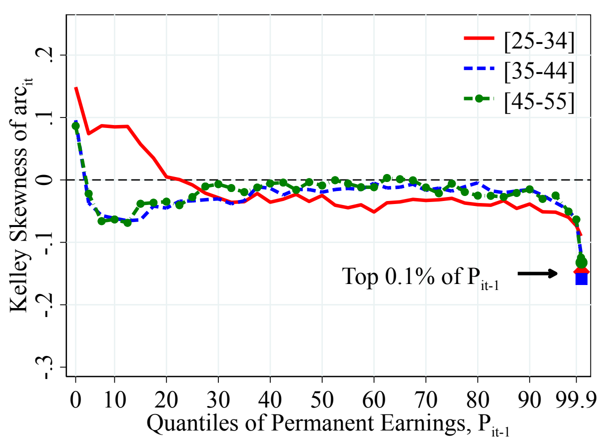
(b) Women



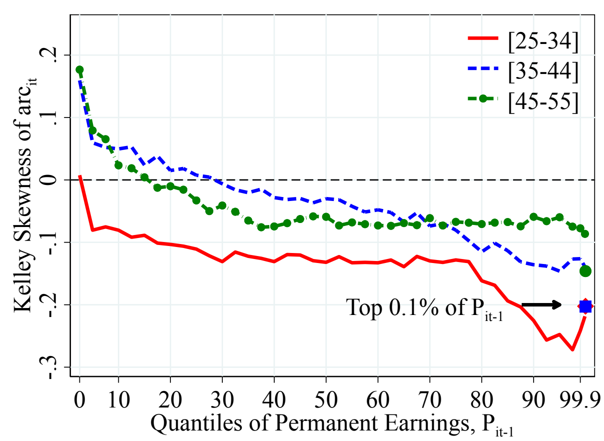
Notes: Figure B.13 shows the P90-P10 of log growth rate of residual earnings for men and women within quantiles of the permanent income distribution, P_{it-1}. In each plot, the solid markers represent P90-P10 for those workers at the top 0.1% of the permanent income distribution for different age groups (diamond for 25 to 34 years old, square for 35 to 44 years old, and circle for 45 to 55 years old). See Section 2 for sample selection and definitions.

Figure B.14 – Kelley Skewness of Earnings Growth by Earnings Level and Age

(a) Men



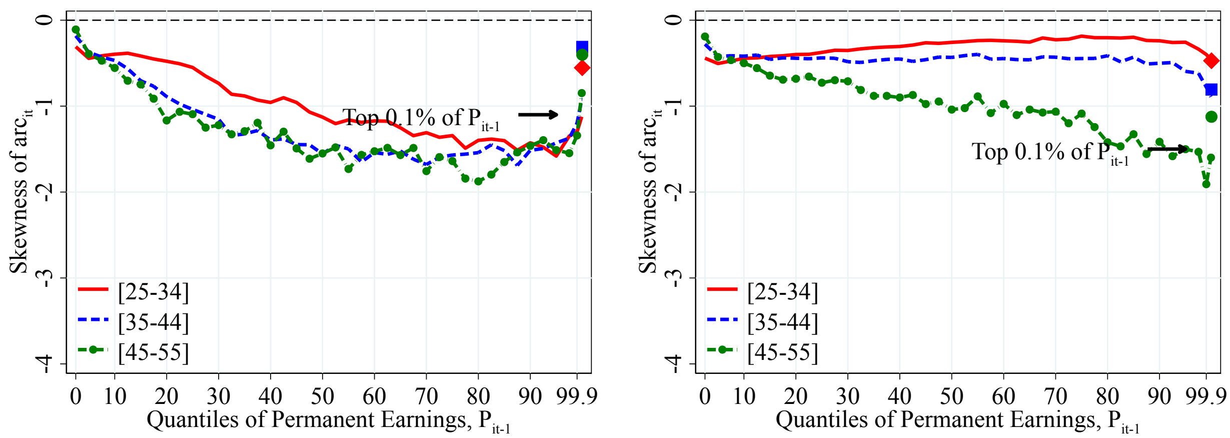
(b) Women



Notes: Figure B.14 shows the Kelley skewness of log growth rate of residual earnings for men and women within quantiles of the permanent income distribution, P_{it-1}. Kelley skewness is defined as $\mathcal{S}_{\mathcal{K}} = ((\text{P90-P50}) - (\text{P50-P10}))\,/\,(\text{P90-P10})$. In each plot, the solid markers represent the Kelley skewness for those workers at the top 0.1% of the earnings distribution for different age groups (diamond for 25 to 34 years old, square for 35 to 44 years old, and circle for 45 to 55 years old). See Section 2 for sample selection and definitions.

Figure B.15 – Skewness of Earnings Growth by Earnings Level and Age

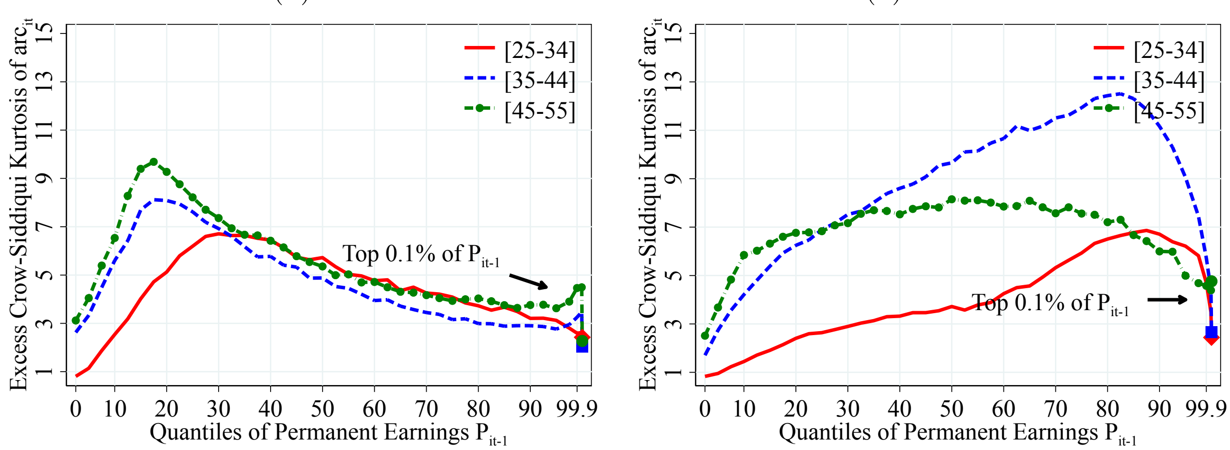
(A) Men (B) Women



Notes: Figure B.15 shows the third standardized moment of log growth rate of residual earnings for men and women with quantiles of the permanent income distribution, P_{it-1}. In each plot, the solid markers represent the corresponding measure of kurtosis for those workers at the top 0.1% of the earnings distribution for different age groups (diamond for 25 to 34 years old, square for 35 to 44 years old, and circle for 45 to 55 years old). See Section 2 for sample selection and definitions.

Figure B.16 – Kurtosis of Earnings Growth by Earnings Level and Age

(A) Men (B) Women



Notes: Figure B.16 shows the excess Crow-Sidiqqui kurtosis of arc-percent earnings growth for men and women with quantiles of the permanent income distribution, P_{it-1}. Excess Crow-Sidiqqui kurtosis is defined as $\mathcal{C}_{\mathcal{K}}$ = (P97.5-P2.5) / (P75-P25) $-$ 2.91 where 2.91 is the value of of the Crow-Siddiqui measure for a Normal distribution. In each plot, the solid markers represent the corresponding measure of kurtosis for those workers at the top 0.1% of the earnings distribution for different age groups (diamond for 25 to 34 years old, square for 35 to 44 years old, and circle for 45 to 55 years old). See Section 2 for sample selection and definitions.

Figure B.17 – Kurtosis of Earnings Growth by Earnings Level and Age

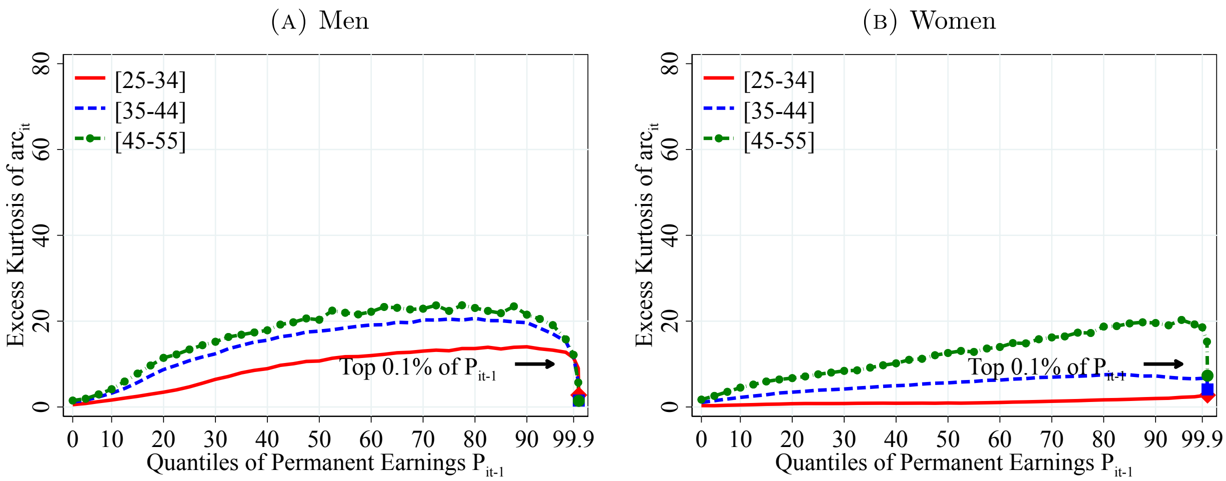
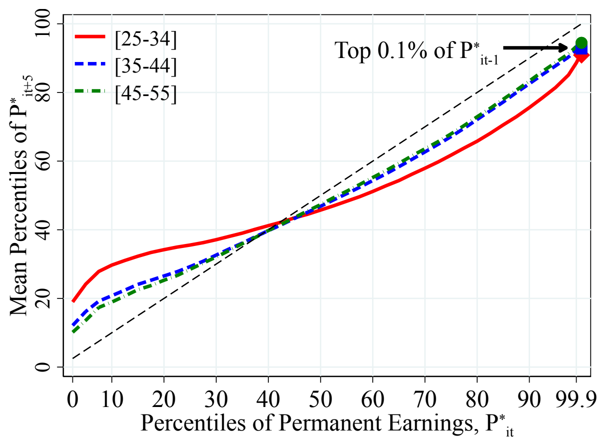


Figure B.17 shows the excess fourth standardized moment of earnings arc-percent changes for men and women with quantiles of the permanent income distribution, P_{it-1}. Excess kurtosis is defined as the value of kurtosis minus 3 which is the corresponding value for a Normal distribution. In each plot, the solid markers represent the corresponding measure of kurtosis for those workers at the top 0.1% of the earnings distribution for different age groups (diamond for 25 to 34 years old, square for 35 to 44 years old, and circle for 45 to 55 years old).

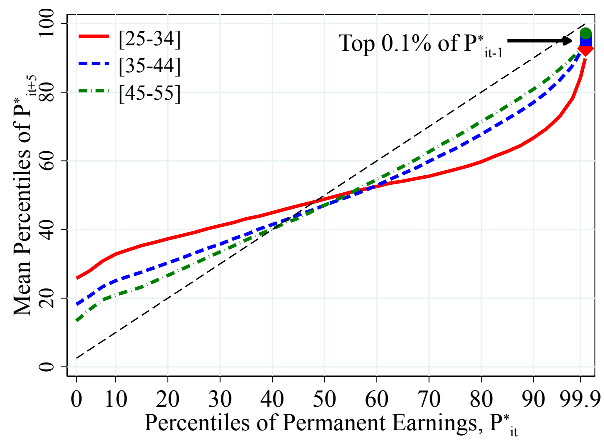
C Additional Figures on Income Mobility

Figure C.1 – Income Mobility: Rank-Rank Measures by Age: Five-Years Change

(a) Men



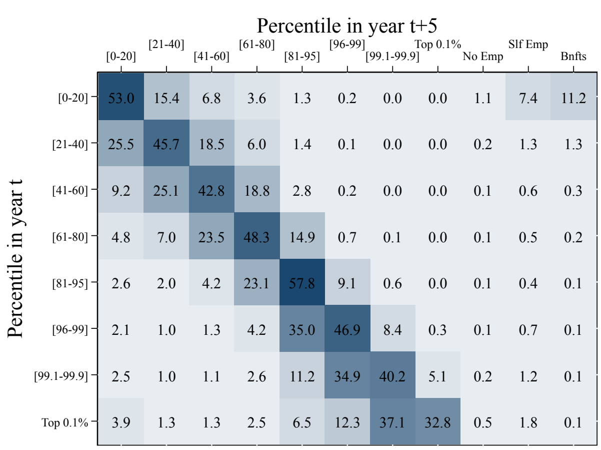
(b) Women



Notes: Figure C.1 shows the average rank obtained by individuals in period $t+5$ in the distribution of (alternative) permanent earnings, P^*_{it+5}, within different percentiles of the distribution of (alternative) permanent earnings in period t, P^*_{it}. To construct this figure, we calculate the average rank in $t+5$ for each year in our sample between 1993 and 2007 (the last years in which a ten-year change can be calculated) for each age group. We then average across all years in our sample.

Figure C.2 – Permanent Earnings Mobility: Transition Matrix of Five-Year Changes

(a) Men



(b) Women

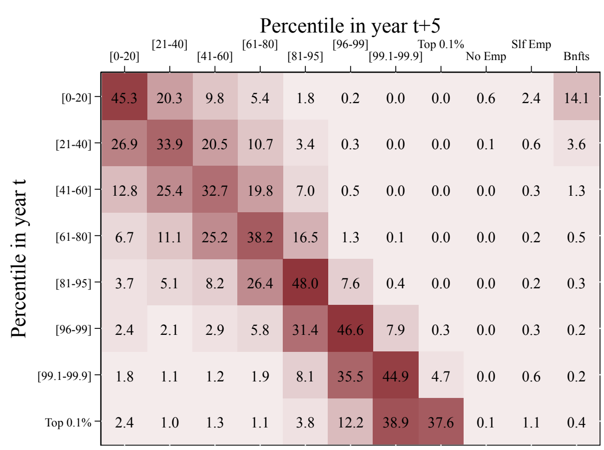


Figure C.2 shows a first-order transition matrix of individuals' permanent earnings between periods t and $t+5$ for a sample of workers between 35 and 44 years old. To construct this figure we calculate permanent earnings for workers between years 1995 and 2007 (the first and last years for which we can calculate permanent earnings and 10-year changes). No Emp. correspond to individuals whose permanent earnings is below the minimum income threshold and do not have significant self employment income or social security benefits in period $t+10$. Slf Emp (Bnfts) corresponds to individuals whose permanent earnings are below the minimum income threshold but the average level of self employment income (benefits) over the last three years is above the minimum income threshold. We then calculate the share of individuals transitioning between the predefined states for each year. Finally, we average the shares across all possible years.

Figure C.3 – Permanent Earnings Mobility: Transition Matrix of Fifteen-Year Changes

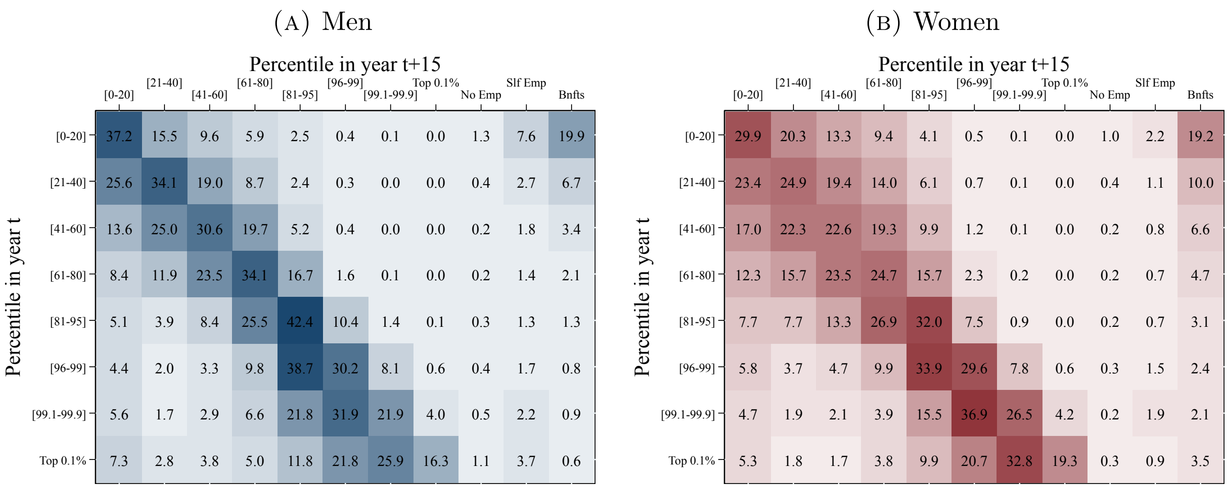


Figure C.3 shows a first-order transition matrix of individuals' permanent earnings between periods t and $t+15$ for a sample of workers between 35 and 44 years old. To construct this figure we calculate permanent earnings for workers between years 1995 and 2007 (the first and last years for which we can calculate permanent earnings and 10-year changes). No Emp. correspond to individuals whose permanent earnings is below the minimum income threshold and do not have significant self employment income or social security benefits in period $t+10$. Slf Emp (Bnfts) corresponds to individuals whose permanent earnings are below the minimum income threshold but the average level of self employment income (benefits) over the last three years is above the minimum income threshold. We then calculate the share of individuals transitioning between the predefined states for each year. Finally, we average the shares across all possible years.

D Additional Figures of Distribution Income Conditional on Parents Income

D.1 Income Inequality and Volatility for Long Panel

FIGURE D.1 – LOG LABOR EARNINGS INEQUALITY LONG SAMPLE

(A) Men (B) Women

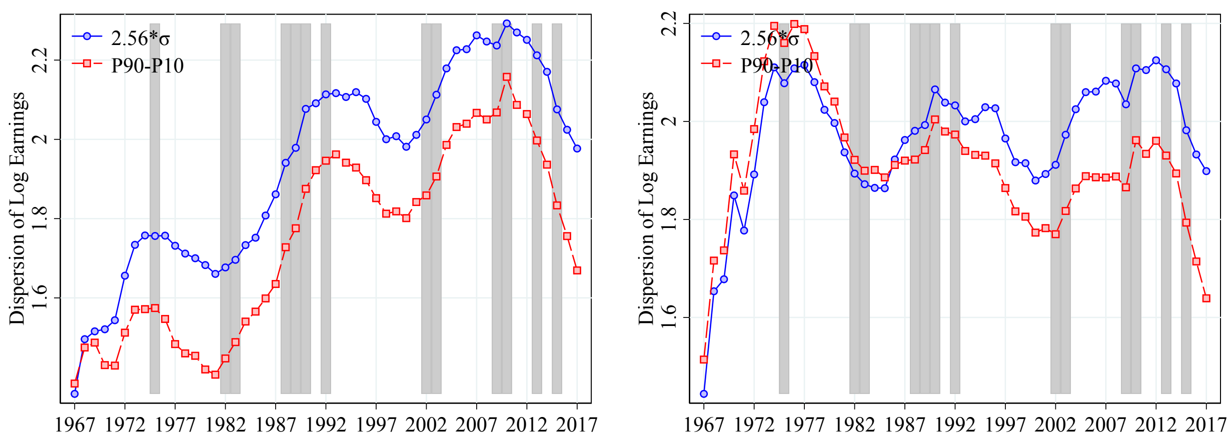
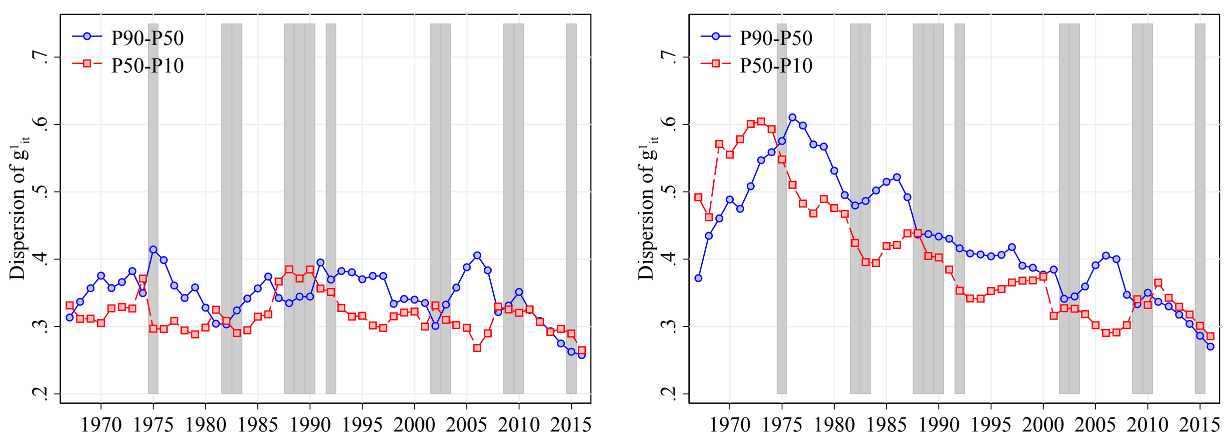


Figure D.1 shows the P90-P10 and 2.56*SD of log income for men and women. The value of 2.56*SD corresponds to the differential between the 10th and the 90th percentiles in a Normal distribution. Shaded areas represent recession years as defined as years with unemployment rate growth 0.4 pp. or more and an output gap of -0.5 or less. Results based on the CS sample. See Section 2 for sample selection and definitions.

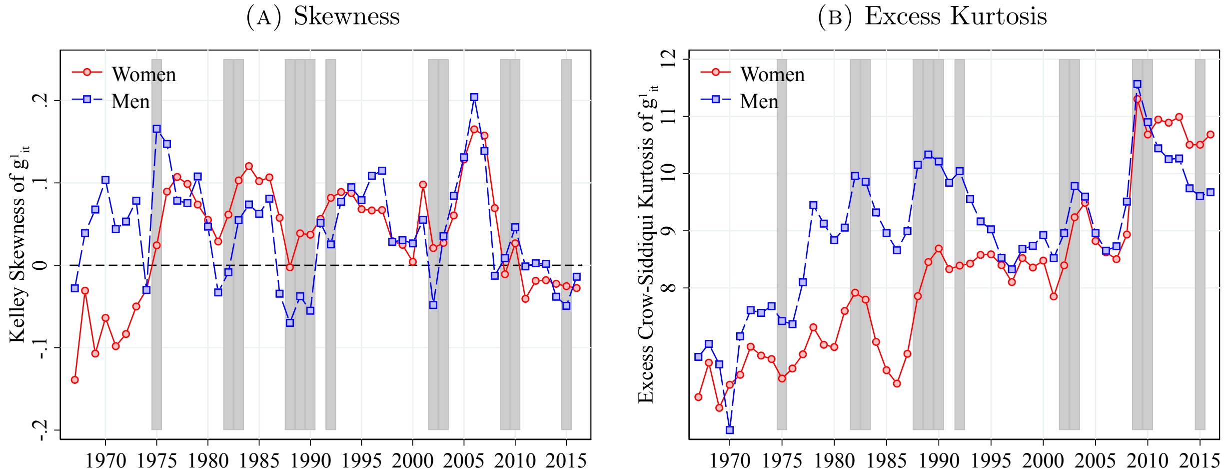
FIGURE D.2 – DISPERSION OF INCOME CHANGES

(A) Men (B) Women



Notes: Figure D.2 shows the 90th-to-50th and 50th-to-10th percentiles differential of income growth for men and women. Shaded areas represent recession years as defined as years with unemployment rate growth 0.4 pp. or more and an output gap of -0.5 or less.

Figure D.3 – Skewness and Kurtosis of Earnings Changes

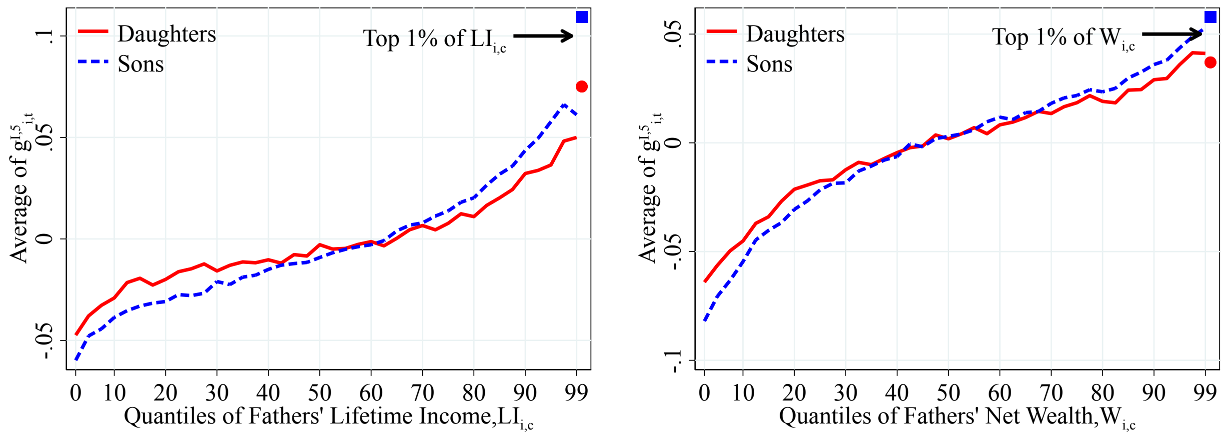


Notes: Figure D.3 shows the Kelley skewness and excess Crow-Siddiqui kurtosis of earnings growth for men and women. The excess Crow-Siddiqui kurtosis is defined as the annual Crow-Siddiqui measure minus 2.91 which is the corresponding value of Crow-Siddiqui for a Normal distribution. Shaded areas represent recession years as defined as years with unemployment rate growth 0.4 pp. or more and an output gap of -0.5 or less.

D.2 Parents and Children Income Dynamics: One- and Five-Year Changes

Figure D.4 – Mean Log Earnings Growth by Fathers Resources

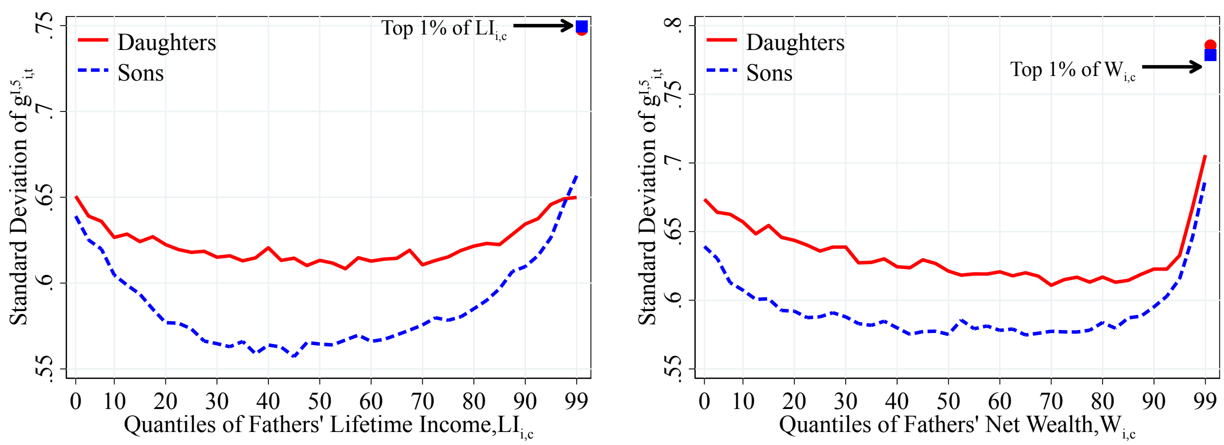
(A) By Lifetime Income | (B) By Net Wealth



Notes: Figure D.4 shows the average of the five-year residual earnings growth for men and women within quantiles of the father's lifetime income distribution (Panel A) and the fathers' households net wealth distribution (Panel B) for a total of 41 quantiles. The top 2.5% of the distribution is further separated in two groups (97.5th to 99th percentiles and 99th percentile and above). In each plot, the lines represent are the average across all years in the sample starting in 1990. The solid markers show the corresponding value among children whose parents were at the top 1% of the corresponding distribution . We estimate residual income growth as the growth rate of the residual of a year-by-year regression of log income on a set of age dummies. We run this regression separately for men and women.

Figure D.5 – Dispersion of Log Earnings Growth by Fathers Resources

(A) By Lifetime Income | (B) By Net Wealth



Notes: Figure D.5 shows the standard of the one-year residual earnings growth for men and women within quantiles of the father's lifetime income distribution (Panel A) and the fathers' households net wealth distribution (Panel B) for a total of 41 quantiles. The top 2.5% of the distribution is further separated in two groups (97.5th to 99th percentiles and 99th percentile and above). In each plot, the lines represent are the average across all years in the sample starting in 1990. The solid markers show the corresponding value among children whose parents were at the top 1% of the corresponding distribution . We estimate residual income growth as the growth rate of the residual of a year-by-year regression of log income on a set of age dummies. We run this regression separately for men and women.

Figure D.6 – Skewness of Log Earnings Growth by Fathers Resources

(a) By Lifetime Income (b) By Net Wealth

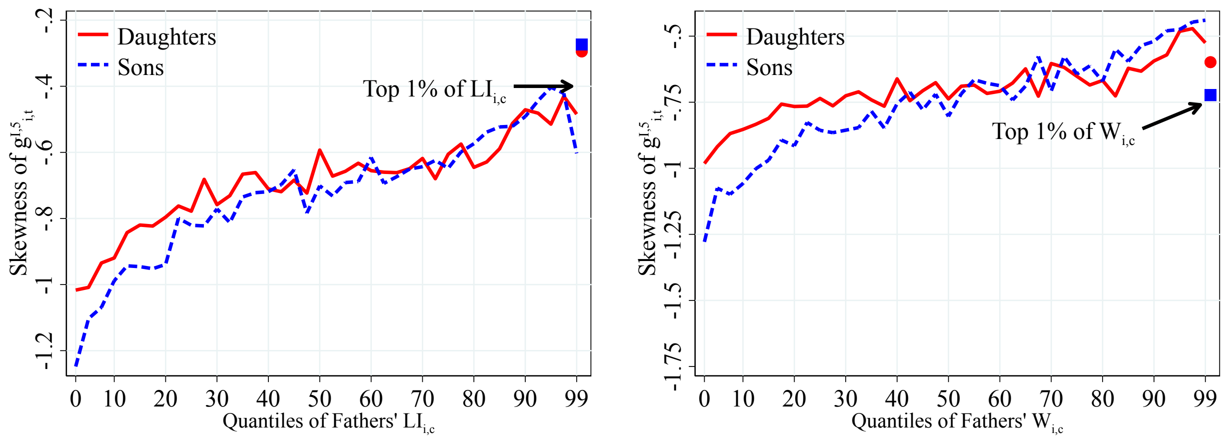


Figure D.6 shows the skewness (third standardized moment) of the one-year residual earnings growth for men and women within quantiles of the father's lifetime income distribution (Panel A) and the fathers' households net wealth distribution (Panel B) for a total of 41 quantiles. The top 2.5% of the distribution is further separated in two groups (97.5th to 99th percentiles and 99th percentile and above). In each plot, the lines represent are the average across all years in the sample starting in 1990. The solid markers show the corresponding value among children whose parents were at the top 1% of the corresponding distribution . We estimate residual income growth as the growth rate of the residual of a year-by-year regression of log income on a set of age dummies. We run this regression separately for men and women.

Figure D.7 – Kelley Skewness of Log Earnings Growth by Fathers Resources

(a) By Lifetime Income (b) By Net Wealth

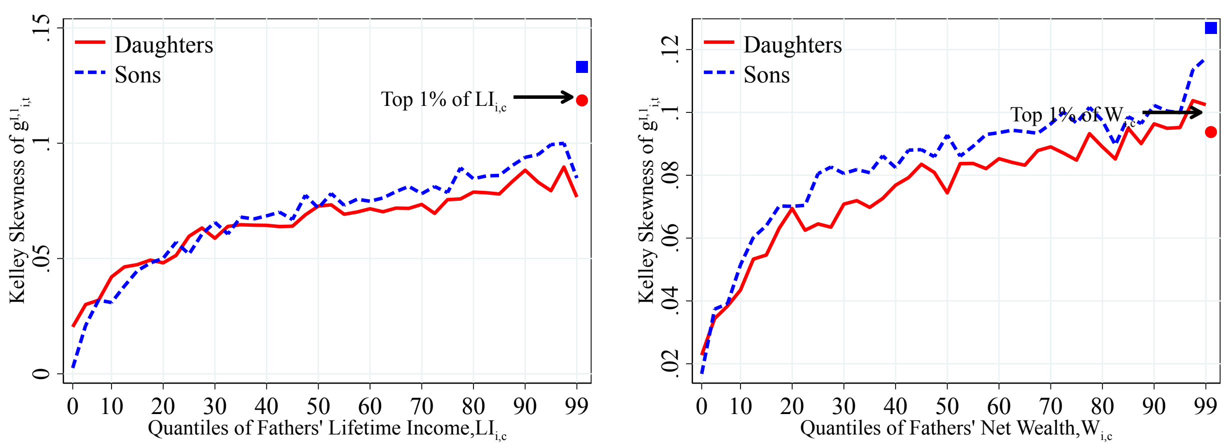


Figure D.7 shows the Kelly skewness of the one-year residual earnings growth for men and women within quantiles of the father's lifetime income distribution (Panel A) and the fathers' households net wealth distribution (Panel B) for a total of 41 quantiles. The top 2.5% of the distribution is further separated in two groups (97.5th to 99th percentiles and 99th percentile and above). In each plot, the lines represent are the average across all years in the sample starting in 1990. The solid markers show the corresponding value among children whose parents were at the top 1% of the corresponding distribution . We estimate residual income growth as the growth rate of the residual of a year-by-year regression of log income on a set of age dummies. We run this regression separately for men and women.

Figure D.8 – Kelley Skewness of Log Earnings Growth by Fathers Resources

(A) By Lifetime Income (B) By Net Wealth

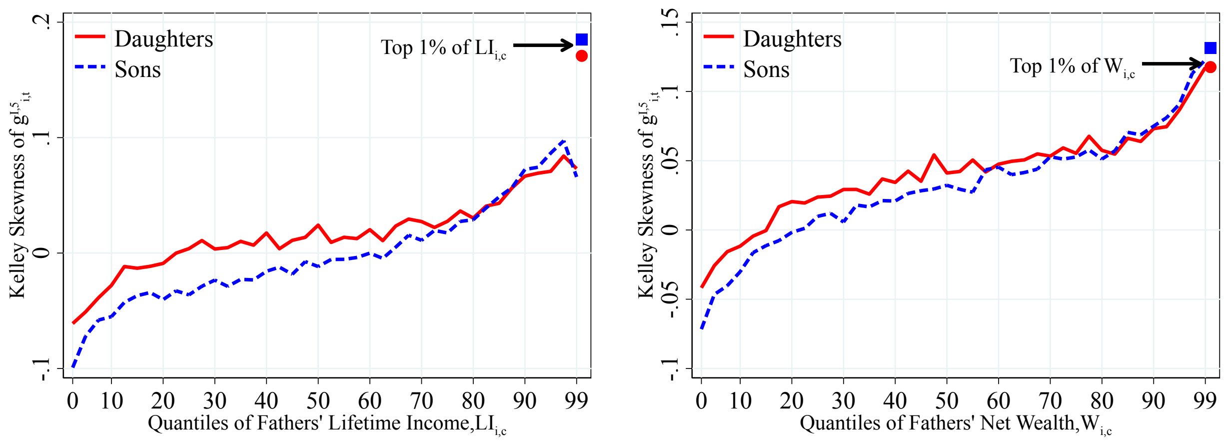


Figure D.8 shows the Kelly skewness of the one-year residual earnings growth for men and women within quantiles of the father's lifetime income distribution (Panel A) and the fathers' households net wealth distribution (Panel B) for a total of 41 quantiles. The top 2.5% of the distribution is further separated in two groups (97.5th to 99th percentiles and 99th percentile and above). In each plot, the lines represent are the average across all years in the sample starting in 1990. The solid markers show the corresponding value among children whose parents were at the top 1% of the corresponding distribution . We estimate residual income growth as the growth rate of the residual of a year-by-year regression of log income on a set of age dummies. We run this regression separately for men and women.

Figure D.9 – Left-Tail Dispersion of Log Earnings Growth By Fathers Resources

(A) By Lifetime Income (B) By Net Wealth

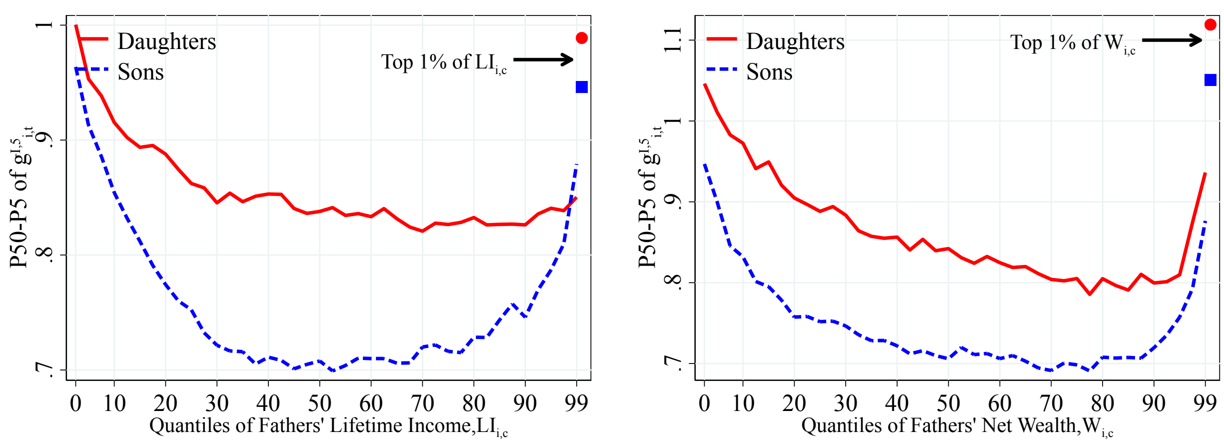


Figure D.9 shows the P50-P5 percentiles differential of the one-year residual earnings growth for men and women within quantiles of the father's lifetime income distribution (Panel A) and the fathers' households net wealth distribution (Panel B) for a total of 41 quantiles. The top 2.5% of the distribution is further separated in two groups (97.5th to 99th percentiles and 99th percentile and above). In each plot, the lines represent are the average across all years in the sample starting in 1990. The solid markers show the corresponding value among children whose parents were at the top 1% of the corresponding distribution . We estimate residual income growth as the growth rate of the residual of a year-by-year regression of log income on a set of age dummies. We run this regression separately for men and women.

Figure D.10 – Right-Tail Dispersion of Log Earnings Growth by Fathers Income

(a) By Lifetime Income　　　(b) By Net Wealth

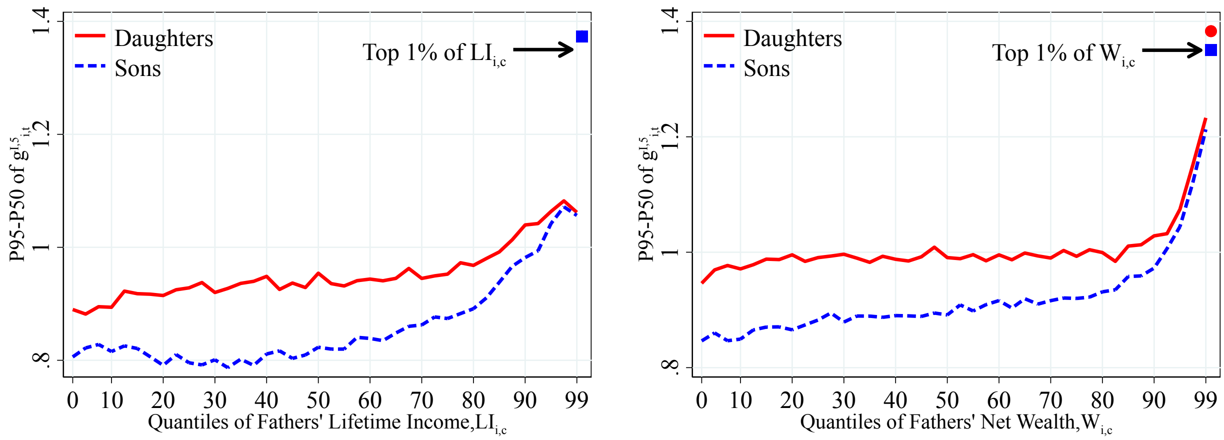


Figure D.10 shows the P95-50 percentiles differential of the one-year residual earnings growth for men and women within quantiles of the father's lifetime income distribution (Panel A) and the fathers' households net wealth distribution (Panel B) for a total of 41 quantiles. The top 2.5% of the distribution is further separated in two groups (97.5th to 99th percentiles and 99th percentile and above). In each plot, the lines represent are the average across all years in the sample starting in 1990. The solid markers show the corresponding value among children whose parents were at the top 1% of the corresponding distribution . We estimate residual income growth as the growth rate of the residual of a year-by-year regression of log income on a set of age dummies. We run this regression separately for men and women.

Figure D.11 – Kurtosis of Log Earnings Growth by Fathers Income

(a) By Father's Lifetime Income　　　(b) By Father's Net Wealth

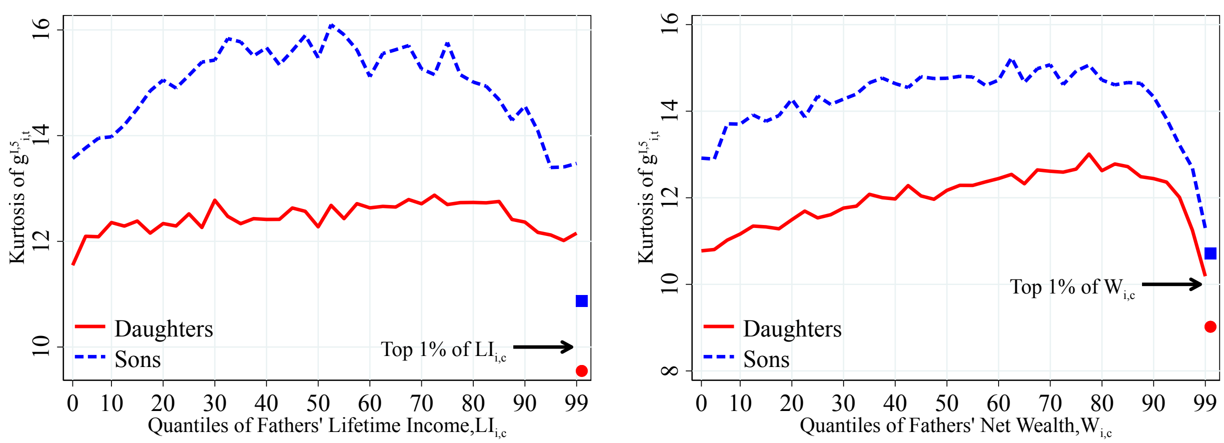
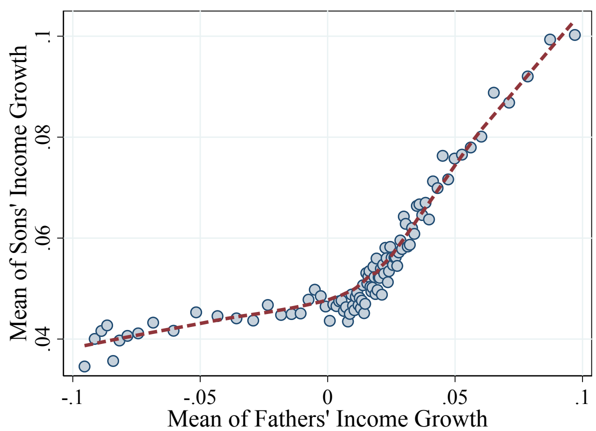


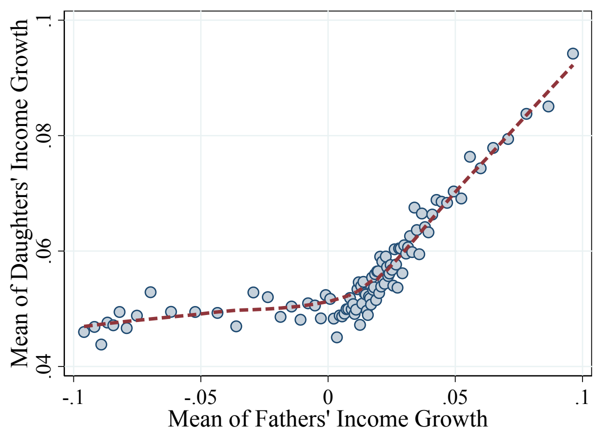
Figure D.11 shows the kurtosis (fourth standardized moment) of the one-year residual earnings growth for men and women within quantiles of the father's lifetime income distribution (Panel A) and the fathers' households net wealth distribution (Panel B) for a total of 41 quantiles. The top 2.5% of the distribution is further separated in two groups (97.5th to 99th percentiles and 99th percentile and above). In each plot, the lines represent are the average across all years in the sample starting in 1990. The solid markers show the corresponding value among children whose parents were at the top 1% of the corresponding distribution . We estimate residual income growth as the growth rate of the residual of a year-by-year regression of log income on a set of age dummies. We run this regression separately for men and women.

Figure D.12 – Average of Income Growth of Fathers and Children

(A) Men



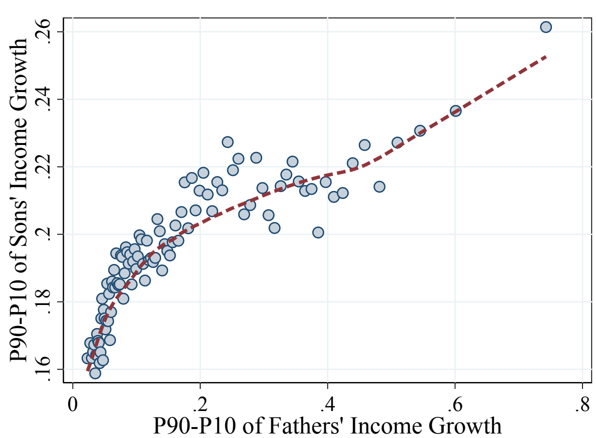
(B) Women



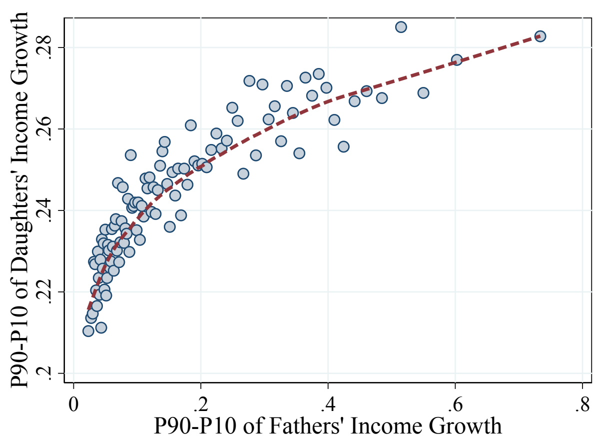
Notes: Figure **??** shows a binned scatter plot of fathers and children mean income growth measured by the individual-level standard deviation. The dashed is the non-linear correlation estimated from a lowess estimator. To scatter plot is based on a sample of 277K fathers-children pairs. The sample is divided in 100 bins.

Figure D.13 – Standard Deviation of Income Growth of Fathers and Children

(A) Men

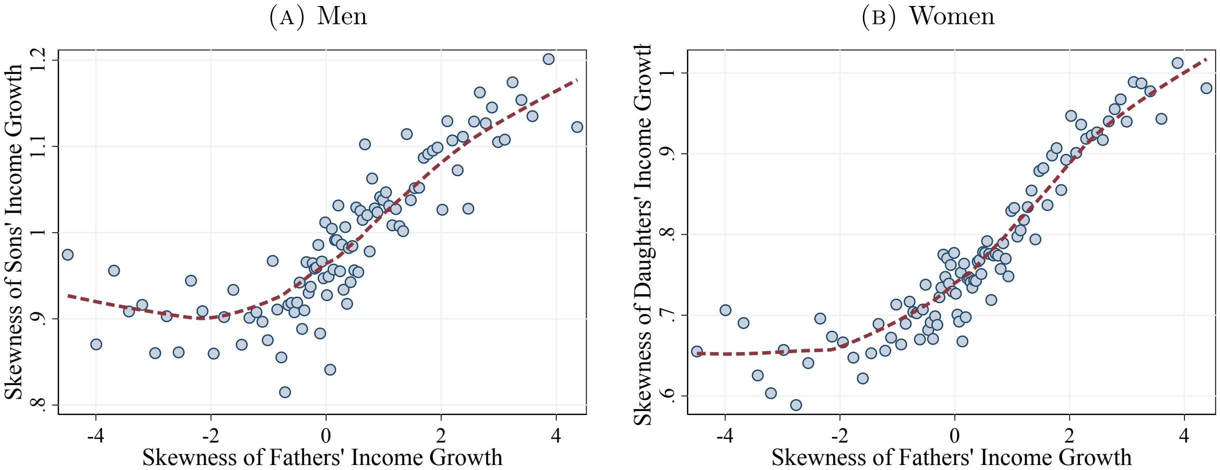


(B) Women



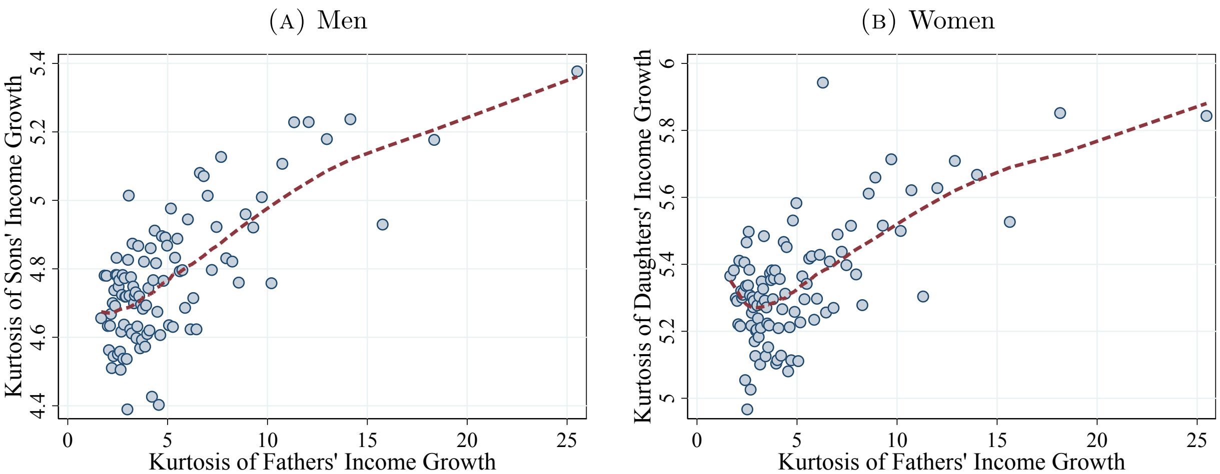
Notes: Figure D.13 shows a binned scatter plot of fathers and children income growth dispersion measured by the individual-level standard deviation. The dashed is the non-linear correlation estimated from a lowess estimator. To scatter plot is based on a sample of 277K fathers-children pairs. The sample is divided in 100 bins.

Figure D.14 – Skewness of Income Growth of Fathers and Children



Notes: Figure D.14 shows a binned scatter plot of fathers and children income growth skewness measured by the individual-level third standardized moment. The dashed is the non-linear correlation estimated from a lowess estimator. To plot is based on a sample of 277K fathers-children pairs. The sample is divided in 130 bins.

Figure D.15 – Crow-Siddiqui of Income Growth of Fathers and Children



Notes: Figure D.15 shows a binscatter plot of fathers and children income growth kurtosis measured by the individual-level Crow-Siddiqui kuetosis. The dashed is the non-linear correlation estimated from a lowess estimator. To plot is based on a sample of 154K fathers-children pairs. The sample is divided in 130 bins. Values in the y-axis residualized from fathers' lifetime income level.

Table D.1 – Determinants of Children's Income Dynamics: More Controls

	(1)	(2)	(3)	(4)	(5)	(6)
	$P50_i^c$		$P9010_i^c$		$\mathcal{S_K}_i^c$	
	Men	Women	Men	Women	Men	Women
$P50_i^f$	0.086***	0.064***	0.320***	0.358***	0.922***	0.810***
	(0.006)	(0.006)	(0.058)	(0.062)	(0.065)	(0.065)
$P9010_i^f$	0.006***	0.005***	0.193***	0.141***	-0.009*	0.028***
	(0.000)	(0.000)	(0.005)	(0.005)	(0.005)	(0.005)
$\mathcal{S_K}_i^f$	0.008***	0.005***	0.050***	0.052***	0.102***	0.082***
	(0.000)	(0.004)	(0.004)	(0.004)	(0.005)	(0.004)
$\log LI_i^c$	0.026***	0.011***	-0.345***	-0.439***	0.082***	0.094***
	(0.000)	(0.000)	(0.003)	(0.003)	(0.003)	(0.003)
$\log LI_i^f$	0.003***	0.002***	0.132***	0.100***	0.026***	0.017***
	(0.000)	(0.000)	(0.004)	(0.004)	(0.004)	(0.004)
$\log W_i^f$	0.001***	0.000***	0.004***	0.004***	0.010***	0.000
	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
$P50_i^c$			1.309***	0.217***	-0.171***	0.217***
			(0.037)	(0.038)	(0.041)	(0.039)
$P9010_i^c$	0.012***	0.002***			-0.026***	-0.024***
	(0.000)	(0.000)			(0.004)	(0.004)
$\mathcal{S_K}_i^c$	-0.001***	0.002***	-0.021***	-0.022***		
	(0.000)	(0.000)	(0.003)	(0.003)		
R^2	0.141	0.033	0.195	0.277	0.036	0.032
N (000's)	79.7	75.6	79.7	75.6	79.7	75.6

Notes: Table II shows the coefficient of a cross-sectional regression of workers-level measures of median lifetime growth, P90-P10 differential, and Kelley Skewness ($\mathcal{S_K}_i$), with the superscript c denoting children and f denoting fathers. Income growth is measure as the one-year log change of a measure of permanent income, calculated as the average income of an individual between years t and $t-2$. In the sample, we consider fathers and children with more than 20 years of data. Lifetime income of fathers and children is calculated as in Equation 1. The measure of lifetime wealth is calculated as the fathers' average wealth between ages 45 and 55 (or the nearest age to this age range for individuals that are observed when they are too young (below 45) or too old (above 55).

Table D.2 – Determinants of Children's Income Dynamics: Centered Moments

	(1)	(2)		(3)	(4)		(5)	(6)
	Mean_i^c			Std.Dev_i^c			Skewness_i^c	
	Men	Women		Men	Women		Men	Women
Mean_i^f	0.157***	0.124***						
	(0.004)	(0.004)						
Std.Dev_i^f				0.144***	0.110***			
				(0.003)	(0.003)			
Skewness_i^f							0.045***	0.047***
							(0.002)	(0.002)
$\log LI_i^c$	0.025***	0.010***		-0.156***	-0.187***		0.057***	0.274***
	(0.000)	(0.000)		(0.001)	(0.001)		(0.009)	(0.009)
$\log LI_i^f$	0.006***	0.0027***		0.058***	0.042***		0.016	0.007
	(0.000)	(0.000)		(0.001)	(0.001)		(0.011)	(0.011)
$\log W_i^f$	0.001***	0.001***		0.002***	0.002***		0.012***	0.017***
	(0.000)	(0.000)		(0.000)	(0.000)		(0.002)	(0.002)
$P50_i^c$				0.361***	-0.057***		-0.776***	-0.370***
				(0.011)	(0.011)		(0.096)	(0.095)
$P9010_i^c$	0.038***	0.022***					-1.170***	-0.807***
	(0.000)	(0.000)					(0.011)	(0.010)
$\mathcal{S}_{\mathcal{K}}{}_i^c$	0.104***	0.107***		-0.039***	-0.038***			
	(0.000)	(0.000)		(0.001)	(0.001)			
$P50_i^f$				0.230***	0.182***		1.581***	2.140***
				(0.019)	(0.019)		(0.165)	(0.163)
$P9010_i^f$	0.016***	0.012***					-0.069***	0.040***
	(0.001)	(0.001)					(0.013)	(0.013)
$\mathcal{S}_{\mathcal{K}}{}_i^f$	0.004***	0.004***		0.026***	0.023***			
	(0.001)	(0.001)		(0.001)	(0.001)			
R2	0.448	0.442		0.207	0.259		0.107	0.0953
N (000s)	155.3	155.3	79.7	75.6	79.7	75.6	79.7	75.6

Notes: Table D.2 shows the coefficient of a cross-sectional regression of workers-level measures of median lifetime growth, P90-P10 differential, and Kelley Skewness ($\mathcal{S}_{\mathcal{K}i}$), with the superscript c denoting children and f denoting fathers. Income growth is measure as the one-year arc-percent change of a measure of permanent income, calculated as the average income of an individual between years t and $t-2$. In the sample, we consider fathers and children with more than 20 years of data. Lifetime income of fathers and children is calculated as in Equation 1. The measure of lifetime wealth is calculated as the fathers' average wealth between ages 45 and 55 (or the nearest age to this age range for individuals that are observed when they are too young (below 45) or too old (above 55).

E The GID Code: Detailed Instructions

E.1 General Directions

This appendix discusses the codes used to generate the sample and the statistics for the core section of the Global Income Dynamics Database Project (GID). The code packet—available in GitHub here—contains seven do-files that execute the initialization of the parameters, execute the sample creation, and produce the figures for the core section of the paper. The packet also contains two auxiliary files used for summary statistics (myprogs.do) and plotting (myplots.do). The codes are written in Stata 13—and tested in Stata 15 and 16—and were designed to produce the statistics listed in the GID Guidelines document, as well as saves the results in CSV files, and creates a large set of figures in PDF. These codes are continually updated for efficiency and, in few cases, for small calculation errors. We strongly suggest regularly downloading the most recent version of the codes.

The codes require researchers to create few folders in their local machines and set few inputs, which reflects the characteristics of the data used in the analysis. The basic steps to run these codes as follows.

1. Create in your local machine the following subfolders (all in lowercase) under the same folder:

   - /do
   - /dta
   - /log
   - /out
   - /figs

   Next, download the provided do files in folder /do and copy the country-specific raw data file in folder /dta. The raw data must be saved in a dta file before running codes. The log files will be saved in the /log folder, the results will be saved under /out, and figures will be saved under /figs.[37]

[37] Notice that the folder /fig will contain several additional subfolders (created by the plotting code), which will orderly save the figures for each section of the core section of the paper.

2. Open 0_Initialize.do in Stata and assign country-specific parameters such as the starting and ending years of the sample, the name and location of the raw dataset, the country's CPI, the exchange rate between the corresponding country and the United States, and so on, for which further instructions are given in Section E.2.

3. Open 1_Gen_Base_Sample.do in Stata, specify the directory of the main folder that contains the above five sub-folders in your local machine and run. This do file renames the variables, does basic sample selection, creates new variables (e.g., log and residual earnings, one-year residual earnings growth), and generates the master_sample.dta, that is a wide-form dataset which will be used in the rest of the do files. The main output of this do file (master_sample.dta) is saved in the folder /dta and contains the following variables (among several others):

   (a) personid: id of the individual used throughout the do files

   (b) male: indicator variable equal to 1 if male and 0 if female

   (c) yob: year of birth of the individual

   (d) yod: year of death of the individual

   (e) educ: indicator variable with education categories

   (f) labor: real labor earnings in levels

   (g) logearn: real labor earnings in log levels

   (h) permearn: permanent income defined as $P_{it-1} = \frac{\sum_{s=t-3}^{t-1} y_{i,s}}{3}$, where $y_{i,s}$ is the real earnings of individual i in period s. Notice that $y_{i,s}$ must be above the minimum income threshold. The value of this threshold must also be specified in the 0_Initialize.do file.

   (i) permearnalt: alternative measure of permanent income, which consider earnings below the minimum income threshold as well.

   (j) researn: residual log earnings

   (k) researn1F: one-year forward residualized log earnings change, g_{it}

   (l) researn5F: five-year forward residualized log earnings change, g_{it}^5

   We provide additional details on the construction of each of these variables in Section E.2.

4. Open 2_DescriptiveStats.do in Stata, specify the directory of the main folder in your local machine, and run. This do file generates a folder under /out, whose name consists of the date the program is run and "Descriptive_Stat."

5. Open 3_Inequality.do in Stata, specify the directory of the main folder in your local machine, and run. This do-file contains cross-sectional moment on income inequality and earnings concentration.

6. Open 4_Volatility.do in Stata, specify the directory of the main folder in your local machine, and run. This do file generates a set of .csv files with the statistics for the section "Key statistics 3: Volatility and Higher-Order Moments."

7. Open 5_Mobility.do in Stata, specify the directory of the main folder in your local machine, and run. This do file generates a set of .csv files with the statistics for the section "Key statistics 4: Mobility."

8. Open 6_Core_Figs.do in Stata, specify the directory of the main folder in your local machine, the directories where the different results are saved (Inequality, Mobility, and so on) and where the figures will be saved. The default is the folder /figs and figures are saved in PDF format.[38]

In the next section, we provide some additional details on each of the codes. All programs are heavily commented, and we have made our best of our effort to make them bug-free. If you find any problem, please let us know so we can update the codes.

E.2 Specific Details on the Codes and Variable Construction

0_Initialize.do

The 0_Initialize.do defines the variable names, time span, and vectors used throughout the codes and allows each team to select some options that best suit their dataset. Given its importance, here we discuss several key details (more comments can be found in the do-file). Lines 5 to 18 of 0_Initialize.do define general variables that must be followed by the teams to generate the core statistics. No change is required in this section. These definitions ensure that the sample used for the core section of the paper is

[38] To plot additional figures that you might be interested in but are not covered in the file 6_Core_Figs.do, you might also need to modify the file myplots.do. If that is the case, we encourage you to contact us before making changes so all the plots maintain a similar format.

comparable across countries. Lines 20 to 100 require the input of the user. Please read in detail.

1. **Unix vs. Windows**. Define whether the machine on which you are running your codes is Unix/Mac (unix=1) or Windows (unix=0).[39]

2. **Wide vs. long format**. Define whether the raw sample is in wide form (wide =1) or long form (wide=0). If it is in long form, the 1_Gen_Base_Sample.do file will convert it to wide form (one row per individual) when creating the dataset master_sample.dta. The rest of the codes are designed to work with this .dta.

    (a) By long format, we mean a dataset in which each observation (row) is an individual-year pair. In other words, workers' observations are stacked, there is one column that defines the unit of time (year) and one column for each variable defining the value of each variable within the year (one column for earnings, one for education, and so on).

    (b) By wide format, we mean a dataset in which each observation (row) is an individual and different columns define different observations for the same individual. In other words, workers' observations are side by side, and there is one column per year defining each variable (one column is the earnings in 2000, a second column is the earnings in 2001, and so on).

3. **Missing values for labor income.** If there are genuine missing values for labor income please set global ${miss_earn} to 1 (lines 33 to 36). If it is set to zero (the default), the code will convert all missing earnings observations to zero. This is particularly important if your raw dataset is in long form and there are no observations for zero labor income in a given year.

4. **Variable names**. Specify the names of the variables in your data set between lines 41 and 48. These variables are the minimum set necessary to generate all the statistics in the Guidelines; hence, each team must make sure the raw data contain these variables. The 1_Gen_Base_Sample.do file then will rename these variable to our choices in the master_sample.dta. This step helps to simplify the code in the rest of the do-files.

[39]Although STATA run on Windows machines corrects the folder separators, just to be on the safe side, we specify whether the separator is "/" or "\", which will then be used to locate the sub-folders.

5. **Variable types.** The do files are written under certain assumptions about the type of variables available in each dataset. We did not attempt to change the format of the variables, hence, each team must make sure that the raw data contains the correct format (i.e. education must be a numerical categorial integer variable, gender must be binary, and so on). Here we describe in detail the variables used in the analysis

   (a) ${personid_var}: Numerical categorical variable. Teams must make sure an individual id appears only one time per year in the sample.

   (b) ${male_var}: Numerical categorial variable that is equal to 1 if the individual is male, 0 if female.

   (c) ${yob_var}: Numerical categorical variable that defines the year of birth of an individual. Teams must make sure this is not missing or changes across different observations of the same individual (if the raw data are in long form). Individuals with missing ${yob_var} will be dropped from the sample.

   (d) ${yod_var}: Numerical categorical variable that defines the year of death of an individual. Teams must make sure this variable does not change across different observations of the same individual. Individuals with missing ${yod_var} will be treated as if they where still alive by the end of the sample.

   (e) ${educ_var}: Numerical categorical variable that defines the education group of an individual. This variable can change across different observations of an individual. There is no restriction on the number of categories this might contain.

   (f) ${labor_var}: Numerical variable that defines the labor earnings of an individual. This variable might contain missing values. Recall that you also need to choose whether the missing observations are set to 0 by setting global ${miss_earn} to 1 or 0 in line 36.

   (g) ${year_var}: Numerical variable that defines the year of the observation if the raw data are in long form.

6. **First and last year.** Specify the first and last year of the sample for which the statistics will be calculated. The sample is assumed to have no gaps in between (all years between ${yrfirst} and ${yrlast} are available).

7. **Density estimation.** Global ${kyear} defines for which years the densities will be calculated. By default, the code calculated the densities in years ending with 0 and 5 (i.e., 1995, 2000, 2005, and so on). In case you want to calculate densities every year, change ${kyear} = 1.

8. **Quantile estimates.** Quantile estimates are mainly used in the 5_Mobility.do do file. See the code for additional details. The global ${nquantiles} defines how many quantiles will be used to divide the distribution of permanent income. The default is 40. The global ${nquantilesalt} does the same for the quantiles of the distribution of alternative permanent income.

9. **Heterogeneity groups.** The global ${hetgroup} specifies what heterogeneous characteristics are considered. By default, the code calculates statistics by gender, education, age, and the cross groups. Additional levels of heterogeneity can be easily incorporated as long as the corresponding variables are passed to the sample.[40]

10. **CPI, min income, and exchange rate.** The matrices cpimat, rmininc, and exrate contain the CPI, the min income threshold, and the exchange rate (nominal) that is used throughout the code. These need to be imputed from ${yrfirst} to ${yrlast} *without* gaps. All nominal variables must be deflated by 2018 prices. Hence, set the global ${cpi2018} equal to the corresponding value the CPI in 2018 for your country.

1_Gen_Base_Sample.do

The 1_Gen_Base_Sample.do code takes the raw data and creates the master sample which is used by the rest of the code to generate the statistics. The master sample is built as a wide-format dataset. If the raw data is in long format, lines 30 to 75 reshape the data to a wide format. Because of the sheer size of the administrative datasets, the codes do not use the Stata reshape routine, which is significantly slower than our code.

Having reshaped the dataset, lines 85 to 130 creates a base sample by transforming nominal values into real values and dropping observations for individuals outside the predefined age range (25 to 55 in the baseline setup). The codes also have the possibility

[40] Check the code myprogs.do for details.

to add a small amount of noise to each observation—in the case this is necessary for disclosure considerations—and transforms to 0 observations that are missing in the sample. The code also creates the basic measure of labor earnings, labor 'yr', which is the real labor earnings in year 'yr'. For simplicity, denote this measure Y_{it}. The resulting dataset is saved base_sample.dta.

The following section calculates three measures used throughout the code: a measure of log real labor earnings, residual labor earnings, and an alternative measure of residual earnings. The first is the log value of real earnings, $\log Y_{it}$, defined for all individuals and periods in which $Y_{i,t}$ is greater than a predefined minimum income threshold. We define a second measure of log earnings, denoted as $\log Y_{it}^c$, which similar to $\log Y_{it}$ but extend the sample to individual observations that are 1/3 above the minimum income threshold. This value is typically defined in the United States as the real value of earnings derived from working full time for one quarter at the federal minimum wage. The specific value can change from country to country and can be changed in the 0_Initialize.do code.

We construct residual earnings by running a year-by-gender regression of $\log Y_{it}$ on a set of age dummies and denote the residuals of this regression as ε_{it}. We construct a second measure by using $\log Y_{it}^c$, with residuals denoted as ε_{it}^c. Finally, we construct a third measure of residual earnings by running a year-gender regression on a set of age and education dummies. The resulting dataset is saved in researn.dta. The code also saves the dummies of the first regression, which capture the average profile of earnings over the life cycle.

Using the residual earnings, we construct a measure of residual earnings growth as $g_{it}^k = \varepsilon_{it+k}^c - \varepsilon_{it}$ with $k \in \{1, 5\}$. In this way, the measure of earnings growth considers individuals that have earnings above the minimum income threshold in period t but can at most be one-third below this value in period k. Furthermore, to account for individuals moving in and out of the labor market, we construct a second measure of earnings growth using the arc-percent method. Specifically, we first calculate a rescaled measure of earnings, $\tilde{Y}_{it} = Y_{it}/\overline{Y_t}$, where $\overline{Y}_t$ is the average real labor income within a year-gender-age group. We then calculate the arc-percent measure as $arc_{it} = \left(\tilde{Y}_{it+k} - \tilde{Y}_{it}\right)/0.5 \times \left(\tilde{Y}_{it+k} + \tilde{Y}_{it}\right)$. Notice that this measure is defined for all observations in the sample, including observations with zero labor earnings.

The last two sections of 1_Gen_Base_Sample.do pertain to the calculation of a measure of individual-level permanent income. We calculate two measures. The first defines permanent earnings in period t as the average of labor earnings, Y_{it}, between years $t-1$ and $t-3$, only for years in which labor earnings are above the minimum income threshold and for individuals whose measure of permanent earnings was calculated using at least two observations of income above the minimum income threshold. Using this average measure, we run a set of regressions on a set of age dummies by year-gender groups. The residuals of these regressions, denoted by P_{it}, is our measure of an individual's permanent earnings. We save this measure in the file permearn.dta.

The second measure of permanent income is constructed as the average labor earnings between period t and $t-2$ using all income measures, including those below the minimum income threshold. This alternative measure of permanent income, denoted by P^*_{it}, is constructed for individuals whose P^*_{it} was calculated using three non-missing observations, and at least one of these observations was above the minimum income threshold. We save this measure in the file permearnalt.dta

The last section of the code collects all of the constructed measures (labor earnings, residual earnings, earnings growth, and permanent earnings) and saves the master_sample.dta, which is used in the rest of the codes.

2_DescriptiveStats.do

The 2_DescriptiveStats.do calculates descriptive statistics for each of the different samples used in the analysis. The samples are divided into three subsamples. The Cross-Sectional sample (CS) considers all year-individual observations with labor income at or above the minimum income threshold. The Longitudinal Sample (LX) considers year-individual observations with non-missing measures of g^1_{it} and g^5_{it}. Notice that, by construction, this second sample considers data between ${yrfirst} to ${yrlast}-5. Finally, the Heterogeneity sample (H) considers year-individual observations with with non-missing measures of g^1_{it}, g^5_{it}, and P_{it}. Hence, this sample considers data between ${yrfirst}+3 to ${yrlast}-5.

For each of these samples, we calculate cross sectional moments of the real labor earnings distribution. To allow a simple cross-country comparison, labor income is transformed into 2018 US dollars using the country-specific average exchange rate of 2018 between the country and the United States. This measure of real labor earnings on dollars is only used in this code and all other measures in the code consider real local currency values. The moments we calculate consist of centered moments (mean, average, skewness, and kurtosis) and detailed percentiles of the distribution across the entire population and within gender groups. Results are saved in a set of CSV files denoted by the sample under analysis. This code also saves a set of CSV files containing the age dummies calculated in 1_Gen_Base_Sample.do.

3_Inequality.do

The 3_Inequality.do calculates cross-sectional moments of the distribution of labor earnings, residual labor earnings, and permanent earnings across the entire distribution and within gender, and within three age groups defined by the following age thresholds: 25 to 34, 35 to 44, and 45 to 55. Cross-sectional moments are calculated using functions bymysum_detail (which calculate centered moments) and bymyPCT (which calculates percentiles).

Using the same data, this code calculates the empirical density of log earnings across the population and for men and women using a k-density estimator of Stata. All parameters, by the number of nodes, for the k-density estimator (i.e., band width, estimator, and so on) are set to their default values. These parameters can be adjusted in the function bymyKDN and bymyKDNmale available in myprograms.do.

The code also calculates measures of income concentration using the function bymyCNT and tail coefficients of the income distribution using the function bymyRAT. The concentration thresholds are predefined (e.g., top 1%, top 10%, bottom 50%, and so on) but can be changed in myprograms.do. Similarly, the cutoffs used to calculate the tail coefficient are predefined but can be changed in myprograms.do.

Finally, the code calculates a set of cross-sectional moments for the distribution of residual earnings which control for education dummies. Note that in case your dataset does not contain an education identifier, this measure of residual earnings—and the corresponding statistics—will coincide with the standard measure of residual earnings.

The code then collects all calculations and saved separated CSV files. In particular, we save cross sectional-moments of different variables using the naming convention L_'vari'_sumstat.csv. Concentration measures are saved in L_earn_con.csv and L_earn_gStPop1.csv. Tail indexes are saved in RI_maleagegp_earn_idex.csv. In case we calculate these measures within different heterogeneity groups, each file will have a suffix for gender (male suffix) or age (agegp suffix). Additional details can be found between lines 274 to 577. The rest of the code collects the cross-sectional statistics and saves a set of CSV files in the folder Inequality.

4_Volatility.do

The 4_Volatility.do code calculates cross-sectional moments of the distribution of one- and five-year residual earnings changes and arc-percent changes across the population and within heterogeneity groups defined by gender and three age groups (defined above). The code also calculates empirical densities of each growth variable using the function bymyKDN and a predefined set of points.

To provide a sense of the concentration of growth observations in different sections of the distribution, the function kurpercentiles calculates the fraction of observations at different intervals of the domain of a given growth measure. These intervals are defined by deviations from the mean of the distribution at the following cutoffs: +/-1%, +/-5%, +/-σ, +/-2σ, and +/-3σ, where σ is the standard deviation of the corresponding distribution.

Lines to 167 to 330 calculate moments of the earnings growth distribution within percentiles of the permanent income distribution. In particular, in each year, individuals are sorted into ${nquantiles} quantiles plus three additional quantiles identifying individuals between the 97.5th and 99th percentiles of the permanent income distribution, and individuals at the top 1% of the permanent income distribution (those above the 99th percentile). Then, within each quantile, we calculate cross-sectional moments of the distribution of one- and five-year residual earnings growth, g_{it}^k. We repeat the same exercise within gender and age groups. The rest of the code collects the cross-sectional statistics and saves a set of CSV files in the folder Volatility.

5_Mobility.do

The 5_Mobility.do code calculates rank-rank measures of income mobility at different horizons (1, 3, 5, 10, and 15 years), across the population, and within gender and age groups. To calculate the rank-rank measure, we first select a given "jump" k and we keep only those individuals with non-missing alternative measure of permanent earnings, P_{it}^*, in periods t and $t+k$. Then, we sort individuals by their level of P_{it}^* and give each individual an index defined by its sorted position divided by the total number of observations. Denote tis continues ranking by c_{it}. This continuous measure is then divided in ${nquantilemob} quantiles. We further separate those individuals between the 97.5th and 99th percentiles, those between 99th and 99.9th percentiles, and those above the 99.9th percentile. Denote this discrete ranking by q_{it}. Then, we sort individuals with respect to P_{it+k}^* where $k \in \{1,3,5,10,15\}$ and we assign a continuous index. Finally, we calculate the average level of c_{it+k} for all individuals i whose ranking in period t is equal to $q_t = q_{it}$. The resulting rankings are then saved in CSV files in the folder Mobility.

6_Core_Figs.do and 7_PaperFigs.do

The _Core_Figs.do and 7_PaperFigs.do codes collect the CSV generated in previous codes and generate a large set of pdf-files. The 6_Core_Figs.do code has a large number of figures separated by team (inequality, mobility, and so on). Each section and each variable saves figures in different folders. The code 7_PaperFigs.do plots a selected set of figures that are meant to provide a baseline set of results for all papers in this issue.