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

Fatih Karahan† Serdar Ozkan‡ Jae Song§

November 20, 2019

Abstract

We study the determinants of lifetime earnings (LE) inequality in the US using administrative balanced panel data between ages 25 and 55 by focusing on the roles of job ladder dynamics and on-the-job learning. Empirically, we first document that lower LE workers switch jobs more often, mainly driven by higher nonemployment risk. Second, average annual earnings growth for job stayers is surprisingly similar at around 2% in the bottom two thirds of the LE distribution, whereas it rises for job switchers almost linearly from zero at the bottom to around 4% at the 90th percentile. Third, top decile LE workers enjoy high earnings growth regardless of job switching. To interpret these facts, we estimate a life-cycle job ladder model with on-the-job learning featuring ex-ante heterogeneity in learning ability as well as in job ladder risk—job loss, job finding, and contact rates. We find that learning ability is Pareto distributed and explains almost all earnings growth heterogeneity above the median LE—the main driver of LE inequality. This feature of the model is also consistent with the Pareto tails of within-age earnings distributions seen in the data. As for the differences below the median LE, we find that 80% of lifetime earnings growth differences would vanish if they had the same ex-ante job ladder risk. We also validated this large ex-ante heterogeneity in job-ladder risk with direct evidence on worker flow rates from the survey data. We find very little role for ex-post productivity and job-ladder shocks in LE inequality.

Keywords: Job ladder, search frictions, life-cycle earnings risk, life time earning inequality, Pareto tails, fractal top inequality

*This project started joint with Samuel Kapon, whom we are grateful. Special thanks to Gerald Ray and Pat Jonas for their help and support. We also thank Rachel Schuh for the excellent research assistance. For insightful comments and discussions, we thank Fatih Guvenen, Henry Hyatt, Rasmus Lentz, Jeremy Lise, Giuseppe Moscarini, Ayşegül Şahin, Chris Tonetti, Ludo Visschers, and David Wiczer. An earlier version of this paper was circulated under the title "Sources of Earnings Growth Heterogeneity" and "Sources of Inequality in Earnings Growth over the Life Cycle." The views expressed herein are those of the authors and do not represent those of the Social Security Administration, the Federal Reserve Bank of New York, or the Federal Reserve System.

†FRB of New York; `fatih.karahan@ny.frb.org; sites.google.com/site/yfatihkarahan/`

‡University of Toronto; `serdar.ozkan@utoronto.ca; www.serdarozkan.me`

§Social Security Administration; `jae.song@ssa.gov`

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 lifetime earnings inequality. In this paper, we study these differences using administrative balanced panel data by focusing on the roles of heterogeneity in ability to accumulate human capital (e.g., Huggett *et al.* (2011)) and in ability to climb up the job ladder due to ex-ante differences in labor market risk such as job loss, job finding rate, or contact rate (e.g. Bagger *et al.* (2014)).[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 to 2013 from the U.S. Social Security Administration (SSA). Using a 10% sample of cohorts born between 1951 to 1957 and with a serious labor force attachment, we first compute each worker's total labor income between ages 25 to 55 and rank them into 50 lifetime earnings (LE) quantiles. Those at the 90th (top 2) percentile of the lifetime earnings 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 top 0.1% (1%) accounting for around 29% of total LE among top 1% (10%) of the population. Furthermore, a 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, whereas median workers experience a two-fold increase and those at the bottom see little to no earnings growth.

To shed light on differences in career paths of different LE groups we next document their life cycle job switching patterns. First, on average over the life cycle around 30% of the bottom LE workers stay with the same employer in two full consecutive years, compared to around 60% above the median. Resonating with these large differences, people at the bottom of the LE distribution work for about 12 different employers between

[1]Some other possible explanations of wage growth heterogeneity include firms and workers learning about workers' abilities (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 of earnings distribution.

the ages of 25 and 55; more than twice as many as those at the top.

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 very 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 2 percent 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 are 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 together imply that differences in lifetime earnings growth between workers in the bottom half of the LE distribution has to be coming from switcher growth differences whereas for workers in the top half stayer growth differences should be the main culprit as they rarely switch employers. Next, building on this intuition we develop a quantitative structural model of job ladder to disentangle the various economic forces that shape the distribution of wage growth of job stayers and switchers. In particular, we develop and estimate a job ladder model in the spirit of Jarosch (2015), Cahuc *et al.* (2006) and Bagger *et al.* (2014). The model features heterogeneity in worker and firm fixed effects, on the job search, employer competition, and idiosyncratic shocks to worker productivity. We add to this framework an explicit life cycle structure in the form of perpetual youth (à la Blanchard (1985) and Yaari (1965)). More importantly, we allow for rich worker heterogeneity in unemployment risk, job finding rate and the contact rate for employed workers, as well as the ability to learn on-the-job, which can be thought of as differences in the returns to experience. 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. These moments include the cross-sectional distribution of earnings changes. Specifically, we target the variance, left-skewness, and excess kurtosis of annual earnings changes conditional on various age and LE groups, and separately for job stayers and switchers. We also target the fraction and wage growth of job stayers and switchers by lifetime earnings groups and over the life cycle. The model matches well the targeted moments.

We argue in Section 4.2 that these targeted moments allow us to identify the importance of human capital and job ladder risk. While all features of the data are relevant, the key insight relies on realizing that the earnings changes for job switchers is very different than that of stayers. If job ladder risk is not important for wages, then one would observe job stayers and switchers to experience a 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 are informative about the nature of the job ladder risk.

One key finding from our estimation is the vast ex-ante heterogeneity in unemployment risk, job finding rate and the contact rate for employed workers, which the model infers from the large variation in the moments over the LE distribution. In particular, we estimate a 10% quarterly job loss risk for young bottom LE workers, compared to around 2% for older workers above the median. Similarly, job finding rates display large differences, ranging from around 25% for young workers at the bottom to 50% among older workers above the median LE. Given the annual nature of the SSA data, we cannot directly test these estimated of labor market flows. Therefore, we analyze data from the Survey of Income and Program Participation (SIPP) to find large differences in job loss and finding rates among workers with different past earnings and over the life cycle. As an external validation, our model can capture this heterogeneity in worker flow rates in the SIPP data. Turning to the offer arrival rate for employed workers, we find that young workers at the bottom of the LE distribution are contacted with around 30% probability in a quarter. The same figure is more than 55% for older top LE workers. In contrast, SIPP data shows that workers with high past earnings are *less* likely to make job-to-job transitions. Surprisingly, 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 by 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. In the model, differences in lifetime earnings are remarkably similar to differences in lifetime wages, implying that wage—rather than employment—differences explain the vast majority in 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 unemployment risk and 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 lifetime wages, we find them to be driven by wage growth over the life cycle, resonating our finding in the data on lifetime earnings inequality and earnings growth. Therefore, we focus on decomposing the differences in wage growth over the life cycle in the last part of the paper. In a series of experiments, we isolate the relative roles of ex ante differences in the job ladder risk and the returns to experience.

Unemployment risk accounts for about 50% of the wage growth differences between the bottom and median. High unemployment rates among low LE workers reduces wage growth by preventing them from accumulating human capital and from climbing the job ladder. We find that the former channel is stronger, 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 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, do not explain much of the wage growth differences above the median.

We find that learning ability is Pareto distributed and explains almost all earnings growth heterogeneity above the median LE but only about 20% of the heterogeneity among the lower half.[2] Thus, a key conclusion is that different economic forces are driving 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. In the model, human capital is capitalized into wages

[2] Along with the Pareto distributed firm productivities, this feature of the model is consistent with the Pareto tails of *within-age* distributions of earnings with tail indices declining from younger ages to older ones as we document from the data. Yet the 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)).

in all firms similarly. Therefore, wage growth always reflects workers' human capital accumulation, regardless whether they stay with the current employer or change jobs. Thus, for workers that enjoy high wage growth regardless of job switching—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 for 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, 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 SSA

Our data is 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). Furthermore, the data are uncapped since 1978. The MEF has a small number of extremely high earnings observations. In each year, we cap (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 (PCE) deflator, taking 2005 as the base year. For background information and detailed documentation of the MEF, see Panis *et al.* (2000) and Olsen and Hudson (2009).

W-2 forms contain another crucial information for our purpose, an employer identification number (EIN), which identifies firms at the level they file their tax returns with the IRS. We use this variable as the unique identifier of a firm, which allows us 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 lower level 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. Given the focus of the paper on lifetime earnings inequality we select individuals born between 1953 and 1960 hence for whom we 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 issue with the SSA data is that, unlike survey data, it doesn't have direct measures of labor force participation. We address this problem by imposing the following restrictions on our sample. We exclude individuals with earnings below a time-varying minimum earnings threshold $Y_{\min,t}$ (e.g. ~\$1,885 in 2010) for i) at least one fourths of their working life, or ii) two or more consecutive years.[3] These two criteria help us exclude early retirees, disabled and those that are out of labor force for other reasons.[4] We also impose workers to be not self employed (iii) for more than one eighth of their working life, or (iv) two or more consecutive years.[5] 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 small self employment income on the side.

[3] $Y_{\min,t}$ is defined as one-fourth of a full-year full-time (13 weeks at 40 hours per week) salary at half of the minimum wage.

[4] 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 looking for jobs the entire time.

[5] 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.

Table I – 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

Table I provides the breakdown of the sample selection. Our initial sample consists of 1,845,640 individuals. Of these individuals, about 18% are self employed in at least one fourth of their working life. Then, about 490,000 individuals are eliminated from the sample, because they do not satisfy the minimum years of employment criterion. We exclude close to 160,000 individuals due to consecutive nonemployment in the third stage, and about 27,000 individuals due to consecutive self employment in the last stage. This procedure leaves us with a final sample of 840,194 individuals for which we have at least 31 years of earnings data.[6]

We compute lifetime earnings as the sum of individuals' W-2 earnings from 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). We find that individuals around the 90th percentile (LE_{45}) earn 3.7 times as much as those around the tenth percentile (LE_5) over their working lives (Table II). This magnitude of inequality is roughly the half of the cross-sectional earnings inequality Guvenen *et al.* (2014b) documented from the SSA data. 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 the LE_5 earning almost twice as much as that of the LE_1, whereas the inequality is much more pronounced at the top of the LE distribution. LE_{50} earns almost 4 times as much as the LE_{45}, and 13 times more

[6]Clearly, our final sample is highly selected as we lose more than half of the initial sample. In Appendix B, we document our key empirical findings for a much broader sample. Our empirical findings from this broader sample are qualitatively very similar to our baseline results.

than LE_5.

In fact, the upper tail of the LE distribution follows a power law, or equivalently, top inequality is fractal in nature: Top 1% (2%) accounts for around 29% of total lifetime income among top 10% (20%) of the population, which is essentially identical to the share of top 0.1% (0.2%) among the top 1% (2%) of the LE distribution (bottom panel of Table II).[7] Importantly and interestingly, this power law also holds in the *cross-sectional* distribution of earnings conditional on age. Moreover, as expected, earnings concentration at the top—measured as the relative earnings share of top 0.1% to top 1%—increases sharply over the life cycle from 0.23 at age 25 to 0.38 at age 55.[8] It is well known that earnings distribution has Pareto tails (e.g., Piketty and Saez (2003); Atkinson *et al.* (2011)) but to our best knowledge this is the first paper to document a power law within each age.

It is well known that, in the U.S., even though earnings inequality starts early in life, differences in earnings growth over the life cycle are key for understanding the overall 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 II. 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 log growth of average earnings between different ages over the LE distribution (see Guvenen *et al.* (2018) for a similar figure with 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 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

[7]The log density and log inverse CDF of (log) lifetime earnings distribution on Figure 17 shows clearly that lifetime earnings has a Pareto tail with a slope coefficient of –2.13.

[8]Appendix A.1 documents the Pareto tails of the cross-sectional distribution of earnings at each age in more detail. In particular, Figure 18 shows the relative income shares over the life cycle and Figures 19 and 20 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.

Table II – Selected inequality measures from the LE distribution

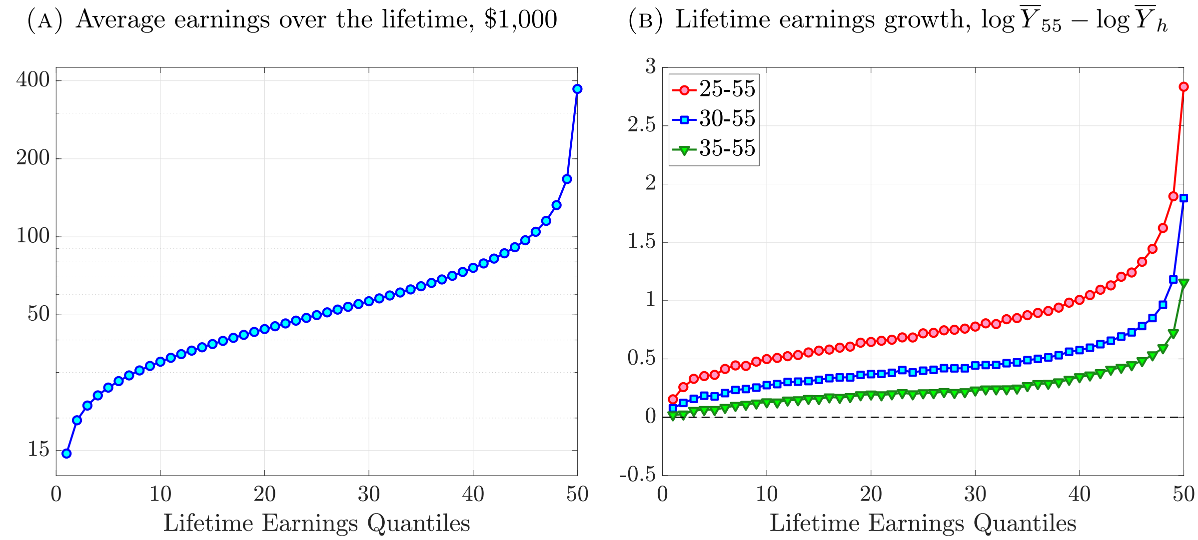
	(1) Total LE	(2) Age 25	(3) Age 40	(4) Age 55
Ratio of lifetime earnings				
LE_{50}/LE_1	25.5	3.3	23.8	47.8
LE_{50}/LE_5	14.2	2.1	12.9	25.0
LE_{50}/LE_{45}	3.8	1.1	3.5	5.2
LE_{45}/LE_5	3.7	2.0	3.7	4.8
LE_{38}/LE_{13}	1.9	1.4	1.9	2.2
LE_5/LE_1	1.8	1.6	1.8	1.9
Relative Top Income Shares				
S(0.1)/S(1)	0.29	0.24	0.31	0.38
S(0.2)/S(2)	0.29	0.24	0.30	0.38
S(1)/S(10)	0.29	0.23	0.30	0.37
S(2)/S(20)	0.28	0.22	0.30	0.36
Pareto tail index	2.2	2.6	2.0	1.7

Note: The first column shows the ratio of lifetime earnings of selected LE quantiles. The next three columns show the ratio of average annual earnings for the same LE groups at different ages.

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 have 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.

It is possible that some of this steep rise in earnings growth at the top might be simply due to transition from school enrollment around the age of 25 to employment in the labor market. For example, top LE individuals might be pursuing graduate degrees around these 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

FIGURE 1 – Heterogeneity in lifetime earnings growth



Note: 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.

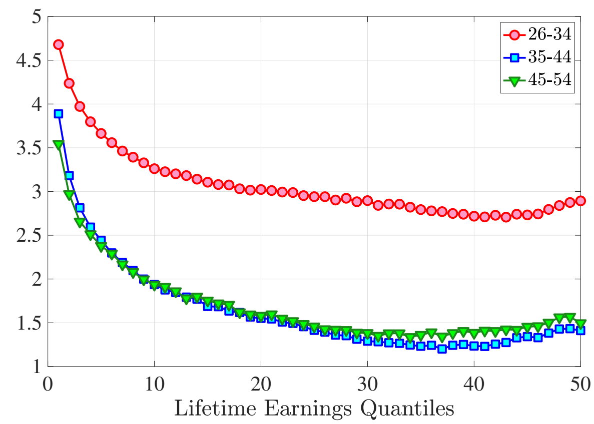
to 55 when schooling is unlikely to matter much. While the magnitudes change, we we still find a steep profile of earnings growth with respect to LE suggesting that low labor supply at age 25 is not the major driver of these patterns.

Striking differences exist even within the top LE group (top two percent), 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 percent overall) averages an 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 21). 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, which we now turn to. 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. In an influential work, Topel and Ward (1992) showed that job transitions account for two thirds of total earnings growth within the

FIGURE 2 – Number of Unique Employers over the Working Life

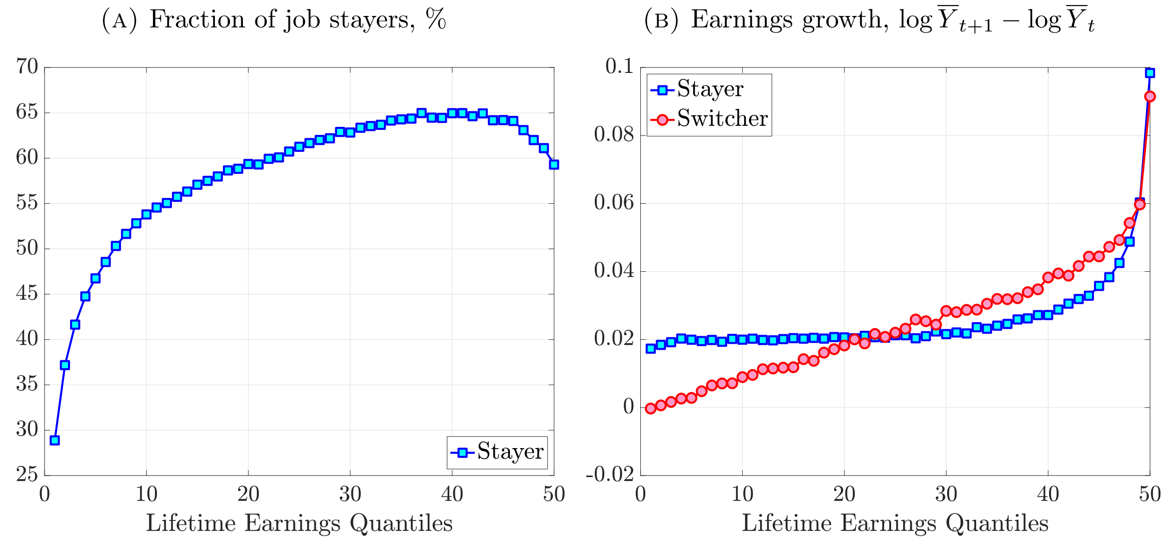


first ten years after labor market entry. Therefore, we start by investigating how the number of (distinct) employers over the working life differ between LE groups.

Figure 2 shows the average number of different employers individuals in each LE group work for during the different stages of the working life. Individuals at the bottom of the LE distribution work for almost 5 different employees on average over the first decade of their career (25–34). This implies that around a half of this group changes employers in any given year, or alternatively, a given worker changes employer 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 the first glance, one might think that low LE individuals switch jobs very often and experience a large earnings growth as a result. However, as we will see next, the nature of switches are very different across the LE groups.

We now document average earnings growth across LE groups for workers that stay with the same employer and for those that 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 typical that a worker may have more than one W-2 in a given year. Moreover, some workers may hold multiple jobs

Figure 3 – Job stayers and switchers



Note: The left panel shows the fraction of workers in each LE group that 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.

concurrently. Together, these issues pose a challenge for a precise identification of job stayers and switchers. Clearly, there is more than one plausible definition for a job stayer, and we opt with 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 $t+1$.[9] This definition ensures that the main employer was the same firm in years t and $t+1$. For example, a worker may have started working with this firm in December of $t-1$ and left the firm in January of $t+2$. This definition would correctly identify this worker as a stayer between 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 that make direct job-to-job transitions, experience nonemployment, and those that come out of nonemployment. We return to this heterogeneity later.

The left panel of figure 3 shows the fraction of job stayers within each LE group, averaged over the working life. Resonating the large differences in the number of different

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

jobs over the life cycle, there is similarly a large heterogeneity in the likelihood of staying with the same firm: Bottom LE individuals are job stayers on average 30% of their life cycle according to the definition above, compared to around 60% above the median. Thus, individuals at higher LE quantiles are more likely to stay with the same firm.

How much does a worker experience an earnings growth when he stays with the same employer versus when he switches jobs? The answer to this question differs widely across the LE distribution (Figure 3b). Several remarks are in order. First, log average earnings growth for job stayers (between years 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 earnings growth of 2% on average when he stays with the same employer. Average earnings growth for stayers increases sharply from LE_{33} onward, reaching around 10% at LE_{50}.

Turning to the earnings growth of job switchers, we find that essentially annual earnings do not increase for job switchers at the bottom of the LE distribution. This earnings growth rises almost linearly with LE 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 are very different throughout the LE distribution. As we argue later, together with much less heterogeneity among job stayers, 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 resulting 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 have to come from the differences in the frequency as well as the nature of job switches. We now investigate the differences in the types of job switches across the LE distribution.

As we discussed before, job switchers are a very heterogeneous group as they include workers who switch jobs directly or because of a job loss (or a quit). The annual nature of the data makes it hard to separate these two types directly. Yet, we argue that one can still classify workers into these groups under a plausible (albeit ad-hoc) assumption based on their realized change in annual earnings. Specifically, we assume that the workers that make a direct job-to-job transition would not be willing to take a wage cut larger than 25% (we later verify this assumption using the estimated model).[10] This allows us to

[10] Jolivet *et al.* (2006) shows 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.

classify switchers that see their earnings decline by more than 25% as "U-switchers," as they have most likely experienced some nonemployment spell in $t + 1$. We label the remaining job switchers as "E-switchers." This group contains workers that make direct job switches as well as those coming out of nonemployment in $t + 1$.

More than 35% of job switches are U-switches for bottom LE workers (Figure 4a). This share declines sharply over the LE distribution, reaches a low of 15% for LE_{40}, before increasing to around 20% for top LE workers. Thus, on average, higher LE individuals are more likely to make job switches involving earnings increases. Investigating 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 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 by more than 60%.

The annual nature of the SSA 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. To alleviate this problem, we construct earnings growth between years t and $t+2$, with the idea that earnings in $t+2$ are less affected by the events in year t or $t+1$.[11] Figure 24 shows that our substantive conclusions hold when we analyze earnings growth over a longer horizon.

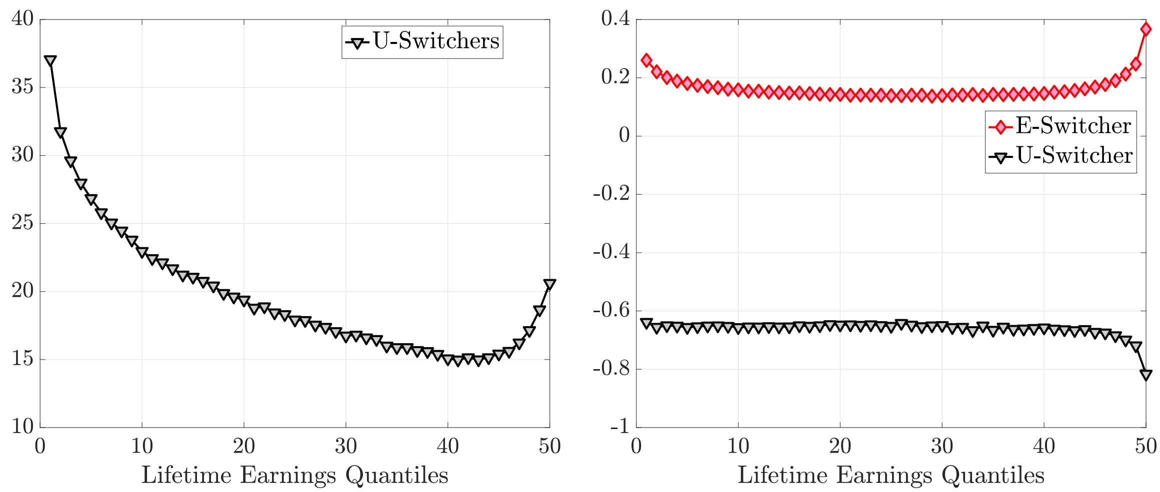
Life-cycle variation So far, our analysis has focused on the heterogeneity across the earnings distribution. There is also significant age variation in job switching and earnings growth patterns. Average life cycle profiles of earnings and job mobility have been extensively documented before. The key advantage of our data over existing work is to allow us to investigate differences in these life cycle profiles between LE groups. Figure 5 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 that stay with the same firm increases and is concave (Figure 5a). This increase is consistent with declining unemployment risk and job mobility that has been extensively documented before (see, for example, Jung and Kuhn (2016)). Interestingly, this profile is shared by all LE groups, though the concavity is a bit more pronounced above the median. Turning to the average earnings growth of job stayers, we find a flat profile below LE_{30} *at all ages*. Moreover, consistent with previous literature, the rate of earnings growth declines

[11] Our approach throughout the paper to deal with such issues is to use the model, where we aggregate simulated quarterly data to annual, and construct moments in a similar fashion.

Figure 4 – E-switchers and U-switchers

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

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



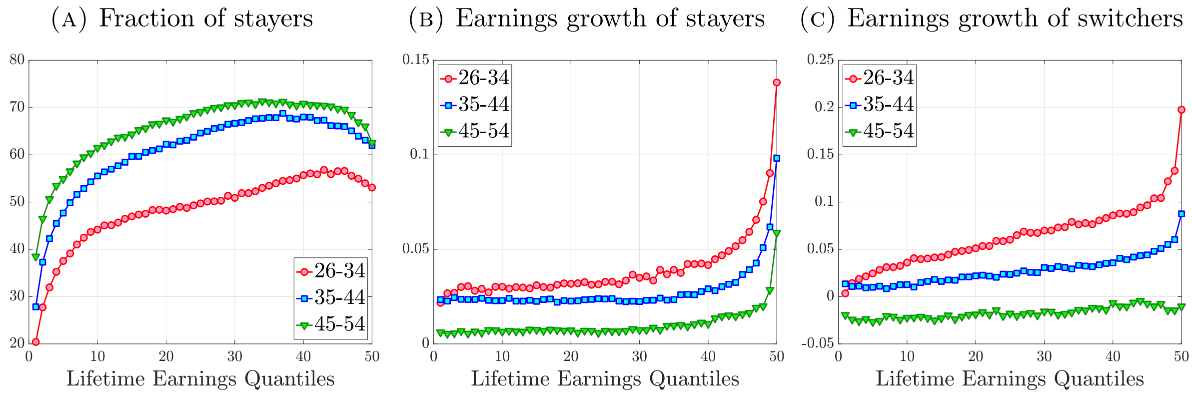
Note: The left panel shows the share of U-switchers among job stayers in each LE group. The right panel plots the log growth of average earnings $\bar{Y}$ between t and $t+1$ for U-switchers and E-switchers separately.

with age.[12] Finally, we investigate the earnings growth of job switchers over the life cycle (Figure 5c). Similar to job stayers, the average earnings growth for job switchers is highest at younger ages (aged 26–34). This growth rate declines sharply over the life cycle, especially for higher LE individuals. Earnings growth of stayers is negative for older individuals (aged 45–54) throughout the LE distribution.

Evidence from SIPP data We acknowledge that our definition of U- and E-switches is based on a somewhat arbitrary cutoff. However, the annual nature of the SSA data does not allow us to come up with a more precise definition. To complement our analysis, and to provide additional evidence for heterogeneity in job loss rates, we turn to data from SIPP. SIPP is a nationally representative sample 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. Labor market transition rates such as unemployment, job finding and job-to-job transition rates can be computed at high frequency.

[12] 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 working life includes individuals up to age 65). See, for example, Heckman (1976).

FIGURE 5 – Life cycle variation



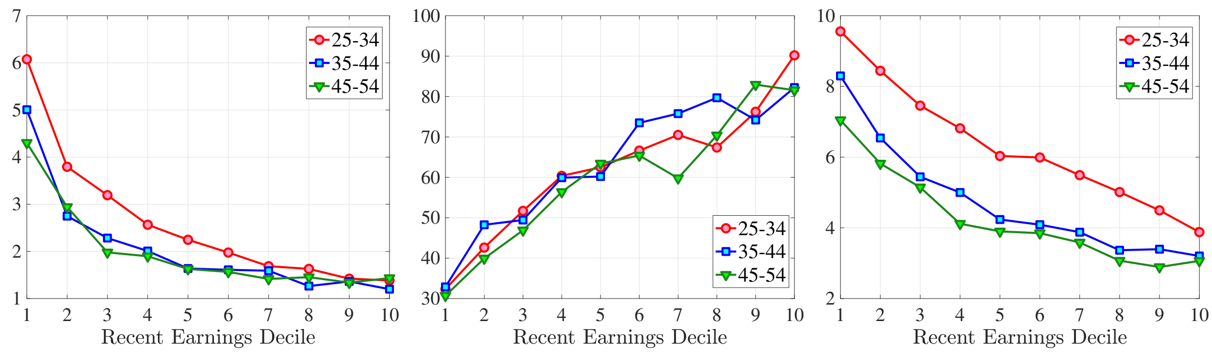
We select a sample of males between ages 25 and 55 with some labor force attachment (further details on SIPP and our sample are relegated to Appendix E). We use the panel structure to rank workers into 10 equally-sized deciles within each age group (25–34, 35–44 and 45–55) based on their recent earnings 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 4-months.

Job loss rates show significant heterogeneity across previous recent earnings deciles for all age groups (Figure 6). For example, unemployment risk at the bottom decile can be almost five times than 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. The middle panel shows that the 4-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 that 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 to the underlying sources. In what follows, we introduce a quantitative model of wages and job turnover with heterogeneity in returns to experience

FIGURE 6 – Labor market flow rates across the income distribution

(A) Unemployment Risk, %　　(B) Job Finding Rate, %　　(C) Contact Rate, %



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. Even though this model endogenously generates some age variation in job mobility and earnings dynamics, it falls short of capturing the magnitudes in the data. Thus, we add an explicit life cycle structure to this framework, which we model 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 that is sold in a perfectly 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:

$$U(\{c_t\}_{t=0}^{\infty}) = \sum_{t=0}^{\infty} \rho^t \log c_t.$$

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

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

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

Here, $\tilde{h}_{it}$ denotes the non-stochastic 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 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.[13] 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. 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 that match with a new firm draw from a productivity distribution CDF $F(p)$ with a support of $\left[\underline{p}, \infty\right]$ common to all workers.[14] 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 given by

$$y_{it} = p_{j(i,t)} + h_{it}.$$

[13] We also tried a version with individual-specific quadratic returns to experience and found negligible heterogeneity.

[14] As in Bagger *et al.* (2014), we assume that the match with the least productive firm generates a surplus that is positive for all workers, so that in equilibrium all meetings turn into a match.

3.1.1 Heterogeneity in Search and Matching

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 in Section 2. These differences 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. Therefore, to account for the sources of earnings growth, we need to explicitly model the differences in job finding rates.

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 potential employer, whose productivity is drawn from the distribution $F(p)$.[15] As Figures 4a and 6c have shown, workers differ in the types and rates of job switches. Of course, our framework can generate qualitatively similar patterns without explicit differences in the contact rates. Namely, 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.

Unemployment risk A job dissolves exogenously with probability $\delta^a(\alpha_i)$ at the end of the period, in which case the worker stays unemployed in the next period and searches for a job. Again, 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. Workers that are hit by separation shocks find a job immediately with probability $\xi\lambda(\alpha_i)$. As we discuss later, our model period is a quarter and a nonnegligible fraction of layoffs 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)). To allow for this in our model, we allow for the possibility of finding a job within the same period.

[15]The contact between an employed worker and a firm triggers a renegotiation between the worker, the incumbent firm and the poaching firm that we explain below.

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 that find jobs negotiate their wage, workers that 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), here we focus on the key equations and how the life-cycle structure affects them. We relegate most of the analytical derivations to Appendix G.

In this model 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 We start by discussing the wages of hires from unemployment. Let $V_0^a\left(h; \mathbb{I}_i\right)$ and $V^a\left(r, h, p; \mathbb{I}_i\right)$ 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 $V^a\left(r, h, p; \mathbb{I}_i\right)$ 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. $V_0^a\left(h; \mathbb{I}_i\right) = V^a\left(0, h, p_{\min}; \mathbb{I}_i\right)$. This assumption implies that an unemployed worker accepts any job offer and simplifies the problem.[16]

The wage bargaining protocol dictates that unemployed workers receive θ share of the expected match surplus, where θ captures the worker's bargaining power.[17] More specifically, the piece rate of a hire from unemployment, r_0, solves

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

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

Poaching When an employed worker is contacted by a firm with productivity p', the incumbent firm and the poacher compete à la Bertrand. As such, the more productive firm outbids the less productive one by offering the worker an expected lifetime utility that the other firm cannot, and in turn obtains his services. We now discuss the wage that arises as a result of this competition.

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

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

Note that as Postel-Vinay and Robin (2002) have shown, such job switches may result with workers accepting wage losses as they anticipate faster wage growth in firms with a higher productivity. These 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

[16] 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)).

[17] 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).

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

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

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. In this case, 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

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

Value functions We are now ready to describe the value functions that are used in the equations above. Here, we describe the problem of young workers ($a = y$), which highlights the perpetual aging feature. The problem of old workers is presented in Appendix G. Equation (6) describes the expected lifetime utility of a young worker with human capital h employed at a firm with productivity p paid at piece rate r.

$$\begin{aligned}
V^y\left(r, h, p; \mathbb{I}_i\right) \quad = \quad & w + \frac{\delta^y}{1+\rho}\left[\left(1 - \xi\lambda_0^y\right) V_0^y\left(h; \mathbb{I}_i\right) + \xi\lambda_0^y \int_{p_{\min}}^{\infty} \theta\mathbb{E}\left[V^y\left(0, h', p; \mathbb{I}_i\right) - V_0^y\left(h; \mathbb{I}_i\right)\right]\right] \\
& + \frac{\lambda_1^y}{1+\rho}\int_p^{\infty} \mathbb{E}\left[\left(1-\theta\right) V^y\left(0, h', p; \mathbb{I}_i\right) + \theta V^y\left(0, h, x; \mathbb{I}_i\right)\right] dF\left(x; \mathbb{I}_i\right) \\
& + \frac{\lambda_1^y}{1+\rho}\int_{q^y(r,h,p;\mathbb{I}_i)}^{p} \mathbb{E}\left[\left(1-\theta\right) V^y\left(0, h', x; \mathbb{I}_i\right) + \theta V^y\left(0, h', p; \mathbb{I}_i\right)\right] dF\left(x; \mathbb{I}_i\right) \\
& + \frac{\gamma}{1+\rho}\mathbb{E}V^o\left(r, h', p; \mathbb{I}_i\right) \\
& + \frac{1}{1+\rho}\left[1 - \delta^y - \gamma - \lambda_1^y \bar{F}\left(q^y\left(r, h, p; \mathbb{I}_i\right)\right)\right] \mathbb{E}V^y\left(r, h', p; \mathbb{I}_i\right)
\end{aligned} \quad (6)$$

The worker collects his wage $e^w = e^{r+h+p}$ and moves to the search and matching stage.[18] Here, one of the following events occur: i) his job is terminated exogenously (he may find a job within the period), ii) he gets an outside offer from a firm with higher productivity and changes employers, iii) the worker gets a counter offer, renegotiates his piece rate and stays with his current employer, iv) he gets old and moves to the next period as an old worker employed at the same firm, and v) he keeps his job, doesn't get an outside offer that triggers a renegotiation and remains young.

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 fix some of the parameters: The quarterly discount rate ρ is set to 0.005 to match 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).

Next, we estimate the remaining parameters by employing the simulated method of moments (SMM). We simulate quarterly data 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.[19] In particular, given the parameters of the model, we simulate a panel of 100,000 individuals. Each individual starts his life unemployed at the age of 23 and remains in the labor force until the age of 55. We discard the first two years of the simulated panel 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 our 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 that experience large annual earnings declines. While idiosyncratic shocks to worker productivity could in principle account for these large losses, we found them to be quantitatively insufficient.

[18]This implies that per-period utility is linear log wage: $u(c) = \log(e^w) = w = r + p + h$.

[19]In earlier literature, this class of models are 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.

To give the model a chance to match this feature, we add a recall option of unemployed workers by their last employers.[20] 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 our model analytically tractable, and keep the estimation computationally feasible.[21]

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 separately. 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 for each LE group. We choose to not target the heterogeneity in lifetime income growth shown in Figure 1. As we discuss in the next section, the model is already identified using these five sets of moments.

Grouping workers We condition each targeted moment on lifetime earnings and age groups. Specifically, we first calculate workers' lifetime earnings as explained in Section 2 and then assign them into 12 percentile groups: 1-4, 5–10, 11--20, ..., 81–90, 91–96, 97–100.[22] 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

[20] Fujita and Moscarini (2017) document that recalls are quite frequent and affect wage dynamics.

[21] Solving for the wage equation analytically, as we do in Appendix G, is not possible when the recall option is recognized during bargaining.

[22] In other words, we aggregate the 50 LE groups used in Section 2 into these 12 percentile groups for estimation.

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.[23] 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 when selecting the sample from the data, we drop workers with two or more consecutive years of zero earnings. However, there are still observations in our data with zero income in a given year. We would like to keep them as part of our estimation as they clearly contain valuable information about the importance of search frictions. For this purpose, we use the arc percent growth measure defined as follows:

$$\frac{Y_{t+1} - Y_t}{\left(Y_{t+1} - Y_t\right)/2},$$

where Y_t is annual earnings. Cross-sectional moments targeted in our estimation are shown in Figure 32.

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 33 and 34 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) of each LE group at age 25. This moment of the data, along with average earnings in other age groups, is shown in Figure 35.

4.2 Identification

Below we provide an informal discussion of identification of our model. We acknowledge that when employing SMM, all parameters are determined jointly within the estimation as most parameters affect more than one aspect of the data. In this section,

[23] In principle, one can target the entire distribution of these earnings changes. However, these distributions are well summarized by the first four centralized moments. An alternative is to target percentiles or percentile-based moments such as the 90-10 differential, Kelley's skewness and Moores kurtosis of the annual earnings change distribution. We have experimented estimating the model by targeting these moments as well and found similar results. We choose to target centralized moments as they are less costly to compute.

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.

Ex-ante worker productivity (α, β) The concave life cycle profile of earnings growth across all workers 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 initial earnings levels of LE groups and their stayer earnings growth (Figure 3b) 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 the the earnings growth of job stayers through outside offers. As we discuss next, we use other features of the data to separately 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.[24] 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 distri-

[24]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).

bution (Figure 3b). 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 generate a *different* profile of earnings growth for stayers and switchers. Later, 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–transitions, those switches that involve a larger than 25% earnings loss, (Figure 4b) are intimately linked to the job loss rate δ. Moreover, this feature of the data is not affected by the rate of direct job-to-job transitions, because such transitions either result in wage increases, or smaller than 25% wage losses, and are therefore counted among E-switches.[25]

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

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

shown in Section 2. Therefore, the age dependence in these parameters are 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 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 ex-ante worker fixed-effect α is normally distributed with mean μ_α and standard deviation σ_α. Furthermore, β is drawn from a Pareto distribution with shape and scale parameters χ_w and ψ_w, respectively, and is correlated with α with the correlation coefficient $\rho_{\alpha\beta}$.[26] The Pareto tail of β captures the "high-growth" worker types that experience a much higher earnings growth than other individuals. 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, job finding rate and the contact rate as a function of α_i 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 follows a Pareto distribution with shape and scale parameters χ_f and ψ_f, respectively.[27] We normalize the scale parameter ψ_f to 1, because one cannot separately identify the scale of the firm distribution 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

[26] 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.

[27] 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.

with standard deviation σ_ε. $\pi(\alpha)$ is modeled as a cubic spline separately for each age group. 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.

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.[28]

Our SMM estimator is defined by the following:

$$\hat{\theta} = \arg\min_{\theta} \boldsymbol{F}(\theta)' W \boldsymbol{F}(\theta), \tag{7}$$

where $\boldsymbol{F}(\theta)$ is a column vector in which deviations of model moments from their empirical target are stacked, that is,

$$\boldsymbol{F}(\theta) = \left[F_1(\theta), ..., F_N(\theta)\right]^T.$$

The weighting matrix, W, is chosen subjectively and reflect our beliefs on the importance of each set of moments.[29] We target a total of 380 moments to estimate 41 parameters.

Numerical method for estimation We minimize the objective value as follows. We first generate 15,000 uniform Sobol (quasi-random) points and compute the objective value for each of these. Then, we select the best 1,000 Sobol points (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 (which is slow but performs well on difficult objectives) and the DFNLS algorithm of Zhang *et al.* (2010), which is much faster but has a higher tendency to be stuck at local minima. We have found that the combination balances speed with reliability and

[28] 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.

[29] The weighting matrix, W, is chosen such that we assign 15% relative weight to the first two sets of moments (i.e., cross-sectional moments of job stayers and switchers) and 30% weight to 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 10% relative weight.

provides good results. In the end, we pick the best parameter estimates out of 1,000 local minima.

5 Estimation results

In this section, we present the key parameter estimates and discuss the model's performance in fitting the targeted moments.

5.1 Parameter estimates

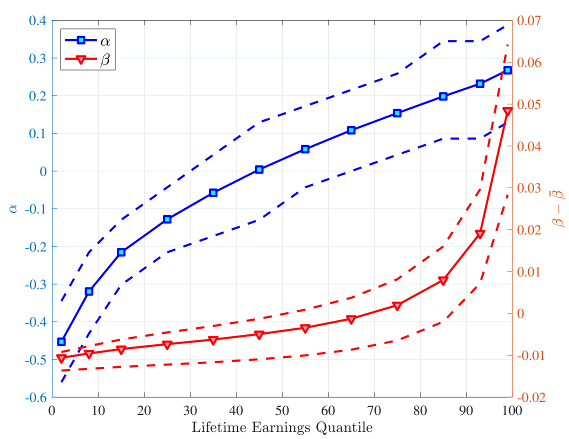
We first discuss our key parameter estimates and relate them to the empirical moments that have a pronounced effect on them.

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 7). α increases almost linearly throughout the LE distribution. Top LE individuals have an α that is more than 60 log points larger than those at the bottom. Moreover, there is a sizable variation within each LE group. The interquartile range (dashed lines in Figure 7) is around 10 log points. Together with this, the standard deviation of α in the entire population is 0.25.

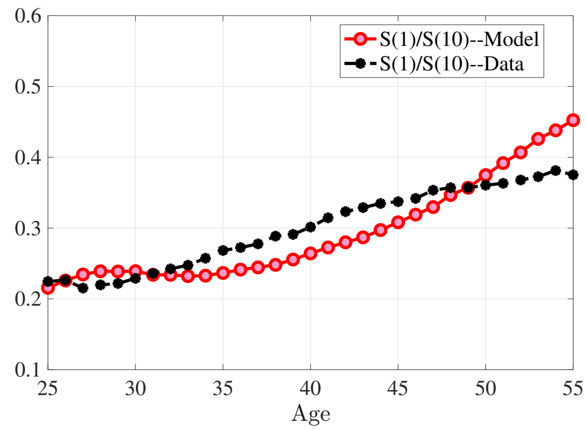
Returns to experience, β, also increases with LE—not surprising given the positive correlation with α of $\rho_{\alpha\beta} = 0.44$—however with a different shape: β is relatively flat

FIGURE 7 – Worker Type and Returns to Tenure



Note: This figure 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.

Figure 8 – Worker Type and Returns to Tenure



Note: This figure shows the ratio of incomes earned by top 1% earners ($S(1)$) relative to top 10% earners ($S(10)$).

in the bottom two thirds of the LE distribution and increases steeply towards the top, essentially mirroring stayer earnings growth in Figure 3b.[30] 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 income share of top earners at each age.

Figure 8 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 around 0.4. This increase is driven by the growing importance of the return heterogeneity in annual earnings. Compared to the data, the relative share in the model increases faster in the last five years. We have also analyzed if the model exhibits power law throughout the life cycle as seen in the data (see Figure 18). We find that after age 35 in the model, the relative income share of top 0.1% in top 1% is similar to that of top 1% in top 10%. Before age 35, earnings concentration in top 1% is larger than earnings concentration in top 10%, implying that the earnings distribution of young workers in the model has thicker tails than that of the data. Overall, these results confirm that earnings follow a

[30] 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)).

power law throughout the working life both in the data and the model.

We would like to point out that the 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". 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. This is larger in magnitude than estimated in Jarosch (2015) using German data. Furthermore, this is not the only channel in our model that contribute to scars from unemployment, which are estimated to be 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 foregone opportunity of accumulating experience.

Heterogeneity in flow rates

Next, we turn to the heterogeneity in labor market flow rates. Figure 9 plots in three panels how the unemployment risk, the job finding rate, and the contact rate vary with calendar age and LE groups.[31] Unemployment risk declines sharply with lifetime earnings up to median LE and is essentially flat for individuals above the median (Figure 9a). 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 stochastic 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

[31] 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.

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 SIPP data shown earlier in Section 2. To this end, we follow the exact same sample construction in model-generated data as we do on SIPP data, and compute flows between labor market states for three age and 10 recent earnings groups. Figure 9b shows how the unemployment risk varies with recent earnings in the SIPP data averaged over the lifecycle along with its model counterpart. While this feature is 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 9c). For example, the quarterly job finding rate increases from around 30% at the bottom for workers aged 25–34 to above 50% at the top for workers aged 45–54. These estimates imply that bottom LE workers aged 25–34 stay unemployed for around 3 quarters, compared to around 2 quarters for top LE workers aged 45–54. Coupled with high unemployment risk for low LE workers, these differences imply large differences in actual experience over the life cycle (Figure 10). In particular, quarters worked over the working life range from around 80 for low LE individuals to 120 at the top. This large heterogeneity, especially below the median, has important implications for earnings growth differences which we discuss later.

Our estimate of the heterogeneity in job finding rates is qualitatively consistent with the external evidence from the SIPP data (Figure 9d). 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 directly measure recalls using survey data from the SIPP, whereas we identify 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 rate to be increasing with respect to lifetime earnings and age, with a range between 15% and 40% (Figure 9e).

While the increasing likelihood of the contact rate with respect to LE seems inconsistent with a declining job-to-job transition rate in SIPP shown in Figure 6c, the model actually captures this pattern well due to an endogenous mechanism. 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 have analyzed data from a supplement to the Survey of Consumer Expectations (SCE), administered by the Federal Reserve Bank of New York. The SCE 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 aged 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 III). We find that contacts increase in previous wages and is quite steep at the top. People in the highest group (workers above 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 those at the bottom (0.43 vs. 0.09, respectively). Moreover, almost all of their contacts are unsolicited.

Life-cycle variation Since there is little life cycle variation in flow rates in the data, we relegate the fit along this dimension to Appendix H.3 Figure 37. Overall, the model captures well the decline in job loss risk and job-to-job transition rate with age. However, there is 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 younger workers in the data.

Table III – 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

Note: Respondents age 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).

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.

Idiosyncratic shocks

Recall that the Gaussian idiosyncratic shocks hit once a year with some probability $\pi(\alpha)$ when workers stay with the same employer. We estimate that the annual probability of experiencing such a shock increases from around 5% to 20% from the bottom LE workers to top. Recall that the distribution of these shocks are identified from the second to fourth order moments of the annual earnings growth for job-stayers. Figure 12 shows that the variance increases above the 20th percentile and kurtosis falls 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 earnings dynamics for job-stayer is mainly driven by the endogenous mechanisms of the job ladder model including recalls.

5.2 Model's fit to the data

In this section, we present and 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 various pieces in the model are informed by the different aspects of the data.

Cross-sectional moments Figure 12 shows the fit of the model to cross-sectional moments. For the clarity of exposition, we suppress the life cycle variation in these

Figure 9 – Labor market flows

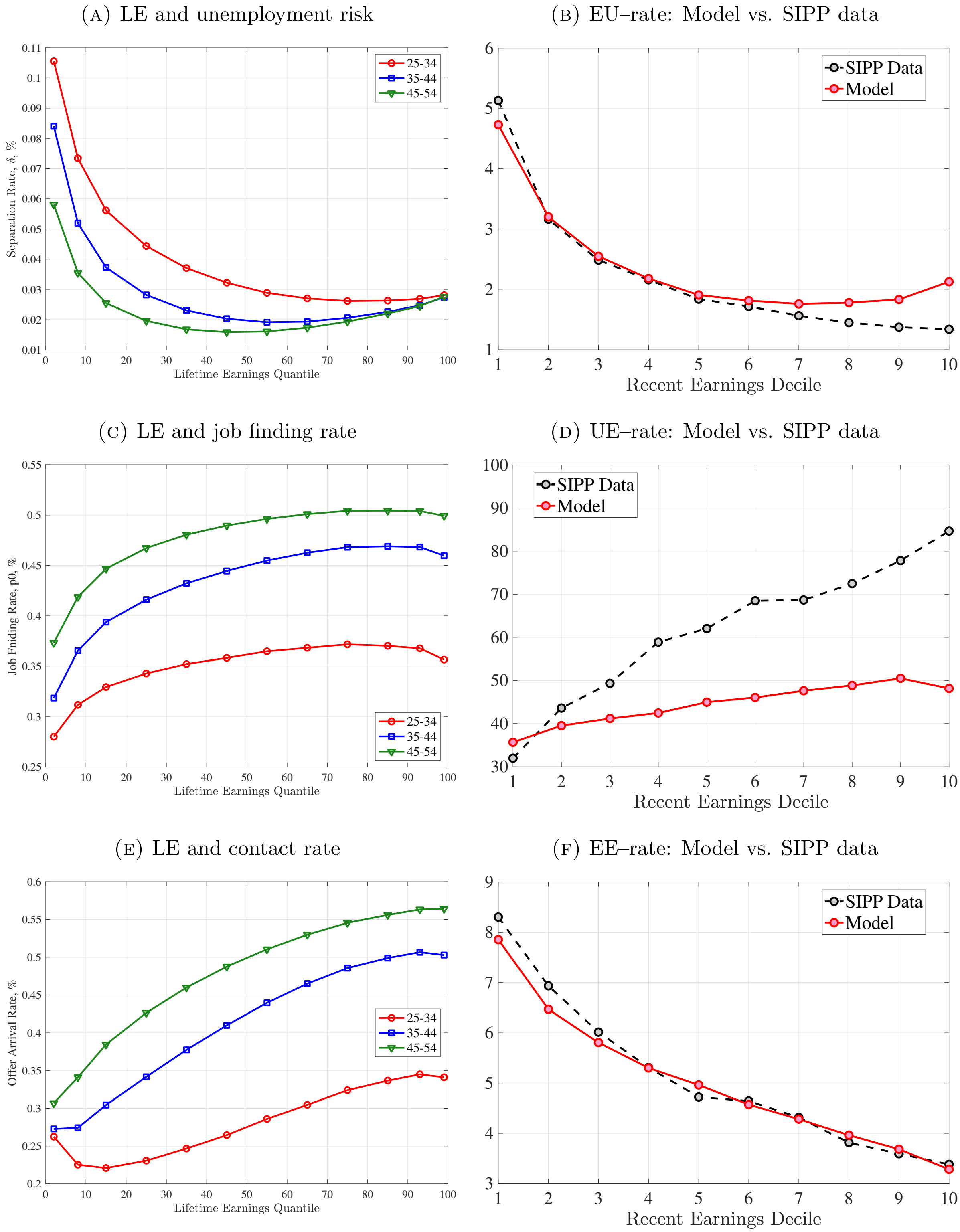


FIGURE 10 – Differences in actual experience

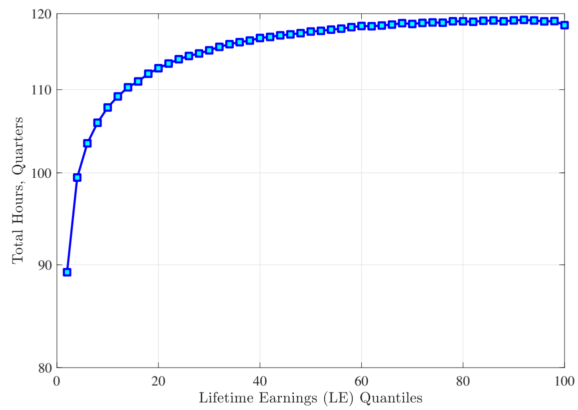
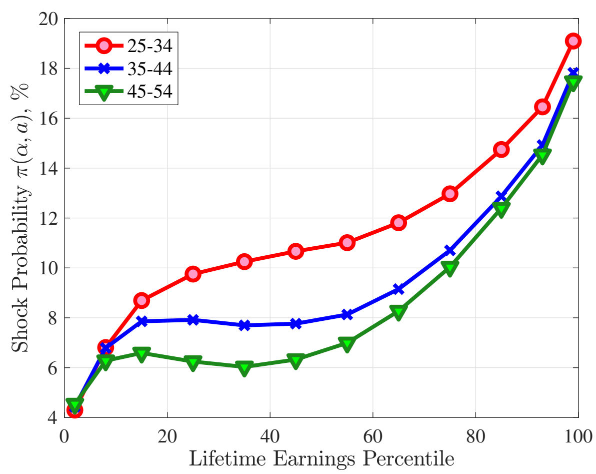


FIGURE 11 – Idiosyncratic Shock probability, $\pi(\alpha, a)$



moments by plotting their averages over three age groups. The fit along the life cycle dimension is shown in Appendix H.3.

The model captures well the standard deviation of earnings changes for job stayers and switchers (Figure 12a). Both in the data and 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 9) combined with an increasing poaching rate (Figure 9e) 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, outside offer that leads to renegotiation and idiosyncratic shocks. Due to their high unemployment risk, the share of recalls is highest for bottom LE individuals, which tends to push up the standard deviation at the bottom. As we move to the right along the LE distribution, unemployment risk fades, but 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 9e and 9f). Moreover, idiosyncratic shocks become more frequent and contribute to the increasing standard deviation for job stayers above the median.

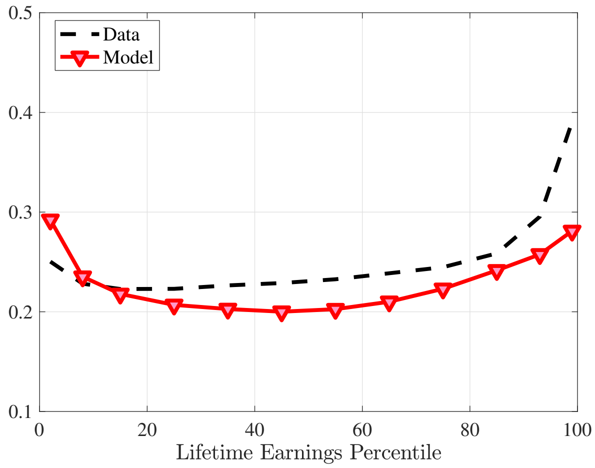
Turning to skewness, we find that the model captures well the essential features of the data (Figures 12c and 12d). First, earnings changes are negatively skewed for both job switchers and stayers. For switchers, the negative skewness is mostly a result of flows into unemployment, which result in the 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 a longer unemployment duration, pushing skewness down at the bottom. On the other hand, job loss is less frequent but much more costly for high LE individuals as they have more search capital and negotiation rents to lose. The latter force dominates and causes skewness of earnings changes to be more negative for job switchers among high LE individuals.

Negative skewness of earnings changes among job stayers is due to recalls, which generate large earnings declines within the same firm. In the absence of recalls, the model cannot generate 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 become more frequent 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.

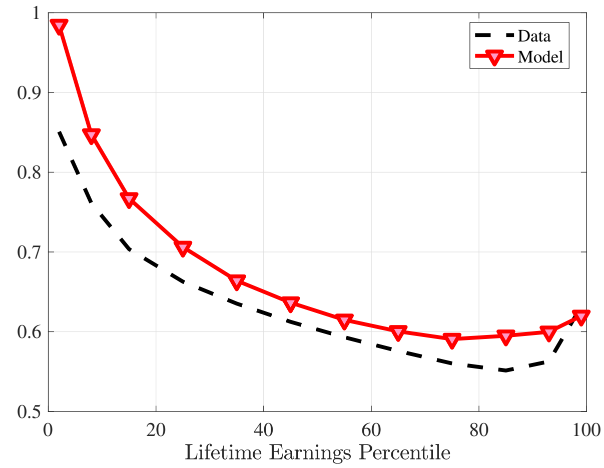
The model is quite successful in matching the extent of kurtosis and its variation over the LE distribution. Kurtosis measures the tendency of a distribution to stay away from $\mu \pm \sigma$ (Moors (1986)). Distributions with excess kurtosis tend to have pointy centers and longer tails relative to a Gaussian distribution. Infrequent events that lead to large earnings 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

FIGURE 12 – Model's fit to cross-sectional moments of earnings changes

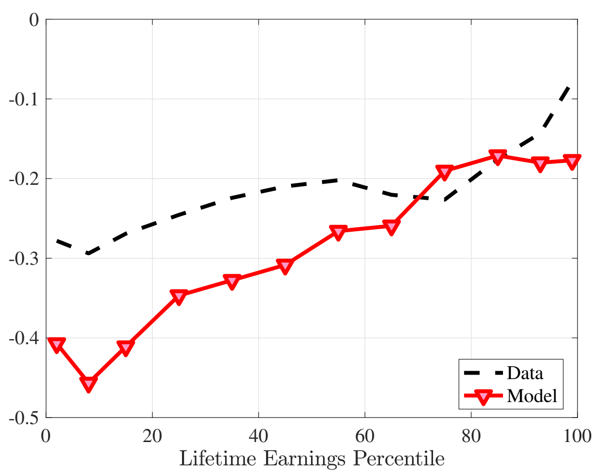
(A) Standard deviation, stayers



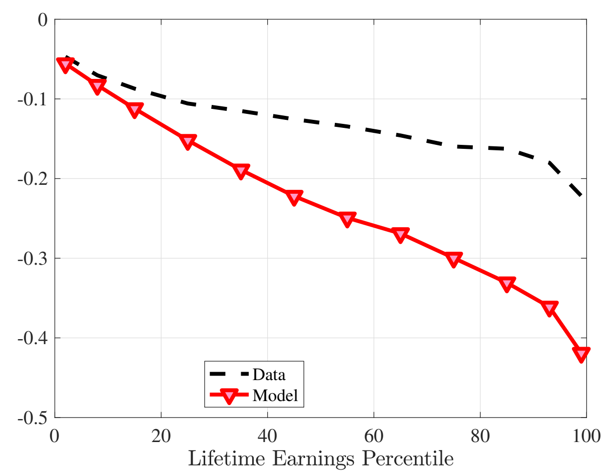
(B) Standard deviation, switchers



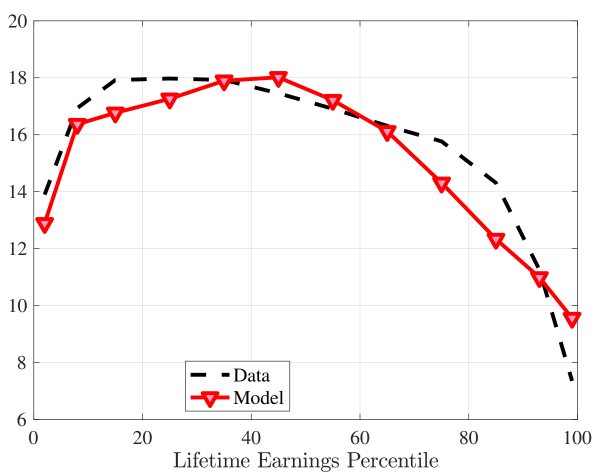
(C) Skewness, stayers



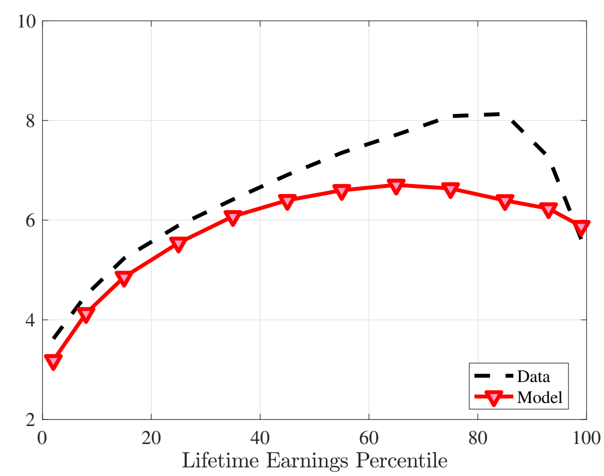
(D) Skewness, switchers



(E) Kurtosis, stayers



(F) Kurtosis, switchers



in the absence of idiosyncratic shocks, earnings changes of job stayers would be a lot more leptokurtic. The idiosyncratic shocks, despite being leptokurtic themselves, help the model bring down kurtosis of job stayers closer to values we see 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.

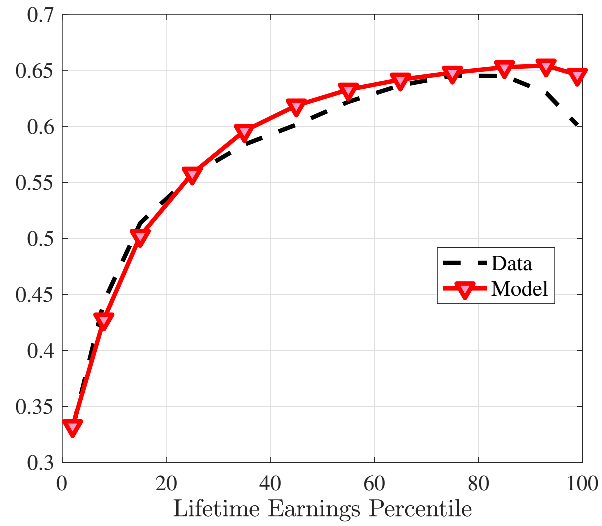
Finally, we investigate the model's fit on cross-sectional moments along the life cycle dimension. Figure 36 shows how the higher-order moments of earnings changes for stayers and switchers vary between three age groups. As in the data, life cycle variation in the model is less pronounced than the variation 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 profile of the share of job stayers by LE quantile in the data (Figure 13). There are few job stayers at the bottom of the distribution 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.

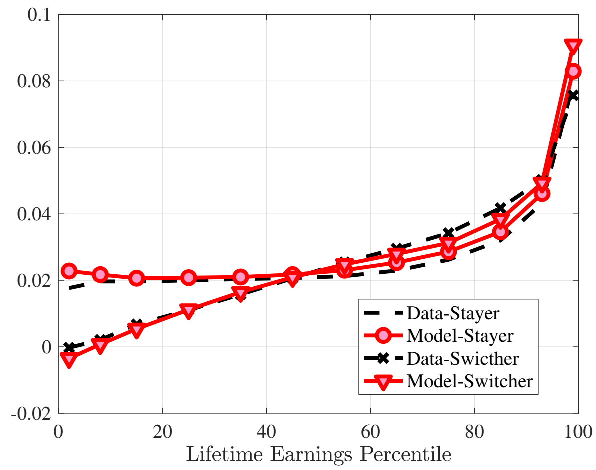
The model also generates overall a realistic average earnings growth for job stayers and switchers throughout the LE distribution (Figure 13b). In particular, there is little heterogeneity among job stayers for the bottom two thirds of the LE distribution, which, as discussed before, is in part due to the relatively flat average profile of returns to 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 7 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 13b), an outside

FIGURE 13 – Model's fit to income growth moments

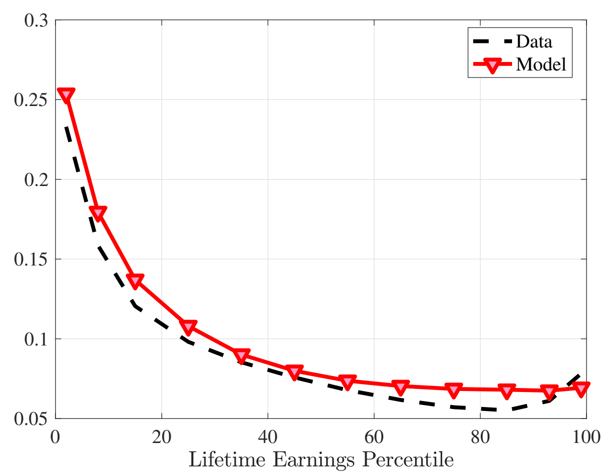
(A) Fraction of stayers



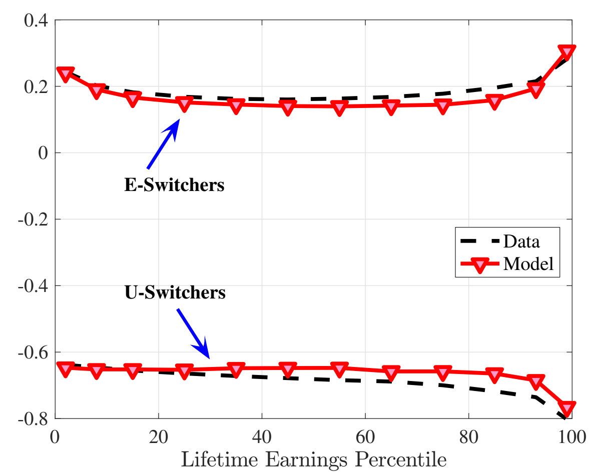
(B) Earnings growth of stayers and switchers



(C) Fraction of E-switchers



(D) Earnings growth of E-switchers and U-switchers



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 13b). In particular, as in the data, there is a large variation throughout the LE distribution, ranging from zero at the bottom to 10 log points. 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 in the model as in the data, increases sharply from around 65% at the bottom of the LE distribution to above 80% (Figure 13c). These shares are slightly below than 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 13d).

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

6 Decomposing lifetime earnings

We now turn to decomposing the differences in lifetime earnings into the various components in the model. The model matches well the distribution of lifetime earnings in the data (Figure 14a). 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 model does much better in matching the distribution 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 IV).

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 over the life cycle? Figure 14a 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

Table IV – 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

Note: 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 individuals in LE_i to those in LE_j.

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 bottom LE is about 25% lower than the median workers. Employment differences above the median are negligible in comparison (Figure 14b).

Before investigating the sources of differences in lifetime wages in detail, we briefly discuss the sources of large differences in employment below the median. In our framework, 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 $\delta^a(\alpha_i = 0)$, the job loss risk of workers with $\alpha_i = 0$, which is the average α of median LE workers, as shown in Figure 7, and compute the resulting distribution of total lifetime employment (in quarters). 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.[32] 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 lose their jobs at the same rate (Figure 14b). When we also eliminate differences in job finding probabilities 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 role to luck in generating sizable employment differences over the working life.[33]

We now turn to the major source of lifetime earnings differences, i.e., wage differentials. First, we establish that wage differences are largely shaped by *wage growth* heterogeneity rather than the initial differences in *levels*. Figure 14a shows that when all sources of wage growth have been turned off, the model generates a wage inequality that is an order of magnitude smaller.[34] Therefore, it is essential to understand why some workers have a much steeper wage profile than others, which is what we turn to now.

Recall that in the model wage growth can differ across individuals due to differences

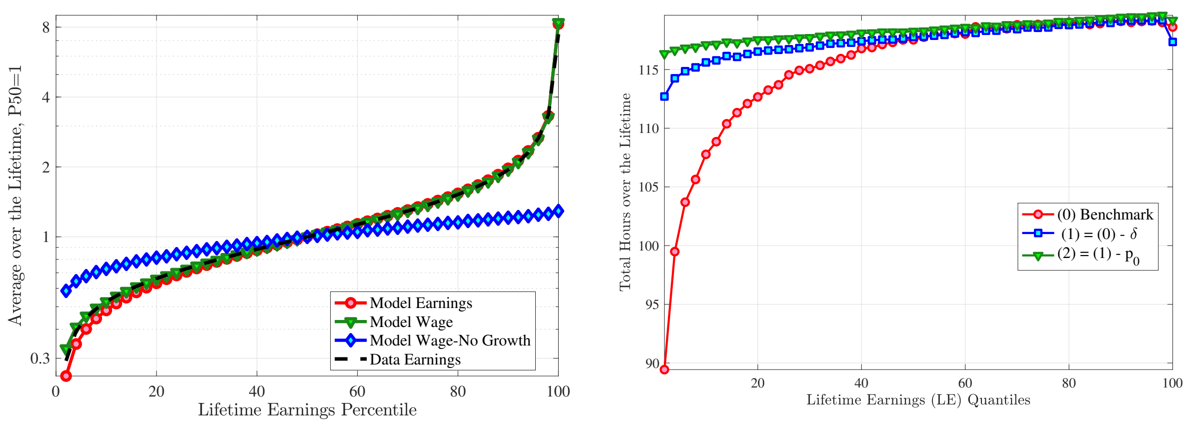
[32] The set of individuals that are grouped to LE_j are the same in each experiment. If we were to regroup, the groups may change in each experiment. Since our goal is to quantify the effect of having a high job loss risk, we do not reassign workers into LE groups based.

[33] One caveat to this result is that our model does not allow unemployment begetting further unemployment risk in the future as in Jarosch (2015). We study the importance of this feature for the US

FIGURE 14 – Lifetime earnings, wages and employment

(A) Lifetime earnings and wages

(B) Total Employment



Note: Lifetime earnings (wages) is defined as total (average) earnings (wages) over an individual's working life. "Model Wage–No Growth" corresponds to an experiment that shuts down the heterogeneity in β, gets rid of search frictions by setting $\delta = 1, \lambda_0 = 1$, $\lambda_1 = 0$ for all individuals, removes idiosyncratic productivity shocks and makes the firm distribution degenerate. In this specification of the model, the only source of wage (and thus earnings) differences is heterogeneity in permanent ability, α. In the figure, 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.

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 each of these factors in driving wage growth differences over the life cycle, we conduct several experiments, where 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 15 shows the lifetime wage growth differences between LE groups for the benchmark estimation as well as for our counterfactual experiments. [35]

We start by removing the differences in unemployment risk, which we accomplish by

data in an ongoing project.

[34]Specifically, we shut down the heterogeneity in β altogether, set $\delta = 1, \lambda_0 = 1$, $\lambda_1 = 0$, and make the firm distribution degenerate. In this specification of the model, the only source of wage differences is the permanent ability α.

[35]Note that we provide a separate decomposition of lifetime earnings growth into wage and hours growth components (Figure 38a). The model matches well the earnings growth heterogeneity in the data. Earnings growth is less dispersed than that of wage growth between LE groups, especially in the bottom half of the LE distribution. The difference is due to a higher employment growth over the life cycle at the bottom of the LE distribution.

shutting off differences in job loss and job finding risk together (by setting $\delta^a(\alpha_i) = \delta^a(0)$ and $\lambda_0^a(\alpha_i) = \lambda_0^a(0)$).[36] 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 median LE (series (1) in Figure 15). Specifically, slightly more than 50% of wage growth differences at the bottom would diminish, if the workers at the bottom ranks of the LE distribution had job loss and job finding rates similar to that of median LE workers. High unemployment rates of low-income individuals—as shown in Figure 10—not only prevent them from accumulating human capital but also lead to depreciation in human capital during unemployment. Furthermore higher incidence of unemployment prevent bottom LE workers from climbing the job ladder. Figure 16 shows the contributions of human capital, search capital and negotiation rents to the wage growth differences between the bottom and median LE workers.[37] 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.[38] 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 the Danish data.

Job loss and job finding differences have a much lower impact at the upper half of the distribution, mainly because they have fairly similar job loss and job finding rates 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.

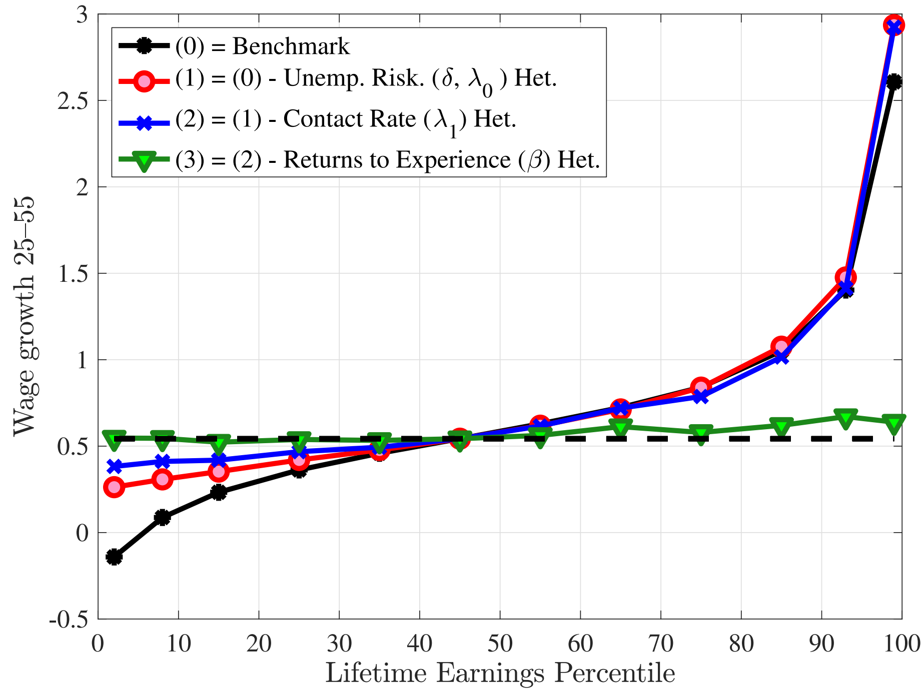
Recall that LE groups also display sizable differences in their contact rates λ_1. In contrast to the job loss and job finding rates, these differences have a smaller effect on lifetime wage growth differences (Figure 15). Specifically, eliminating differences

[36] We analyzed the roles of job loss and job finding components separately as well. These results are reported in Appendix H.3.

[37] 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 log piece rate $\mathbb{E}\left[r_{i,55} - r_{i,25}\right]$.

[38] Lower growth in negotiation capital for top LE workers is because these workers are employed at more productive firms, who 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 15 – Decomposing differences in lifetime wage growth

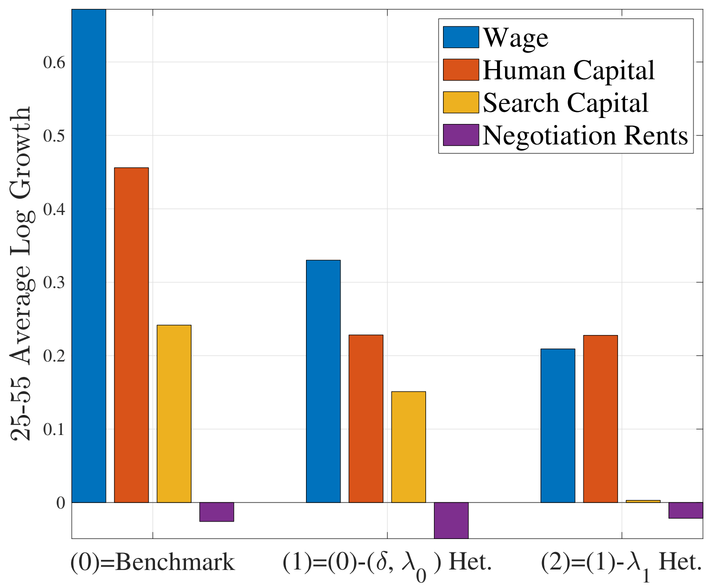


Note: This figures shows the decomposition of the heterogeneity in lifetime wage growth into the job ladder and human capital components. The black line shows the wage growth in the benchmark, the circled-red line plots wage growth when differences in ex-ante job loss and job finding rates δ, λ_0 have been turned off, the blue line eliminates differences in the contact rate λ_1, and the black dashed line corresponds to the case, where we also eliminate differences in the returns to experience, β.

in 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 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 15). 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 some-

Figure 16 – Growth Differences between the Median and Bottom LE



Note: This figures shows the decomposition of the differences in average lifetime log wage growth between bottom and median LE workers into human capital, search capital and negotiation rents. The left panel shows these differences in the benchmark, the middle panel plots differences when ex-ante heterogeneity in job loss and job finding rates δ, λ_0 have been turned off, the panel eliminates the heterogeneity in the contact rate λ_1,.

what 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 7).

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 worker's human capital accumulation, regardless whether he stays with the current employer or switches to a new one. Therefore, if the data shows high wage growth for a group of workers relative to the median regardless of job switching, the model attributes this to a high returns to experience. This is the case in the data for higher LE individuals (Figure

3b).

The difference in earnings growth between stayers and switchers are 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 median.

7 Conclusion

This paper investigates the causes of the large heterogeneity in lifetime earnings. Differences in earnings growth over the working life is an important contribution to lifetime earnings inequality. Using detailed administrative data from SSA records, we show that the earnings growth is surprisingly similar for the bottom 80% 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 the rate at which individuals accumulate human capital are key for inequality at the upper half of the distribution. Differences in ex-ante job ladder risk—job loss, job finding and contact rates heterogeneities—are the most important factor quantitatively.

References

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

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.

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.

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.

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 US 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*.

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), S11–44.

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.

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

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 us. 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 danish 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.

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., Technical report, mimeo.

Piketty, T. and Saez, E. (2003). Income inequality in the united states: 1913-1998. *Quarterly Journal of Economics*, **118** (1), 1–39.

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 us unemployment rate so much lower? *NBER macroeconomics annual*, **13**, 11–61.

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.

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 U.S. Administrative Data from 1974 to 2004*. Tech. rep., 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.

A Moments for Top Earners

A.1 Pareto Tails of the Earnings Distribution

Figure 17 – Pareto Tails in the Top 5% of Lifetime Earnings Distribution

(a) Log Density of Lifetime Earnings   (b) Log Inverse CDF of Lifetime Earnings

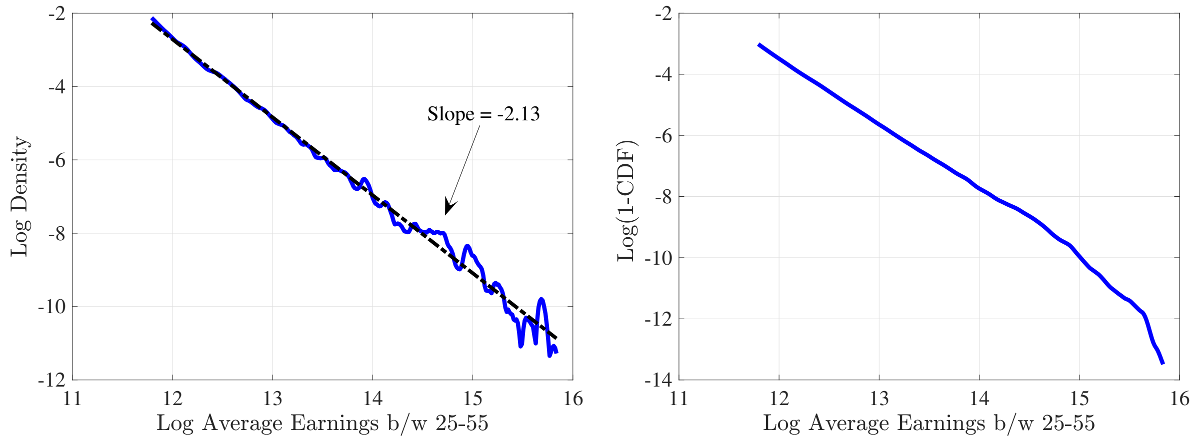
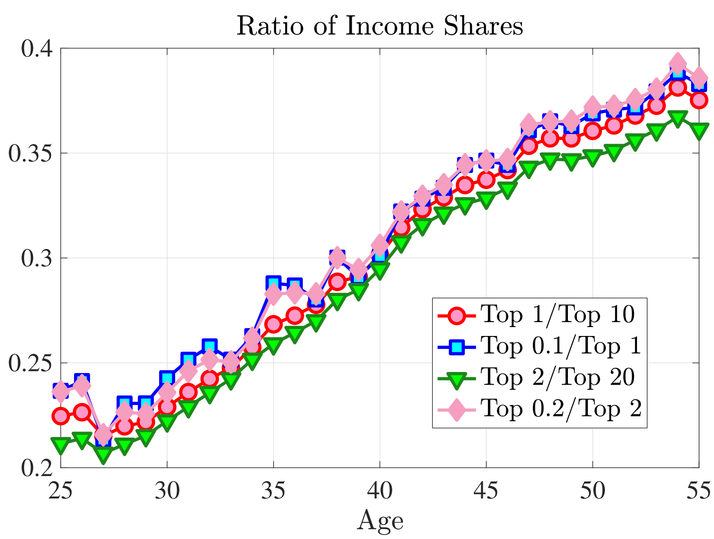


Figure 18 – Ratios of Top Income Shares



A.2 Earnings Growth of Top Earners

FIGURE 19 – Log Density of Top 5% of within Age Earnings Distribution

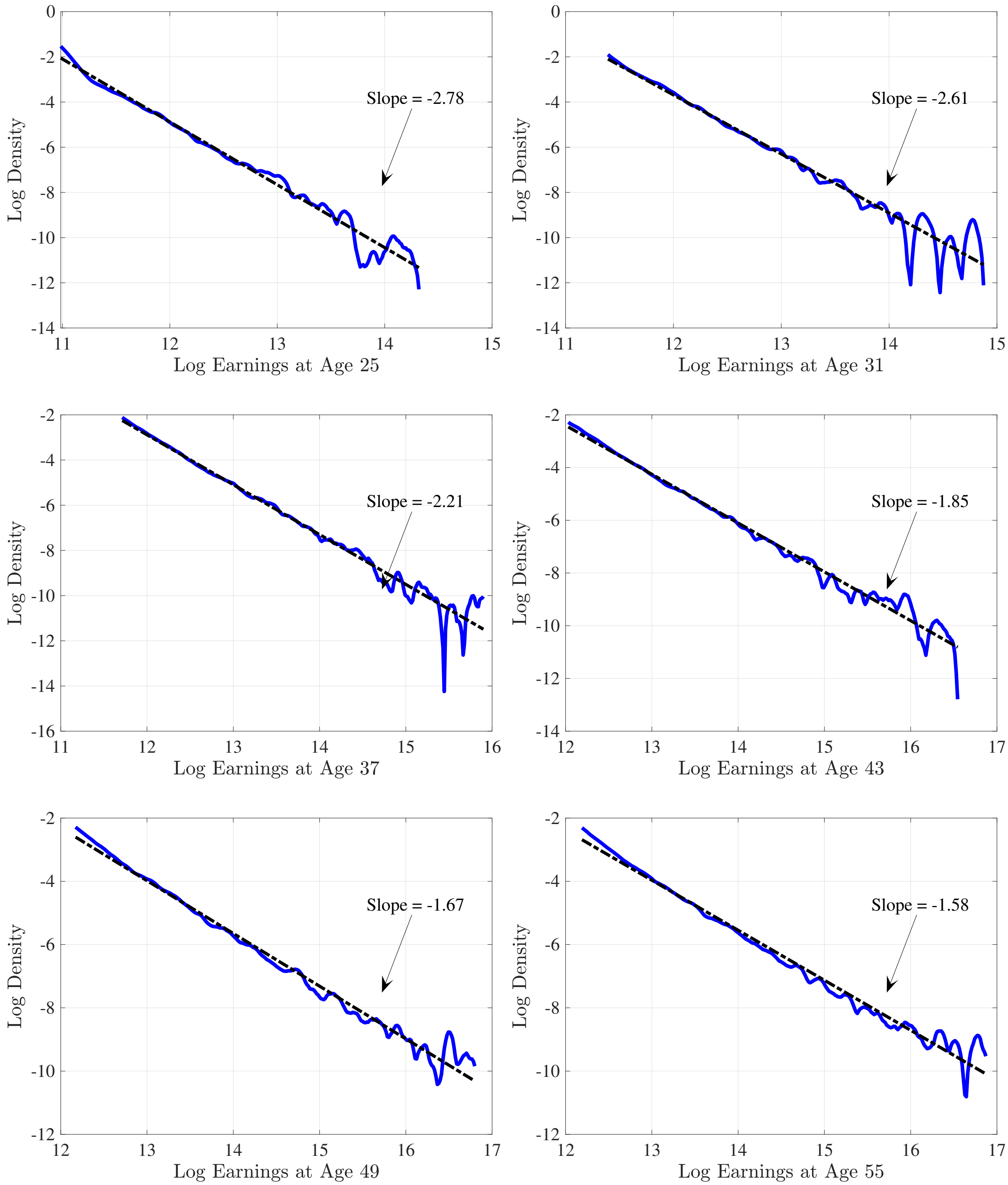


Figure 20 – Log Inverse CDF of Top 5% of within Age Earnings Distribution

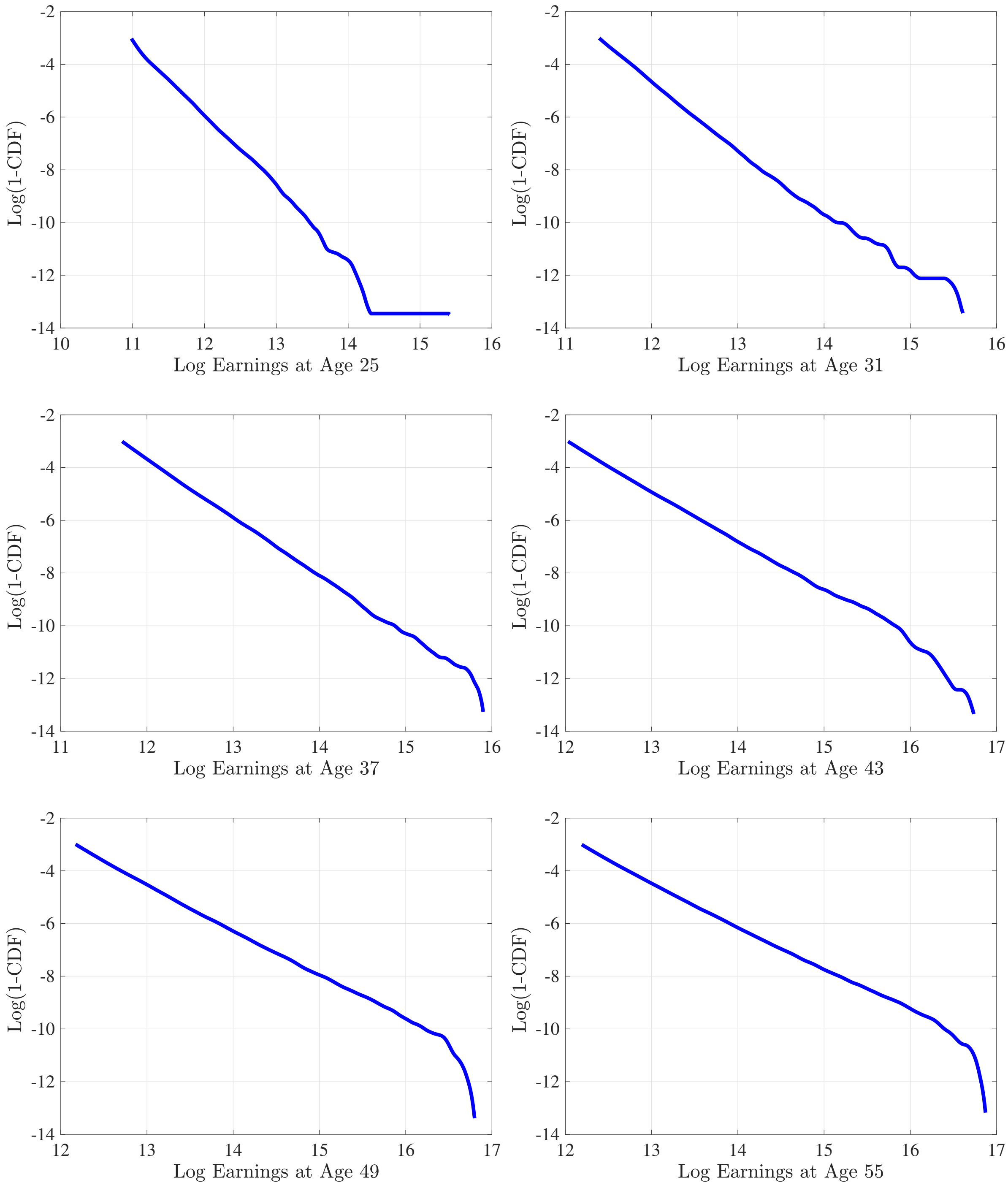
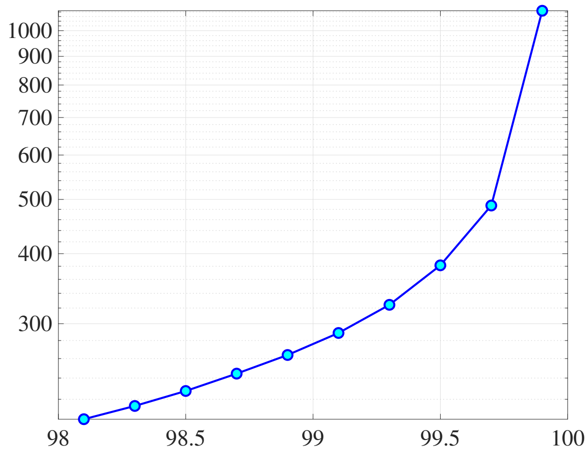
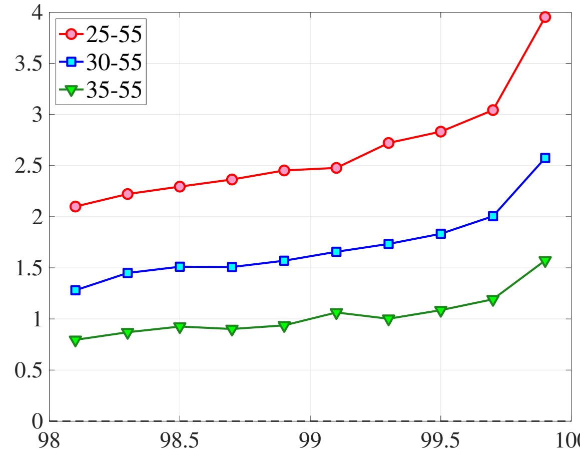


Figure 21 – Heterogeneity in lifetime earnings growth

(A) Average earnings over the lifetime, \$1,000



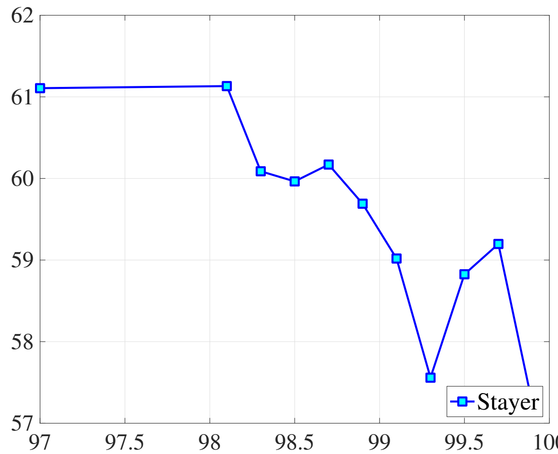
(B) Lifetime earnings growth, $\log \overline{Y}_{55} - \log \overline{Y}_h$



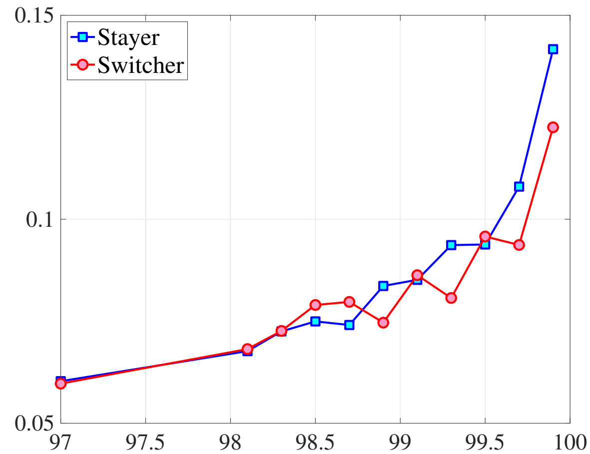
Note: The left panel shows the average annual earnings over the life cycle for each LE group. The right panel shows the log difference of average earnings $\bar{Y}$ between age 55 and various ages over the LE distribution.

Figure 22 – Job stayers and switchers

(A) Fraction of job stayers, %

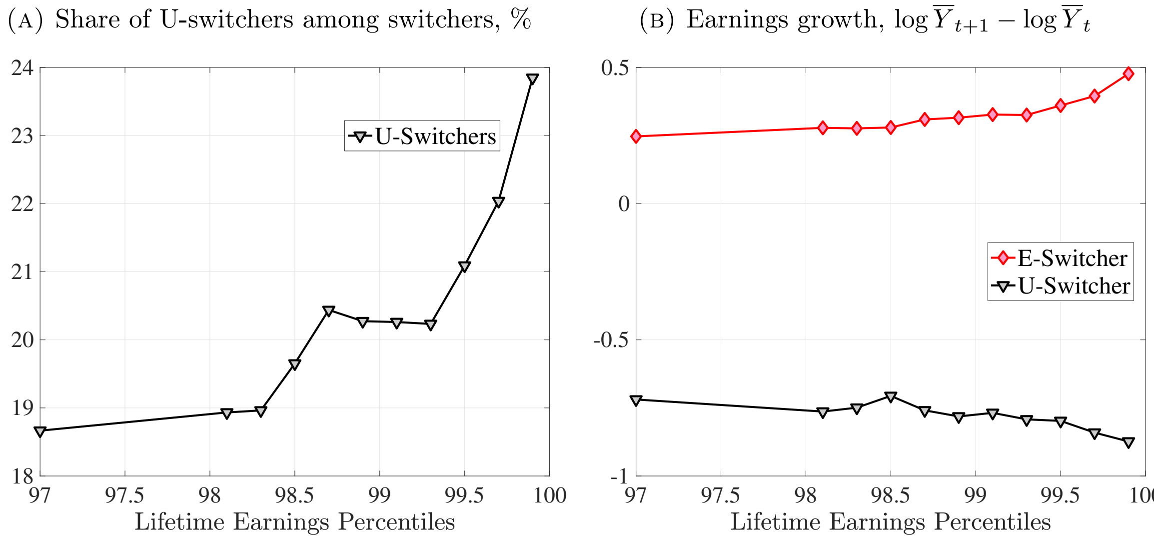


(B) Earnings growth, $\log \overline{Y}_{t+1} - \log \overline{Y}_t$



Note: The left panel shows the fraction of workers in each LE group that 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 23 – E-switchers and U-switchers



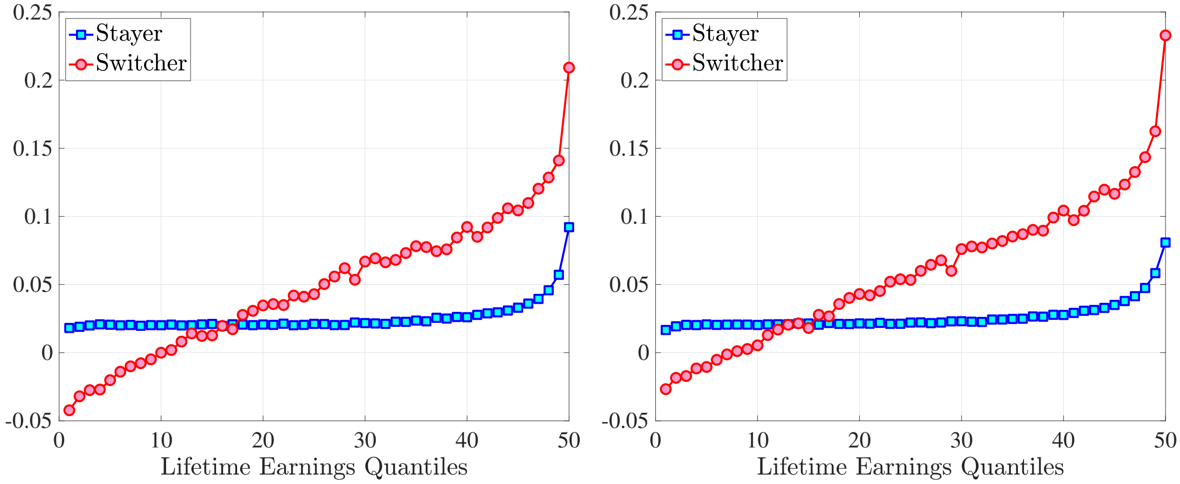
Note: The left panel shows the share of U-switchers among job stayers in each LE group. The right panel plots the log growth of average earnings $\bar{Y}$ between t and $t+1$ for U-switchers and E-switchers separately.

B Moments for a Broader Sample

C Earnings Growth Using Full-Year Employment

Figure 24 – Job stayers and switchers

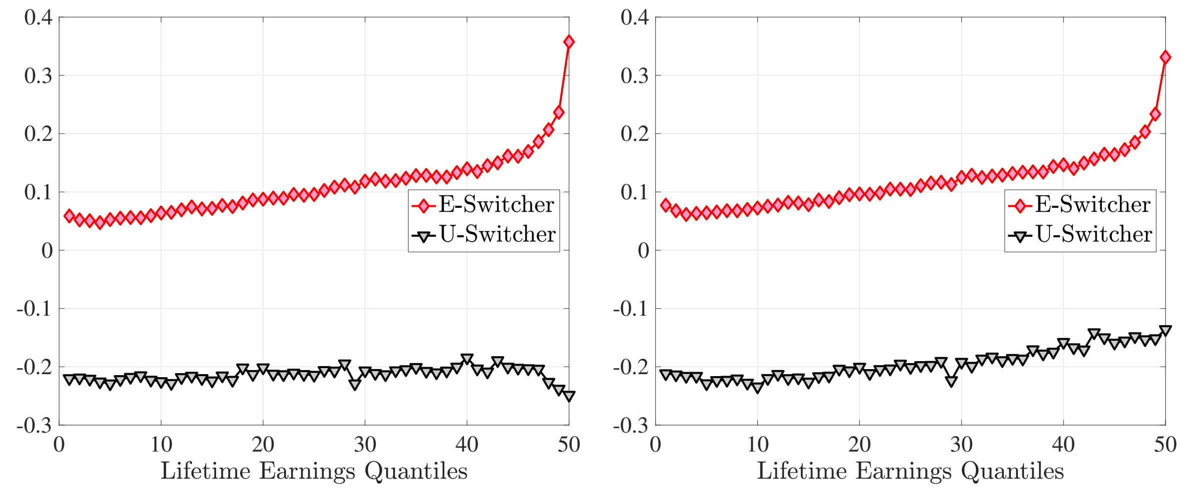
(A) Log average earnings growth, $\log \overline{Y}_{t+1} - \log \overline{Y}_t$ (B) Average log earnings growth, $\mathbb{E}\left[y_{t+1} - y_t\right]$



Note: The left panel shows the fraction of workers in each LE group that 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 25 – E-switchers and U-switchers

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



Note: 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.

D Percentiles and moments of earnings growth for job stayers and nonstayers

In section 2, we documented several facts about the mean annual earnings growth of job stayers and switchers. This section documents investigates the distribution of earnings changes for these groups by lifetime earnings percentiles. Figures 26–28 plot the standard deviation, skewness and kurtosis of earnings growth for stayers and nonstayers. Each figure consists of two panels. The left panel show the relevant moment of earnings growth in a specific age, whereas the right panel shows the distribution of earnings growth over a decade for various decades of the working life (25–35, 36–45, and 46–55). Figures 29–30 show the differences between selected percentiles of these distributions.

FIGURE 26 – Standard deviation of wage growth: Stayers and Nonstayers

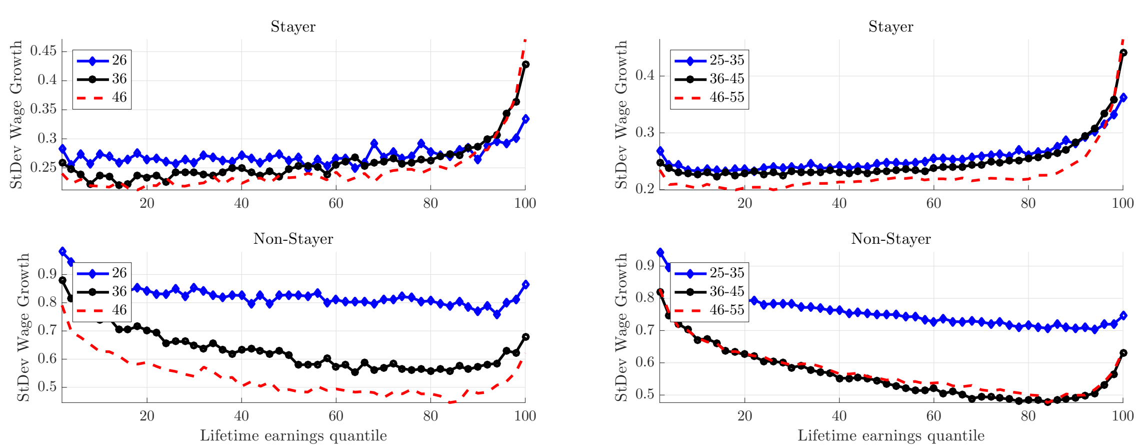


Figure 27 – Skewness of wage growth: Stayers and Nonstayers

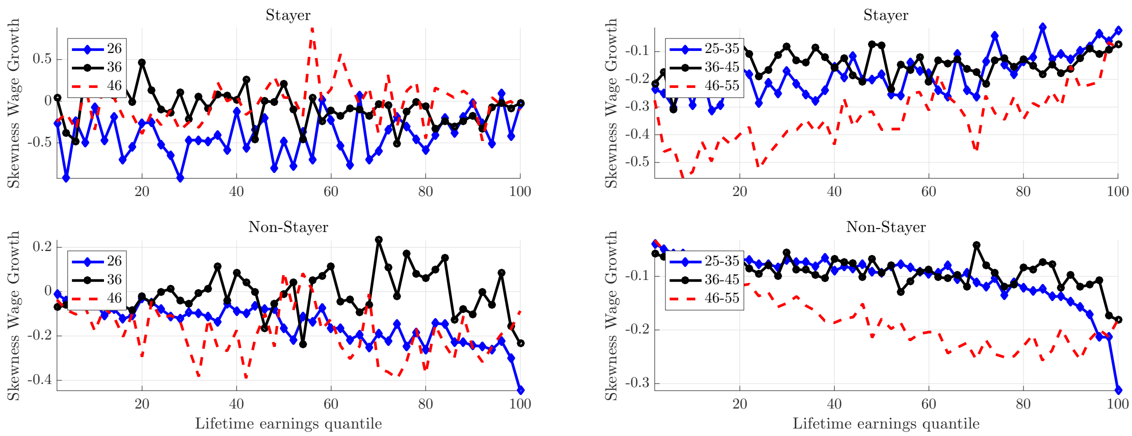


Figure 28 – Kurtosis of wage growth: Stayers and Nonstayers

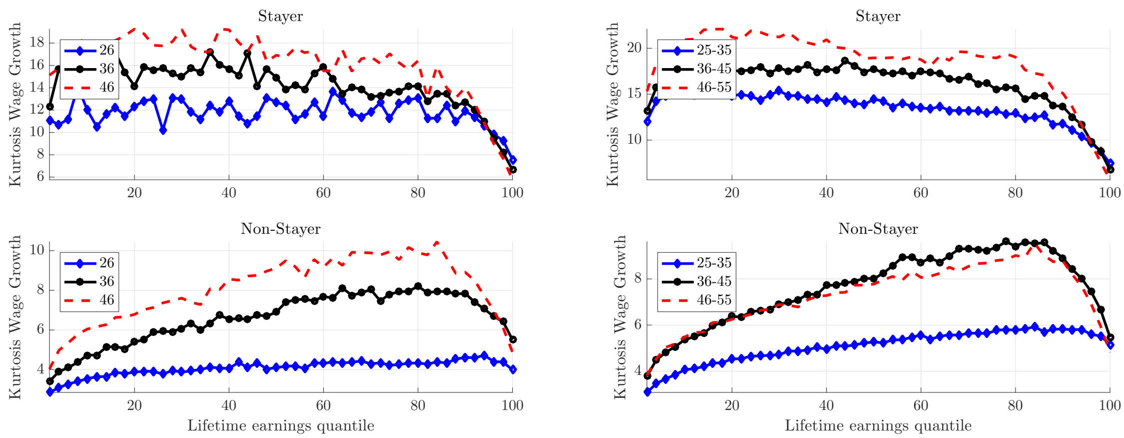


Figure 29 – P99–P1 of wage growth: Stayers and Nonstayers

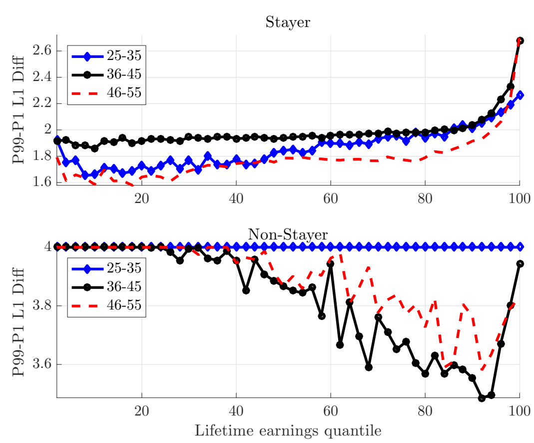


FIGURE 30 – P95–P5 of wage growth: Stayers and Nonstayers

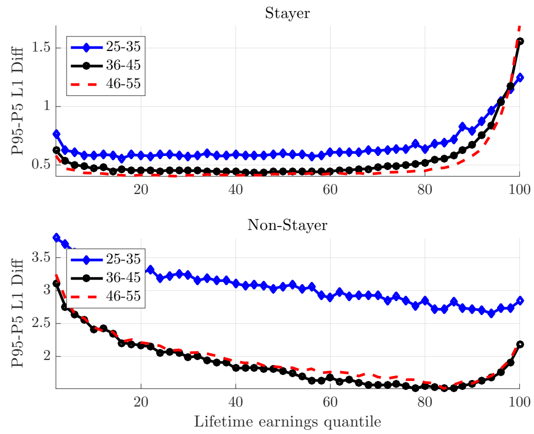
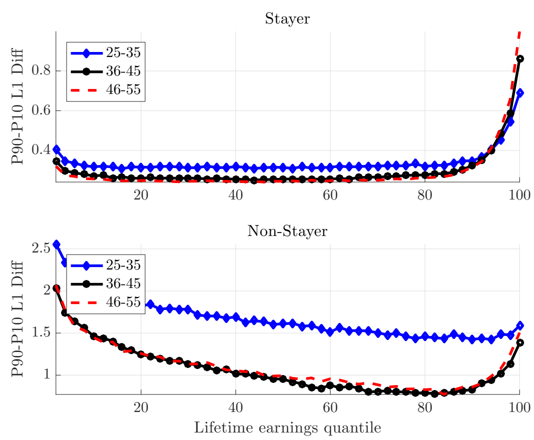


FIGURE 31 – P90–P10 of wage growth: Stayers and Nonstayers



E Survey of Income and Program Participation

There are two important drawbacks of the SSA data. The first is its annual frequency, which doesn't allow us to see higher frequency movements in earnings. The second is that it doesn't 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 consists 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.[39] We also use the SIPP to compute labor market flow rates for individuals by educational attainment.

E.0.1 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 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.

E.0.2 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.

F Survey of Consumer Expectations

[Describe SCE here generally and add the special survey used. Cite Aysegul's work and borrow some descriptions from there.]

G Model derivations

To the baseline model in Bagger et al., we add a recall option. Let λ_r denote the probability of recall. The superscript "o" refers to old workers, m refers to middle age, and y refers to young individuals. We start by deriving everything for old workers and then proceed backwards.

[39] 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.

The value function is as follows:

$$
\begin{aligned}
V^o(r,h_t,p) &= w_t + \frac{\delta^o}{1+\rho}V_0^o(h_t) \\
&+\frac{\lambda_1^o}{1+\rho}\int_p^{\bar{p}} E_t\left[(1-\theta)V^o(0,h_{t+1},p)+\theta V^o(0,h_{t+1},x)\right]dF(x) \\
&+\frac{\lambda_1^o}{1+\rho}\int_{q^o(r,h_t,p)}^{p} E_t\left[(1-\theta)V^o(0,h_{t+1},x)+\theta V^o(0,h_{t+1},p)\right]dF(x) \\
&+\frac{1}{1+\rho}\left[1-\delta^o-\lambda_1^o\bar{F}(q^o(r,h_t,p))\right]E_tV^o(r,h_{t+1},p) \qquad (8)
\end{aligned}
$$

Integrating (8) by parts, we obtain

$$
\begin{aligned}
V^o(r,h_t,p) &= w_t + \frac{\delta^o}{1+\rho}V_0^o(h_t) \\
&+\frac{1}{1+\rho}E_t\Bigg\{(1-\delta^o)V^o(r,h_{t+1},p) \\
&+\lambda_1^o\theta\int_p^{\bar{p}}\frac{\partial V^o}{\partial x}(0,h_{t+1},x)\bar{F}(x)dx \\
&+\lambda_1^o(1-\theta)\int_{q^o(r,h_t,p)}^{p}\frac{\partial V^o}{\partial x}(0,h_{t+1},x)\bar{F}(x)dx\Bigg\} \qquad (9)
\end{aligned}
$$

Applying (9) with $r=0$, and noting that $q(0,h_t,p)=p$, we get

$$
\begin{aligned}
V^o(0,h_t,p) &= p+h_t+\frac{\delta^o}{1+\rho}V_0^o(h_t) \\
&+\frac{1}{1+\rho}E_t\Bigg\{(1-\delta^o)V^o(0,h_{t+1},p) \\
&+\lambda_1^o\beta\int_p^{\bar{p}}\frac{\partial V^o}{\partial x}(0,h_{t+1},x)\bar{F}(x)dx\Bigg\}
\end{aligned}
$$

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

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

This expression, upon collecting terms yields

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

Substituting this expression back in (9), and letting $C^o(p) \equiv \frac{1}{\rho+\delta^o+\lambda_1^o\theta\bar{F}(x)}$, we get

$$\begin{aligned} V^o(r,h_t,p) &= w_t + \frac{\delta^o}{1+\rho}V_0^o(h_t) \\ &+\frac{1}{1+\rho}E_t\Big\{(1-\delta^o)V^o(r,h_{t+1},p) \\ &+\lambda_1^o\theta\int_p^{\bar{p}}(1+\rho)C^o(x)\bar{F}(x)\,dx \\ &+\lambda_1^o(1-\theta)\int_{q^o(r,h_t,p)}^{p}(1+\rho)C^o(x)\bar{F}(x)\,dx\Big\} \end{aligned} \tag{10}$$

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

$$E_tV^o(r,h_{t+1},p) = E_t\{V^o(0,h_{t+1},q^o) + \theta[V^o(0,h_{t+1},p) - V^o(0,h_{t+1},q^o)]\} \tag{11}$$

We first rewrite this as follows:

$$E_tV^o(r,h_{t+1},p) - V^o(0,h_{t+1},q^o) = \theta E_t\{V^o(0,h_{t+1},p) - V^o(0,h_{t+1},q^o)\}$$

Substituting (10) into this, and rearranging terms, we obtain

$$\begin{aligned} &r+p-q^o(r,h_t,p) + \frac{1-\delta^o}{1+\rho}[V^o(r,h_{t+2},p) - V^o(0,h_{t+2},q^o)] \\ &\quad -\lambda_1^o\theta\int_{q^o}^{p}\frac{(1+\rho)\bar{F}(x)}{\rho+\delta^o+\lambda_1^o\theta\bar{F}(x)}dx \\ &\quad +\frac{\lambda_1^o(1-\theta)}{1+\rho}\int_{q^o(r,h_t,p)}^{p}\frac{(1+\rho)\bar{F}(x)}{\rho+\delta^o+\lambda_1^o\theta\bar{F}(x)}dx \\ &= E_t\Big\{\theta[p-q^o(r,h_t,p)] \\ &\quad +\theta\frac{1-\delta^o}{1+\rho}[V^o(0,h_{t+2},p) - V^o(0,h_{t+2},q^o) \\ &\quad -\lambda_1^o\theta^2\int_{q^o}^{p}\frac{(1+\rho)\bar{F}(x)}{\rho+\delta^o+\lambda_1^o\theta\bar{F}(x)}dx \end{aligned}$$

Rearranging terms, we obtain

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

Substituting the last term with (11) and using the law of iterated expectations, we get

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

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

$$r=-(1-\theta)\left[p-q^o(r,p)\right]-\lambda_1^o(1-\theta)^2\int_{q^o(r,p)}^{p}C^o(x)\,\bar{F}(x)\,dx \tag{12}$$

Solution for middle-age workers

The value function for middle-age workers is as follows:

$$\begin{aligned}
V^m(r,h_t,p) &= w_t + \frac{\delta^m}{1+\rho} V_0^m(h_t) \\
&+ \frac{\lambda_1^m}{1+\rho} \int_p^{\bar{p}} E_t\left[(1-\theta)V^m(0,h_{t+1},p) + \theta V^m(0,h_{t+1},x)\right] dF(x) \\
&+ \frac{\lambda_1^m}{1+\rho} \int_{q^m(r,h_t,p)}^{p} E_t\left[(1-\theta)V^m(0,h_{t+1},x) + \theta V^m(0,h_{t+1},p)\right] dF(x) \\
&+ \frac{\gamma^o}{1+\rho} E_t V^o(r,h_{t+1},p) \\
&+ \frac{1}{1+\rho}\left[1-\delta^m-\gamma^o-\lambda_1^m \bar{F}\left(q^m(r,h_t,p)\right)\right] E_t V^m(r,h_{t+1},p)
\end{aligned} \tag{13}$$

Integrating (13) by parts, we obtain

$$\begin{aligned}
V^m(r,h_t,p) &= w_t + \frac{\delta^m}{1+\rho} V_0^m(h_t) \\
&+ \frac{1}{1+\rho} E_t \Bigg\{ (1-\delta^m-\gamma^o) V^m(r,h_{t+1},p) \\
&+ \lambda_1^m \theta \int_p^{\bar{p}} \frac{\partial V^m}{\partial x}(0,h_{t+1},x)\bar{F}(x)\,dx \\
&+ \lambda_1^m (1-\theta) \int_{q^m(r,h_t,p)}^{p} \frac{\partial V^m}{\partial x}(0,h_{t+1},x)\bar{F}(x)\,dx \\
&+ \gamma^o E_t V^o(r,h_{t+1},p) \Bigg\}
\end{aligned} \tag{14}$$

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

$$\begin{aligned}
V^m(r,h_t,p) &= w_t + \frac{\delta^m}{1+\rho}V_0^m(h_t) \\
&+\frac{1}{1+\rho}E_t\Bigg\{(1-\delta^m-\gamma^o)V^m(r,h_{t+1},p) \\
&+\lambda_1^m\theta\int_p^{\bar{p}}\frac{\partial V^m}{\partial x}(0,h_{t+1},x)\bar{F}(x)\,dx \\
&+\lambda_1^m(1-\theta)\int_{q^m(r,h_t,p)}^{p}\frac{\partial V^m}{\partial x}(0,h_{t+1},x)\bar{F}(x)\,dx \\
&+\gamma^o\Bigg(w_{t+1}+\frac{\delta^o}{1+\rho}V_0^o(h_{t+2}) \\
&+\frac{1}{1+\rho}\Bigg\{(1-\delta^o)V^o(r,h_{t+2},p) \\
&+\lambda_1^o\theta\int_p^{\bar{p}}\frac{(1+\rho)\bar{F}(x)}{\rho+\delta^o+\lambda_1^o\theta\bar{F}(x)}dx \\
&+\lambda_1^o(1-\theta)\int_{q^o(r,h_{t+1},p)}^{p}\frac{(1+\rho)\bar{F}(x)}{\rho+\delta^o+\lambda_1^o\theta\bar{F}(x)}dx\Bigg\}\Bigg)\Bigg\}
\end{aligned} \tag{15}$$

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

$$\begin{aligned}
V^m(0,h_t,p) &= p+h_t + \frac{\delta^m}{1+\rho}V_0^m(h_t) \\
&+\frac{1}{1+\rho}E_t\Bigg\{(1-\delta^m-\gamma^o)V^m(0,h_{t+1},p) \\
&+\lambda_1^m\theta\int_p^{\bar{p}}\frac{\partial V^m}{\partial x}(0,h_{t+1},x)\bar{F}(x)\,dx \\
&+\lambda_1^m(1-\theta)\int_{q^m(0,h_t,p)=p}^{p}\frac{\partial V^m}{\partial x}(0,h_{t+1},x)\bar{F}(x)\,dx \\
&+\gamma^o\Bigg(w_{t+1}+\frac{\delta^o}{1+\rho}V_0^o(h_{t+1}) \\
&+\frac{1}{1+\rho}\Bigg\{(1-\delta^o)V^o(0,h_{t+2},p) \\
&+\lambda_1^o\theta\int_p^{\bar{p}}\frac{(1+\rho)\bar{F}(x)}{\rho^o+\delta^o+\lambda_1^o\theta\bar{F}(x)}dx \\
&+\lambda_1^o(1-\theta)\int_{q^o(0,h_{t+1},p)=p}^{p}\frac{(1+\rho)\bar{F}(x)}{\rho+\delta^o+\lambda_1^o\theta\bar{F}(x)}dx\Bigg\}\Bigg)\Bigg\}
\end{aligned}$$

This boils down to

$$
\begin{aligned}
V^m(0,h_t,p) &= p+h_t+\frac{\delta^m}{1+\rho}V_0^m(h_t) \\
&+\frac{1}{1+\rho}E_t\Bigg\{(1-\delta^m-\gamma^o)V^m(0,h_{t+1},p) \\
&+\lambda_1^m\theta\int_p^{\bar{p}}\frac{\partial V^m}{\partial x}(0,h_{t+1},x)\bar{F}(x)\,dx \\
&+\gamma^o\Bigg(p+h_{t+1}+\frac{\delta^o}{1+\rho}V_0^o(h_{t+1}) \\
&+\frac{1}{1+\rho}\Bigg\{(1-\delta^o)V^o(0,h_{t+2},p) \\
&+\lambda_1^o\theta\int_p^{\bar{p}}\frac{(1+\rho)\bar{F}(x)}{\rho+\delta^o+\lambda_1^o\theta\bar{F}(x)}dx\Bigg)\Bigg\}\Bigg\}
\end{aligned}
$$

Differentiating this with respect to p, we get

$$
\begin{aligned}
\frac{\partial V^m}{\partial p}(0,h_t,p) &= 1 \\
&+\frac{1}{1+\rho}E_t\Bigg\{(1-\delta^m-\gamma^o)\frac{\partial V^m}{\partial p}(0,h_{t+1},p) \\
&-\lambda_1^m\theta\frac{\partial V^m}{\partial p}(0,h_{t+1},p)\bar{F}(p) \\
&+\gamma^o\Bigg(1+\frac{1}{1+\rho}\Bigg\{(1-\delta^o)\frac{\partial V^o}{\partial p}(0,h_{t+2},p) \\
&-\lambda_1^o\theta\frac{(1+\rho)\bar{F}(p)}{\rho+\delta^o+\lambda_1^o\theta\bar{F}(p)}\Bigg\}\Bigg)\Bigg\}
\end{aligned}
$$

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

$$\begin{aligned}\frac{\partial V^m}{\partial p}(0,h_t,p) &= 1\\ &+\frac{1}{1+\rho}E_t\Bigg\{(1-\delta^m-\gamma^o)\frac{\partial V^m}{\partial p}(0,h_{t+1},p)\\ &-\lambda_1^m\theta\frac{\partial V^m}{\partial p}(0,h_{t+1},p)\bar{F}(p)\\ &+\gamma^o\Bigg(1+\frac{1}{1+\rho}\Bigg\{(1-\delta^o)\frac{1+\rho}{\rho+\delta^o+\lambda_1^o\theta\bar{F}(p)}\\ &-\lambda_1^o\theta\frac{(1+\rho)\bar{F}(p)}{\rho+\delta^o+\lambda_1^o\theta\bar{F}(p)}\Bigg\}\Bigg)\Bigg\}\end{aligned}$$

Collecting terms, we get

$$\begin{aligned}\frac{\partial V^m}{\partial p}(0,h_t,p) &= 1\\ &+\frac{1}{1+\rho}E_t\Bigg\{\left(1-\delta^m-\gamma^o-\lambda_1^m\theta\bar{F}(p)\right)\frac{\partial V^m}{\partial p}(0,h_{t+1},p)\\ &+\gamma^o\Bigg(1+\frac{1}{1+\rho}\Bigg\{(1-\delta^o)\frac{1+\rho}{\rho+\delta^o+\lambda_1^o\theta\bar{F}(p)}\\ &-\lambda_1^o\theta\frac{(1+\rho)\bar{F}(p)}{\rho+\delta^o+\lambda_1^o\theta\bar{F}(p)}\Bigg\}\Bigg)\Bigg\}\end{aligned}$$

Assume that $\frac{\partial V^m}{\partial p}(0,h_t,p)$ is independent of h (to be verified later in some sense). This means we can drop the expectation operator on the left hand side. Then, we get

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

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

$$
\frac{1+\frac{\gamma^o}{1+\rho}\left(1+\frac{1}{1+\rho}\left(1-\delta^o-\lambda_1^o\theta\bar{F}\left(p\right)\right)\frac{\partial V^o}{\partial p}\right)}{1-\frac{1-\delta^m-\gamma^o-\lambda_1^m\theta\bar{F}(p)}{1+\rho}}
$$

$$
\begin{aligned}
\frac{\partial V^m}{\partial p}\left(0,h_t,p\right) &= \frac{1+\frac{\gamma^o}{1+\rho}\left(1+\frac{1}{1+\rho}\left\{\left(1-\delta^o-\lambda_1^o\theta\bar{F}\left(p\right)\right)\frac{\partial V^o}{\partial p}\right\}\right)}{1-\frac{1-\delta^m-\gamma^o}{1+\rho}+\frac{\lambda_1^m\theta\bar{F}(p)}{1+\rho}} \\
&= \left(1+\rho\right)\frac{1+\frac{\gamma^o}{1+\rho}\left(1+\frac{1}{1+\rho}\left\{\left(1-\delta^o\right)\frac{1+\rho}{\rho+\delta^o+\lambda_1^o\theta\bar{F}(p)}-\lambda_1^o\theta\frac{(1+\rho)\bar{F}(p)}{\rho+\delta^o+\lambda_1^o\theta\bar{F}(p)}\right\}\right)}{\rho+\delta^m+\gamma^o+\lambda_1^m\theta\bar{F}\left(p\right)} \\
&= \left(1+\rho\right)\frac{1+\frac{\gamma^o}{1+\rho}\left(1+\frac{1-\delta^o}{\rho+\delta^o+\lambda_1^o\theta\bar{F}(p)}-\frac{\lambda_1^o\theta\bar{F}(p)}{\rho+\delta^o+\lambda_1^o\theta\bar{F}(p)}\right)}{\rho+\delta^m+\gamma^o+\lambda_1^m\theta\bar{F}\left(p\right)} \\
&= \left(1+\rho\right)\frac{1+\frac{\gamma^o}{\rho+\delta^o+\lambda_1^o\theta\bar{F}(p)}}{\rho+\delta^m+\gamma^o+\lambda_1^m\theta\bar{F}\left(p\right)} \\
&\equiv \left(1+\rho\right)C^m\left(p\right) \qquad (16)
\end{aligned}
$$

Then, we substitute the expression for $\frac{\partial V^m}{\partial x}(0, h_t, x)$ into (15), we obtain

$$\begin{aligned}
V^m(r, h_t, p) &= w_t + \frac{\delta}{1+\rho} V_0^m(h_t) \\
&+ \frac{1}{1+\rho} E_t \Bigg\{ (1 - \delta^m - \gamma^o) V^m(r, h_{t+1}, p) \\
&+ \lambda_1^m \theta (1+\rho) \int_p^{\bar{p}} C^m(x) \bar{F}(x)\, dx \\
&+ \lambda_1^m (1-\theta)(1+\rho) \int_{q^m(r,h_t,p)}^{p} C^m(x) \bar{F}(x)\, dx \\
&+ \gamma^o \Bigg(w_{t+1} + \frac{\delta^o}{1+\rho} V_0^o(h_{t+1}) \\
&+ \frac{1}{1+\rho} \Bigg\{ (1-\delta^o) V^o(r, h_{t+1}, p) \\
&+ \lambda_1^o \theta (1+\rho) \int_p^{\bar{p}} C^o(x) \bar{F}(x)\, dx \\
&+ \lambda_1^o (1-\theta)(1+\rho) \int_{q^o(r,h_{t+1},p)}^{p} C^o(x) \bar{F}(x)\, dx \Bigg\} \Bigg) \Bigg\}
\end{aligned} \tag{17}$$

Now, to obtain the equation that implicitly defines $q^m()$, we need to combine (17) with (11) and arrange terms. But first, we rewrite equation (11).

$$E_t \{V^m(r, h_{t+1}, p) - V^m(0, h_{t+1}, q^m(r, h_t, p))\} = \theta E_t \{V^m(0, h_{t+1}, p) - V^m(0, h_{t+1}, q^m(r, h_t, p))\}.$$

Combining equation (17) with the expression above, and rearranging terms, we obtain

$$
\begin{aligned}
E_t\left\{V^m\left(r,h_{t+1},p\right)-V^m\left(0,h_{t+1},q^m(r,h_t,p)\right)\right\} = \\
p+r-q^m(r,h_t,p) \\
+\frac{1-\delta^m-\gamma^o}{1+\rho}E_t\left[V^m\left(r,h_{t+2},p\right)-V^m\left(0,h_{t+2},q^m(r,h_t,p)\right)\right] \\
-\lambda_1^m\beta\int_{q^m(r,h_t,p)}^{p}C^m\left(x\right)\bar{F}\left(x\right)dx \\
+\lambda_1^m\left(1-\beta\right)\int_{q^m(r,h_t,p)}^{p}C^m\left(x\right)\bar{F}\left(x\right)dx \\
+\frac{\gamma^o}{1+\rho}\left[p+r-q^m(r,h_t,p)\right] \\
+\frac{\gamma^o\left(1-\delta^o\right)}{\left(1+\rho\right)^2}E_t\left[V^o\left(r,h_{t+2},p\right)-V^o\left(0,h_{t+2},q^m\right)\right] \\
-\frac{\gamma^o\lambda_1^o\beta}{1+\rho}\int_{q^m}^{p}C^o\left(x\right)\bar{F}\left(x\right)dx \\
\left.+\frac{\gamma^o\lambda_1^o\left(1-\beta\right)}{1+\rho}\int_{q^o(r,h_{t+1},p)}^{p}C^o\left(x\right)\bar{F}\left(x\right)dx\right) = \\
\beta E_t\left\{V^m\left(0,h_{t+1},p\right)-V^m\left(0,h_{t+1},q^m(r,h_t,p)\right)\right\} = \\
\beta\left[p-q^m(r,h_t,p)\right]+\beta\frac{\gamma^o}{1+\rho}\left[p-q^m(r,h_t,p)\right] \\
+\beta\frac{1-\delta^m-\gamma^o}{1+\rho}E_t\left[V^m\left(0,h_{t+2},p\right)-V^m\left(0,h_{t+2},q^m(r,h_t,p)\right)\right] \\
-\lambda_1^m\beta^2\int_{q^m}^{p}C^m\left(x\right)\bar{F}\left(x\right)dx \\
\left.+\beta\frac{\gamma^o\left(1-\delta^o\right)}{\left(1+\rho\right)^2}E_t\left[V^o\left(0,h_{t+2},p\right)-V^o\left(0,h_{t+2},q^m(r,h_t,p)\right)\right]\right) \\
-\frac{\gamma^o\lambda_1^o\beta^2}{1+\rho}\int_{q^m}^{p}C^o\left(x\right)\bar{F}\left(x\right)dx
\end{aligned}
$$

We now collect terms and obtain

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

Noting that 1) $E_{t}\left[\beta V^{m}(0,h_{t+2},p)-V^{m}(r,h_{t+2},p)\right]$ equals $-(1-\beta)\,E_{t}V^{m}\left(0,h_{t+2},q^{m}(r,h_{t+1},p)\right)$, 2) and $E_{t}\left[\beta V^{o}(0,h_{t+2},p)-V^{o}(r,h_{t+2},p)\right]$ equals $-(1-\beta)\,E_{t}V^{o}\left(0,h_{t+2},q^{o}(r,h_{t+1},p)\right)$, and plugging these to the expression above, we obtain

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

Further rearranging and algebra yields,

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

Note that as in the Bagger et al. paper, we ignore solutions that depend on h and focus on deterministic solutions. For us, this means that the next to last line evaluates to 0. Since this also implies that the functions q^m and q^o depend on h in a trivial way, we drop those from the notation. We should solve the following equation for q^m.

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

H Estimation

H.1 Targeted moments in the estimation

H.2 Numerical method for estimation

H.3 Additional results

FIGURE 32 – Cross-sectional moments of earnings growth for job stayers and switchers

FIGURE 33 – Fraction of job stayers, E- and U-switchers by LE and age groups

FIGURE 34 – Earnings growth of job stayers, E- and U-switchers by LE and age groups

FIGURE 35 – Earnings levels by LE groups at selected ages

FIGURE 36 – Model fit on cross-sectional moments by age

FIGURE 37 – Model vs. Data: Labor market flows with age variation

(A) EU–rate: Model vs. SIPP data (B) UE–rate: Model vs. SIPP data (C) EE–rate: Model vs. SIPP data

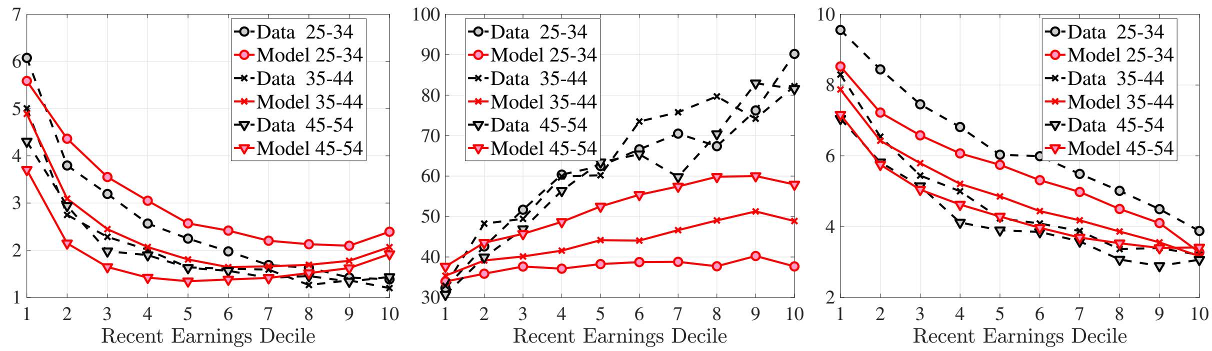
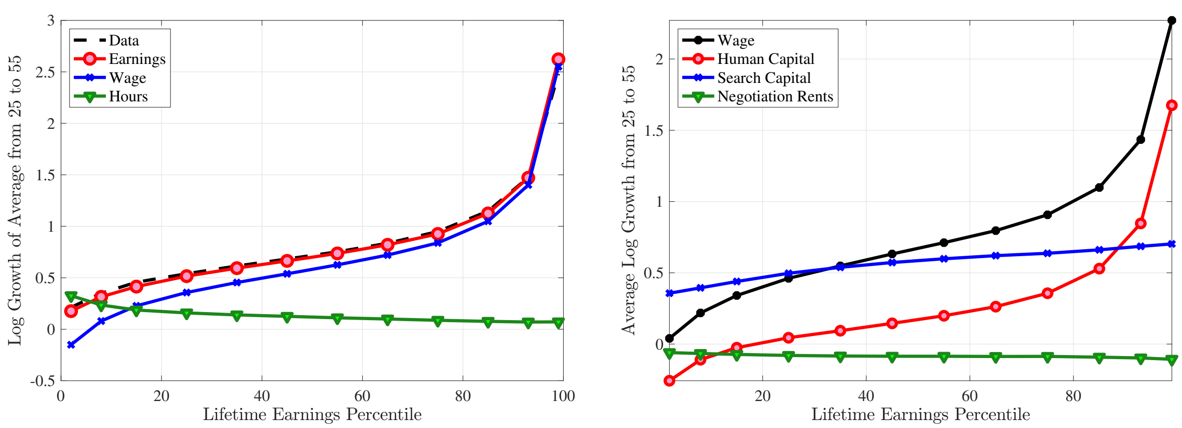


FIGURE 38 – Decomposing Earnings and Wage Growth

(A) Earnings, wage and hours growth (B) Human capital, search capital, negotiation rents



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