Federal Reserve Bank of New York
Staff Reports

Anatomy of Lifetime Earnings Inequality: Heterogeneity in Job Ladder Risk vs. Human Capital

Fatih Karahan
Serdar Ozkan
Jae Song

Staff Report No. 908
December 2019

Anatomy of Lifetime Earnings Inequality:
Heterogeneity in Job Ladder Risk vs. Human Capital
Fatih Karahan, Serdar Ozkan, and Jae Song
Federal Reserve Bank of New York Staff Reports, no. 908
December 2019
JEL classification: E24, J24, J31

Abstract

We study the determinants of lifetime earnings (LE) inequality in the United States, for which differences in lifetime earnings growth are key. Using administrative data and focusing on the roles of job ladder dynamics and on-the-job learning, we document that 1) lower LE workers change jobs more often, mainly driven by higher nonemployment; 2) earnings growth for job stayers is similar at around 2 percent in the bottom two-thirds of the LE distribution, whereas for job switchers it rises with LE; and 3) top LE workers enjoy high earnings growth regardless of job switching. We estimate a job ladder model with on-the-job learning featuring ex ante heterogeneity in learning ability and job ladder risk—job loss, job finding, and contact rates. We find that learning ability differences explain almost all earnings growth heterogeneity above the median, whereas ex ante heterogeneity in job ladder risk accounts for 80 percent of LE growth differences below the median.

Key words: job ladder, inequality, Pareto tails

Karahan: Federal Reserve Bank of New York (email: fatih.karahan@ny.frb.org). Ozkan: University of Toronto (email: serdar.ozkan@utoronto.ca). Song: Social Security Administration (email: jae.song@ssa.gov). This paper was previously circulated under the titles “Sources of Earnings Growth Heterogeneity” and “Sources of Inequality in Earnings Growth over the Life Cycle.” The authors give special thanks to Gerald Ray and Pat Jonas for their help and support and to Samuel Kapon, with whom they started this project. They also thank Brendan Moore and Rachel Schuh for excellent research assistance. For insightful comments and discussions, the authors thank Fatih Guvenen, Chris Huckfeldt, Henry Hyatt, Gregor Jarosch, Rasmus Lentz, Jeremy Lise, Giuseppe Moscarini, Aysegül Sahin, Benjamin Schoefer, Chris Tonetti, Ludo Visschers, and David Wiczer. The views expressed in this paper are those of the authors and do not necessarily reflect the position of the Social Security Administration, the Federal Reserve Bank of New York, or the Federal Reserve System.

To view the authors’ disclosure statements, visit
https://www.newyorkfed.org/research/staff_reports/sr908.html.

1 Introduction

There are large differences in lifetime earnings (LE) among workers in the United States (Guvenen *et al.* 2017, 2014a). Even though inequality starts early in life, the striking differences in earnings growth over the life cycle are key for understanding the LE distribution. In this paper, we study these differences using administrative balanced panel data by focusing on the roles of two dimensions of heterogeneity: the ability to i) accumulate human capital and ii) climb the job ladder due to differences in job loss risk, job finding rate, and contacts from other potential employers.[1] Our goal is to quantify the importance of each of these factors by studying empirically and quantitatively the differences in the career paths of workers with different lifetime earnings.

In our empirical analysis, we use a confidential employer-employee matched panel of earnings histories of male workers between 1978 and 2013 from the U.S. Social Security Administration (SSA). Using a 10% sample of workers born between 1953 and 1960, we first compute workers' total labor income over the ages of 25 to 55, and rank them into 50 LE quantiles. Those at the 90th percentile (top 2%) of the LE distribution earn about 3.7 (14) times that of those at the 10th percentile. The inequality is much more pronounced at the top end of the LE distribution, which follows a power law with the top 0.1% (1%) accounting for around 29% of total LE among the top 1% (10%) of the population. Importantly, the vast majority of these differences is due to earnings growth heterogeneity. For example, top LE earners see their earnings rise by more than 17-fold between the ages of 25 and 55, median LE workers experience a two-fold increase, and those at the bottom see essentially no earnings growth.

To shed light on differences in the career paths of different LE groups we next document their job switching patterns. First, on average only 30% of the bottom LE workers stay with the same employer in two full consecutive years, compared to around 60% for workers above the median. Resonating with these large differences, people at the bottom of the LE distribution work for about 12 different employers between the ages of 25 and 55, more than twice as many as those at the top.

[1] There is a long line of literature studying the fanning out of inequality over the life cycle dating back to seminal papers by Mincer (1974), Heckman (1976), Deaton and Paxson (1994). Some explanations of wage growth heterogeneity include human capital accumulation à la Ben Porath (Huggett *et al.* 2011), learning about workers' ability (e.g., Jovanovic 1979; Pastorino 2019; Gibbons and Waldman 1999), workers selecting into positions via "tournaments" (Lazear and Rosen 1981), or workers sorting into jobs according to their comparative advantage (Lise *et al.* 2016). See Neal and Rosen (2000) for a comprehensive review of theories on earnings distribution.

Next, we investigate average earnings growth for job stayers and switchers across the LE distribution. We find that average growth for job stayers is surprisingly similar at around 2% in the bottom two thirds of the LE distribution, and increases steeply in the top tercile, reaching around 10% at the top quantile of the LE distribution. As for the average earnings growth of job switchers, we find much larger heterogeneity across the LE quantiles: It rises almost linearly from zero for the bottom LE quantile to around 4% for the 90th percentile, after which it accelerates to 10% for the top LE individuals. This large heterogeneity indicates that the nature of job switches is very different throughout the LE distribution. In fact, we argue that more than 35% of job switches are a result of a significant unemployment spell for bottom LE individuals, compared to only around 15% for the 90th percentile workers.

These facts imply that differences in lifetime earnings growth in the bottom half of the LE distribution are coming from growth differences of job switchers, whereas stayer growth differences should be the main culprit in the upper half, as high LE workers rarely switch employers. Building on this intuition, we develop a structural model to disentangle the economic forces that shape the distributions of the earnings growth of job stayers and switchers across the LE distribution. Specifically, we estimate a job ladder model with two-sided heterogeneity in the spirit of Cahuc *et al.* (2006) and Bagger *et al.* (2014). The model features learning on the job, on the job search, employer competition, and idiosyncratic shocks to worker productivity. We add to this framework a life-cycle structure in the form of perpetual youth. Importantly, we allow for rich worker heterogeneity in unemployment risk, the job finding rate and the contact rate for employed workers, as well as the ability to learn on the job. Finally, the model features recalls for unemployed workers by their last employers (Fujita and Moscarini 2017).

We estimate this model by targeting a rich set of moments from the SSA data. Specifically, we target the variance, skewness and kurtosis of the cross-sectional distributions of earnings changes conditional on age and LE, and separately for job stayers and switchers. We also target the fraction and wage growth of job stayers and switchers by LE groups and over the life cycle. We argue in Section 4.2 that these moments—which the model captures well—allow us to identify the importance of human capital and job ladder risk throughout the LE distribution. The key insight relies on realizing that the earnings changes for job switchers are very different than those of stayers: If job ladder risk is not important for wages, then one would observe that job stayers and switchers experience similar growth on average, driven largely by their returns to experience. By the same token, the differences in the distributions of wage changes between stayers and switchers over the LE distribution are informative about the nature of the job ladder risk faced by these groups.

One key finding from our estimation is the vast ex-ante heterogeneity in job ladder risk. We estimate a quarterly job loss risk of 9% for bottom LE workers, compared to 2% above the median. Similarly, job finding rates display large differences, ranging from around 30% at the bottom to 50% above the median LE. Given the annual nature of the SSA data, we cannot directly test these estimates. Instead, we use the Survey of Income and Program Participation (SIPP) to document large differences in job loss and finding rates among workers with different past earnings and over the life cycle—quantitatively consistent with the estimated model. Turning to the contact rate for employed workers, we find that bottom LE workers are contacted with 30% probability in a quarter versus 55% for the top. However, SIPP data show that high-earnings workers are *less* likely to make job-to-job transitions. Our model matches this feature of the data as well, because, despite getting more outside offers, high LE workers work for high-productivity firms on average and can rarely be poached. To directly test this mechanism, we analyze data from a special supplement to the Survey of Consumer Expectations (SCE), which collects information—among other things—on the contacts employed workers receive from other firms. We find that people with higher past earnings are contacted more frequently in the data, consistent with the estimates in the model.

Given that the estimated model provides a good account of the career trajectories of workers by LE groups, we use it to decompose the differences in lifetime earnings. First, we find that wage—rather than employment—differences explain the vast majority of LE inequality. The only exception is inequality at the bottom half, where employment differences also play some role because bottom LE workers work about 25% less than the median. Higher ex-ante job loss rate and a lower job finding rate for bottom LE workers explain almost all of these employment differences. Employment differences above the median are negligible in comparison.

Turning to differences in lifetime wages, we find them to be driven by wage growth over the life cycle, resonating with our empirical findings on earnings inequality and earnings growth. In a series of experiments, we isolate the relative roles of ex-ante differences in the job ladder risk and the returns to experience. Heterogeneity in unemployment risk accounts for about 50% of the wage growth differences between the bottom and the median. High unemployment rates among low LE workers reduce wage growth by preventing them from accumulating human capital and from climbing the job ladder with the former channel accounting for about 60% of the total effect. Differences in contact rates also have a nonnegligible effect on wage growth heterogeneity. Eliminating these differences closes an additional 20% of the wage growth gap between the bottom and the median by allowing low LE workers to move to better firms. Importantly, while the differences in unemployment risk and the contact rate are important at the bottom half of the LE distribution, they do not explain much of the heterogeneity above the median.

We find that learning ability is Pareto distributed and explains almost all earnings growth heterogeneity above median LE but only about 20% of it among the lower half. Along with Pareto-distributed firm productivities, the model is consistent with the Pareto tails of *within-age* earnings distributions with tail indices declining over the life cycle, which we document in the data. Yet the typical models of top income inequality deliver a Pareto distribution only in the entire population but not within age (see Gabaix *et al.* 2016; Gabaix 2009; Jones and Kim 2018).

A key conclusion of our study is that different economic forces are driving the inequality in different parts of the LE distribution. While bottom LE workers experience a low wage growth relative to the median throughout their working life primarily due to a poor labor market experience, workers at the upper half see a high wage growth primarily because they get very high returns to experience. These quantitative findings resonate with the patterns of average income growth of job stayers and switchers in the data. For workers who enjoy high wage growth regardless of job switching—the top LE group in the data—the model assigns a high returns to experience. If instead earnings growth is lower for a group, such as low LE workers, when they change employers compared to staying, the model attributes a bigger role to job ladder risk.

The rest of the paper is organized as follows. Section 2 presents the data and the stylized facts. Section 3 describes the model, Section 4 discusses its structural estimation, and Section 5 presents the estimation results. Section 6 provides the decomposition of lifetime earnings and Section 7 concludes.

2 Empirical Analysis

In this section, we document several stylized facts that motivate and guide our analysis of lifetime earnings inequality. Most of our analysis is based on administrative data from the Social Security Administration (SSA), but we also use data from the Survey of Income and Program Participation.

2.1 The MEF data

Our data are drawn from the Master Earnings File (MEF) of the U.S. Social Security Administration records. The MEF is the main source of earnings data for the SSA and contains information for every individual in the United States who was ever issued a Social Security number. Basic demographic variables available in the MEF are date of birth, place of birth, sex, and race. The earnings data are derived from the employee's W-2 forms, which U.S. employers have been legally required to send to the SSA since 1978. The measure of labor earnings is annual and includes all wages and salaries, bonuses, and exercised stock options as reported on the W-2 form (Box 1). The MEF has a small number of extremely (uncapped) high earnings observations. In each year, we winsorize observations above the 99.999th percentile in order to avoid potential problems with these outliers. We convert nominal earnings records into real values using the personal consumption expenditure deflator, taking 2005 as the base year. For detailed documentation of the MEF, see Panis *et al.* (2000) and Olsen and Hudson (2009).

W-2 forms contain another crucial piece of information for our purpose, an employer identification number (EIN), which identifies firms at the level at which they file their tax returns with the IRS. We use this variable to follow each worker's career path at an annual frequency. Note that an EIN is a different concept than an "establishment," which typically represents a single geographic facility of the firm. Two caveats are worth mentioning regarding the use of EINs to identify firms. First, an EIN is not always the same as the parent firm, because some large firms choose to file taxes at a level lower than the parent firm (see Song *et al.* 2018). Second, firms may change their EINs, for example, due to ownership changes (see Haltiwanger *et al.* 2014).

Sample selection and the construction of lifetime incomes We construct a 10% sample from the MEF based on the randomly assigned last four digits of (a confidential transformation of) the SSN. We select individuals born between 1953 and 1960, for whom we therefore have 31 years of data between ages 25 and 55 (referred to as a worker's lifetime). Furthermore, we work with a sample of *wage and salary* workers with a strong labor market attachment because the mechanisms we investigate speak to labor market participants. One drawback is that the MEF does not have direct measures of labor force participation. We address this problem by excluding individuals with earnings below a time-varying minimum earnings threshold $Y_{\min,t}$—25% of a full-year full-time salary at half the minimum wage, e.g. ≈\$1,885 in 2010—for i) at least one fourth of their working life, or ii) two or more consecutive years. These two criteria help us exclude early retirees, the disabled and those who are out of the labor force for other reasons.[2] We also drop workers that are self-employed (iii) for more than one eighth of their working life, or (iv) for two or more consecutive years.[3] These restrictions exclude workers who choose self-employment as their career path, and yet keep those who rely on self-employment income during unemployment spells, as well as payroll workers with a small self-employment income on the side. This procedure reduces our sample from 1,845,640 individuals to 840,194 for whom we have at least 31 years of earnings data.[4]

We compute lifetime earnings as the sum of individuals' W-2 earnings from ages 25 to 55. This measure is then used to assign workers into 50 equally sized lifetime earnings quantiles. We let LE_j for $j = 1, \ldots, 50$ denote the jth quantile of the LE distribution.

2.2 Stylized facts on lifetime earnings growth

We start by documenting lifetime earnings inequality (Figure 1a). Individuals around the 90th percentile (LE_{45}) earn 3.7 times as much as those around the 10th percentile (LE_5) over their working lives (Table I). This inequality is roughly half the cross-sectional earnings inequality in Guvenen *et al.* (2014b): The ratio of the 90th percentile to the 10th percentile of the annual earnings distribution hovered around 8 throughout our sample period. LE differences are relatively muted at the bottom, with LE_5 earning almost twice as much as the LE_1. Inequality is more pronounced at the top. LE_{50} earns almost 4 times as much as LE_{45}, and 13 times more than LE_5.

In fact, the upper tail of the LE distribution follows a power law; i.e., top inequality is fractal in nature: The top 1% (2%) accounts for around 29% of total lifetime income among the top 10% (20%) of the population, which is essentially identical to the share of the top 0.1% (0.2%) among the top 1% (2%) (bottom panel of Table I).[5] Importantly and interestingly, this power law also holds in the *cross-sectional* distribution of earnings

[2] Note that a nonemployment spell of at least two full *calendar* years implies a significantly longer nonemployment duration. Given the duration dependence of job finding rates in the literature (Jarosch and Pilossoph 2018), a worker with such a long nonemployment spell is unlikely to have been unemployed and looking for jobs the entire time.

[3] A worker is defined to be self-employed if he has self-employment income above the minimum earnings threshold $Y_{\min,t}$ and more than 10% of his annual total earnings.

[4] Clearly, our final sample is highly selective (see Table IV for a detailed breakdown of the sample selection). Appendix A.3 documents the key empirical findings for a much broader sample and finds them to be qualitatively similar to our baseline results.

[5] The log density and log inverse CDF of (log) lifetime earnings distribution on Figure A.1 show clearly that lifetime earnings have a Pareto tail with a slope of –2.13.

TABLE I – Selected inequality measures from the LE distribution

Ratio of lifetime earnings					Relative top income shares				
	LE	25	40	55		LE	25	40	55
LE_{50}/LE_1	25.5	3.3	23.8	47.8	S(0.1)/S(1)	0.29	0.24	0.31	0.38
LE_{50}/LE_5	14.2	2.1	12.9	25.0	S(0.2)/S(2)	0.29	0.24	0.30	0.38
LE_{50}/LE_{45}	3.8	1.1	3.5	5.2	S(1)/S(10)	0.29	0.23	0.30	0.37
LE_{45}/LE_5	3.7	2.0	3.7	4.8	S(2)/S(20)	0.28	0.22	0.30	0.36
LE_{38}/LE_{13}	1.9	1.4	1.9	2.2	Pareto index, ζ	2.2	2.6	2.0	1.7
LE_5/LE_1	1.8	1.6	1.8	1.9					

Notes: Left panel: First column shows the ratio of lifetime earnings of selected LE quantiles. The next three columns show the ratio of average annual earnings at different ages. Right panel shows the ratios of the share of top incomes for the lifetime earnings distribution and the cross-sectional earnings distribution at various ages. The bottom row reports the tail index ζ of each distribution, specified by the following Pareto CDF: $P[x > w]\ Cw^{-\zeta}$.

conditional on age. Moreover, as expected, earnings concentration at the top—measured as the relative earnings share of the top 0.1% to the top 1%—increases sharply over the life cycle from 0.23 at age 25 to 0.38 at age 55.[6] It is already established that the earnings distribution has Pareto tails (Piketty and Saez 2003; Atkinson *et al.* 2011), but to the best of our knowledge, ours is the first paper to document a power law within age.

It is well known that in the U.S. differences in earnings growth over the life cycle are key for understanding the inequality in lifetime earnings (see, for example, Haider 2001, Heathcote *et al.* 2005, Huggett *et al.* 2011, Kaplan 2012). This fanning out of earnings over the life cycle is also visible in the second to fourth columns of Table I. The ratio of average earnings of LE_{45} to LE_5 increases from 2.0 at age 25 to 3.7 at age 40 and reaches 4.8 by the age of 55. Again, the differences are much larger at the top of the LE distribution: The ratio of annual earnings of LE_{50} and LE_5 increases from 2.1 at age 25 to 12.9 at age 40 and reaches 25.0 by the age of 55.

To better illustrate this point, Figure 1 shows the log growth of average earnings between different ages over the LE distribution (see Guvenen *et al.* 2018 for a similar figure from a broader sample). We compute the log growth of average earnings between ages h_1 and h_2 ($\log \overline{Y}_{h_2,j} - \log \overline{Y}_{h_1,j}$) by differencing the average earnings across all workers in those LE and age cells. This growth measure allows us to include workers with zero

[6]Appendix A.2.1 documents the Pareto tails of the cross-sectional distribution of earnings at each age in more detail. In particular, Figure A.2 shows the relative income shares over the life cycle and Figures A.3 and A.4 show the log density and the inverse CDF of the right tail of the cross-sectional earnings distribution at different ages. Log density is linear in the tails at all ages, confirming the Pareto tail, and the slope gets closer to 1 in absolute value, which points to rising concentration at the top.

earnings. We use this measure whenever we refer to earnings growth. The results are similar qualitatively for the average of log earnings growth, which excludes observations less than the minimum income threshold. Earnings growth is positively related to the level of lifetime earnings, which is not surprising, since, all else the same, one should expect the higher growth individuals to rank at the top of the distribution. However, the quantitative magnitudes are striking: The top LE earners (LE_{50}) see their earnings rise by more than 17-fold between the ages of 25 and 55, median workers experience a two-fold increase, whereas those at the bottom see little to no earnings growth (around 16%). These large differences in earnings growth make an unmistakable contribution to the level of lifetime earnings inequality. In other words, while there are initial differences in the earnings of a cohort when they enter the labor market, the fanning out that occurs over the next 30 years or so is at least as important for lifetime inequality.

Some of this steep rise in earnings growth at the top could simply be due to transition from school to employment in the labor market. For example, top LE individuals might be pursuing graduate degrees around earlier ages. While the lack of education data does not allow us to answer this question directly, Figure 1 plots earnings growth between the ages of 30 and 55 and 35 and 55 when schooling is unlikely to matter much. While the magnitudes change, we still find a steep profile of earnings growth with respect to LE, suggesting that low labor supply at age 25 is not the major driver of these patterns.

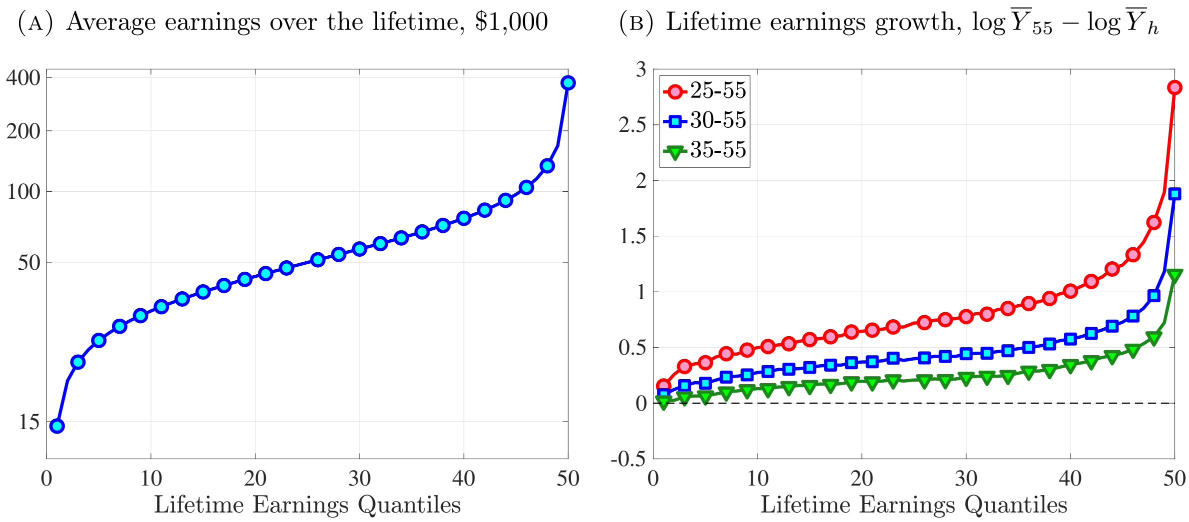
Striking differences exist even within the top LE group (top 2%), both in terms of the level of lifetime earnings and earnings growth over their career. For example, the top decile within this group (top 0.2% overall) averages annual earnings of over $\$1,000,000$, compared to $\$200,000$ for the bottom decile. Over the ages 25 to 55, annual earnings grow around 700% at the bottom decile, compared to more than 5000% at the top (Figure A.5). These features are striking but not surprising given that the income distribution is well characterized by a Pareto distribution (Piketty and Saez 2003).

2.3 Career paths by lifetime earnings

A natural immediate question is: What accounts for these large differences in earnings growth? To this end, we investigate the differences in labor market experiences between LE groups. Earlier work has shown that job mobility is important for earnings growth over the life cycle (Topel and Ward, 1992). Therefore, we start by investigating how the number of (distinct) employers over the working life differs between LE groups.

Individuals at the bottom of the LE distribution work for almost 5 different employers

FIGURE 1 – Heterogeneity in lifetime earnings growth

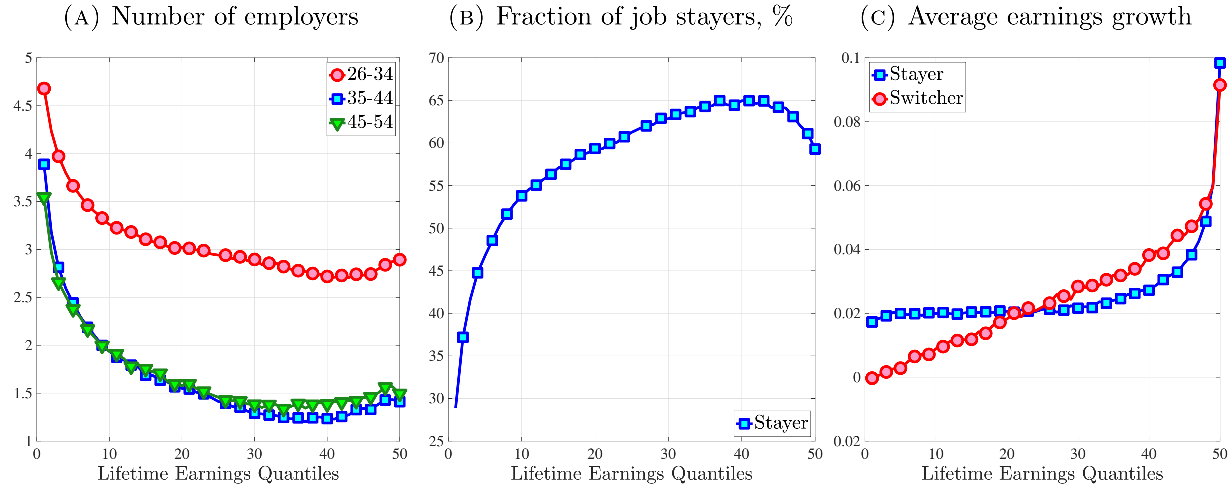


Notes: The left panel shows the average annual earnings over the life cycle by LE group. The right panel shows the log difference of average earnings $\bar{Y}$ between age 55 and various ages over the LE distribution. For clarity we use one marker for every other LE quantile.

on average over the first decade of their careers (ages 25–34, Figure 2a). This implies that around half of this group changes employers in any given year, or alternatively, a given worker changes employers on average once every other year. More interestingly, the number of unique employers drops sharply to around 2 until the median, and stays roughly constant in the upper half of the LE distribution. As workers age, job switching declines throughout the LE distribution. Interestingly, there is still a fair bit of job switching after age 35 at the bottom ranks of the LE distribution. While top workers work for around 1.5 different employers per decade after age 35, bottom workers still end up working for 3.5 employers on average, not much lower than the number of distinct employers in the first decade of their careers. At first glance, one might think that low LE individuals switch jobs very often and experience large earnings growth as a result. As we will see next, the nature of switches is very different across the LE groups.

We now document the average earnings growth across LE groups for workers who stay with the same employer and for those who change jobs. The SSA dataset contains a unique employer identification number (EIN) for each job that a worker holds in a given year. Given the annual frequency of the data, it is possible for a worker to have more than one W-2 in a given year. Moreover, some workers may hold multiple jobs concurrently. Given the lack of exact job spells, these issues pose a challenge for a precise classification of job stayers and switchers. There is more than one plausible definition for a job stayer,

FIGURE 2 – Job stayers and switchers



Notes: The left panel shows the number of distinct employers employers over the working life by LE. The middle panel shows the fraction of workers in each LE group who are job stayers according to our definition, calculated for each age and averaged over the working life. The right panel plots the log growth of average earnings $\bar{Y}$ between t and $t+1$ for job stayers and switchers separately by LE, again, averaged over t over the working life.

and we opt for a conservative one. Specifically, we call a worker a job stayer between years t and $t+1$ if i) he has income from the same employer in years $t-1$, t, $t+1$, and $t+2$; ii) his income in years t and $t+1$ is above the minimum income threshold for that year; and iii) this employer accounts for at least 90% of his total labor income in years t and $t+1$.[7] This definition ensures that the main employer was the same firm in years t and $t+1$. We label all other workers as job switchers. Note that according to this definition, switchers are a very heterogeneous group and consist of people who make direct job-to-job transitions, those who experience nonemployment, and those who come out of nonemployment. We return to this heterogeneity later.

The left panel of Figure 2 shows the fraction of job stayers within each LE group, averaged over the working life. Resonating with the large differences in the number of different jobs over the life cycle, there is similarly a large heterogeneity in the likelihood of staying with the same firm: Bottom LE individuals stay with the same firm on average for 30% of their working life, compared to around 60% above the median. Thus, individuals at higher LE quantiles are more likely to stay with the same firm.

How much of an earnings growth does a worker experience when he stays with the

[7]We find similar results when we impose the condition that the main employer accounts for at least 50% of the total income. Results are available upon request.

same employer versus when he switches jobs? The answer differs widely across the LE distribution (Figure 2c). For job *stayers* log average earnings growth (between t and $t+1$) is surprisingly similar (at around 2%) in the bottom two thirds of the LE distribution. This implies that any worker below LE_{33} experiences an annual earnings growth of 2% on average when he works for the same employer. Average earnings growth for stayers increases sharply from LE_{33} onward, reaching around 10% at LE_{50}.

Turning to job *switchers*, we find that average annual earnings growth is essentially zero at the bottom of the LE distribution and rises almost linearly to around 4% for LE_{45}, after which it accelerates to 10% for the top LE individuals. This large heterogeneity indicates that the nature of job switches is very different throughout the LE distribution. As we argue later, this is a critical aspect of the data for understanding the different forces behind the earnings growth of job stayers and switchers—and eventually the differences in lifetime earnings growth. For example, given the little heterogeneity among job stayers below the median LE, it is clear that the pronounced differences in lifetime earnings growth below the median are due to the differences in the frequency and nature of job switches. We now investigate these differences.

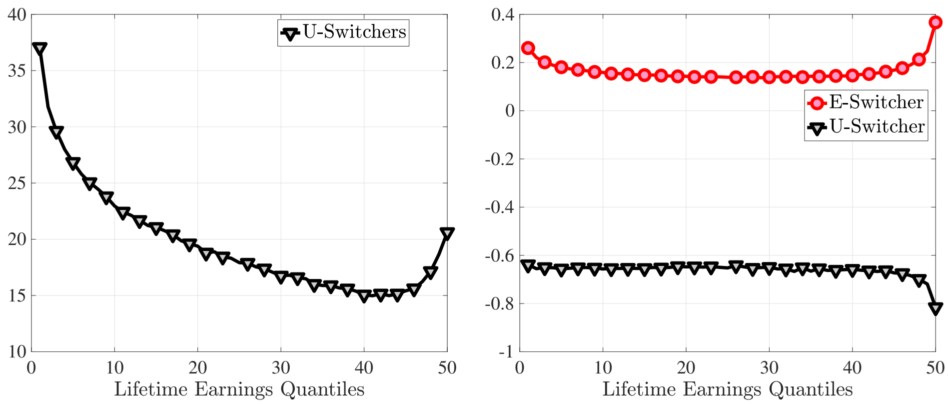
As we discussed before, job switchers are a very heterogeneous group as they include workers who switch jobs directly or due to a job loss (or a quit). The annual nature of the data does not allow us to separate these directly. Yet, we argue that the earnings growth distribution of switchers are informative about the nature of switches. For example, switchers who see their earnings decline by more than 25% have most likely experienced some nonemployment spell in $t+1$. Thus, we classify such workers as "U-switchers," and the remaining job switchers as "E-switchers." The latter contains workers that make direct job switches as well as those coming out of nonemployment in $t+1$.[8]

More than 35% of job switches are U-switches for bottom LE workers (Figure 3a). This share declines sharply over the LE distribution and reaches a low of 15% for LE_{40}, before increasing to 20% for top LE workers. Thus, on average, higher LE individuals are more likely to make job switches involving earnings increases. Investigating the average earnings growth associated with each type of switch, we find large differences between E- and U-switches, but little variation across the LE distribution (except for the bottom

[8] Jolivet *et al.* (2006) show that a sizable portion of direct job-to-job transitions indeed involve wage cuts. Sorkin (2018) argues that some of these differences can be traced to amenity differences across firms. Tanaka *et al.* (2019) link the earnings declines from direct job switchers to labor force dynamics at both the origin and destination firms.

FIGURE 3 – E-switchers and U-switchers

(A) Share of U-switchers, %　　(B) Earnings growth, $\log \overline{Y}_{t+1} - \log \overline{Y}_t$



Notes: The left panel shows the share of U-switchers among job stayers in each LE group averaged over the life cycle. The right panel plots the log growth of average earnings $\bar{Y}$ between t and $t+1$ for U-switchers and E-switchers by LE.

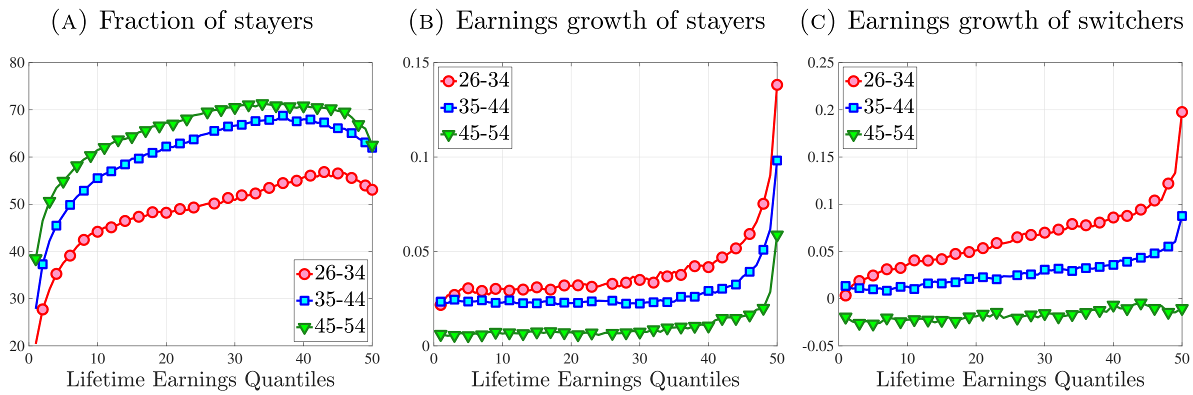
and the top end). On average, an E-switch is associated with an earnings increase of larger than 15%, whereas a U-switch is associated with a decline of more than 60%.

The annual nature of our data limits the analysis of the earnings changes of job switchers. If a worker becomes unemployed some time in year t or $t+1$, then his earnings in $t+1$ may reflect earnings from a short-term job in that year.[9] To alleviate this problem, we construct (normalized/average) earnings growth between the years when a worker is full-year employed in the same firm before and after the switch. Our substantive conclusions hold when we analyze earnings growth over a longer horizon (Figure A.9).

Life-cycle variation There is significant age variation in job switching and earnings growth patterns that have been extensively documented before. The key advantage of our data over existing work is to allow us to investigate differences in these life-cycle profiles between LE groups. Figure 4 plots the fraction of stayers and the earnings growth of job stayers and switchers for three stages of the working life. Several remarks are in order. First, the fraction of workers who stay with the same firm increases and is concave. This increase is consistent with declining unemployment risk and job mobility documented before (see Topel and Ward 1992; Jung and Kuhn 2016). Interestingly, this profile is shared by all LE groups, though the concavity is more pronounced above the median. Turning to the average earnings growth of job stayers, we find a flat profile below LE_{30}

[9] Our approach throughout the paper to dealing with such issues is using the estimated model where we aggregate simulated quarterly earnings to annual, and construct moments in a similar fashion.

FIGURE 4 – Life cycle variation



at all ages. Moreover, consistent with the existing literature, the rate of earnings growth declines with age.[10] Finally, we investigate the earnings growth of job switchers over the life cycle (Figure 4c). Similar to that for job stayers, the average earnings growth for job switchers is highest at younger ages (ages 26–34). This growth rate declines sharply over the life cycle, especially for higher LE individuals. The earnings growth of stayers is negative for older individuals (aged 45–54) throughout the LE distribution.

Evidence from SIPP data Our definition of U- and E-switches is based on a somewhat arbitrary cutoff, and the annual nature of the SSA data does not allow us to come up with a more precise definition. To provide additional evidence for the heterogeneity in job loss rates, we turn to data from SIPP. SIPP is a nationally representative sample of households. The data consist of monthly observations in overlapping panels with length between 2.5 and 4 years, with the first panel conducted in 1984. Each panel is conducted in waves, interviewing households every four months about the prior four months. Job loss, job finding, and job-to-job transition rates, can be computed at a high frequency.

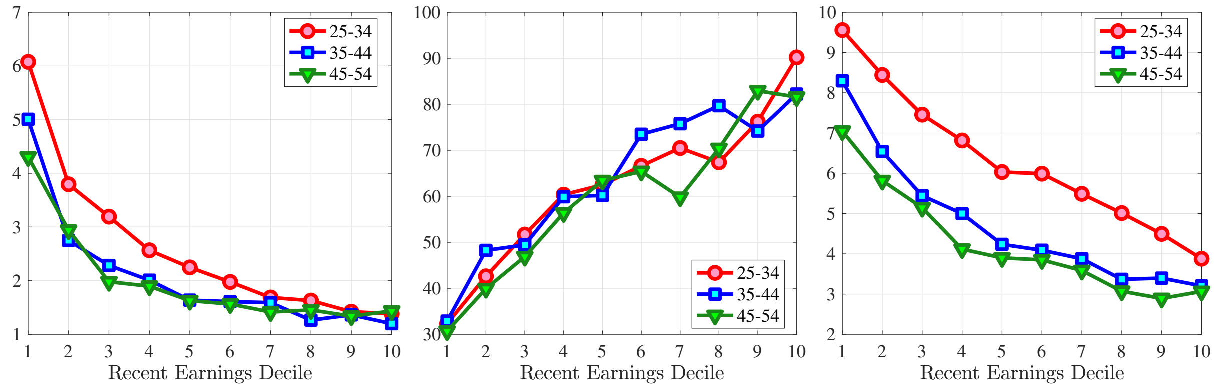
We select a sample of males (ages 25–55) with some labor force attachment (further details are in Appendix B). We use the panel structure to rank workers into 10 equally sized deciles within each age group (25–34, 35–44 and 45–55) based on their recent earnings (RE) over the past three years. Next, we compute the job loss (EU), job finding (UE), and job-to-job (EE) transition rates for each group over the next four months.

Job loss rates show significant heterogeneity across previous recent earnings deciles for all age groups (Figure 5). For example, unemployment risk at the bottom decile can

[10] This feature is not specific to job stayers; a large body of work finds the life-cycle profile of earnings to be concave and hump shaped when individuals up to age 65 are included (e.g, Heckman (1976)).

FIGURE 5 – Labor market flow rates across the income distribution

(A) Unemployment Risk, %　(B) Job Finding Rate, %　(C) Job-to-job Transition Rate, %



be almost five times that of the top decile. There are also marked differences over the life cycle, with young workers much more exposed to unemployment than older ones. Our finding indicates that workers with low wages are much less likely to find stable jobs, and would arguably therefore not be able to move to better jobs, as that requires clinging on to the job ladder. Earlier literature has emphasized the life-cycle variation in unemployment risk (Jung and Kuhn 2016; Shimer 2001). We find that between-RE variation in job loss rates is an order of magnitude larger. The middle panel shows that the four-month job finding probabilities (UE rates) are strongly increasing with the level of past earnings. This rate is around 30% for young workers (25-34) with low earnings, and increases monotonically up to 90%. Moving to job-to-job transition rates over a four-month period (right panel), we find that these are as high as 10% for young workers with low earnings, decline with recent earnings, and are about 4% for the top decile.

We have documented several facts regarding the careers of individuals who end up in different parts of the LE distribution. While these facts are useful for describing the various components of earnings growth heterogeneity, they do not suffice to provide a structural interpretation of the underlying sources. In what follows, we introduce a quantitative model of wages and job turnover with heterogeneity in returns to experience and job ladder risk. Namely, a structural model will allow us to disentangle the various economic forces that shape the distribution of wage changes of job stayers and switchers.

3 Model

We build on Bagger *et al.* (2014) as it features a tractable framework to study the role of job search and learning on the job in generating wage growth. Despite endogenously generating some age variation in job mobility and earnings dynamics, this model falls short of capturing the magnitudes in the data. Thus, we incorporate a life-cycle structure to this framework as perpetual youth à la Blanchard (1985) and Yaari (1965).

3.1 Environment

The economy is populated by heterogeneous workers and firms that produce a single consumption good sold in a competitive market. Workers can be employed or unemployed, and search for jobs in a frictional labor market, both on and off the job. They start life as young (y) and become old (o) with probability γ. They have preferences with log per-period utility over consumption, and discount future periods at rate ρ:

$$U(\{c_t\}_{t=0}^{\infty}) = \sum_{t=0}^{\infty} \left(\frac{1}{1-\rho}\right)^t \log c_t.$$

There is no inter-temporal savings technology that allows workers to smooth their consumption. This assumption along with the log preferences greatly simplifies computation.

Worker productivity Each worker enters the labor market with no experience and accumulates human capital as he gains actual experience from employment. The human capital of worker i in period t is given by

$$\begin{aligned} h_{it} &= \tilde{h}_{it} + \varepsilon_{it}, \quad \tilde{h}_{i0} = \alpha_i \\ \tilde{h}_{it} &= \begin{cases} \tilde{h}_{it-1} - \varsigma & \text{if unemployed} \\ \tilde{h}_{it-1} + \beta_i + \zeta\left(\tau_{it}^2 - (\tau_{it}-1)^2\right) & \text{if employed} \end{cases} \end{aligned} \tag{1}$$

Here, $\tilde{h}_{it}$ denotes the nonstochastic component of human capital. Its level at the beginning of the working life, $\tilde{h}_{i0}$, is determined by the worker's type α_i, which reflects permanent heterogeneity in productivities due to differences in initial conditions such as innate ability and education, and can be thought of as a worker's type. Human capital accumulates as the worker gains actual experience τ_{it} through employment. The rate of human capital accumulation has a worker-specific linear component β_i, potentially correlated with α_i, and a common quadratic component ζ.[11] In Huggett *et al.* (2011) individual-specific growth rates of human capital arise as a result of different investment choices due to the heterogeneity in productivities in the production of human capital.

[11] We also tried a version with individual-specific ζ and did not find significant heterogeneity.

Our model captures this heterogeneity through exogenous differences in returns to experience. When a worker is unemployed, his human capital $\tilde{h}_{it}$ depreciates at a constant rate ς. Finally, ϵ_{it} is an idiosyncratic shock to worker productivity whose distribution depends on worker type α_i and age, and captures the residual sources of variation not modeled in our framework. We specify the process for ϵ_{it} in Section 4 in detail.

Firm distribution and production technology The workers who meet a new firm draw from a productivity distribution CDF $F(p)$ with a support of $\left[\underline{p}, \infty\right]$ common to all workers. A worker with human capital h_{it}, who works for a firm with permanent productivity $p_{j(i,t)}$, produces a homogeneous good according to a log-linear production function: The log-output y_t is given by $y_{it} = p_{j(i,t)} + h_{it}$.

3.1.1 Heterogeneity in search and matching

Unemployment risk A job dissolves exogenously with probability $\delta^a(\alpha_i)$, in which case the worker searches for a job. We model separation rates to be heterogeneous across workers of different types and ages. This heterogeneity is needed to capture the declining unemployment risk by the wage and age of workers discussed in Section 2.

Job finding rate An unemployed worker of age $a \in \{y, o\}$ with permanent ability α_i meets a firm with probability $\lambda_0^a(\alpha_i)$, which captures ex-ante heterogeneity in job finding rates. This heterogeneity is motivated by our findings from the SSA and SIPP data and are potentially important for wage growth over the life cycle, as workers with a high job finding rate will work for more years, end up accumulating more human capital and, on average, work for more productive firms. To account for the sources of earnings growth, we explicitly model the differences in job finding rates. Workers who are hit by separation shocks find a job immediately with probability $\xi\lambda_0^a(\alpha_i)$. As we discuss later, our model period is a quarter, and a nonnegligible fraction of laid-off workers find a job within three months (Abrahám *et al.* 2016). Moreover, there is evidence of transitions that look like direct job-to-job switches but are actually involuntary (Jolivet *et al.* 2006). Thus, we allow for the possibility of finding a job within the same period.

Search on the job While employed, workers search for better jobs and with probability $\lambda_1^a(\alpha_i)$ receive an outside offer from another employer, whose productivity is drawn from the distribution $F(p)$, triggering a renegotiation between two firms that we explain below. As Figures 3a and 5c have shown, workers differ in the types and rates of job switches. Our framework can generate qualitatively similar patterns without explicit differences in the contact rates: High-wage workers—employed on average by more productive firms—are less likely to get an offer that beats their current employer. This reduces their job-to-job transition rate even if they receive counteroffers at the same rate as low-wage workers. Similarly, as workers get older, they settle into higher paying jobs and are less likely to move. However, our estimation shows that this endogenous mechanism is insufficient to explain the quantitative differences in the data.

Timing of events At the beginning of each period, the productivity shocks are drawn and workers' human capital is updated according to equation 1. Next, output is produced and wages are paid. There is no inter-temporal savings device, so workers consume their wages. At the end of the period, search and matching shocks are realized: Unemployed workers who find jobs negotiate their wage, workers who receive an outside offer renegotiate their wages or switch employers, and employed workers that draw separation shocks become unemployed. They may find a job immediately or have to wait for the next period to search. Aging occurs stochastically at the end of the period with probability γ and is mutually exclusive from the labor market shocks.

3.2 Wage determination

In this section, we briefly explain the wage bargaining protocol. Since the framework is already discussed in Bagger *et al.* (2014), we focus on the key equations and how the life-cycle structure affects them. Analytical derivations are provided in Appendix C.

Wages are specified as piece-rate contracts. In particular, if a worker with human capital h works for a firm of productivity p at a piece rate of $R = e^r \leq 1$, he receives a log wage w of $w = r + p + h$. Here R, the contractual piece rate, is determined endogenously. Upon meeting with a firm, the worker bargains over this piece rate R, which is not updated until the worker meets with another firm.

We now describe how this piece rate is determined for workers with different labor market states. First, let's define $\mathbb{I}_i \equiv \{\alpha_i, \beta_i\}$ as the vector of individual-specific state variables capturing ex-ante (fixed) heterogeneity. Note that as we discussed above, $\mathbb{I}_i$ pins down the individual-specific worker flow rates as well as the firm distribution, i.e., $\{\delta^y(\alpha_i), \delta^o(\alpha_i), \lambda_0^y(\alpha_i), \lambda_0^o(\alpha_i), \lambda_1^y(\alpha_i), \lambda_1^o(\alpha_i)\}$. The value functions introduced below are individual specific and thus a function of $\mathbb{I}_i$ in addition to other state variables.

Hires from unemployment Let $V_0^a(h; \mathbb{I}_i)$ and $V^a(r, h, p; \mathbb{I}_i)$ denote the expected lifetime utility of an unemployed worker i with human capital h at age a, and when he is employed at a firm with productivity p at piece rate e^r, $r < 0$, respectively. We define $V^a(r, h, p; \mathbb{I}_i)$ below and assume that the value of unemployment is equivalent to employment in the least productive firm of type $p_{\min}$ extracting the entire match surplus, i.e., $V_0^a(h; \mathbb{I}_i) = V^a(0, h, p_{\min}; \mathbb{I}_i)$. This assumption implies that an unemployed worker accepts any job offer and simplifies the problem.[12]

The wage bargaining protocol dictates that unemployed workers receive θ share of the expected match surplus, where θ captures the worker's bargaining power.[13] More specifically, the piece rate of a hire from unemployment, r_0, solves

$$\mathbb{E}V^a(r_0, h', p; \mathbb{I}_i) = V_0^a(h; \mathbb{I}_i) + \theta\mathbb{E}\left[V^a(0, h', p; \mathbb{I}_i) - V_0^a(h; \mathbb{I}_i)\right]. \tag{2}$$

The worker's surplus from the match is the increase in expected lifetime utility from unemployment to a state where he is paid his entire output ($r = 0$). Thus, when an unemployed is hired, the firm offers a piece rate that increases his expected lifetime utility by θ share of this surplus. In equation (2), the expectation is with respect to ϵ_{t+1}.

Poaching When a worker is contacted by a firm with productivity p', the incumbent firm and the poacher compete. The more productive firm outbids the less productive one and hires the worker. We now discuss the wage that arises as a result of this competition.

There are several cases to consider. First, suppose that the poacher has higher productivity; $p' > p$. Then, the poacher hires the worker by paying a piece rate r' that increases the worker's value by θ–share of the surplus generated by the match:

$$\mathbb{E}V^a(r', h', p'; \mathbb{I}_i) = \mathbb{E}\left\{V^a(0, h', p; \mathbb{I}_i) + \theta\left[V^a(0, h', p'; \mathbb{I}_i) - V^a(0, h', p; \mathbb{I}_i)\right]\right\}. \tag{3}$$

Note that as Postel-Vinay and Robin (2002) have shown, job switches may result in workers accepting wage losses, as they anticipate faster wage growth in higher productivity firms. Wage losses upon job switches are a prominent feature of the data.

Second, let's consider a case in which the poacher has lower productivity than the current employer. Bertrand competition implies that the incumbent firm retains the worker, possibly by adjusting the worker's piece rate. This new piece rate offers the worker the maximum value he could attain working at firm p', i.e., the value associated

[12] This assumption is typical in this class of models and is justified by the high empirical job acceptance rate of the unemployed (Van den Berg 1990).

[13] Even though this assumption is not a direct outcome of a Nash bargaining solution, Bagger *et al.* (2014) and Cahuc *et al.* (2006) argue that this protocol can be micro-founded as the equilibrium of a strategic bargaining game adapted from Rubinstein (1982).

with $r = 0$ ($R = 1$), and a θ–share of the additional surplus generated by the offer. In this case, the new piece rate r' solves the following equation:

$$\mathbb{E}V^a(r', h', p; \mathbb{I}_i) = \mathbb{E}\{V^a(0, h', p'; \mathbb{I}_i) + \theta[V^a(0, h', p; \mathbb{I}_i) - V^a(0, h', p'; \mathbb{I}_i)]\}. \quad (4)$$

Note that in contrast to other models of on-the-job search such as Burdett and Mortensen (1998a) and Hubmer (2018), this model generates potentially large and leptokurtic increases in wages for job stayers, which is prevalent in the data (Guvenen *et al.* 2018).

In some cases, the productivity of the poacher may be so low that the new offer does not generate any additional surplus and therefore does not trigger a change in the piece rate. Then, the worker discards the offer. Let $q^a(r, h, p; \mathbb{I}_i)$ denote this threshold firm productivity such that offers from firms with $p' \leq q^a(r, h, p; \mathbb{I}_i)$ are discarded. q^a solves

$$\mathbb{E}V^a(r, h', p; \mathbb{I}_i) = \mathbb{E}\{V^a(0.h', q^a; \mathbb{I}_i) + \theta[V^a(0, h', p; \mathbb{I}_i) - V^a(0, h', q^a; \mathbb{I}_i)]\}. \quad (5)$$

Value functions Our model adds perpetual youth to Bagger *et al.* (2014). Therefore the value functions governing the problems of the young and old workers differ only due to stochastic aging. Here, we do not provide the equations and the detailed derivation of the wage equations, and show them in Appendix C instead.

4 Estimation

We now use this model to estimate the contributions of the heterogeneity in the worker flow rates and the ability to accumulate human capital to the differences in earnings growth over the life cycle. To this end, we first exogenously set four parameters: The quarterly discount rate ρ is set to 0.005 to match the annual rate of 2%; workers' bargaining power θ is set to 0.4 following Bagger *et al.* (2014); the quarterly aging probability γ is set to 1/60 so that a worker becomes old on average in 15 years; and the reallocation probability ξ is set to 0.4 following Abrahám *et al.* (2016).

We estimate the remaining parameters using the simulated method of moments (SMM). We simulate quarterly data of 100,000 individuals and create a model-based matched employer-employee panel mimicking the SSA sample, which is then used for computing the model counterparts of our targets.[14] Each individual starts unemployed

[14] In the earlier literature, this class of models is typically estimated using higher frequency data (e.g., Bagger *et al.* 2014 and Cahuc *et al.* 2006). Such data with a long enough panel to construct lifetime incomes are not available for the U.S.

at the age of 23 and remains in the labor force until 55. We discard the first two years and use ages 25 to 55. We aggregate quarterly data to annual observations. Importantly, we subject the model to the same sample selection criteria used to construct our SSA sample and compute the model counterparts of targeted moments.

Before turning to our targets, we discuss one feature that we add to the model to obtain a better fit to the data. In the data, there are many job stayers who experience large declines in annual earnings. While idiosyncratic shocks to worker productivity could in principle account for these large losses, we found them to be quantitatively insufficient. To give the model a chance to match this feature, we add a recall option of unemployed workers by their last employers (Fujita and Moscarini 2017). We do this in a very simple way: When an unemployed worker receives a new job offer, with probability λ_r this offer comes from the workers' last employer. As Fujita and Moscarini (2017) show, the recall option changes the wages of recalled workers as it affects the value of a job to a worker. However, we assume that when unemployed workers negotiate with a potential firm, both the worker and the firm ignore the worker's option of a recall in case of future unemployment in the bargaining. We make this assumption in order to keep the estimation computationally feasible because solving for the wage equation analytically is not possible when the recall option is recognized during bargaining.

4.1 Targeted moments

We target five sets of moments. The first two are about the cross-sectional distribution of earnings changes for job stayers and switchers. The third and fourth have to do with the fraction of job stayers, E–switchers, and U–switchers and their average annual earnings growth, respectively. Finally, we target average earnings at age 25 by LE group. We choose to not target the heterogeneity in lifetime income growth. As we argue in the next section, the model is already identified using these five sets of moments.

Grouping workers We condition each targeted moment on LE and age groups. Specifically, we calculate workers' lifetime earnings as in Section 2 and assign them into 12 percentile groups: 1–4, 5–10, 11–20, ..., 81–90, 91–96, 97–100. Furthermore, to capture the life-cycle variation we group workers into three age groups: 25–34, 35–44, 45–54.

Cross-sectional moments of earnings growth As documented in Guvenen *et al.* (2018), earnings changes are highly leptokurtic and left skewed. This shape of the earnings change distribution is broadly consistent with job ladder models: Most workers see little change but a small share experience a large swing due to unemployment, a job-to-job transition or an outside offer, which in turn may lead to a left-skewed and leptokurtic distribution. Hubmer (2018) shows that a job search model as in Burdett and Mortensen (1998b) can generate a plausible distribution of earnings changes as well as how that distribution varies between income groups.

Based on these insights, we target the mean, standard deviation, skewness, and kurtosis of annual earnings changes for job stayers and switchers separately.[15] Rather than conditioning workers based on their earnings over the past five years as in Guvenen *et al.* (2018), we condition them based on their lifetime earnings. We find that the variation over past earnings is qualitatively similar to the variation over lifetime earnings.

An issue when computing growth rates is dealing with zero earnings. Recall that in our sample, we drop workers with two or more consecutive years of zero earnings. However, there are still observations with no income in a given year. We would like to keep them as they contain information about the importance of search frictions. For this purpose, we use the arc percent growth measure defined as $2(Y_{t+1} - Y_t)/(Y_{t+1} - Y_t)$, where Y_t is annual earnings. Targeted cross-sectional moments are shown in Figure D.2.

Average income growth moments Next, we target the fraction and average income growth of job stayers, E–switchers, and U–switchers by three age and 12 LE groups. The details of how these moments are constructed are discussed in Section 2.3. Figures D.1 and D.3 show these moments by age and the targeted LE groups.

Average earnings at age 25 Finally, we target the average real earnings (in 2010 dollars) by LE group at age 25. This moment of the data is shown in Figure D.4.

4.2 Identification

Below we provide an informal discussion of identification of our model. We acknowledge that when SMM is employed, all parameters are determined jointly within the estimation as most parameters affect more than one aspect of the data. In this section, our goal is to show that each feature of the model has a pronounced effect on at least one unique moment targeted in the estimation. Namely, there is at least one unique feature of the data that informs each ingredient of the model. This identification discussion also justifies the selected targeted moments presented in the previous section.

[15] An alternative is to target percentiles or percentile-based moments such as the 90-10 differential, Kelley's skewness, and Moors' kurtosis, which we have experimented with and found similar results. We target centralized moments as they are less costly to compute.

Ex-ante worker productivity (α, β) The concave average life-cycle profile of earnings growth is informative about the average experience profile of worker productivity, driven in the model by the mean of the joint (α, β) distribution and the common quadratic term ζ. The differences in the initial earnings levels of LE groups and their stayer earnings growth (Figure 2c) help us pin down the variance-covariance matrix of the joint distribution of α and β. Note that the distribution of firm productivities also has a first-order effect on the initial earnings dispersion as well as on the earnings growth of job stayers through outside offers. As we discuss next, we use other features of the data to identify the distribution of firm productivities.

Firm productivity distribution In the estimation of job ladder models, identifying the distribution of firm productivities is a key challenge. There are several approaches to estimate this distribution using matched employer-employee data.[16] For example, Postel-Vinay and Robin (2002), Cahuc *et al.* (2006) and Bagger *et al.* (2014) use data on firms' value added or profitability to back out the firm distribution. We cannot implement this method as our dataset doesn't contain any direct information on value added or profitability. Barlevy (2008) shows that under appropriate conditions the wage gains of job switchers could identify the offer distribution nonparametrically, even in the presence of unobserved worker heterogeneity. Bagger and Lentz (2014) use poaching patterns between firms to rank firms with respect to their productivity. More recently, Bonhomme *et al.* (2017) develop a new approach to classify firms into discrete groups using a k-means algorithm. We follow a different approach based on the differences between average earnings growth for job stayers and switchers over the LE distribution.

The key insight for identifying the firm productivity distribution relies on realizing that the earnings growth for job switchers is very different than that of stayers, with stayer growth exhibiting relatively little heterogeneity at the bottom two thirds of the LE distribution and switchers showing much larger differences throughout the LE distribution (Figure 2c). If there was no job ladder to be climbed (i.e., the firm distribution was degenerate), then the average earnings growth of switchers and stayers would look very similar as they would both be mainly driven by the differences in β. Job ladder dynamics through the shape of the firm distribution, on the other hand, help the model

[16] Other papers have also used only worker-side data to estimate such models by relying on the distribution of wages coming out of unemployment to identify the wage offer distribution; e.g., Bontemps *et al.* (1999) and Lise (2013). This approach is not reliable in an environment with worker heterogeneity as shown in Barlevy (2008).

generate a *different* profile of earnings growth for stayers and switchers. We confirm this insight in our simulations.

Heterogeneity in worker flow rates ($\delta^a(\alpha), \lambda_0^a(\alpha), \lambda_1^a(\alpha), \lambda_r$) Our strategy relies on identifying these flow rates separately for each LE and age group and then linking the LE groups to ex-ante worker type α.

U–switches, those that involve a larger than 25% earnings loss (Figure 3b), are intimately linked to the job loss rate δ. Moreover, their frequency is not affected by the rate of job-to-job transitions, because such transitions result in either wage increases or wage losses smaller than 25%, and are therefore counted among E-switches.[17]

Turning to the job finding rate λ_0, this rate determines how long a given unemployment spell lasts. Therefore, it has a pronounced effect on the average earnings loss of U–switchers along with the possible wage decline associated with falling off the job ladder. The latter is determined by the shape of the firm distribution, whose empirical underpinning is discussed above.

Finally, the stayer probability is given by a combination of the job loss rate δ and the offer arrival rate for the employed λ_1 as well as the recall rate λ_r. The key feature that identifies the recall rate is the left skewness of earnings growth for job stayers. In the model, stayer growth distribution is dramatically *right skewed* in the absence of recalls. Having already identified δ and λ_r, stayer probability can now be used to pin down λ_1.

Idiosyncratic shocks (ϵ) These shocks are residuals of earnings growth not explained by the structural features of the model. Our simulations show that the structural features of the model can explain well the earnings distribution of job switchers. Thus, we use the higher-order moments of the distribution of earnings changes for job stayers to identify the parameters of idiosyncratic shocks.

Age dependence in parameters Key moments identifying the flow rates and the distribution of idiosyncratic shocks have strong age variation in the data, as we have shown in Section 2. Therefore, the age dependence in these parameters is identified from the age variation in targeted moments.

4.3 Estimation methodology

In this section we first explain the functional form assumptions concerning the worker and firm distributions as well as the flow rates. While our identification strategy does

[17]In our simulations, less than 0.2% of direct job-to-job switches lead to a wage cut larger than 25%.

not require specific functional forms, these assumptions allow us to have more statistical power and keep the estimation computationally feasible. Next, we describe the SMM objective function along with the computational method used for estimation.

Functional forms The worker fixed-effect α is normally distributed with mean μ_α and standard deviation σ_α. β is Pareto distributed with shape and scale parameters χ_w and ψ_w, respectively, and is correlated with α with the correlation coefficient $\rho_{\alpha\beta}$.[18] The Pareto tail of β captures the "high-growth" worker types who experience a much higher earnings growth than other individuals. Polachek *et al.* (2015) estimate individual-specific learning abilities in a human capital production function and find this distribution to be fat tailed. Gabaix *et al.* (2016) argue that these workers, as opposed to a random growth mechanism, are key for explaining the rising top income inequality.

We model the heterogeneity in worker flow rates as a function of worker type α and age. In particular, we use a cubic spline to model unemployment risk, the job finding rate, and the contact rate as a function of $\alpha_i - \mu_\alpha$ for each age group. We experimented with the number of points for each flow rate and concluded that three points for each age group was flexible enough for job finding and contact rates, whereas unemployment risk required 5 points for each age to fit the heterogeneity in the data.

Finally, we assume that firm productivity is Pareto distributed with shape and scale parameters χ_f and ψ_f, respectively.[19] We normalize the scale parameter ψ_f to 1, as one cannot separately identify ψ_f and the mean of the α distribution.

We assume that the idiosyncratic shocks hit once a year with some probability $\pi(\alpha)$ when workers stay with the same employer (every four periods in the model). They are modeled as an i.i.d process with innovations drawn from a Gaussian distribution with standard deviation σ_ε. $\pi(\alpha)$ is modeled as a cubic spline separately for each age group.[20]

[18] We also estimated a version of our model with Gaussian β and have found that a fat-tailed distribution such as Pareto helps the model better match the very large earnings growth of top LE groups relative to the median. We revisit this choice later in the context of estimation results in Section 5.

[19] We have experimented with log-normally distributed firm productivity and found that a Pareto fits the data better. Hubmer (2018) uses a different search model along the lines of Burdett and Mortensen (1998b) and reaches a similar conclusion.

[20] We have also experimented with alternative specifications where switchers also experience idiosyncratic shocks but we have found that the earnings dynamics of switchers—specifically, the second-to-fourth order moments of their annual earnings growth—are well captured by the endogenous mechanisms in the model. When we model the idiosyncratic shocks to be an AR(1) process we have found the persistence to be low around 0.50.

SMM objective function Let d_n for $n = 1, ..., N$ denote a generic empirical moment, and let $m_n(\theta)$ be the corresponding model moment that is simulated for a given vector of model parameters, θ. We minimize the sum of squared deviation between the data and the simulated moment.[21] Our SMM estimator is defined by the following:

$$\hat{\theta} = \arg\min_{\theta} \boldsymbol{F}(\theta)' W \boldsymbol{F}(\theta), \quad \boldsymbol{F}(\theta) = [F_1(\theta), ..., F_N(\theta)]^T \tag{6}$$

where $\boldsymbol{F}(\theta)$ is a column vector in which deviations of model moments from their empirical target are stacked. The weighting matrix W reflects our beliefs on the importance of each set of moments.[22] We target a total of 380 moments to estimate 41 parameters.

Numerical method for estimation We minimize the objective value as follows. We generate 15,000 uniform Sobol (quasi-random) points, compute the objective value for each of these, and select the best 1,000 (ranked by the objective value), each of which is used as an initial guess for the local minimization stage. This stage is performed with a mixture of Nelder-Mead's downhill simplex algorithm and the DFNLS algorithm of Zhang *et al.* (2010). In the end, we pick the best parameter estimates out of 1,000 local minima. For more details and discussion of this algorithm see Guvenen *et al.* (2018).

5 Estimation Results

We now present the key parameter estimates and discuss the model's fit to the data.

5.1 Parameter estimates

We first discuss the key parameter estimates and relate them to the moments that inform them the most. The full set of estimates are presented in Appendix D.

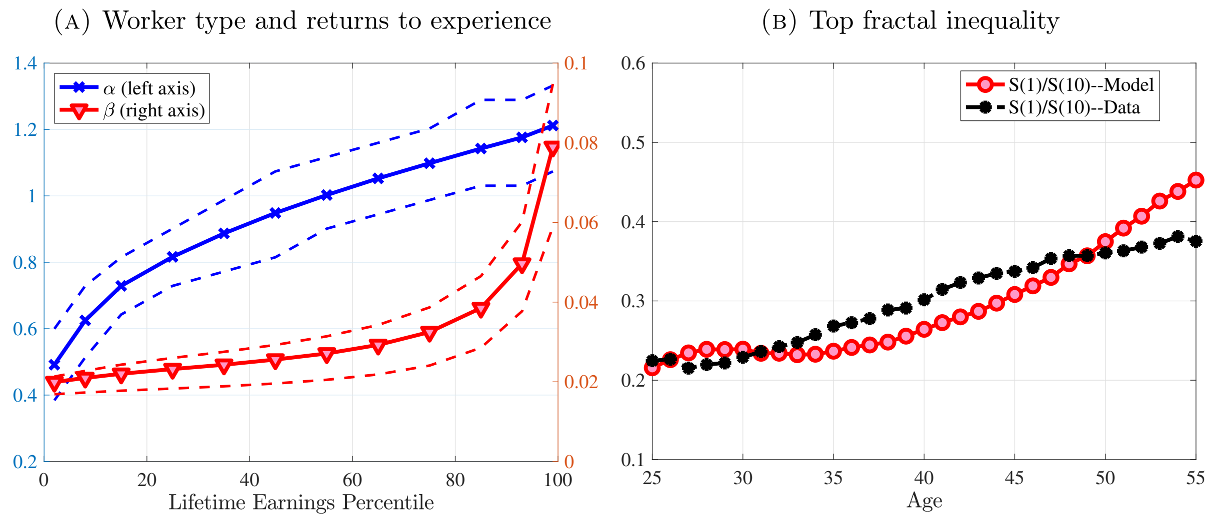
Distribution of α and β

We start by investigating the heterogeneity in permanent ability α_i and the returns to experience β_i. There are large differences in the values of α and β across the LE distribution (Figure 6a). α increases almost linearly throughout the LE distribution. Top

[21] Average earnings at age 25 have a much larger scale than all the other moments targeted in the estimation. To deal with potential issues that could arise for the large variation in the scales of the moments, we construct the deviation for this moment as the arc percentage deviation from the target.

[22] The weighting matrix, W, is chosen such that we assign a 15% relative weight to the first two sets of moments (i.e., cross-sectional moments of job stayers and switchers) and a 30% weight to the third and fourth sets of moments (i.e., the fraction of job stayers, EE—switchers, and EUE—switchers and their average wage growth). And finally, the last set of moments (average earnings at age 25 for each LE group) is given a 10% relative weight.

FIGURE 6 – Worker types and fractal inequality



Notes: The left panel shows the mean of the distributions of α and β by LE groups. Both distributions have been demeaned to have mean zero in the overall distribution. The right panel shows the ratio of incomes earned by the top 1% earners ($S(1)$) relative to the top 10% earners ($S(10)$).

LE individuals have an α that is more than 60 log points larger than that of those at the bottom. Moreover, there is a sizable variation within each LE group. The interquartile range (dashed lines in Figure 6a) is around 10 log points. Together with this, the standard deviation of α in the entire population is 0.25.

Return to experience, β, also increases with LE—not surprising given its positive correlation with α of $\rho_{\alpha\beta} = 0.44$—however with a different shape: β is relatively flat in the bottom two-thirds and increases steeply towards the top, essentially mirroring stayer earnings growth in Figure 2c.[23] Clearly, the variation of β over the LE distribution is dictated largely by the shape of its distribution in the population, which is assumed to be Pareto. This assumption, along with a Pareto firm productivity distribution, has important implications for the income share of top earners at each age.

Figure 6b shows the relative earnings share of the top 1% of the cross-sectional earnings distribution in the top 10%. While not targeted in the estimation, the model tracks this moment fairly well from age 25 to 50, increasing from around 0.2 to 0.4. This increase is driven by the growing importance of the return heterogeneity in annual

[23]Note that the interquartile range of β also increases from less then 0.005 at the bottom to more than 0.03 at the top, which is a direct feature of the fat tail of the Pareto distribution. The standard deviation of β in the population is estimated to be 0.017, in line with the estimates in the literature using different methodologies and datasets (Huggett *et al.* 2011 and Guvenen 2009).

earnings. Compared to the data, the relative share in the model increases faster in the last five years. We have also analyzed whether the model exhibits power law throughout the life cycle as in the data (Figure A.2). We find that after age 35 in the model, the relative income share of the top 0.1% in the top 1% is similar to that of the top 1% in the top 10%. Before age 35, the earnings concentration in the top 1% is larger than the earnings concentration in the top 10%, implying that the earnings distribution of young workers in the model has thicker tails than that in the data. These results confirm that earnings follow a power law throughout the working life in the data and in the model.

Typical models of top income inequality deliver a Pareto distribution only in the entire population but not within each age (e.g., see Gabaix *et al.* 2016; Gabaix 2009; Jones and Kim 2018). This is because in these models income grows exponentially over time and the distribution of experience is also exponential because of a Poisson displacement process. These two features together generate an exponentially distributed log income, or a Pareto distribution for income (see Jones and Kim (2018) for a more detailed discussion). However, the distribution of log income within each age is Gaussian in the random growth setting or in a process with normally distributed "growth types" (see also Guvenen *et al.* (2014a) who argue that several features of the MEF data are not consistent with a mechanism of top inequality through the accumulation of random returns over long periods of time). Our model has two ingredients that generate Pareto tails at each age: a Pareto firm and a Pareto β distribution. The former is key for generating Pareto tails early in life and the latter helps capture an increasing concentration of earnings at the top, both of which are clear features of the data.

Human capital depreciation We estimate human capital depreciation to be around 1.5% on a quarterly basis, larger in magnitude than estimated in Jarosch (2015) using German data. This is not the only channel in our model that contributes to scars from unemployment, which are large and persistent (Von Wachter *et al.* 2009, Krolikowski 2017). In our setting, an unemployed worker loses search capital, negotiation rents as well as the forgone opportunity of accumulating experience.

Heterogeneity in flow rates

Figure 7 plots in three panels how the unemployment risk, the job finding rate, and the contact rate vary with calendar age and LE groups.[24] Unemployment risk declines

[24] We would like to remind that the old (o) and young (y) ages in the model do not correspond to the *calendar* age in our simulations. Due to the stochastic aging process, there are old (o) workers in the model, even at earlier ages in the simulation.

sharply with lifetime earnings up to median LE and is essentially flat for individuals above the median (Figure 7a). Consistent with previous work, we find the unemployment risk to be significantly higher for young workers (see Shimer 1998 and Jung and Kuhn 2016). Our estimates for δ^a imply that aging leads to settling into more stable jobs. Jarosch (2015) endogenizes this dynamics using a two-dimensional job ladder model and uses life-cycle variation in the data to quantify the importance of job stability. An important feature of our results, as well the SIPP data, is that the life-cycle variation in unemployment risk is dwarfed by the differences between income groups.

To see if the estimated heterogeneity in unemployment risk is consistent with the data, we investigate how the model fits the evidence from the SIPP data (Section 2). To this end, we follow the exact same sample construction in the model-generated data as we do in the SIPP data, and compute flows between labor market states for three age and 10 recent earnings groups. Figure 7b shows how the unemployment risk varies with recent earnings in the SIPP data averaged over the life cycle along with its model counterpart. While not explicitly targeted in the estimation, the model captures remarkably well the extent of variation in the data, except for the top decile, where there is a slight uptick in the model-based UE rate but not in its empirical counterpart.

We estimate the job finding rate to be increasing with LE and age (Figure 7c). For example the quarterly job finding rate increases from around 30% at the bottom for workers ages 25–34 to above 60% at the top for workers ages 45–54. These estimates imply that bottom LE workers ages 25–34 stay unemployed for around 3 quarters, compared to less than 2 quarters for top LE ages 45–54. Coupled with a high unemployment risk for low LE workers, these differences imply large differences in actual experience over the life cycle (Figure 10b). In particular, quarters worked over the working life range from 90 for low LE individuals to 120 at the top. This large heterogeneity, especially below the median, has implications for earnings growth differences that we discuss later.

Our estimate of the heterogeneity in job finding rates is qualitatively consistent with the external evidence from the SIPP data (Figure 7d). While the model generates an increasing pattern of job finding rates with respect to recent earnings, the variation is much less pronounced compared to the data.

We estimate that 12.5% of unemployed workers return after the jobless spell to their last employer ($\lambda_r = 0.125$). This recall probability is lower than the 40% measured in Fujita and Moscarini (2017). They measure recalls directly using survey data from the

SIPP, whereas we infer them indirectly to match the left tail of the earnings growth distribution of job stayers.

Turning to the contact rate for employed workers, we find this to be increasing with lifetime earnings and age, with a range between 15% and 40% (Figure 7e). While the increasing likelihood of the contact rate with LE seems inconsistent with a declining job-to-job transition rate in the SIPP data (Figure 5c), the model actually captures this pattern well endogenously. High LE workers get a lot of offers, climb the job ladder fast and work for high-productivity firms that are hard to poach from. Thus, the success of the model in matching the SIPP evidence is due to high LE workers rejecting most of the contacts since they are already employed at firms with a higher surplus.

To inspect this mechanism in the data, we analyze data from a supplement to the Survey of Consumer Expectations (SCE), which is a monthly, nationally representative survey of roughly 1,300 individuals that asks respondents about their expectations about various aspects of the economy. The special supplement asks a variety of questions that are tailored to an individual's employment status, prior work history and job search behavior (see Faberman *et al.* 2017 for more details.). Importantly for our purposes, it asks about the number of employer contacts and job offers received, and how those contacts and offers arose; i.e., whether they were the result of traditional search methods or whether they came about through a referral or an unsolicited employer contact. To keep the analysis similar, we take a sample of employed respondents ages 25–55, and group them into five bins based on their average wages over the last year. We then report for each group the average number of contacts they received from other potential employers (Table II). We find that contacts increase in previous wages and are quite high at the top. People in the highest group (workers above the 95th percentile) are contacted around 0.43 times per month. This is more than two times larger compared to workers in the lowest quantile, consistent with the underlying mechanism in the model. Moreover, inspecting unsolicited contacts, those that were not initiated by the employee, we find much larger differences. For top earners, contacts are almost five times more likely than for those at the bottom (0.43 vs. 0.09, respectively). Moreover, essentially all of their contacts are unsolicited.

Life-cycle variation Since flow rates vary little over the life-cycle in the data, we relegate this dimension to Appendix D.2 Figure D.6. Overall, the model captures well the decline in job loss risk and job-to-job transition rate with age. However, there is

Table II – Subjective contact rate

Recent earnings groups	1-25%	26-50%	51-75%	76-94%	95+%
Total Number of Contacts	0.18	0.18	0.13	0.26	0.43
Unsolicited Contacts	0.09	0.02	0.04	0.11	0.43

Notes: Respondents ages 25-55. Individuals who report 25 or more contacts in the last 4 weeks are dropped from the sample. We assign zero contacts for those reporting a positive number of contacts but none corresponding with either (i) an employer directly online or through email, (ii) an employer directly through other means, including in-person, or (iii) an employment agency or career center (including a career center at a school or university).

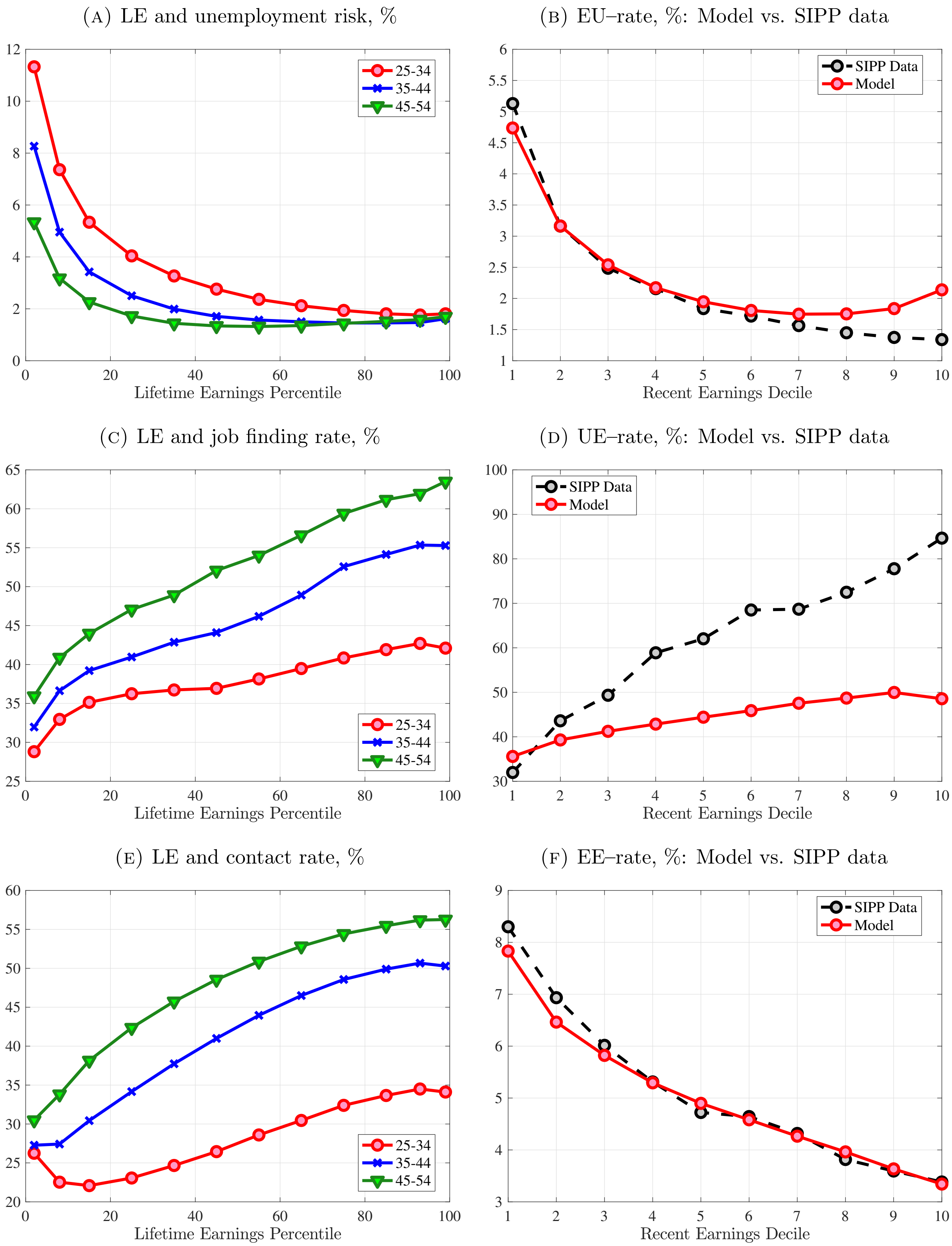
little age variation in the data in job finding rates for unemployed workers, whereas the model estimates are systematically higher for older workers. As discussed earlier, the job finding rate is identified from the left tail of the switcher earnings growth, which is shorter for old workers than for younger workers in the data.

Related work Lentz *et al.* (2018) implement the finite mixture approach of Bonhomme *et al.* (2019) to estimate a model of wage dynamics and employment mobility with lots of heterogeneity using Danish data. Consistent with our results, they find that layoff rates are strongly decreasing in mean wage, especially so for low-tenure workers. Moreover, they estimate contact rates for employed workers to be increasing in worker type. These findings are also consistent with Bagger *et al.* (2014). Cairo and Cajner (2017) show that more educated workers have similar job finding rates but much lower and less volatile separation rates than their less educated peers. These results mirror our findings on differences in flow rates between high and low LE workers.

Idiosyncratic shocks

Recall that idiosyncratic shocks hit job stayers once a year with probability $\pi(\alpha)$. We estimate that the probability of experiencing such a shock increases with LE from around 5% to 20%. Recall that the distribution of these shocks is identified from the second-to-fourth order moments of the annual earnings growth for job stayers. Figure 8 shows that the variance increases above the 20th percentile and kurtosis decreases above the 40th percentile of the LE distribution. These patterns require idiosyncratic shocks to be more likely for higher LE workers. In the lower end of the LE distribution idiosyncratic shocks matter less and therefore the earnings dynamics for job stayers are mainly driven by the endogenous mechanisms of the job ladder model including recalls. This finding is consistent with evidence from Norway that large earnings changes are mostly driven by wage changes for high earners (Halvorsen *et al.* 2019).

Figure 7 – Labor market flows



5.2 Model's fit to the data

We now discuss the model's performance in fitting the targeted moments. In doing so, we also discuss the economic forces behind the model's fit, which helps us further understand how the pieces in the model are informed by the different aspects of the data.

Cross-sectional moments Figure 8 shows the fit of the model to cross-sectional moments. For the clarity of exposition, we suppress the life cycle variation and plot averages over three age groups. The fit along the life-cycle is shown in Appendix D.2.

The model captures well the standard deviation of earnings changes for job stayers and switchers (Figure 8a). Both in the data and in the model, job switchers have a higher standard deviation throughout the LE distribution. In the model, big changes to earnings happen when people switch jobs because of a job loss. The declining unemployment risk (Figure 7a) combined with an increasing poaching rate (Figure 7e) implies that a higher share of job switchers at the bottom go through unemployment as opposed to direct job switches, and explains why the standard deviation is higher at the bottom compared to the rest of the distribution. The profile flattens out because there is much less variation in the unemployment risk above the median.

For job stayers, earnings changes are driven by job loss followed by a recall, an outside offer that leads to renegotiation and idiosyncratic shocks. Due to their high job loss rates, the share of recalls is highest at the bottom, which tends to push up the standard deviation at the bottom. As we move to the right along the LE distribution, unemployment risk fades, the prevalence of outside offers increases, and a larger share of such offers result in the worker staying with the same employer, and getting a large raise (Figures 7e and 7f). Moreover, idiosyncratic shocks become more frequent and contribute to the increasing standard deviation for job stayers above the median.

Turning to skewness, we find that the model captures well the essential features of the data (Figures 8c and 8d). First, earnings changes are negatively skewed for both job switchers and stayers. For switchers, the negative skewness is mostly a result of flows into unemployment, which result in the worker losing the position on the job ladder and human capital depreciation throughout the spell of unemployment. The decreasing profile of skewness (increasing negative skewness) is a result of two offsetting forces. On the one hand, human capital depreciation is stronger for low LE individuals due to longer unemployment durations, pushing skewness down at the bottom. On the other hand, job loss is less frequent but more costly for high LE individuals as they have more search capital and negotiation rents to lose. The latter force dominates and causes the skewness of earnings changes to be more negative for job switchers among high LE individuals.

Recalls generate large earnings declines within the same firm. In the absence of recalls, the model cannot generate a negative skewness for job stayers. As we move to the right of the LE distribution, the left tail shrinks as recalls become less frequent. The right tail expands, because outside offers arrive more often and are more likely to result in wage renegotiation. Both forces combined result in a milder negative skewness for job stayers at higher LE percentiles.

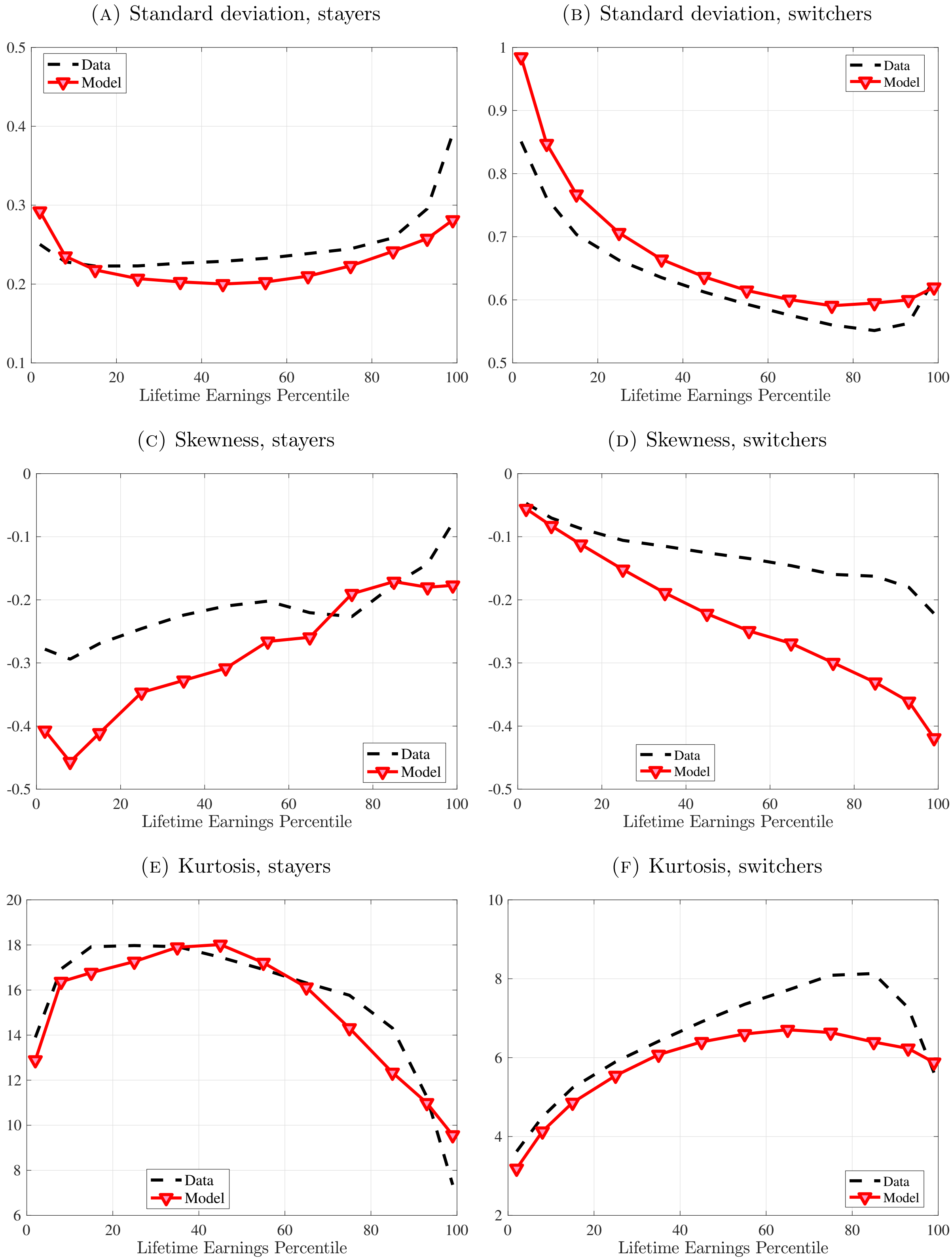
The model is quite successful in matching the extent of kurtosis and its variation over the LE distribution. Kurtosis measures the tendency of a distribution to stay away from $\mu \pm \sigma$ (Moors 1986). Distributions with excess kurtosis tend to have pointy centers and longer tails relative to a Gaussian one. Infrequent events that lead to large changes, such as outside offers and unemployment spells followed by recalls, are the leading sources of excess kurtosis for job stayers. In fact, they are so strong that without idiosyncratic shocks, earnings changes would be a lot more leptokurtic. The idiosyncratic shocks, despite being leptokurtic themselves, help the model bring down the kurtosis of job stayers closer to values in the data. Earnings changes of job switchers are also leptokurtic in the model and the data, but to a lesser degree compared to job stayers.

Finally, we investigate the model's fit on cross-sectional moments along the life-cycle dimension. Figure D.2 shows how the higher-order moments of earnings changes for stayers and switchers vary between three age groups. As in the data, life-cycle variation in the model is less pronounced than the variation between LE groups. Overall, we conclude that the model does fairly well in capturing the essential moments of earnings changes for job stayers and switchers across the LE distribution and over the life cycle.

Income growth moments Next, we study job stayers and switchers. The model reproduces remarkably well the increasing share of job stayers by LE quantile in the data (Figure 9). There are few job stayers at the bottom due to high flow rates into unemployment. The share of job stayers essentially follows the unemployment risk along the LE distribution, increasing up to around the 70th percentile and stabilizing thereafter.

The model also generates overall a realistic average earnings growth for job stayers and switchers throughout the LE distribution (Figure 9b). In particular, there is little heterogeneity among job stayers for the bottom two thirds of the LE distribution, which, as discussed before, is in part due to the relatively flat average profile of returns to

FIGURE 8 – Model's fit to cross-sectional moments of $\frac{Y_{t+1}-Y_t}{(Y_{t+1}-Y_t)/2}$

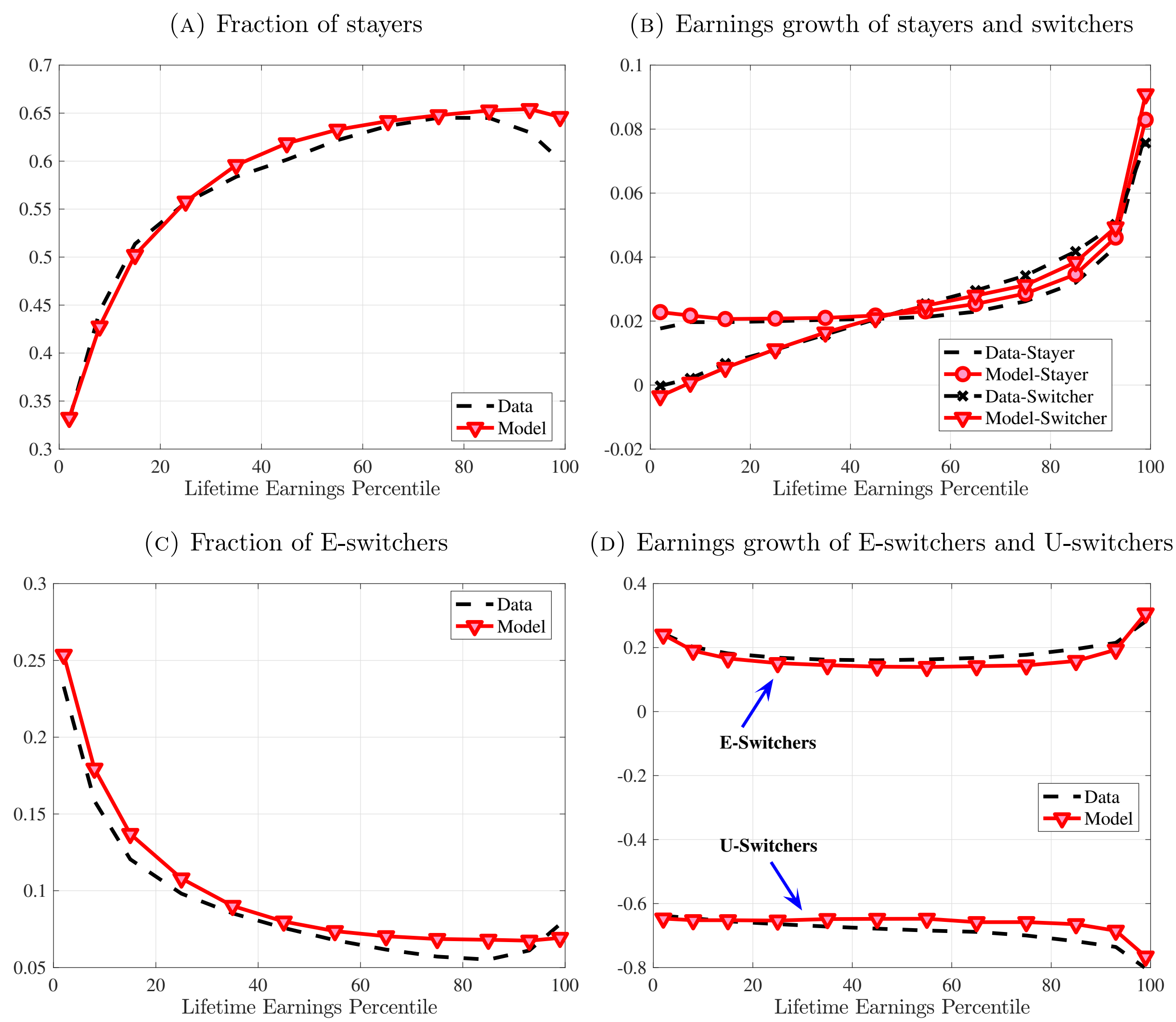


experience (β) in each LE group. Earnings growth of job stayers has a component due to human capital accumulation, governed by β, and a component due to the job ladder, through outside offers that lead to wage increases on the job. As Figure 6a shows, the former component is basically flat for two thirds of the distribution with a very small positive slope. Yet, the earnings growth of stayers in the model is *higher* at the low end of the distribution compared to the 20th percentile. This feature has to do with the second component, which is stronger at the low end. This result may seem surprising because bottom LE individuals have the lowest contact rates when employed. However, given their high unemployment risk, employed workers at the bottom tend to also have a lower piece rate as they frequently lose their job before they receive many outside offers and can negotiate a better piece rate. A lower piece rate implies that, conditional on staying with the same firm (which is the group we consider in Figure 9b), an outside offer is more likely to lead to wage renegotiation. Thus, there are two competing forces determining the effect of the job ladder risk at the bottom: a lower contact rate and a higher share of those contacts that lead to wage growth. It turns out that the latter is stronger at the bottom compared to the 20th percentile of the LE distribution.

Turning to job switchers, the model captures well their average earnings growth (Figure 9b). In particular, there is a large variation throughout the LE distribution, ranging from zero at the bottom to 9%. Moreover, consistent with the data, most of this heterogeneity is due to compositional differences among job switchers. The share of E-switchers among all switchers, defined the same way as in the data, increases sharply from around 65% at the bottom of the LE distribution to above 80% (Figure 9c). These shares are slightly below those in the data but capture remarkably well the variation along the LE dimension. Finally, consistent with the data, there is much less between-group heterogeneity in the earnings growth of E-switchers and U-switchers (Figure 9d).

Thus, we conclude that the estimated job ladder model captures quite well the key features of the careers of individuals in different parts of the lifetime earnings distribution.

Figure 9 – Model's fit to income growth moments



6 Decomposing Lifetime Earnings

The model matches well the distribution of lifetime earnings in the data (Figure 10a). Looking at the top of the distribution, LE_{50} earns around 4.19 times as much as LE_{45} in the model, slightly overstating this ratio in the data (3.83). The fit is much better below LE_{45}. For example, LE_{45} earns about 1.97 times as much as LE_{25} in the model, compared to 1.94 in the data. Moreover, the ratio of LE_5 to LE_1 is 1.80 in the model, slightly below its empirical counterpart of 1.92 (Table III).

To what extent are these large differences in lifetime earnings driven by differences in wages earned by different individuals as opposed to differences in employment rates

Table III – Lifetime earnings differences across LE groups

	LE_{50}/LE_{45}	LE_{45}/LE_{25}	LE_{25}/LE_{5}	LE_{5}/LE_{1}
Data	3.83	1.94	1.90	1.80
Model	4.19	1.97	2.07	1.92

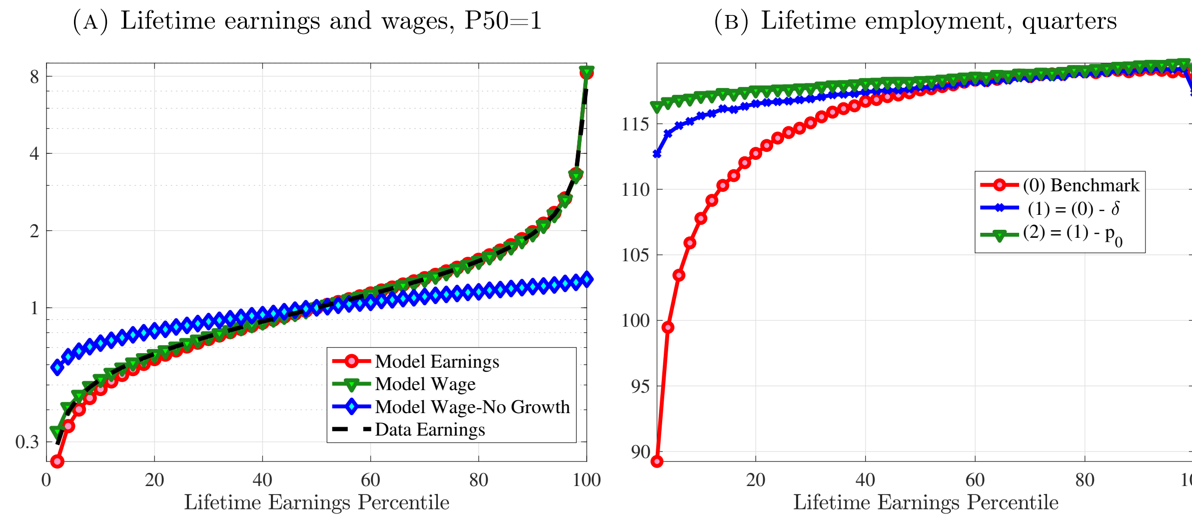
Notes: This table reports the differences in lifetime earnings between different groups of individuals in the model and the data. LE_i/LE_j is computed as the ratio of average lifetime earnings of the individuals in LE_i to those in LE_j.

over the life cycle? Figure 10a plots the earnings and wage differences in the model by normalizing the median group to 1. This figure shows that wage—rather than employment—differences explain the vast majority of LE inequality. Differences in average wages over the life cycle are remarkably similar to the lifetime earnings differences, except below the 25th percentile, where differences in employment (measured as the number of quarters worked over the working life) play an important role. For example, employment of workers at the bottom of LE is about 25% lower than that of the median workers. Employment differences above the median are negligible in comparison (Figure 10b).

Before investigating the sources of lifetime wage differences, we briefly discuss the sources of employment differences below the median. These differences arise due to ex-ante heterogeneity in unemployment risk and job finding rates as well as the ex-post job ladder risk; i.e., ex-ante similar workers experiencing different job loss and job finding shocks. To measure their relative roles, we first shut down ex-ante heterogeneity in job loss risk by endowing all individuals with $\delta^a(0)$, the job loss risk of workers with $\alpha_i - \mu_\alpha = 0$, which is roughly the average value for median LE workers (Figure 6a), and compute the resulting distribution of total lifetime employment. In doing so (and in all experiments that follow), we keep the rankings of workers, and thus the composition of LE groups, unchanged from the baseline. Therefore, the differences between this experiment and the baseline are only due to the differences in ex-ante job loss risk, δ.

We find that employment differences between the bottom and top LE decline sharply from around 25% to 7% when all workers have the same job loss rate (Figure 10b). When we further eliminate differences in job finding rates by setting $\lambda_0^a(\alpha)$ to $\lambda_0^a(0)$ for all workers, employment differences decline further: Throughout their lifetime, bottom LE individuals work only 3% less than those at the top. The remaining differences are entirely due to the ex-post realizations; i.e., luck. Our estimation thus attributes little

Figure 10 – Lifetime earnings, wages and employment



Notes: "Model Wage–No Growth" corresponds to an experiment that shuts down the heterogeneity in β, eliminates search frictions ($\delta = 1, \lambda_0 = 1, \lambda_1 = 0$), removes idiosyncratic shocks and makes the firm distribution degenerate. In this specification, the only source of wage (and earnings) differences is heterogeneity in permanent ability, α. Each series is normalized so that it takes a value of 1 for the median group. Therefore, the values reflect differences relative to median LE individuals.

role to luck in generating sizable lifetime employment differences.[25]

We now turn to wage differentials. First, we establish that these are largely shaped by *wage growth* heterogeneity rather than by the initial differences in *levels*. Figure 10a shows that when all sources of wage growth have been turned off and only the differences in permanent ability α are allowed for, the model generates a wage inequality that is an order of magnitude smaller. Therefore, it is essential to understand why some workers have a much steeper wage profile than others.

Recall that in the model wage growth can differ across individuals due to differences in the ability to accumulate human capital and climb the job ladder. The latter contains ex-ante and ex-post differences in unemployment, and the quality and quantity of offers on and off the job. To assess the relative roles of these factors in lifetime wage growth differences, we shut down each component one after another, until we eliminate all differences, again keeping the composition of the LE groups the same with our benchmark. Figure 11a shows the lifetime wage growth differences between LE groups for the benchmark estimation as well as for the counterfactual experiments.[26]

[25]One caveat is that in our model current unemployment does not beget future unemployment as in Jarosch (2015).

[26]We provide a separate decomposition of lifetime earnings growth into wage and hours growth

We start by eliminating the differences in unemployment risk, which we accomplish by shutting off differences in job loss and job finding rates together ($\delta^a(\alpha_i) = \delta^a(0)$ and $\lambda_0^a(\alpha_i) = \lambda_0^a(0)$). Heterogeneity in unemployment risk has a marked effect on wage growth differences between the bottom and median LE workers, and to a lesser degree above the median (series (1) in Figure 11a). Specifically, slightly more than 50% of wage growth differences at the bottom would diminish, if the workers at the bottom had job loss and job finding rates similar to those of the median LE workers. High unemployment rates of low-income individuals (Figure 10b) not only prevent them from accumulating human capital but also lead to depreciation in human capital during unemployment. Furthermore, a higher incidence of unemployment prevents bottom LE workers from climbing the job ladder. Figure 11b shows the contributions of human capital, search capital and negotiation rents to the wage growth differences between the bottom and median LE workers.[27] Differences in human capital accumulation account for almost 70% of the wage growth differences between these two groups. The remaining difference is essentially due to the accumulation of search capital (i.e., working for more productive firms). The contribution of the negotiation capital is slightly negative, meaning that workers at the bottom experience larger growth in their piece rate compared to those at the median.[28] Eliminating unemployment risk brings down the differences in human and search capital accumulation, and thus differences in wage growth, by around 65%. These findings are overall consistent with those in Bagger *et al.* (2014) from Danish data.

Job loss and job finding differences matter much less at the upper half of the distribution, mainly because these workers have fairly similar unemployment risk to begin with. Because top earners have a slightly higher job loss risk than median workers, eliminating this difference would actually *raise* their income growth further by around 20 log points.

Recall that LE groups also display sizable differences in their contact rates λ_1. In contrast to the job loss and job finding rates, these differences have a smaller effect on lifetime wage growth differences (Figure 11a). Specifically, eliminating differences in

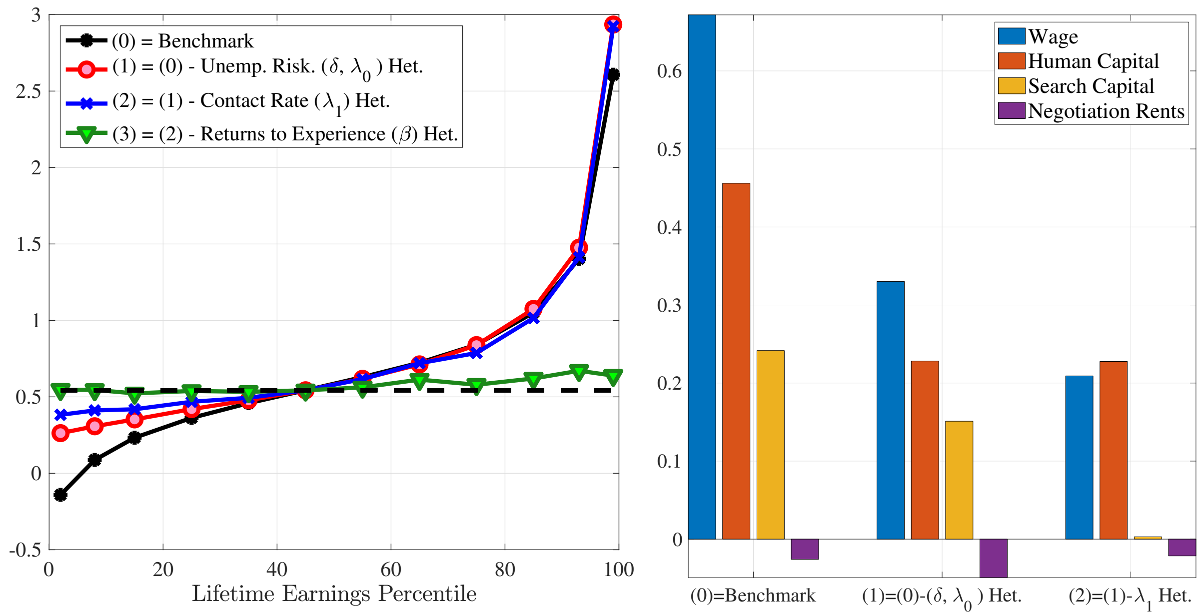
components (Figure D.7a). The model matches well the earnings growth heterogeneity in the data. Earnings growth is less dispersed than wage growth between LE groups, especially in the bottom half of the LE distribution, due to higher lifetime employment growth at the bottom of the LE distribution.

[27] The growth in search capital is measured as the log change in firm productivity $\mathbb{E}\left[p_{j(i,55)} - p_{j(i,25)}\right]$, and the growth in negotiation rents is measured as the change in the log piece rate $\mathbb{E}\left[r_{i,55} - r_{i,25}\right]$.

[28] Lower growth in negotiation capital for top LE workers occurs because these workers are employed at more productive firms, which enjoy a stronger monopsony power as they are hard to poach from. Gouin-Bonenfant *et al.* (2018) argue that this channel is key for understanding the decline in aggregate labor share, whereas it plays a smaller role in our findings.

Figure 11 – Decomposing wage growth between ages 25 and 55

(a) Determinants of wage growth heterogeneity (b) Log wage growth: Median vs. Bottom LE



Notes: The left panel decomposes the differences in lifetime wage growth into the job ladder and human capital components. The black line shows the wage growth in the benchmark, the circled red line shows wage growth without ex-ante job loss and job finding rate differences, the blue line eliminates the differences in the contact rate, and the black dashed line corresponds to the case with no heterogeneity in returns to experience. The right panel decomposes the differences in average lifetime log wage growth between bottom and median LE workers into human capital, search capital and negotiation rents.

contact rates would close an additional 20% of the wage growth gap between the bottom and the median, with essentially no effect at the top. All of this effect is due to the closing of search capital differences. Namely, endowing bottom LE individuals with the contact rate of median LE individuals allows bottom LE workers to climb to better jobs. These two experiments show that eliminating differences in job ladder risk can go a long way in ameliorating the labor market experiences of bottom LE workers and eliminate more than 70% of the differences in wage growth with median LE workers.

Next, we turn to the role of heterogeneity in returns to experience. To this end, we assign all workers' β to the average, which eliminates all heterogeneity in wage growth differences except for the differences due to idiosyncratic productivity shocks and the realizations of labor market shocks (Figure 11a). A couple of remarks are in order. First, luck—the realizations of idiosyncratic productivity and job ladder shocks— plays a negligible role in lifetime wage growth. On average, above median LE individuals are somewhat more lucky, but this has a very small quantitative effect. Second, eliminating differences in returns to experience has an effect across the entire LE distribution, with the largest effect on top LE earners. Together with the fact that job ladder risk plays a small role for top earners, we conclude that the reason why top earners experience a much larger wage growth than the median is primarily because they have a much higher pace of human capital accumulation according to our estimation (Figure 6a).

Which feature of the data tells the model that human capital accumulation is more important at the upper half of the LE distribution and vice versa at the bottom half? While all targeted moments are informative, we argue that the differences between income growth of job stayers and switchers are key. To see this, note that human capital is capitalized into wages in all firms. Therefore, wage growth always reflects a worker's human capital accumulation, regardless of whether he stays with the current employer or switches to a new one. If the data show a high wage growth for a group of workers relative to median workers regardless of job switching, as is the case in the data for higher LE individuals (Figure 2c), the model infers a high returns to experience.

The difference in earnings growth between stayers and switchers is informative about the role of job ladder risk. If a group of workers experience lower growth when switching than they do when they stay with the same employer, the model rationalizes this by inferring a poor job ladder, due to a high job loss or a low job finding rate. At the bottom of the LE distribution, job switchers experience much smaller earnings growth compared to stayers, consistent with our finding that the job ladder component is more important for explaining their lower lifetime growth relative to the median.

7 Conclusion

This paper investigates the causes of the large heterogeneity in lifetime earnings. Differences in earnings growth over the working life are important for lifetime earnings inequality. Using detailed administrative data from SSA records, we show that earnings growth is surprisingly similar for the bottom two-thirds of the LE distribution when they stay with the same employer. Differences arise when workers change employers, with earnings growth rising with LE. Moreover, top LE individuals experience a much larger earnings growth relative to the median regardless of whether they remain with the same employer or switch to a new one. We use these facts along with other facts on job switching and unemployment rates to estimate a job ladder model featuring human capital accumulation and "lots of heterogeneity." Our results show that differences in returns to experience are key for inequality at the upper half of the distribution. Differences in ex-ante job ladder risk—job loss, job finding, and contact rate heterogeneities—are the most important factor for the bottom. These differences have important implications for the design of unemployment insurance policies.

An emerging literature studies the effects of firms' power in setting wages. Firms can hire and retain workers at wages lower than the competitive fringe if they are large in a market (Berger *et al.* 2019; Jarosch *et al.* 2019) or if they do not face much competition from other employers, either due to contractual restrictions on job mobility (Johnson *et al.* 2019) or other frictions. One interpretation of the estimated differences in outside contacts is about employers' ability to restrict poaching. Through this lens, our results suggest that firms are better able to restrict poaching for low-skill workers and have more power over them. This interpretation is consistent with Caldwell and Danieli (2018), who find much less competitive pressure for low-skill workers.

Lastly, our analysis has focused on worker differences in job ladder risk. Some of the differences in job stability and outside contacts could be a characteristic of jobs rather than workers. Jarosch (2015) focuses on firm heterogeneity in job stability. More broadly, firms might contribute to wage growth heterogeneity by providing different learning environments (Gregory 2019). To fully understand the role of firms and workers in wage inequality, a unified approach is necessary, which we leave for future work.

References

Abrahám, A., Alvarez-Parra, F. and Forstner, S. (2016). The effects of moral hazard on wage inequality in a frictional labor market. *Unpublished working paper*.

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.

Bagger, J., Fontaine, F., Postel-Vinay, F. and Robin, J.-M. (2014). Tenure, experience, human capital and wages: A tractable equilibrium search model of wage dynamics. *American Economic Review*, **104** (6), 1551–1596.

— and Lentz, R. (2014). An empirical model of wage dispersion with sorting. *The Review of Economic Studies*.

Barlevy, G. (2008). Identification of search models using record statistics. *The Review of Economic Studies*, **75** (1), 29–64.

Berger, D. W., Herkenhoff, K. F. and Mongey, S. (2019). *Labor Market Power*. Tech. rep., National Bureau of Economic Research.

Blanchard, O. J. (1985). Debt, deficits, and finite horizons. *Journal of Political Economy*, **93** (2), 223–247.

Bonhomme, S., Lamadon, T. and Manresa, E. (2017). A distributional framework for matched employer employee data. *Unpublished manuscript, University of Chicago*.

—, — and — (2019). A distributional framework for matched employer employee data. *Econometrica*, **87** (3), 699–739.

Bontemps, C., Robin, J.-M. and Van den Berg, G. J. (1999). An empirical equilibrium job search model with search on the job and heterogeneous workers and firms. *International Economic Review*, **40** (4), 1039–1074.

Burdett, K. and Mortensen, D. T. (1998a). Wage differentials, employer size, and unemployment. *International Economic Review*, **39** (2), 257–73.

— and — (1998b). Wage differentials, employer size, and unemployment. *International Economic Review*, pp. 257–273.

Cahuc, P., Postel-Vinay, F. and Robin, J.-M. (2006). Wage bargaining with on-the-job search: Theory and evidence. *Econometrica*, **74** (2), 323–364.

Cairo, I. and Cajner, T. (2017). Human capital and unemployment dynamics: Why more educated workers enjoy greater employment stability. *The Economic Journal*, **128** (609), 652–682.

Caldwell, S. and Danieli, O. (2018). Outside options in the labor market. *Unpublished manuscript*.

Deaton, A. and Paxson, C. (1994). Intertemporal choice and inequality. *Journal of Political Economy*, **102** (3), 437–67.

Faberman, R. J., Mueller, A. I., Şahin, A. and Topa, G. (2017). *Job search behavior among the employed and non-employed*. Tech. rep., National Bureau of Economic Research.

Fujita, S. and Moscarini, G. (2017). Recall and unemployment. *American Economic Review*, **107** (12), 3875–3916.

Gabaix, X. (2009). Power laws in economics and finance. *Annu. Rev. Econ.*, **1** (1), 255–294.

—, Lasry, J.-M., Lions, P.-L. and Moll, B. (2016). The dynamics of inequality. *Econometrica*, **84** (6), 2071–2111.

Gibbons, R. and Waldman, M. (1999). Careers in organizations: Theory and evidence. *Handbook of Labor Economics*, **3**, 2373–2437.

Gouin-Bonenfant, E. *et al.* (2018). Productivity dispersion, between-firm competition and the labor share. In *Society for Economic Dynamics. Meeting Papers*, vol. 1171.

Gregory, V. (2019). Firms as learning environments: Implications for earnings dynamics and job search. *Unpublished Manuscript*.

Guvenen, F. (2009). An empirical investigation of labor income processes. *Review of Economic Dynamics*, **12** (1), 58–79.

—, Kaplan, G. and Song, J. (2014a). *The glass ceiling and the paper floor: Gender differences among top earners, 1981-2012*. Tech. rep., National Bureau of Economic Research.

—, —, — 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.

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

Haider, S. J. (2001). Earnings instability and earnings inequality of males in the United States: 1967-1991. *Journal of Labor Economics*, **19** (4), 799–836.

Haltiwanger, J., Hyatt, H. R., McEntarfer, E., Sousa, L. and Tibbets, S. (2014). Firm age and size in the longitudinal employer-household dynamics data. *US Census Bureau Center for Economic Studies Paper No. CES-WP-14-16*.

HALVORSEN, E., HOLTER, H., OZKAN, S. and), K. S. (2019). *Dissecting Idiosyncratic Income Risk: Evidence from Norwegian Registry Data,*. Tech. rep., Technical report, mimeo.

HEATHCOTE, J., STORESLETTEN, K. and VIOLANTE, G. L. (2005). Two views of inequality over the life cycle. *Journal of the European Economic Association*, **3** (2-3), 765–775.

HECKMAN, J. J. (1976). A life-cycle model of earnings, learning, and consumption. *Journal of political economy*, **84** (4, Part 2), S9–S44.

HUBMER, J. (2018). The job ladder and its implications for earnings risk. *Review of Economic Dynamics*, **29**, 172–194.

HUGGETT, M., VENTURA, G. and YARON, A. (2011). Sources of lifetime inequality. *American Economic Review*, **101** (7), 2923–54.

JAROSCH, G. (2015). *Searching for Job Security and the Consequences of Job Loss*. Working paper, Princeton University.

—, NIMCZIK, J. S. and SORKIN, I. (2019). *Granular Search, Market Structure, and Wages*. Tech. rep., National Bureau of Economic Research.

— and PILOSSOPH, L. (2018). Statistical Discrimination and Duration Dependence in the Job Finding Rate. *The Review of Economic Studies*.

JOHNSON, M. S., LAVETTI, K. and LIPSITZ, M. (2019). The labor market effects of legal restrictions on worker mobility. *Available at SSRN 3455381*.

JOLIVET, G., POSTEL-VINAY, F. and ROBIN, J.-M. (2006). The empirical content of the job search model: Labor mobility and wage distributions in Europe and the U.S. In *Structural Models of Wage and Employment Dynamics*, Emerald Group Publishing Limited, pp. 269–308.

JONES, C. I. and KIM, J. (2018). A Schumpeterian model of top income inequality. *Journal of Political Economy*, **126** (5), 1785–1826.

JOVANOVIC, B. (1979). Job matching and the theory of turnover. *Journal of Political Economy*, **87** (5), 972–90.

JUNG, P. and KUHN, M. (2016). Earnings losses and labor mobility over the life cycle. *Journal of the European Economic Association*.

KAPLAN, G. (2012). Inequality and the life cycle. *Quantitative Economics*, **3** (3), 471–525.

KROLIKOWSKI, P. (2017). Job ladders and earnings of displaced workers. *American Economic Journal: Macroeconomics*, **9** (2), 1–31.

LAZEAR, E. P. and ROSEN, S. (1981). Rank-order tournaments as optimum labor contracts. *Journal of Political Economy*, **89** (5), 841–864.

LENTZ, R., PIYAPROMDEE, S., ROBIN, J.-M. *et al.* (2018). On worker and firm heterogeneity in wages and employment mobility: Evidence from anish register data. *PIER Discussion Papers*, **91**.

LISE, J. (2013). On the job search and precautionary savings. *Review of Economic Studies*, **80**, 1086–1113.

—, MEGHIR, C. and ROBIN, J.-M. (2016). Matching, sorting and wages. *Review of Economic Dynamics*, **19**, 63–87.

MINCER, J. (1974). *Schooling, Experience and Earnings*. New York: Columbia University Press.

MOORS, J. J. A. (1986). The meaning of kurtosis: Darlington reexamined. *The American Statistician*,

40, 283–284.

NEAL, D. and ROSEN, S. (2000). Theories of the distribution of earnings. *Handbook of Income Distribution*, **1**, 379–427.

OLSEN, A. and HUDSON, R. (2009). Social Security Administration's master earnings file: Background information. *Social Security Bulletin*, **69** (3), 29–46.

PANIS, C., EULLER, R., GRANT, C., BRADLEY, M., PETERSON, C. E., HIRSCHER, R. and STINBERG, P. (2000). *SSA Program Data User's Manual*. Baltimore, MD: Social Security Administration.

PASTORINO, E. (2019). *Careers in firms: the role of learning and human capital*. Tech. rep., mimeo.

PIKETTY, T. and SAEZ, E. (2003). Income inequality in the United States: 1913-1998. *Quarterly Journal of Economics*, **118** (1), 1–39.

POLACHEK, S. W., DAS, T. and THAMMA-APIROAM, R. (2015). Micro-and macroeconomic implications of heterogeneity in the production of human capital. *Journal of Political Economy*, **123** (6), 1410–1455.

POSTEL-VINAY, F. and ROBIN, J.-M. (2002). Equilibrium wage dispersion with worker and employer heterogeneity. *Econometrica*, **70** (6), 2295–2350.

RUBINSTEIN, A. (1982). Perfect equilibrium in a bargaining model. *Econometrica: Journal of the Econometric Society*, pp. 97–109.

SHIMER, R. (1998). Why is the U.S. unemployment rate so much lower? *NBER Macroeconomics Annual*, **13**, 11–61.

— (2001). The impact of young workers on the aggregate labor market. *The Quarterly Journal of Economics*, **116** (3), 969–1007.

SONG, J., PRICE, D. J., GUVENEN, F., BLOOM, N. and VON WACHTER, T. (2018). Firming up inequality. *The Quarterly Journal of Economics*, **134** (1), 1–50.

SORKIN, I. (2018). Ranking firms using revealed preference. *The Quarterly Journal of Economics*, **133** (3), 1331–1393.

TANAKA, S., WARREN, L. and WICZER, D. (2019). *Earnings growth, job flows and churn*. Tech. rep., mimeo.

TOPEL, R. H. and WARD, M. P. (1992). Job mobility and the careers of young men. *The Quarterly Journal of Economics*, **107** (2), 439–79.

VAN DEN BERG, G. J. (1990). Search behaviour, transitions to non-participation and the duration of unemployment. *The Economic Journal*, **100** (402), 842–865.

VON WACHTER, T., SONG, J. and MANCHESTER, J. (2009). Long-term earnings losses due to mass layoffs during the 1982 recession: An analysis using us administrative data from 1974 to 2004. *Unpublished paper, Columbia University.*

YAARI, M. E. (1965). Uncertain lifetime, life insurance, and the theory of the consumer. *The Review of Economic Studies*, **32** (2), 137–150.

ZHANG, H., CONN, A. R. and SCHEINBERG, K. (2010). A derivative-free algorithm for least-squares minimization. *SIAM Journal on Optimization*, **20** (6), 3555–3576.

Supplemental Online Appendix

NOT FOR PUBLICATION

A Additional Results from the MEF

A.1 Sample Selection

Our initial sample consists of 1,845,640 individuals (Table IV). About 18% are self-employed in at least one fourth of their working life. About 490,000 are eliminated, as they do not satisfy the minimum years of employment criterion. We exclude close to 160,000 (27,000) individuals due to consecutive nonemployment(self-employment). This procedure leaves us with a final sample of 840,194 individuals for whom we have at least 31 years of earnings data.

A.2 Moments for Top Earners

A.2.1 Pareto Tails of the Earnings Distribution

FIGURE A.1 – Pareto tails in the top 5% of lifetime earnings distribution

(A) Log density of lifetime earnings (B) Log Inverse CDF of Lifetime Earnings

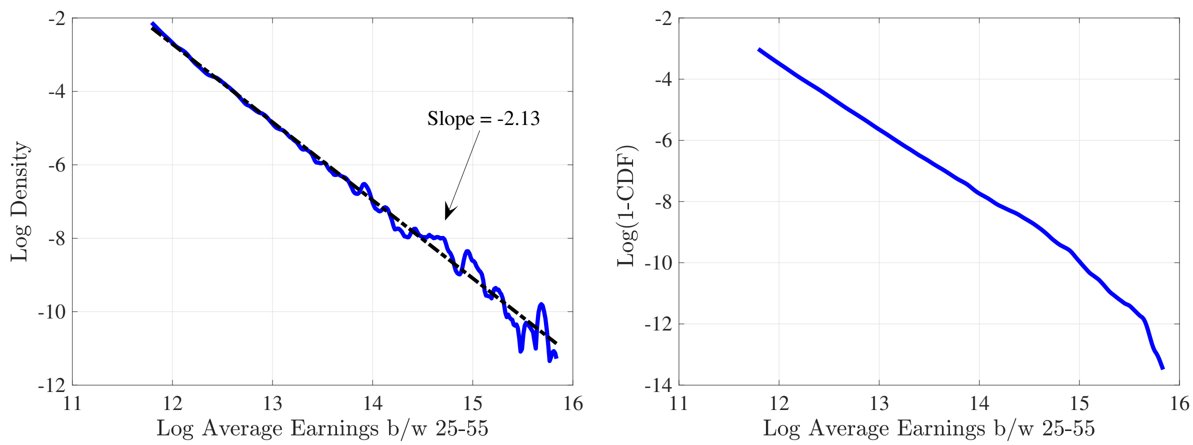


TABLE IV – Sample selection

	# individuals dropped	Size after selection
Initial sample		1,845,640
# yrs self-employed	326,822	1,518,818
# yrs employed	489,504	1,029,314
consecutive nonemployment	161,420	867,894
consecutive self-employment	27,700	840,194

Figure A.2 – Ratios of top income shares

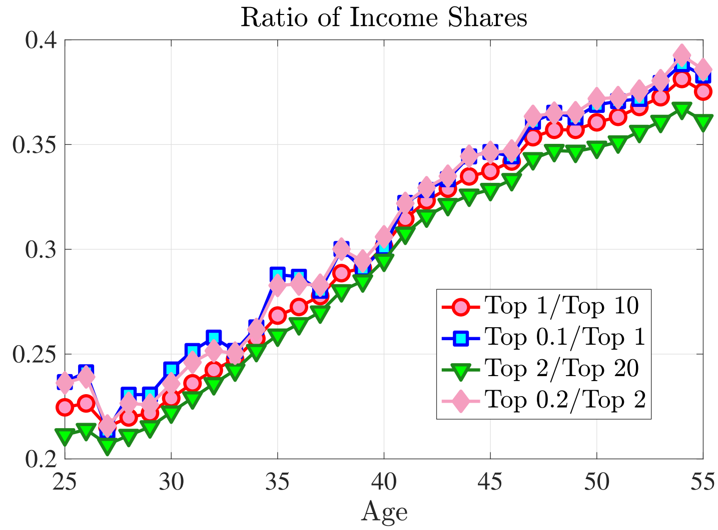


Figure A.3 – Log density of top 5% of within-age earnings distribution

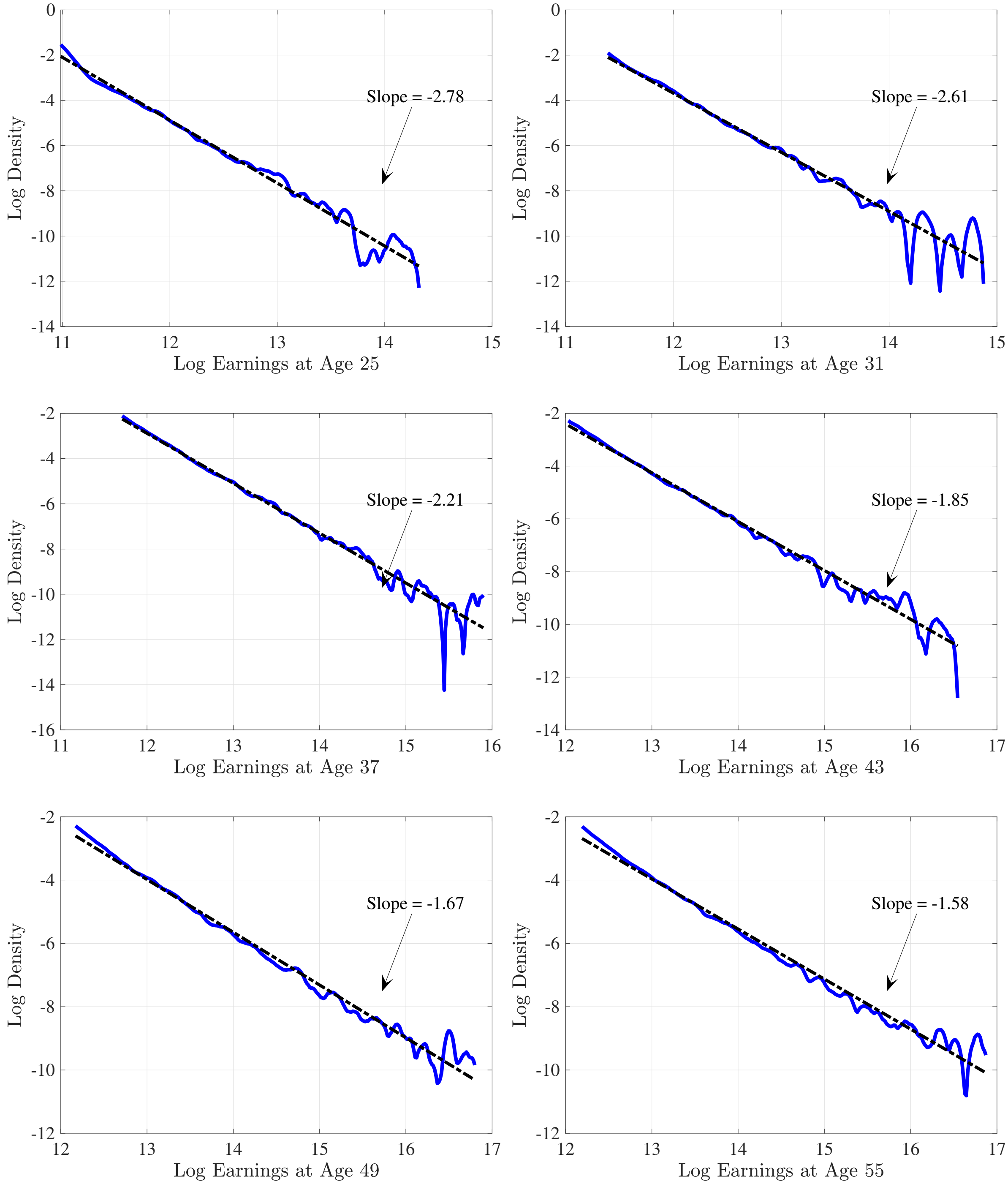
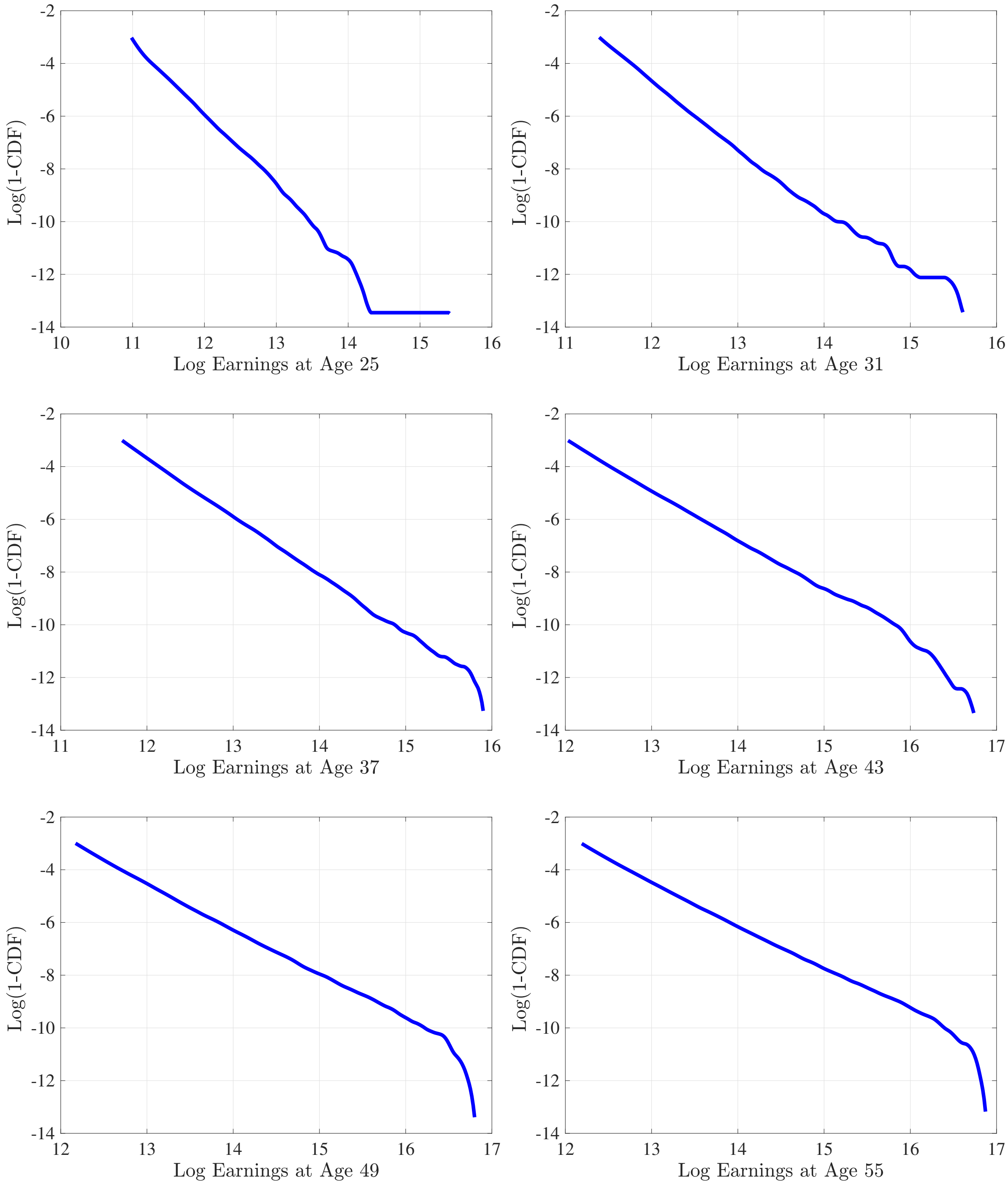


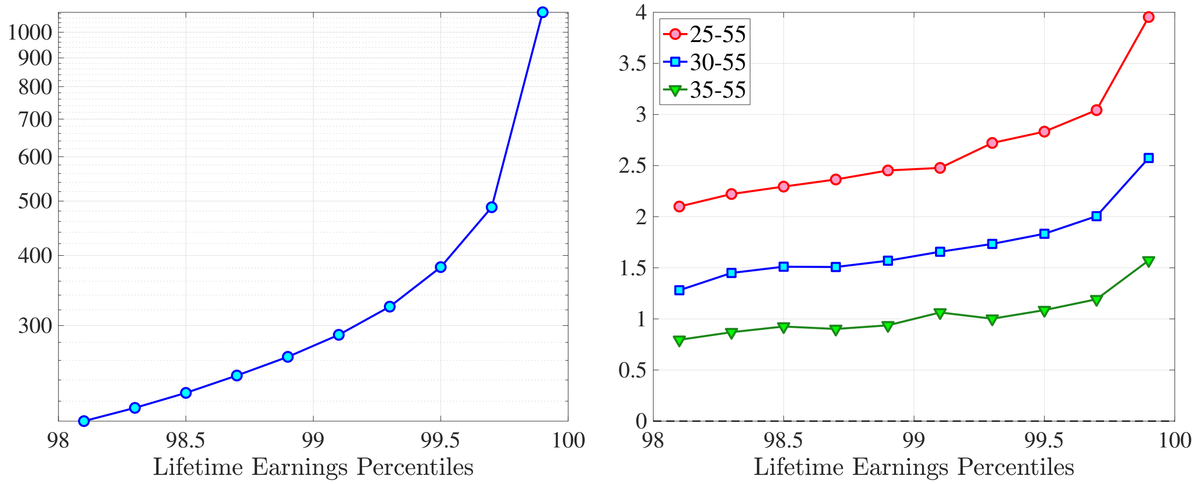
Figure A.4 – Log inverse CDF of top 5% of within-age earnings distribution



A.2.2 Earnings Growth of Top Earners

FIGURE A.5 – Heterogeneity in lifetime earnings growth

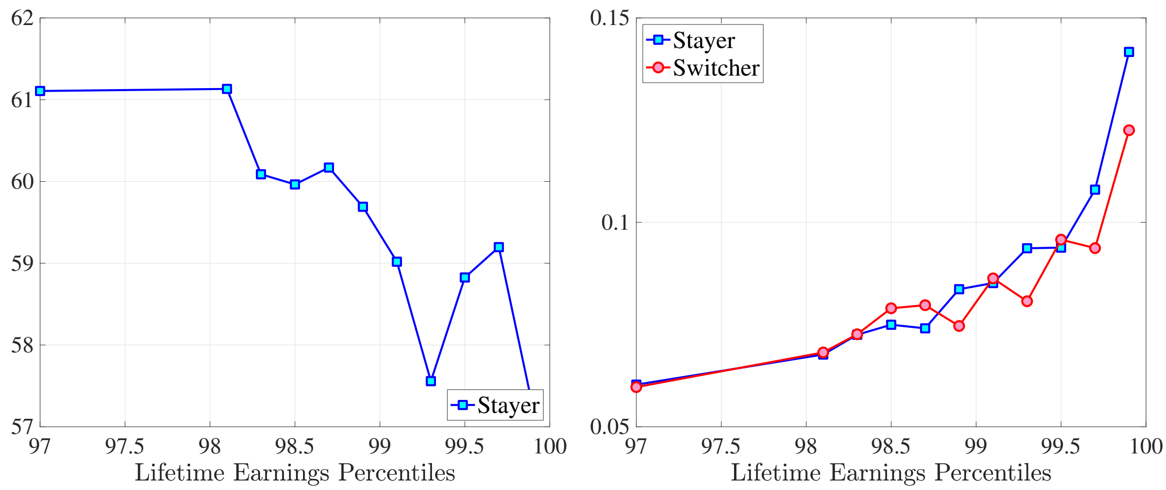
(A) Average earnings over the lifetime, \$1,000 (B) Lifetime earnings growth, $\log \overline{Y}_{55} - \log \overline{Y}_h$



Notes: The left panel shows the average annual earnings over the life cycle for each LE group. The right panel shows the log difference of average earnings $\bar{Y}$ between age 55 and various ages over the LE distribution.

FIGURE A.6 – Job stayers and switchers

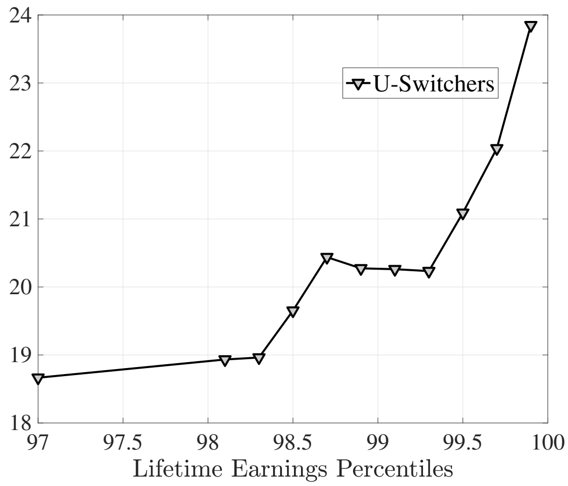
(A) Fraction of job stayers, % (B) Earnings growth, $\log \overline{Y}_{t+1} - \log \overline{Y}_t$



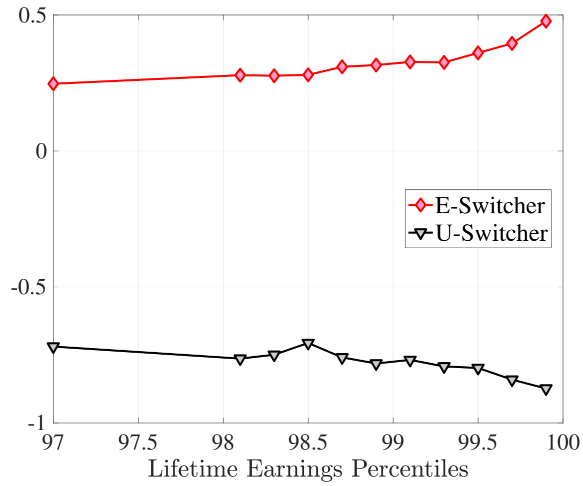
Notes: The left panel shows the fraction of workers in each LE group who are job stayers according to our definition, calculated for each age and averaged over the working life. The right panel plots the log growth of average earnings $\bar{Y}$ between t and $t+1$ for job stayers and switchers separately, again, averaged over t over the working life.

FIGURE A.7 – E-switchers and U-switchers

(A) Share of U-switchers among switchers, %



(B) Earnings growth, $\log \overline{Y}_{t+1} - \log \overline{Y}_t$



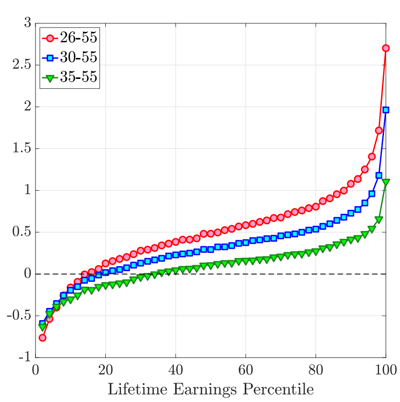
Notes: The left panel shows the share of U-switchers among job stayers in each LE group. The right panel plots the log growth of average earnings $\bar{Y}$ between t and $t+1$ for U-switchers and E-switchers separately.

A.3 Moments for a Broader Sample

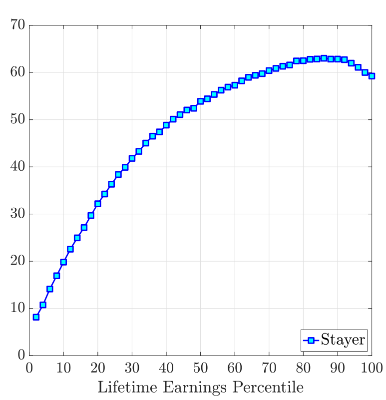
We select individuals for whom we have 33 years of data between ages 25 and 60 over 1978 and 2013. Furthermore, we exclude individuals who do not have earnings above the time-varying minimum earnings threshold for at least 15 years or who are self-employed for more than 8 years over their life cycle.

FIGURE A.8 – Moments from the Broader Sample

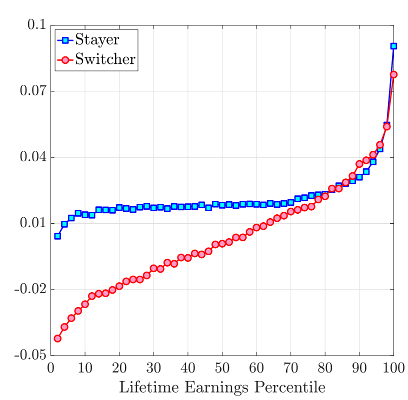
(A) Lifetime earnings growth, $\log \overline{Y}_{55} - \log \overline{Y}_h$



(B) Fraction of job stayers, %



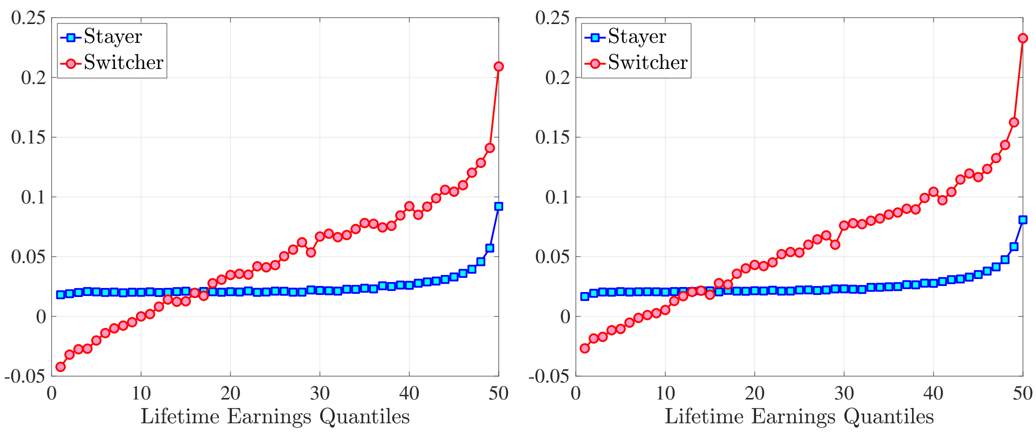
(C) Earnings growth, $\log \overline{Y}_{t+1} - \log \overline{Y}_t$



A.4 Earnings Growth Using Full-Year Employment

FIGURE A.9 – Job stayers and switchers

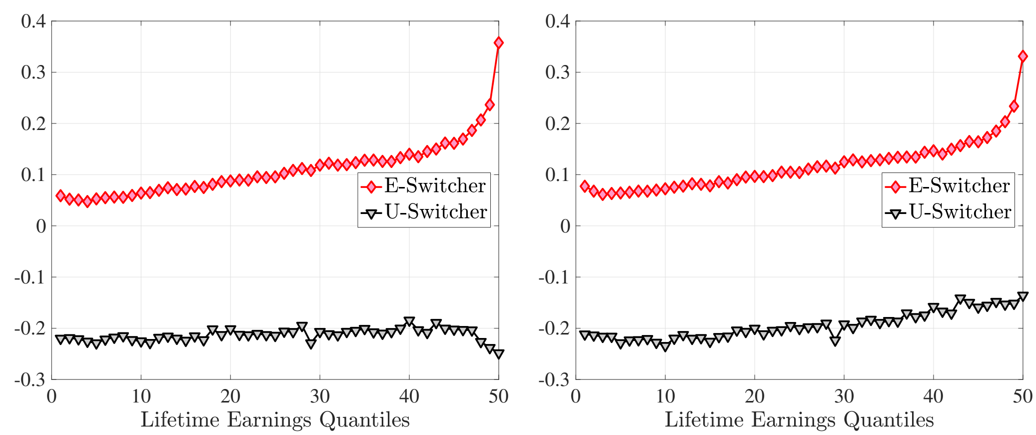
(A) Log average growth, $\log \overline{Y}_{t+1} - \log \overline{Y}_t$ (B) Average log growth, $\mathbb{E}\left[y_{t+1} - y_t\right]$



Notes: The left panel shows the fraction of workers in each LE group who are job stayers according to our definition, calculated for each age and averaged over the working life. The right panel plots the log growth of average earnings $\bar{Y}$ between t and $t+1$ for job stayers and switchers separately, again, averaged over t over the working life.

FIGURE A.10 – E-switchers and U-switchers

(A) Log average growth, $\log \overline{Y}_{t+1} - \log \overline{Y}_t$ (B) Average log growth, $\mathbb{E}\left[y_{t+1} - y_t\right]$



Notes: The left panel shows the share of U-switchers among job stayers in each LE group. The right panel plots the log growth of average earnings $\bar{Y}$ between t and $t+1$ for U-switchers and E-switchers separately.

B Survey of Income and Program Participation

There are two important drawbacks to the SSA data. The first is their annual frequency, which doesn't allow us to see higher frequency movements in earnings. The second is that they do not allow us to condition the outcomes on the labor market status of workers. To supplement the facts documented in the previous section, we use data from the Survey of Income and Program Participation (SIPP), a nationally representative sample of U.S. households. The data consist of monthly observations in overlapping panels with length between 2.5 and 4 years, with the first panel conducted in 1984. Each SIPP panel is conducted in waves, interviewing households every four months about the prior four months. Using data on labor force status, employment rates and labor market transition rates can be computed at a monthly frequency from the SIPP. Similarly, using individual income data, we are able to investigate how these flow rates vary with the level of earnings.[29] We also use the SIPP to compute labor market flow rates for individuals by educational attainment.

B.1 Sample

The SIPP sample is selected in a way that mirrors (to the extent possible) the SSA sample construction. We select males between the ages of 25 and 55. We convert nominal monthly wage data to real using the personal consumption expenditure (PCE) deflator, using 2010 as the base year. We require people to have prior data for at least 36 months and construct their previous income , by summing their monthly real wage over the past 32 months. We residualize this past income by regressing its logarithm on a full set of age and year dummies. We assign individuals into deciles based on this residual.

B.2 Heterogeneity in Labor Market Flows

We compute rates of three types of labor market flows, EU, UE and EE, over a four-month period to deal with seam bias documented in previous work. Observations that report UNU or NUN over three consecutive months are recoded as UUU and NNN, respectively. We use the employer ID to construct job-to-job transitions.

[29] We cannot rank people by their lifetime earnings, since in the SIPP we don't observe the entire earnings history of individuals. Therefore, we condition workers by their average past wages.

C Model Derivations

To the baseline model in Bagger *et al.* (2014), we add a recall option and stochastic aging. Let λ_r denote the probability of recall for unemployed workers. The superscripts y and o refer to young and old workers, respectively. Young workers become old with probability γ. We start by deriving the wage equation for old workers and proceed backwards to solve the same for young workers. These derivations follow closely those in Bagger *et al.* (2014).

Solving piece rates for old workers

Let $V^o(r,h,p)$ denote the value function of an old worker with human capital h employed at a firm with productivity p at (log-) piece rate r. Note that we are suppressing the dependence of the value functions and labor market transitions (δ, λ_0, λ_1) on individual ability α_i and return to experience β_i. For ease of notation, we also suppress the functional form for wages ($w = r + p + h$). Finally, we let $\kappa^o = \xi\lambda_0^o$ and $\kappa^y = \xi\lambda_0^y$ denote the rate at which workers that lose their job in a period find another job immediately within the same period. $V^o(r,h,p)$ is given by

$$
\begin{aligned}
V^o(r,h,p) \quad = \quad & w + \frac{\delta^o(1-\kappa^o)}{1+\rho}V_0^o(h) + \frac{\kappa^o}{1+\rho}\int_p^{\bar{p}} \mathbb{E}\left[(1-\theta)V_0(h) + \theta V^o(0,h',x)\right]dF(x) \\
& + \frac{\lambda_1^o}{1+\rho}\int_p^{\bar{p}} \mathbb{E}\left[(1-\theta)V^o(0,h',p) + \theta V^o(0,h',x)\right]dF(x) \\
& + \frac{\lambda_1^o}{1+\rho}\int_{q^o(r,h,p)}^{p} \mathbb{E}\left[(1-\theta)V^o(0,h',x) + \theta V^o(0,h',p)\right]dF(x) \\
& + \frac{1}{1+\rho}\left[1-\delta^o-\lambda_1^o\bar{F}\left(q^o(r,h,p)\right)\right]\mathbb{E}V^o(r,h',p). \qquad (7)
\end{aligned}
$$

Integrating (7) by parts, we obtain

$$
\begin{aligned}
V^o(r,h,p) \quad = \quad & w + \frac{\delta^o}{1+\rho}V_0^o(h) + \frac{1}{1+\rho}\mathbb{E}\Big\{(1-\delta^o)V^o(r,h',p) \\
& + \lambda_1^o\theta\int_p^{\bar{p}} \frac{\partial V^o}{\partial x}(0,h',x)\bar{F}(x)\,dx \\
& + \lambda_1^o(1-\theta)\int_{q^o(r,h,p)}^{p} \frac{\partial V^o}{\partial x}(0,h',x)\bar{F}(x)\,dx \\
& + \delta\kappa^o\theta\int_{p_{\min}}^{\bar{p}} \frac{\partial V^o}{\partial x}(0,h',x)\bar{F}(x)\,dx\Big\} \qquad (8)
\end{aligned}
$$

Applying (8) with $r=0$, and noting that $q(0,h,p)=p$, we get

$$\begin{aligned} V^o(0,h,p) &= p+h+\frac{\delta^o}{1+\rho}V_0^o(h)+\frac{1}{1+\rho}\mathbb{E}\Bigg\{(1-\delta^o)V^o(0,h',p) \\ &+\lambda_1^o\beta\int_p^{\bar{p}}\frac{\partial V^o}{\partial x}(0,h',x)\bar{F}(x)\,dx \\ &+\delta\kappa^o\theta\int_{p_{\min}}^{\bar{p}}\frac{\partial V^o}{\partial x}(0,h',x)\bar{F}(x)\,dx\Bigg\} \end{aligned}$$

Then, we differentiate this expression with respect to p, to obtain:

$$\frac{\partial V^o}{\partial p}(0,h,p) = 1+\left[\frac{1-\delta^o-\lambda_1^o\theta\bar{F}(p)}{1+\rho}\right]\frac{\partial V^o}{\partial p}(0,h',p).$$

This expression, upon collecting terms yields

$$\frac{\partial V^o}{\partial p}(0,h,p)=\frac{1+\rho}{\rho+\delta^o+\lambda_1^o\theta\bar{F}(p)}.$$

Substituting this expression back in (8), and letting $C^o(p)\equiv\frac{1}{\rho+\delta^o+\lambda_1^o\theta\bar{F}(x)}$, we get

$$\begin{aligned} V^o(r,h,p) &= w+\frac{\delta^o}{1+\rho}V_0^o(h) \\ &+\frac{1}{1+\rho}\mathbb{E}\Bigg\{(1-\delta^o)V^o(r,h',p) \\ &+\lambda_1^o\theta\int_p^{\bar{p}}(1+\rho)C^o(x)\bar{F}(x)\,dx \\ &+\lambda_1^o(1-\theta)\int_{q^o(r,h,p)}^{p}(1+\rho)C^o(x)\bar{F}(x)\,dx \\ &+\delta\kappa^o\theta\int_{p_{\min}}^{\bar{p}}(1+\rho)C^o(x)\bar{F}(x)\,dx\Bigg\} \end{aligned} \tag{9}$$

Note that q^o is defined by the following indifference condition:

$$\mathbb{E}V^o(r,h',p)=\mathbb{E}\left\{V^o(0,h',q^o)+\theta\left[V^o(0,h',p)-V^o(0,h',q^o)\right]\right\} \tag{10}$$

We first rewrite this as follows:

$$\mathbb{E}V^o(r,h',p)-V^o(0,h',q^o)=\theta\mathbb{E}\left\{V^o(0,h',p)-V^o(0,h',q^o)\right\}$$

Substituting (9) into this, and rearranging terms, we obtain

$$
\begin{aligned}
&r + p - q^o(r,h,p) + \frac{1-\delta^o}{1+\rho}\left[V^o(r,h'',p) - V^o(0,h'',q^o)\right] \\
&-\lambda_1^o\theta\int_{q^o}^{p}\frac{(1+\rho)\bar{F}(x)}{\rho+\delta^o+\lambda_1^o\theta\bar{F}(x)}dx + \frac{\lambda_1^o(1-\theta)}{1+\rho}\int_{q^o(r,h,p)}^{p}\frac{(1+\rho)\bar{F}(x)}{\rho+\delta^o+\lambda_1^o\theta\bar{F}(x)}dx = \\
&\mathbb{E}\left\{\theta\left[p - q^o(r,h,p)\right] + \theta\frac{1-\delta^o}{1+\rho}\left[V^o(0,h'',p) - V^o(0,h'',q^o)\right]\right\} \\
&-\lambda_1^o\theta^2\int_{q^o}^{p}\frac{(1+\rho)\bar{F}(x)}{\rho+\delta^o+\lambda_1^o\theta\bar{F}(x)}dx
\end{aligned}
$$

Rearranging terms, we obtain

$$
\begin{aligned}
r &= -(1-\theta)\left[p - q^o(r,h,p)\right] - \lambda_1^o(1-\theta)^2\int_{q(r,h,p)}^{p}\frac{(1+\rho)\bar{F}(x)\,dx}{\rho+\delta^o+\lambda_1^o\theta\bar{F}(x)} \\
&+\frac{1-\delta^o}{1+\rho}\mathbb{E}\left[(1-\theta)V^o(0,h'',q(r,h_t,p)) + \theta V^o(0,h'',p) - V^o(r,h'',p)\right]
\end{aligned}
$$

Substituting the last term with (10) and using the law of iterated expectations, we get

$$
\begin{aligned}
r &= -(1-\theta)\left[p - q^o(r,h,p)\right] - \lambda_1^o(1-\theta)^2\int_{q(r,h,p)}^{p}\frac{(1+\rho)\bar{F}(x)\,dx}{\rho+\delta^o+\lambda_1^o\theta\bar{F}(x)} \\
&+\frac{(1-\delta^o)(1-\theta)}{1+\rho}\mathbb{E}\left[V^o(0,h'',q^o(r,h,p)) - V^o(0,h'',q^o(r,h',p))\right] \\
&= -(1-\theta)\left[p - q^o(r,h,p)\right] - \lambda_1^o(1-\theta)^2\int_{q^o(r,h,p)}^{p}\frac{\bar{F}(x)\,dx}{\rho+\delta^o+\lambda_1^o\theta\bar{F}(x)} \\
&-\frac{(1-\delta^o)(1-\theta)}{1+\rho}\mathbb{E}\int_{q^o(r,h,p)}^{q^o(r,h',p)}\frac{\partial V^o}{\partial p}(0,h'',x)\,dx \\
&= -(1-\theta)\left[p - q^o(r,h,p)\right] - \lambda_1^o(1-\theta)^2\int_{q^o(r,h,p)}^{p}\frac{\bar{F}(x)\,dx}{\rho+\delta^o+\lambda_1^o\theta\bar{F}(x)} \\
&-\frac{(1-\delta^o)(1-\theta)}{1+\rho}\mathbb{E}\int_{q^o(r,h,p)}^{q^o(r,h',p)}\frac{dx}{\rho+\delta^o+\lambda_1^o\theta\bar{F}(x)}.
\end{aligned}
$$

We look for a deterministic solution (constant with respect to h). This solution is implicitly defined by

$$
r = -(1-\theta)\left[p - q^o(r,p)\right] - \lambda_1^o(1-\theta)^2\int_{q^o(r,p)}^{p}C^o(x)\bar{F}(x)\,dx \tag{11}
$$

Solving piece rates for young workers

The value function for young workers is as follows:

$$
\begin{aligned}
V^y(r,h,p) &= w + \frac{\delta^y(1-\kappa^y)}{1+\rho}V_0^y(h) \\
&+ \frac{\kappa^y}{1+\rho}\int_p^{\bar{p}} \mathbb{E}\left[(1-\theta)V_0(h) + \theta V^y(0,h',x)\right]dF(x) \\
&+ \frac{\lambda_1^y}{1+\rho}\int_p^{\bar{p}} \mathbb{E}\left[(1-\theta)V^y(0,h',p) + \theta V^y(0,h',x)\right]dF(x) \\
&+ \frac{\lambda_1^y}{1+\rho}\int_{q^y(r,h,p)}^{p} \mathbb{E}\left[(1-\theta)V^y(0,h',x) + \theta V^y(0,h',p)\right]dF(x) \\
&+ \frac{\gamma^o}{1+\rho}\mathbb{E}V^o(r,h',p) \\
&+ \frac{1}{1+\rho}\left[1-\delta^y-\gamma-\lambda_1^y\bar{F}(q^y(r,h,p))\right]EV^y(r,h',p) \qquad (12)
\end{aligned}
$$

Integrating (12) by parts, we obtain

$$
\begin{aligned}
V^y(r,h,p) &= w + \frac{\delta^y}{1+\rho}V_0^y(h) + \frac{1}{1+\rho}\mathbb{E}\Bigg\{(1-\delta^y-\gamma)V^y(r,h',p) \\
&+ \lambda_1^y\theta\int_p^{\bar{p}} \frac{\partial V^y}{\partial x}(0,h',x)\bar{F}(x)\,dx \\
&+ \lambda_1^y(1-\theta)\int_{q^y(r,h,p)}^{p} \frac{\partial V^y}{\partial x}(0,h',x)\bar{F}(x)\,dx \\
&+ \delta\kappa^y\theta\int_{p_{\min}}^{\bar{p}} \frac{\partial V^y}{\partial x}(0,h',x)\bar{F}(x)\,dx + \gamma\mathbb{E}V^o(r,h',p)\Bigg\} \qquad (13)
\end{aligned}
$$

Substituting the expression for V^o we derived earlier, we get

$$\begin{aligned} V^y(r,h,p) &= w+\frac{\delta^y}{1+\rho}V_0^y(h)+\frac{1}{1+\rho}\mathbb{E}\Bigg\{(1-\delta^y-\gamma)V^y(r,h',p) \\ &+\lambda_1^y\theta\int_p^{\bar{p}}\frac{\partial V^y}{\partial x}(0,h',x)\bar{F}(x)\,dx+\lambda_1^y(1-\theta)\int_{q^y(r,h,p)}^{p}\frac{\partial V^y}{\partial x}(0,h',x)\bar{F}(x)\,dx \\ &+\delta\kappa^y\theta\int_{p_{\min}}^{\bar{p}}\frac{\partial V^y}{\partial x}(0,h',x)\bar{F}(x)\,dx \\ &+\gamma^o\Bigg(w'+\frac{\delta^o}{1+\rho}V_0^o(h'')+\frac{1}{1+\rho}\Bigg\{(1-\delta^o)V^o(r,h'',p) \\ &+\lambda_1^o\theta\int_p^{\bar{p}}\frac{(1+\rho)\bar{F}(x)}{\rho+\delta^o+\lambda_1^o\theta\bar{F}(x)}dx+\lambda_1^o(1-\theta)\int_{q^o(r,h',p)}^{p}\frac{(1+\rho)\bar{F}(x)}{\rho+\delta^o+\lambda_1^o\theta\bar{F}(x)}dx \\ &+\delta\kappa^o\theta\int_{p_{\min}}^{\bar{p}}\frac{(1+\rho)}{\rho+\delta^o+\lambda_1^o\theta\bar{F}(x)}\bar{F}(x)\,dx\Bigg\}\Bigg)\Bigg\} \end{aligned} \tag{14}$$

Now, evaluating this at $r=0$, we get

$$\begin{aligned} V^y(0,h,p) &= p+h+\frac{\delta^y}{1+\rho}V_0^y(h_t)+\frac{1}{1+\rho}\mathbb{E}\Bigg\{(1-\delta^y-\gamma)V(0,h',p) \\ &+\lambda_1^y\theta\int_p^{\bar{p}}\frac{\partial V^y}{\partial x}(0,h',x)\bar{F}(x)\,dx \\ &+\lambda_1^y(1-\theta)\int_{q^y(0,h,p)=p}^{p}\frac{\partial V^y}{\partial x}(0,h',x)\bar{F}(x)\,dx \\ &+\delta\kappa^y\theta\int_{p_{\min}}^{\bar{p}}\frac{\partial V^y}{\partial x}(0,h',x)\bar{F}(x)\,dx \\ &+\gamma\Bigg(w'+\frac{\delta^o}{1+\rho}V_0^o(h') \\ &+\frac{1}{1+\rho}\Bigg\{(1-\delta^o)V^o(0,h'',p)+\lambda_1^o\theta\int_p^{\bar{p}}\frac{(1+\rho)\bar{F}(x)}{\rho^o+\delta^o+\lambda_1^o\theta\bar{F}(x)}dx \\ &+\lambda_1^o(1-\theta)\int_{q^o(0,h',p)=p}^{p}\frac{(1+\rho)\bar{F}(x)}{\rho+\delta^o+\lambda_1^o\theta\bar{F}(x)}dx \\ &+\delta\kappa^o\theta\int_{p_{\min}}^{\bar{p}}\frac{(1+\rho)}{\rho+\delta^o+\lambda_1^o\theta\bar{F}(x)}\bar{F}(x)\,dx\Bigg\}\Bigg)\Bigg\} \end{aligned}$$

This boils down to

$$
\begin{aligned}
V^y(0,h,p) &= p+h+\frac{\delta^y}{1+\rho}V_0^y(h)+\frac{1}{1+\rho}\mathbb{E}\Bigg\{(1-\delta^y-\gamma)V^y(0,h',p) \\
&+\lambda_1^y\theta\int_p^{\bar{p}}\frac{\partial V^y}{\partial x}(0,h',x)\bar{F}(x)\,dx+\delta\kappa^y\theta\int_{p_{\min}}^{\bar{p}}\frac{\partial V^y}{\partial x}(0,h',x)\bar{F}(x)\,dx \\
&+\gamma\Bigg(p+h'+\frac{\delta^o}{1+\rho}V_0^o(h')+\frac{1}{1+\rho}\Bigg\{(1-\delta^o)V^o(0,h'',p) \\
&+\lambda_1^o\theta\int_p^{\bar{p}}\frac{(1+\rho)\bar{F}(x)}{\rho+\delta^o+\lambda_1^o\theta\bar{F}(x)}dx+\delta\kappa^o\theta\int_{p_{\min}}^{\bar{p}}\frac{(1+\rho)}{\rho+\delta^o+\lambda_1^o\theta\bar{F}(x)}\bar{F}(x)\,dx\Bigg)\Bigg\}\Bigg\}
\end{aligned}
$$

Differentiating this with respect to p, we ge

$$
\begin{aligned}
\frac{\partial V^y}{\partial p}(0,h,p) &= 1+\frac{1}{1+\rho}\mathbb{E}\Bigg\{(1-\delta^y-\gamma)\frac{\partial V^y}{\partial p}(0,h',p)-\lambda_1^y\theta\frac{\partial V^y}{\partial p}(0,h',p)\bar{F}(p) \\
&+\gamma\Bigg(1+\frac{1}{1+\rho}\Bigg\{(1-\delta^o)\frac{\partial V^o}{\partial p}(0,h'',p)-\lambda_1^o\theta\frac{(1+\rho)\bar{F}(p)}{\rho+\delta^o+\lambda_1^o\theta\bar{F}(p)}\Bigg\}\Bigg)\Bigg\}
\end{aligned}
$$

Plugging the expression for $\frac{\partial V^o}{\partial p}$ into here, we obtain

$$
\begin{aligned}
\frac{\partial V^y}{\partial p}(0,h,p) &= 1+\frac{1}{1+\rho}\mathbb{E}\Bigg\{(1-\delta^y-\gamma)\frac{\partial V^y}{\partial p}(0,h',p)-\lambda_1^y\theta\frac{\partial V^y}{\partial p}(0,h',p)\bar{F}(p) \\
&+\gamma\Bigg(1+\frac{1}{1+\rho}\Bigg\{(1-\delta^o)\frac{1+\rho}{\rho+\delta^o+\lambda_1^o\theta\bar{F}(p)}-\lambda_1^o\theta\frac{(1+\rho)\bar{F}(p)}{\rho+\delta^o+\lambda_1^o\theta\bar{F}(p)}\Bigg\}\Bigg)\Bigg\}
\end{aligned}
$$

Collecting terms, we get

$$
\begin{aligned}
\frac{\partial V^y}{\partial p}(0,h,p) &= 1+\frac{1}{1+\rho}\mathbb{E}\Bigg\{\left(1-\delta^y-\gamma-\lambda_1^y\theta\bar{F}(p)\right)\frac{\partial V^y}{\partial p}(0,h',p) \\
&+\gamma\Bigg(1+\frac{1}{1+\rho}\Bigg\{(1-\delta^o)\frac{1+\rho}{\rho+\delta^o+\lambda_1^o\theta\bar{F}(p)} \\
&-\lambda_1^o\theta\frac{(1+\rho)\bar{F}(p)}{\rho+\delta^o+\lambda_1^o\theta\bar{F}(p)}\Bigg\}\Bigg)\Bigg\}
\end{aligned}
$$

Assume that $\frac{\partial V^y}{\partial p}(0,h,p)$ is independent of h (since we are looking for a solution independent of h). This means we can drop the expectation operator on the left-hand side. Then, we get

$$\begin{aligned}\frac{\partial V^y}{\partial p} &= 1+\frac{1}{1+\rho}\Big\{\left(1-\delta^y-\gamma-\lambda_1^y\theta\bar{F}(p)\right)\frac{\partial V^y}{\partial p}\\&+\gamma\left(1+\frac{1}{1+\rho}\left(1-\delta^o-\lambda_1^o\theta\bar{F}(p)\right)\frac{\partial V^o}{\partial p}\right)\Big\}\end{aligned}$$

$$\begin{aligned}\frac{\partial V^y}{\partial p}(0,h,p) &= \frac{1+\frac{\gamma}{1+\rho}\left(1+\frac{1}{1+\rho}\Big\{\left(1-\delta^o-\lambda_1^o\theta\bar{F}(p)\right)\frac{\partial V^o}{\partial p}\Big\}\right)}{1-\frac{1-\delta^y-\gamma}{1+\rho}+\frac{\lambda_1^y\theta\bar{F}(p)}{1+\rho}}\\
&=(1+\rho)\frac{1+\frac{\gamma}{1+\rho}\left(1+\frac{1}{1+\rho}\Big\{(1-\delta^o)\frac{1+\rho}{\rho+\delta^o+\lambda_1^o\theta\bar{F}(p)}-\lambda_1^o\theta\frac{(1+\rho)\bar{F}(p)}{\rho+\delta^o+\lambda_1^o\theta\bar{F}(p)}\Big\}\right)}{\rho+\delta^y+\gamma+\lambda_1^y\theta\bar{F}(p)}\\
&=(1+\rho)\frac{1+\frac{\gamma}{1+\rho}\left(1+\frac{1-\delta^o}{\rho+\delta^o+\lambda_1^o\theta\bar{F}(p)}-\frac{\lambda_1^o\theta\bar{F}(p)}{\rho+\delta^o+\lambda_1^o\theta\bar{F}(p)}\right)}{\rho+\delta^y+\gamma+\lambda_1^y\theta\bar{F}(p)}\\
&=(1+\rho)\frac{1+\frac{\gamma}{\rho+\delta^o+\lambda_1^o\theta\bar{F}(p)}}{\rho+\delta^y+\gamma+\lambda_1^y\theta\bar{F}(p)}\equiv(1+\rho)\,C^y(p) \qquad (15)\end{aligned}$$

Then, we substitute the expression for $\frac{\partial V^y}{\partial x}(0,h,x)$ into (14), and we obtain

$$\begin{aligned}V^y(r,h,p) &= w+\frac{\delta^y}{1+\rho}V_0^y(h_t)+\frac{1}{1+\rho}\mathbb{E}\Big\{(1-\delta^y-\gamma)V^y(r,h',p)\\
&+\lambda_1^y\theta(1+\rho)\int_p^{\bar{p}}C^y(x)\bar{F}(x)\,dx+\lambda_1^y(1-\theta)(1+\rho)\int_{q^y(r,h,p)}^{p}C^y(x)\bar{F}(x)\,dx\\
&+\delta\kappa^y\theta(1+\rho)\int_{p_{\min}}^{\bar{p}}C^y(x)\bar{F}(x)\,dx\\
&+\gamma\Big(w'+\frac{\delta^o}{1+\rho}V_0^o(h')+\frac{1}{1+\rho}\Big\{(1-\delta^o)V^o(r,h',p)\\
&+\lambda_1^o\theta(1+\rho)\int_p^{\bar{p}}C^o(x)\bar{F}(x)\,dx+\lambda_1^o(1-\theta)(1+\rho)\int_{q^o(r,h',p)}^{p}C^o(x)\bar{F}(x)\,dx\\
&+\delta\kappa^o\theta\int_{p_{\min}}^{\bar{p}}C^o(x)\bar{F}(x)\,dx\Big\}\Big)\Big\} \qquad (16)\end{aligned}$$

Now, to obtain the equation that implicitly defines q^m, we need to combine (16) with

(10) and arrange terms. But first, we rewrite equation (10).

$$\mathbb{E}\left\{V^y\left(r,h',p\right)-V^y\left(0,h',q^y\left(r,h,p\right)\right)\right\}=\theta\mathbb{E}\left\{V^y\left(0,h',p\right)-V^y\left(0,h',q^y\left(r,h,p\right)\right)\right\}.$$

Combining equation (16) with the expression above, and rearranging terms, we obtain

$$\begin{aligned}
\mathbb{E}\left\{V^y\left(r,h',p\right)-V^y\left(0,h',q^y(r,h,p)\right)\right\} = & \, p+r-q^y(r,h,p) \\
& +\frac{1-\delta^y-\gamma}{1+\rho}\mathbb{E}\left[V^y\left(r,h'',p\right)-V^y\left(0,h'',q^y(r,h,p)\right)\right] \\
& -\lambda_1^y\beta\int_{q^y(r,h,p)}^{p}C^y\left(x\right)\bar{F}\left(x\right)dx \\
& +\lambda_1^y\left(1-\beta\right)\int_{q^y(r,h,p)}^{p}C^y\left(x\right)\bar{F}\left(x\right)dx \\
& +\frac{\gamma}{1+\rho}\left[p+r-q^y(r,h,p)\right] \\
& +\frac{\gamma\left(1-\delta^o\right)}{\left(1+\rho\right)^2}\mathbb{E}\left[V^o\left(r,h'',p\right)-V^o\left(0,h',q^y\right)\right] \\
& -\frac{\gamma\lambda_1^o\beta}{1+\rho}\int_{q^y}^{p}C^o\left(x\right)\bar{F}\left(x\right)dx \\
& \left.+\frac{\gamma\lambda_1^o\left(1-\beta\right)}{1+\rho}\int_{q^o(r,h',p)}^{p}C^o\left(x\right)\bar{F}\left(x\right)dx\right) = \\
\beta\mathbb{E}\left\{V^y\left(0,h',p\right)-V^y\left(0,h',q^y(r,h,p)\right)\right\} = & \, \beta\left[p-q^y(r,h,p)\right]+\beta\frac{\gamma}{1+\rho}\left[p-q^y(r,h,p)\right] \\
& +\beta\frac{1-\delta^y-\gamma}{1+\rho}\mathbb{E}\left[V^y\left(0,h'',p\right)-V^y\left(0,h'',q^y(r,h,p)\right)\right] \\
& -\lambda_1^y\beta^2\int_{q^y}^{p}C^y\left(x\right)\bar{F}\left(x\right)dx \\
& \left.+\beta\frac{\gamma\left(1-\delta^o\right)}{\left(1+\rho\right)^2}\mathbb{E}\left[V^o\left(0,h'',p\right)-V^o\left(0,h'',q^y(r,h,p)\right)\right]\right) \\
& -\frac{\gamma\lambda_1^o\beta^2}{1+\rho}\int_{q^y}^{p}C^o\left(x\right)\bar{F}\left(x\right)dx
\end{aligned}$$

We now collect terms and obtain

$$
\begin{aligned}
r\left(1+\frac{\gamma}{1+\rho}\right) = & -\left(1+\frac{\gamma}{1+\rho}\right)(1-\beta)\left[p-q^{y}(r,h,p)\right] \\
& -\lambda_{1}^{y}(1-\beta)^{2}\int_{q^{y}(r,h,p)}^{p} C^{y}(x)\,\bar{F}(x)\,dx \\
& +\frac{\gamma\lambda_{1}^{o}\beta}{1+\rho}(1-\beta)\int_{q^{y}(r,h,p)}^{p} C^{o}(x)\,\bar{F}(x)\,dx \\
& -\frac{\gamma\lambda_{1}^{o}(1-\beta)}{1+\rho}\int_{q^{o}(r,h',p)}^{p} C^{o}(x)\,\bar{F}(x)\,dx \\
& +\frac{1-\delta^{y}-\gamma}{1+\rho}\Big[(1-\beta)\,V^{y}\left(0,h'',q^{y}(r,h,p)\right) \\
& \qquad +\beta V^{y}(0,h'',p)-V^{y}(r,h'',p)\Big] \\
& +\frac{\gamma(1-\delta^{o})}{(1+\rho)^{2}}\Big[(1-\beta)\,V^{o}\left(0,h'',q^{y}(r,h,p)\right) \\
& \qquad +\beta V^{o}(0,h'',p)-V^{o}(r,h',p)\Big]
\end{aligned}
$$

Noting that 1) $\mathbb{E}\left[\beta V^{y}(0,h'',p)-V^{y}(r,h'',p)\right]$ equals $-(1-\beta)\,\mathbb{E}V^{y}\left(0,h'',q^{y}(r,h',p)\right)$, 2) and $\mathbb{E}\left[\beta V^{o}(0,h'',p)-V^{o}(r,h'',p)\right]$ equals $-(1-\beta)\,\mathbb{E}V^{o}\left(0,h'',q^{o}(r,h',p)\right)$, and plugging these into the expression above, we obtain

$$
\begin{aligned}
r\left(1+\frac{\gamma^{o}}{1+\rho}\right) = & -\left(1+\frac{\gamma^{o}}{1+\rho}\right)(1-\beta)\left[p-q^{y}\right] \\
& -\lambda_{1}^{y}(1-\beta)^{2}\int_{q^{y}(r,h,p)}^{p} C^{y}(x)\,\bar{F}(x)\,dx \\
& +\frac{\gamma\lambda_{1}^{o}\beta}{1+\rho}(1-\beta)\int_{q^{y}(r,h,p)}^{p} C^{o}(x)\,\bar{F}(x)\,dx \\
& -\frac{\gamma\lambda_{1}^{o}(1-\beta)}{1+\rho}\int_{q^{o}(r,h',p)}^{p} C^{o}(x)\,\bar{F}(x)\,dx \\
& +\frac{1-\delta^{y}-\gamma}{1+\rho}(1-\beta)\,\mathbb{E}\Big[V^{y}\left(0,h'',q^{y}(r,h,p)\right)-V^{y}\left(0,h'',q^{y}(r,h'',p)\right)\Big] \\
& +\frac{\gamma(1-\delta^{o})}{(1+\rho)^{2}}(1-\beta)\,\mathbb{E}\Big[V^{o}\left(0,h'',q^{y}(r,h,p)\right)-V^{o}\left(0,h'',q^{o}(r,h',p)\right)\Big]
\end{aligned}
$$

Further rearranging and algebra yields,

$$
\begin{aligned}
r\left(1+\frac{\gamma}{1+\rho}\right) = & -\left(1+\frac{\gamma}{1+\rho}\right)(1-\beta)\left[p-q^{y}(r,h,p)\right] \\
& -\lambda_1^y (1-\beta)^2 \int_{q^y(r,h,p)}^{p} C^y(x)\,\bar{F}(x)\,dx \\
& +\frac{\gamma\lambda_1^o\beta}{1+\rho}(1-\beta)\int_{q^y(r,h,p)}^{p} C^o(x)\,\bar{F}(x)\,dx \\
& -\frac{\gamma\lambda_1^o(1-\beta)}{1+\rho}\int_{q^o(r,h',p)}^{p} C^o(x)\,\bar{F}(x)\,dx \\
& -\frac{1-\delta^y-\gamma}{1+\rho}(1-\beta)\,\mathbb{E}\int_{q^y(r,h,p)}^{q^y(r,h',p)} \frac{\partial V^y}{\partial x}(0,h'',x)\,dx \\
& -\frac{\gamma(1-\delta^o)}{(1+\rho)^2}(1-\beta)\,\mathbb{E}\int_{q^y(r,h,p)}^{q^o(r,h',p)} \frac{\partial V^o}{\partial x}(0,h'',x)\,dx
\end{aligned}
$$

Recall that we ignore solutions that depend on h and look for deterministic solutions instead. This means that the next to last line evaluates to 0. Since this also implies that the functions q^y and q^o depend on h in a trivial way, we drop those from the notation. Equation (17) can be solved numerically to obtain q^y.

$$
\begin{aligned}
r\left(1+\frac{\gamma}{1+\rho}\right) = & -\left(1+\frac{\gamma}{1+\rho}\right)(1-\beta)\left[p-q^{y}(r,p)\right] \\
& -\lambda_1^y (1-\beta)^2 \int_{q^y(r,p)}^{p} C^y(x)\,\bar{F}(x)\,dx \\
& +\frac{\gamma\lambda_1^o\beta}{1+\rho}(1-\beta)\int_{q^y(r,p)}^{p} C^o(x)\,\bar{F}(x)\,dx \\
& -\frac{\gamma\lambda_1^o}{1+\rho}(1-\beta)\int_{q^o(r,p)}^{p} C^o(x)\,\bar{F}(x)\,dx \\
& -\frac{\gamma(1-\delta^o)}{1+\rho}(1-\beta)\int_{q^y(r,p)}^{q^o(r,p)} C^o(x)\,\bar{F}(x)\,dx
\end{aligned} \tag{17}
$$

D Estimation

Table D.1 shows the parameter estimates.

Table D.1 – Parameter values

Parameter	Value	Explanation
g_t, constant	0.94	Deterministic profile
g_t, linear	0.31	Deterministic profile
g_t, quadratic	-0.06	Deterministic profile
σ_α	0.26	Worker type variance
χ_w	6.92	Shape parameter of β
ψ_w	0.09	Scale parameter of β
$\sigma_{\alpha\beta}$	0.44	Correlation b/w α and β
χ_F	6.3	Shape parameter of firm productivity
ψ_F	1.0	Scale parameter of firm productivity
σ_ϵ	0.51	Variance of idiosyncratic productivity shocks
λ_r		Recall productivity
ξ	0.4	Reallocation probability

D.1 Targeted moments in the estimation

In section 2, we show the fit of the model to selected targets by LE averaged over age groups. In this section we now show the fit by age and LE.

Figure D.1 – Fraction of job stayers, E- and U-switchers by LE and age groups

(A) Fraction of stayers

(B) Fraction of E-switchers

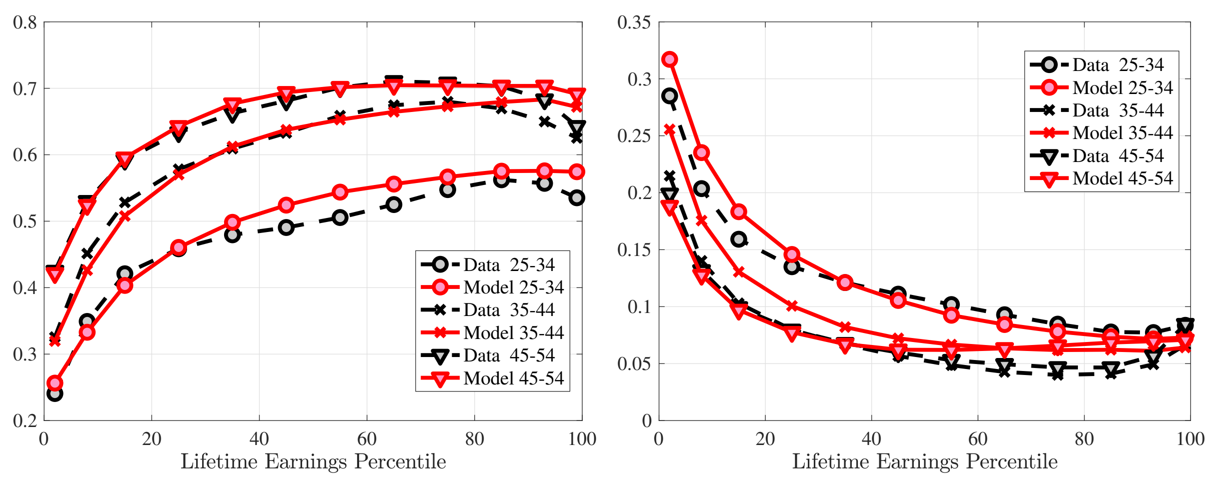


Figure D.2 – Cross-sectional moments of earnings growth for job stayers and switchers

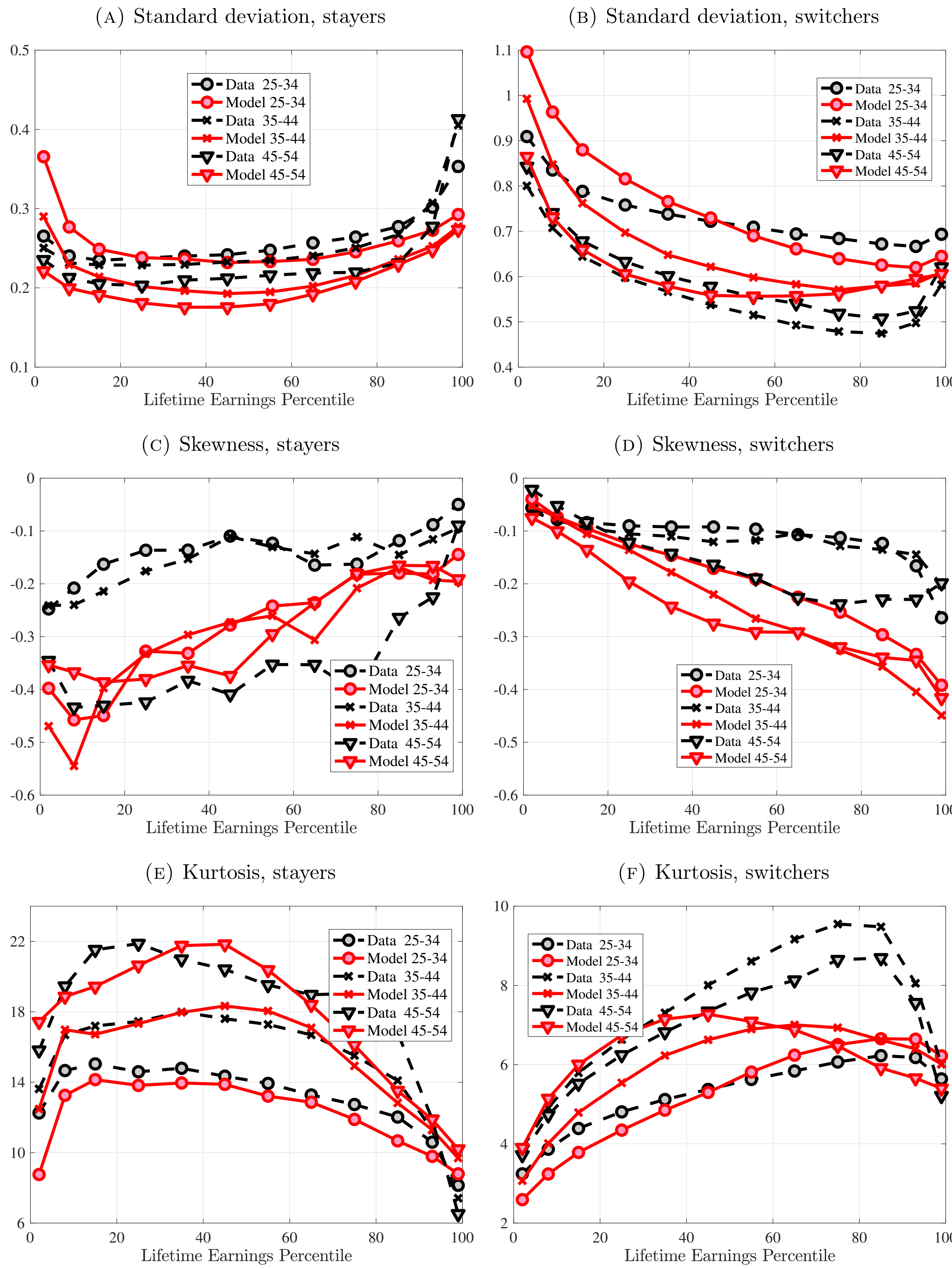
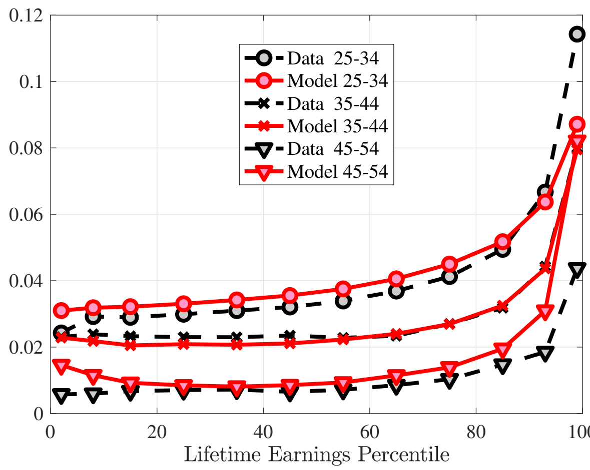
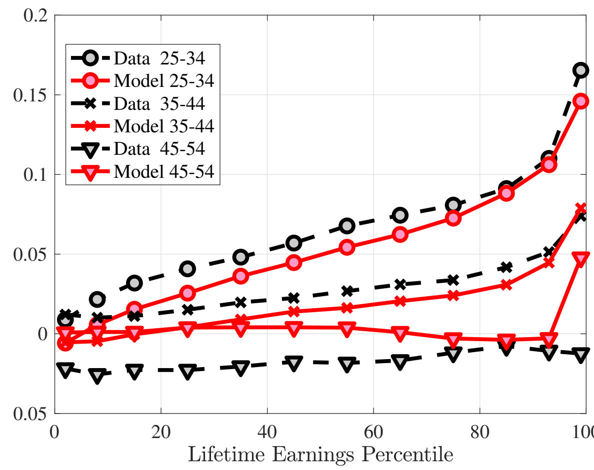


Figure D.3 – Earnings growth of job stayers, E- and U-switchers by LE and age groups

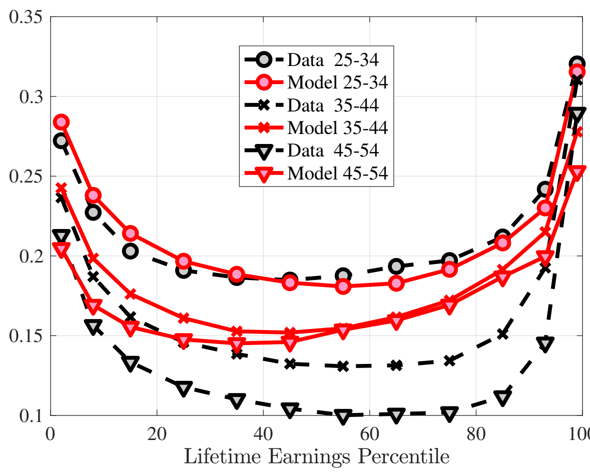
(a) Earnings growth of stayers



(b) Earnings growth of switchers



(c) Earnings growth of E-switchers



(d) Earnings growth of U-switchers

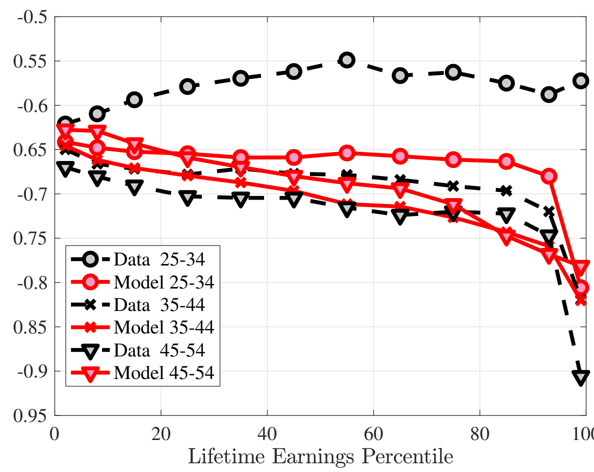
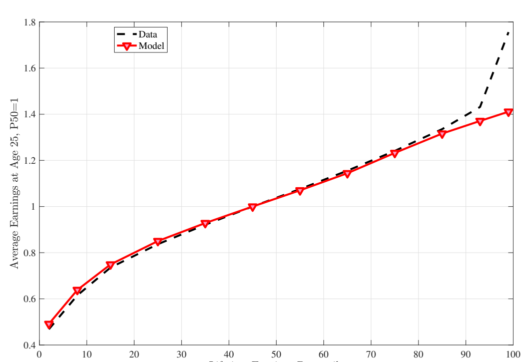


Figure D.4 – Earnings levels by LE groups at age 25



D.2 Additional results

Figure D.5 – Idiosyncratic Shock probability

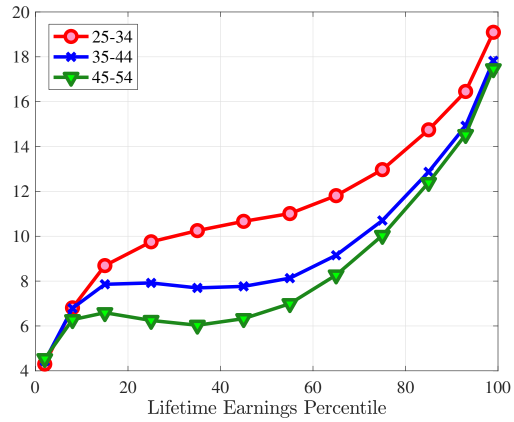
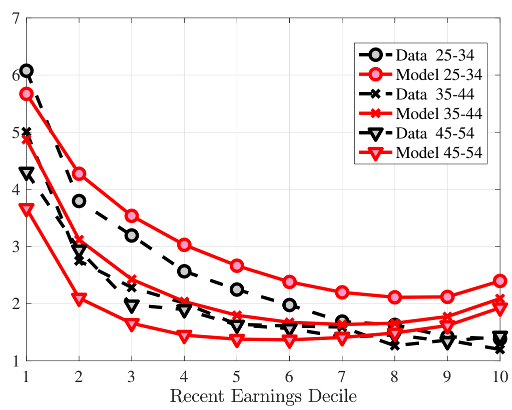
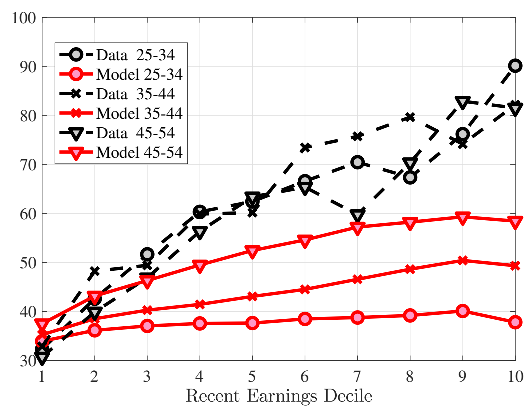


Figure D.6 – Model vs. SIPP Data: Labor market flows with age variation

(a) EU–rate: Model vs. SIPP



(b) UE–rate: Model vs. SIPP



(c) EE–rate: Model vs. SIPP

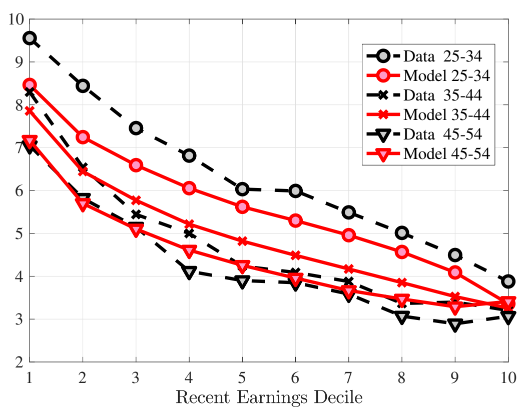
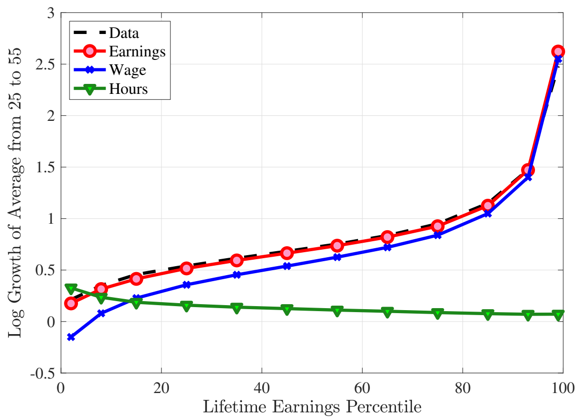
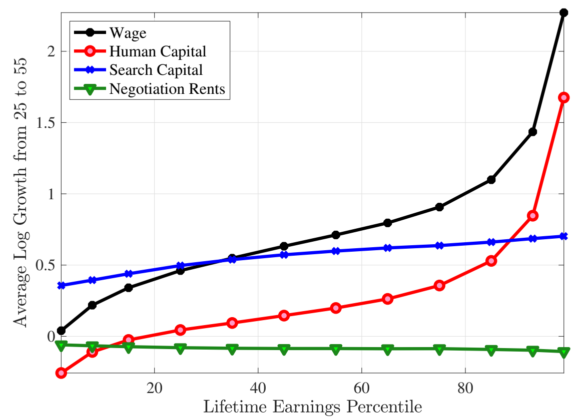


Figure D.7 – Decomposing Earnings and Wage Growth

(A) Earnings, wage and hours growth



(B) Human capital, search capital, negotiation rents



Notes: Notice that the wage growth in the left panel is the log growth of average and in the right panel it is the average log growth of wage. This is because the decomposition in the right panel is only possible when log growth is decomposed.