Earnings losses and labor mobility over the life cycle*

Philip Jung† Moritz Kuhn‡

October, 2017

Abstract

Large and persistent earnings losses following displacement have adverse consequences for the individual worker and the macroeconomy. Leading models cannot explain their size and disagree on their sources. Two mean-reverting forces make earnings losses transitory in these models: search as an upward force allows workers to climb back up the job ladder, and separations as a downward force make nondisplaced workers fall down the job ladder. We show that job stability at the top rather than search frictions at the bottom is the main driver of persistent earnings losses. We provide new empirical evidence on heterogeneity in job stability and develop a life-cycle search model to explain the facts. Our model offers a quantitative reconciliation of key stylized facts about the U.S. labor market: large worker flows, a large share of stable jobs, and persistent earnings shocks. We explain the size of earnings losses by dampening the downward force. Our new explanation highlights the tight link between labor market mobility and earnings dynamics. Regarding the sources, we find that over 85% stem from the loss of a particularly good job at the top of the job ladder. We apply the model to study the effectiveness of two labor market policies, retraining and placement support, from the Dislocated Worker Program. We find that both are ineffective in reducing earnings losses in line with the program evaluation literature.

JEL: E24, J63, J64

Keywords: Life-cycle labor market mobility, Job tenure, Earnings Losses, Worker- and match-specific skills

*We thank seminar participants at various institutions and conferences for useful comments. We especially thank Rudi Bachmann, Christian Bayer, Steven Davis, Georg Duernecker, Mike Elsby, Fatih Guvenen, Marcus Hagedorn, Berthold Herrendorf, Andreas Hornstein, Philipp Kircher, Tom Krebs, Lars Ljungqvist, Iourii Manovskii, Giuseppe Moscarini, Daniel Sullivan, Gianluca Violante, and Ludo Visschers for many suggestions and insightful comments. The usual disclaimer applies.

†TU Dortmund University and IZA, philip.jung@tu-dortmund.de, 44221 Dortmund, Germany.

‡University of Bonn, CEPR, and IZA, mokuhn@uni-bonn.de, 53113 Bonn, Germany.

1 Introduction

Large and persistent earnings losses following job displacement are a prime source of income risk in macroeconomic models (Rogerson and Schindler (2002)). They amplify the costs of business cycles (Krebs (2007), Krusell and Smith (1999)) and increase the persistence of unemployment after adverse macroeconomic shocks (Ljungqvist and Sargent (1998)). Understanding their size and sources is important for macroeconomic policies. However, leading models of the labor market do not provide much guidance, emphasizing different sources and accounting only for small and transitory earnings losses (Davis and von Wachter (2011)). The inability of existing models to account for large and persistent earnings losses calls for an explanation.

This paper offers an explanation based on an estimated structural life-cycle search and matching model of the U.S. economy. It is built around the observation that both an upward and a downward force prevent earnings losses from looming large in most models. The upward force is search. Displaced workers who fall off the job ladder can search on and off the job, trying to climb back up. Search frictions prevent an immediate catch-up, but, given the large job-to-job transition rates observed in the data, search is a powerful mean-reverting mechanism. The downward force is separations at the top of the job ladder. Short match durations due to high separation rates quickly make a currently nondisplaced worker look similar to a displaced worker. These two forces induce mean reversion of the earnings process and make earnings losses transitory and short-lived in most search models.

To explain persistent earnings losses, this paper shifts the emphasis away from displaced workers' inability to recover after displacement and toward the job stability of nondisplaced workers' employment paths. We provide empirical evidence on job stability and heterogeneity in worker mobility by age and tenure based on the Current Population Survey (CPS). We show that the coexistence of large worker turnover (Shimer (2012)) with a large share of stable jobs (life-time jobs in Hall (1982)) dampens the downward force but keeps the upward force in place. This turns the job ladder into a mountain hike that requires free climbing at the bottom but offers a fixed-rope route at the top. Reaching the top takes long, but once workers arrive at the top, the hike becomes a convenient and secure walk. The economic rationale for this job ladder is simple and intuitive: employers and employees in high-surplus jobs agree on high wages and low separation rates, in both cases because of a high surplus. We provide empirical evidence supporting such a negative correlation between wages and separation rates using data from the Survey of Income and Program Participation (SIPP).

Focusing on the earnings paths of nondisplaced workers at the top of the job ladder rather than displaced workers offers a new perspective on the actual size of earnings losses. It also sheds new light on the sources of earnings losses and how they matter for policy. We show that estimators of earnings losses pioneered by Jacobson et al. (1993) and today's standard in the literature have a sizable selection effect due to their construction of the control group of nondisplaced workers. We decompose the sources of earnings losses and find that up to 30% of the estimated earnings losses result from a selection effect, 20% from increased job instability, and 50% from lower wages. Decomposing wage losses further, we find that more than 85% stem from the loss of a particularly good job, meaning a fall from the top of the job ladder. We discuss how our findings matter for active labor market policy. We use the model to study the effectiveness of retraining and placement support programs of the Dislocated Worker Program of the Workforce Investment Act. We find very limited scope for active labor market policies to reduce earnings losses, mirroring the findings from the empirical program evaluation literature (Card et al. (2010)). Our structural model offers a clear reason for this failure: active labor market policy operates on search frictions and could foster mean reversion by making displaced workers recover to the average. However, we argue that active policy cannot affect the downward force that makes nondisplaced workers look so different from the average.

Our emphasis on the evolution of nondisplaced workers' earnings paths rather than the recovery path of displaced workers makes our explanation distinct from previous attempts to explain earnings losses. Existing attempts focus on dampening the upward force of search for better jobs, either by adding search frictions directly or by introducing deterioration of job prospects due to displacement. Explanations based on the deterioration of accumulated experience or skills during unemployment (Ljungqvist and Sargent (2008)) struggle to endogenously account for worker mobility because workers are very reluctant to switch jobs in the presence of large expected skill losses (den Haan et al. (2005)). This explanation also has to rule out subsequent skill accumulation on the job to avoid mean reversion. Others, as we do, point toward the loss of a particularly good job as an explanation for earnings losses (Low et al. (2010)). Falling down the job ladder subsequently leads to more frequent job losses, more unemployment, and job instability (Stevens (1997) and Pries (2004)). Recent explanations in the same spirit can be found in Krolikowski (2017), who makes the job ladder very long, and Jarosch (2014), who makes the job ladder slippery. All of these explanations have in common that they attempt to prevent displaced workers from climbing up the job ladder. However, while frictions to move upward must also exist for our explanation to work, we show that shutting down the downward force is a crucial step for slowing down mean reversion and accounting for large and persistent earnings losses. Without job stability at the top of the job ladder, alternative explanations are likely to fail because the job ladder is a powerful mechanism for mean reversion (Low et al. (2010), Hornstein et al. (2011)). High job stability in high-wage jobs is a key ingredient in generating persistent earnings differences. Our new explanation highlights the tight link between labor market mobility and earnings dynamics.

Our model features heterogeneity in job stability with stable jobs at the top of the job ladder. It jointly accounts for high labor market mobility and persistent earnings losses. To account for high labor market mobility, we need a high degree of transferability of skills in the labor market, and to account for persistent earnings losses, we need jobs at the top of the job ladder that are very stable. The highlighted mechanism explains the inability of most existing labor market models to generate large and persistent earnings losses. They do not account for heterogeneity in job stability but impose a single separation rate across jobs, matching average mobility uniformly along the job ladder. Hence, workers rotate continuously out of good jobs, which results in earnings losses that are highly transitory and short-lived.

We develop a search and matching model that accounts for life-cycle effects and has various sources of skill heterogeneity and on-the-job search. Search is directed (Menzio and Shi (2011)), and wage and mobility choices are efficiently bargained (den Haan et al. (2000a)). The model not only captures the empirical facts on tenure and wages as in Moscarini (2005) but also accounts for the mobility pattern by tenure and age, adding to a recently growing strand of the literature on life-cycle labor market models.[1] Introducing life-cycle dynamics is crucial for our explanation because it copes with the nonstationary dynamics of tenure by age that we document, and it helps to disentangle the relative importance of different components of the skill accumulation process. We explain how we exploit heterogeneity in worker mobility by

[1]Examples for life-cycle models are Menzio et al. (2016), Cheron et al. (2013), and Esteban-Pretel and Fujimoto (2014).

age and tenure to identify model parameters as alternative to an identification relying on wage dynamics and wage heterogeneity.

Regarding mobility, the model accounts for high average worker mobility even for older workers (Farber (1995)), a large fraction of stable jobs (Hall (1982)), and frequent job changes during the first 10 years of working life (Topel and Ward (1992)). Regarding earnings dynamics, the model accounts for a declining age profile of wage gains after job changes and substantial early career wage growth due to job changes (Topel and Ward (1992)), large returns to tenure estimated using the methodology advocated in Topel (1991) and small returns to tenure estimated using the methodology advocated in Altonji and Shakotko (1987), permanent earnings shocks as in Heathcote et al. (2010), and large and persistent earnings losses following job displacement as in Couch and Placzek (2010), Davis and von Wachter (2011), and von Wachter et al. (2009).[2] The model also generates the empirically observed cross-sectional wage inequality that existing models struggle to explain (Hornstein et al. (2011)). Hence, our model not only speaks to the empirical literature studying earnings losses but also offers a quantitative reconciliation of key stylized facts about the U.S. labor market: the coexistence of large worker flows, a large share of stable jobs, and earnings dynamics with large and persistent shocks.

The quantitative success with respect to the size of the earnings losses allows us to quantify the sources of earnings losses. We implement an empirical estimator within our model and decompose earnings losses using counterfactual experiments that are only possible in a structural model. One source is a selection effect in the empirical estimator. We construct an ideal counterfactual experiment of "twin" workers using characteristics unobserved by the econometrician to make workers identical except for the displacement event. We find a sizable upward bias of 30% in estimated earnings losses. While the possibility of bias is well known, its quantitative size could only be localized within a range. Our findings close this gap. Although we emphasize job stability at the top of the job ladder and along the counterfactual employment path of displaced workers, we demonstrate that the assumption on the counterfactual employment path imposed in the empirical implementation strategy is too strong. Once we control for this selection effect, we use the twin experiment to measure the reduction in earnings resulting from lower average employment in the group of displaced workers relative to the group of nondisplaced workers. In

[2]Early contributors to the earnings loss literature are Ruhm (1991) and Stevens (1997). Farber (1999) provides an early survey.

our decomposition, this extensive margin effect accounts for 20%. As a result, direct skill losses account for the remaining 50%, what we call the wage loss effect. We adopt the empirical approach in Stevens (1997) based on data from the Panel Study of Income Dynamics (PSID) and demonstrate that our model-based decomposition is in line with empirical estimates. Given that the empirical earnings loss estimates are an input to many calibrated macroeconomic models, our findings suggest some caution in using the empirical findings at face value.

Our decomposition can go further because we observe in the model the evolution of skills of displaced and nondisplaced workers. We use this information to study whether the extensive margin and the wage loss effect arise from the loss of worker-specific skills or from the loss of a particularly good match. We find that match-specific skill losses account for more than 85% of both effects, therefore justifying the statement that earnings losses are the result of the loss of a particularly good job rather than the deterioration of worker-specific skills.

Our finding on the skill losses is highly relevant for the design of active labor market programs and motivates our policy analysis. We look at two policy pillars, retraining and placement support, of the Dislocated Worker Program of the Workforce Investment Act. We consider worker-specific skill losses as losses that can be restored via retraining, whereas match-specific skill losses need to be restored via placement support that improves the match between workers and jobs by supporting labor market search. Within our model, we implement a stylized retraining and placement support program and find that both programs are ineffective. Retraining will not help much because worker-specific skill losses account for only a small fraction of the earnings losses. Placement support remains ineffective because even if placement support could create six job offers per month (roughly the equivalent of one year of search in our model) and bring the worker back to the average match quality of the worker's cohort, the resulting earnings losses would be reduced by only one-fourth and would remain large and persistent. Hence, active policy might help to remove frictions and foster mean reversion by making displaced workers recover to the average but it cannot affect the downward force that makes nondisplaced workers persistently different from the average. It is the missing downward force due to job stability at the top that drives the persistence of earnings losses.

We proceed as follows: In Section 2, we perform an empirical analysis of worker mobility and job stability. Section 3 develops our life-cycle model of worker mobility and explains the identification of model parameters based on worker mobility. Section 4 discusses the model fit for worker mobility and presents the fit for untargeted earnings dynamics. Section 5 estimates the earnings losses following job displacement from the model and decomposes them. Section 6 studies labor market policies to counteract the adverse consequences of worker displacement. Section 7 concludes.

2 Empirical Analysis

Facts about average worker mobility have been widely documented (e.g. Shimer (2012) and Fallick and Fleischman (2004)). We highlight four facts documenting substantial heterogeneity in worker mobility: (1) transition rates from employment to non-employment and job-to-job transitions decline by age; (2) conditioning on tenure and looking at newly hired workers, transition rates decline by age, but the decline is much smaller than the unconditional decline by age; (3) despite large average transition rates, mean tenure increases linearly with age, showing that many jobs are very stable; (4) wages and separations are strongly negatively correlated, implying that high-wage jobs are more stable.

2.1 Data

Our analysis is based on U.S. data from the monthly Current Population Survey (CPS) files and the Occupational Mobility and Job Tenure supplements for the period 1980 to 2007.[3] In contrast to alternative data sources, the CPS offers large representative cross sections of workers and provides a long time dimension covering several business cycles. This fact allows us to abstract from business cycle fluctuations in transition rates by averaging transition rates over time. Tenure information is not available in the monthly CPS files but only in the irregular Occupational Mobility and Job Tenure supplements. We merge this information with the basic monthly files to construct transition rates by tenure.[4] We follow Shimer (2012) and Fallick and Fleischman (2004) in constructing worker flows. Job-to-job transitions and all transitions out of employment end tenure. To avoid overstating job stability, we take as the separation rate the sum of the transition rate to unemployment and out of the labor force. We relegate details on the data and construction of transition rate and tenure profiles to appendix A.1.

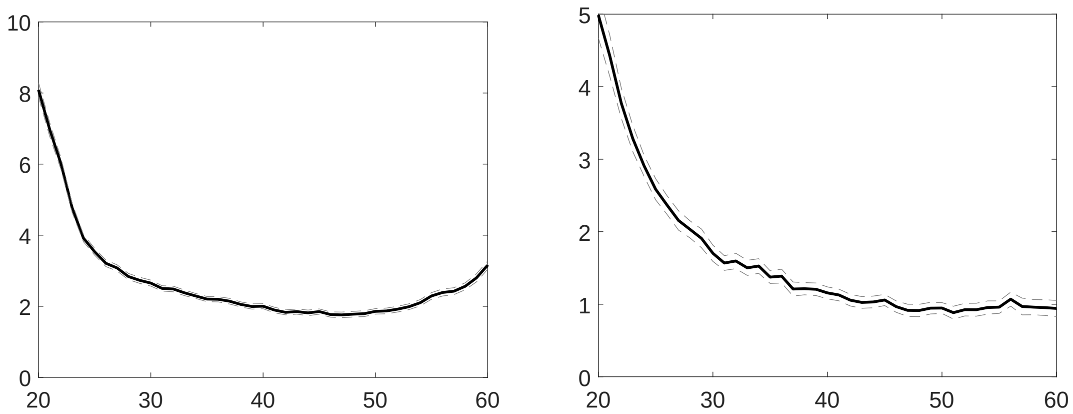
[3] December 2007 marks the beginning of the latest NBER recession. Since this recession marks a pronounced break in the time series of the transition rates, we exclude this time period from our sample.

[4] Tenure information from the supplement files has been used before to document a large share of highly stable jobs in the U.S. labor market (Hall (1982), Farber (1995, 2008), Diebold et al. (1997)).

2.2 Worker mobility and job stability

Figure 1 depicts age heterogeneity in monthly separation and job-to-job transition rates. Both transition rates fall with age. Most of the decrease in transition rates by age takes place between the ages of 20 and 30. This initial period is followed by 25 years of stable transition rates.[5] Separations drop from an initial high of 8% to a low of around 2%, and job-to-job transitions from an initial high of 5% to a low of about 1%. Even during the stable years between ages 30 and 50, approximately 3% of workers leave employers each month. Confidence bands around the profiles indicate that both profiles are tightly estimated.

Figure 1: Empirical age transition rate profiles



(a) Separation rate by age

(b) Job-to-job rate by age

Notes: Age profiles for separation and job-to-job rates. The gray dashed lines show confidence bands using $-/+2$ standard deviations. Transition rates are monthly. Standard deviations are bootstrapped using 10,000 repetitions from the pooled sample stratified by age. The horizontal axis shows age in years, and the vertical axis shows transition rates in percentage points.

The average transition rates by age mask further heterogeneity. Figure 2(a) shows that mean and median tenure increase almost linearly with age. If transition rates were uniform in the population and equal to the 3% of workers who leave employers between ages 30 and 50 every month, then mean tenure would converge to slightly less than 3 years, well below the observed 11 years of tenure at age 50. This shows that even conditional on age, there is large heterogeneity in transition rates. Again, confidence bands show that these profiles are tightly estimated.

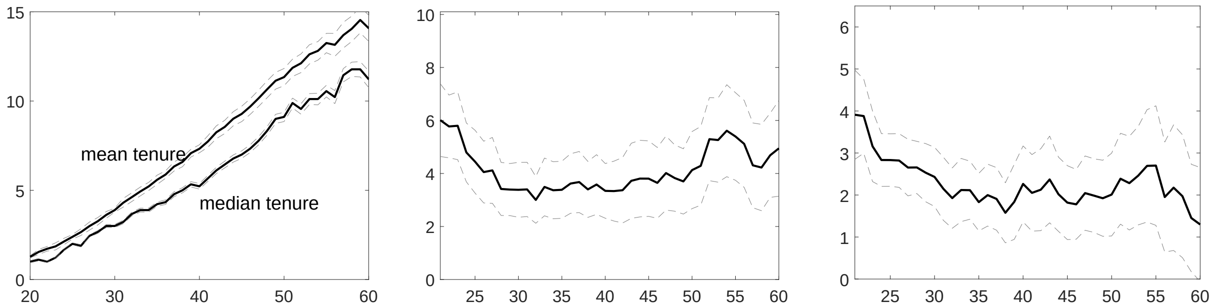
Next, we look at newly hired workers.[6] Considering newly hired workers helps to further

[5] Starting at the age of about 55, separation rates start to increase as workers leave the labor force.

[6] We refer to newly hired workers as those with one year of job tenure. This group is composed of both workers coming from other employers and non-employment. We use moving age windows with a range $+/-$ 2 years

unmask heterogeneity in worker mobility. We refer to age profiles for newly hired workers for simplicity as "newly hired age profiles". Figure 2 plots separation and job-to-job newly hired age profiles together with confidence bands. Two points are important. First, separation (figure 2(b)) and job-to-job (figure 2(c)) newly hired age profiles decline with age. As for the age profiles in figure 1, the decline is concentrated in the first 10 years in the labor market. Second, the decline by age for newly hired workers is about half of the unconditional decline by age. The separation rate declines by about 2.5 percentage points, and the job-to-job transition rate declines by about 1.7 percentage points in comparison to the unconditional 5 percentage points and 3 percentage points decline by age, respectively.[7]

Figure 2: Tenure by age and transition rates by age for newly hired workers



(a) Mean tenure by age

(b) Separation rate by age for new hires

(c) Job-to-job rate by age for new hires

Notes: Panel 2(a) shows mean and median tenure in years by age. The gray dashed lines show confidence bands using $-/+2$ standard deviations. Standard deviations are bootstrapped using 10,000 repetitions from the pooled sample stratified by age. The horizontal axis shows age in years, and the vertical axis shows tenure in years. Panels 2(b) and 2(c) show separation and job-to-job transition rates by age for newly hired workers. Newly hired workers are workers with one year of tenure. The gray dashed lines show confidence bands using $-/+2$ standard deviations. Transition rates are monthly. Standard deviations are bootstrapped using 10,000 repetitions from the pooled sample. The horizontal axis shows age in years starting at age 21, and the vertical axis shows transition rates in percentage points.

This evidence, together with the linear increase in tenure by age, points toward considerable heterogeneity in job stability. While wage heterogeneity has been studied extensively, much less attention has been paid to quantitatively account for the substantial heterogeneity in job stability in models of the labor market. Typically, models of the labor market are designed to explain and study average labor market flows. Our empirical analysis highlights a large share of stable jobs and substantial heterogeneity in worker mobility. As we document next,

centered at each age to construct age profiles.

[7]Here we consider age profiles starting at age 21, anticipating our theoretical model below where workers enter the labor market at age 20 and can have accumulated only one year of tenure at age 21.

this heterogeneity in job stability correlates strongly negatively with wages. We document that high-wage jobs are also very stable.

2.3 Job stability and wages

When studying the connection between wages and job stability, we want to explore whether high-wage jobs today are less likely to separate in the future. For this, we need individual-level panel data to observe future transitions to non-employment given the current wage. We therefore resort to data from the 2004 Survey of Income and Program Participation (SIPP).[8] We construct h-month separation rates. The h-month separation rate is the share of workers who are employed today but who separate at least once within the next h months into non-employment. We consider 4- and 12-month separation rates.[9] We explore the relationship between wages and job stability using two approaches. First, we run a regression of the h-month separation rate $\pi^h_{i,t}$ on log wages $\log(w_{i,t})$ and age dummies $\gamma^a_{i,t}$:

$$\pi^h_{i,t} = \beta \log(w_{i,t}) + \gamma^a_{i,t} + \varepsilon_{i,t},$$

where i indexes individuals and t calendar time. To focus on matches with high separation rates, we also run the regression for newly hired workers only.[10] Table 1 shows the coefficient β from the regressions. We find that coefficients are negative and significant at the 1% level in all specifications.

The coefficient β varies for the different specifications between -0.04 and -0.08. This implies that a 10% higher wage leads to a 0.4 to 0.6 percentage points lower separation rate over 4 months and a 0.7 to 0.8 percentage points lower separation rate over 12 months. This effect is economically significant, given an average separation rate of around 2 percentage points at age 40.

Second, we use residuals from a regression of log wages on age and group workers according to their residuals in wage deciles. We plot separation rates by wage decile in Figure 3. Looking

[8] In the CPS, the panel dimension is very limited, and wage information is only available at the last interview (outgoing rotation group). SIPP data have been used to analyze labor market flows before. The example most closely related to the current paper is Menzio et al. (2016). We provide details on the SIPP data in Section A.2 of the appendix.

[9] We choose 4-month separation rates as our baseline for two reasons. First, it allows us to deal with the well-known problem of seam bias in the SIPP. Second, by looking at a longer time horizon, we capture sufficiently many separations from stable jobs that have very few separations from month to month.

[10] We consider all workers in their initial month on a job as newly hired for this regression.

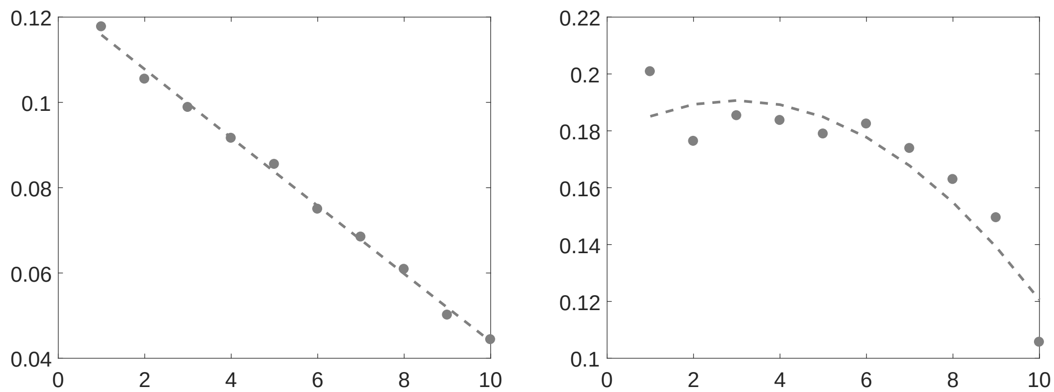
Table 1: Regression coefficients of separation rates on log wages

separation horizon (h)	4 months	12 months
All workers	-0.0392	-0.0668
std. error	(0.0004)	(0.0005)
Newly hired workers	-0.0548	-0.0822
std. error	(0.0016)	(0.0019)

Notes: Regression coefficient β from regression of 4-(12-)month separation rate on log wages and further controls. First row shows regression coefficient from regression with all workers and the corresponding standard errors. Second row shows regression coefficient when only newly hired workers are considered in the regression and the corresponding standard errors.

at all workers in figure 3(a), we find that between the lowest and the highest decile separation rates differ by a factor of almost 3 (0.12 vs. 0.04). In figure 3(b), we show the same wage-job stability relationship but look only at newly hired workers. Again we find a strongly negative relationship. Separation rates decline by roughly 30% across wage deciles (0.18 to 0.12).[11]

Figure 3: Wages and job stability



(a) All workers (b) Newly hired workers

Notes: Separation rates over a 4-month horizon by wage decile using SIPP data. The left panel shows separation rates for all workers as gray circles, and the dashed line shows a quadratic polynomial approximation to the data. The right panel shows separation rates for newly hired workers as gray circles, and the dashed line shows a quadratic polynomial approximation to the data. Workers are grouped in wage deciles using wage residuals. Wage deciles are on the horizontal axis. The vertical axis shows 4-month separation rates. See text for further details.

The next Section develops a structural life-cycle model with two-dimensional skill heterogeneity to account for the documented heterogeneity in worker mobility. The model also features the documented correlation between wages and job stability. By contrast, most existing models

[11] We tried using more control variables in the first-step regression and looked at shorter and longer horizons for the separation rate, and found that the negative relation between wages and job stability is robust.

assume that separations happen exogenously and thereby feature no correlation between wages and separation rates. Heterogeneity in job stability and the correlation with wages will be instrumental in generating large and persistent earnings losses, as we show in Section 5. In online appendix I, we use a simple example with two types to explain the intuition behind the tight link between earnings losses and heterogeneity in job stability.

3 Model

We develop a life-cycle labor market model in the search and matching tradition. For the most part, the building blocks of our model follow a large strand of the literature. Deviations are designed to capture the heterogeneity in labor market mobility and job stability outlined above. We describe the model here and relegate a discussion of our modeling assumptions to online appendix II.1. A detailed derivation of all equations can be found in online appendix II.2.

Time is discrete. There is a continuum of mass 1 of finitely lived risk-neutral agents and a positive mass of risk-neutral firms. Firms and workers discount the future at rate $\beta < 1$. Workers participate for T periods in the labor market followed by T_R periods of retirement. Each firm has the capacity to hire a single worker, and we refer to a worker-firm pair as a match. Agents differ by age a, a vector of skills x, and employment state $\varepsilon = \{e, n\}$ with e for employment and n for non-employment. We use primes to denote variables in the next period. In a slight abuse of notation, we drop primes if variables do not change between periods.

Each period is divided into four stages: bargaining, separation, production, and search. At the bargaining stage, each match bargains jointly about when to separate into non-employment, the amount of wages to be paid if the production stage is reached, and when to accept a job offer from another firm at the search stage. We assume generalized Nash bargaining over the total match surplus, which leads to individually efficient choices. Separations happen after the bargaining stage, job-to-job transitions and transitions from non-employment into employment happen at the search stage, and we assume that a worker's labor market status is observed at the production stage. Vacancy posting by firms is directed to submarkets of worker types $\{\varepsilon, a, x\}$. There is free entry to submarkets, and a matching function determines contact rates in each submarket.

3.1 Skill Process

The skill vector is $x = \{x_w, x_m\}$ where x_w is the skill level of the worker and x_m is the quality of the match. We assume that match-specific skills x_m are drawn at the beginning of a match according to a probability distribution $g(x_m)$ where g is taken to be a discrete approximation to the normal density with (exponential) mean normalized to 1 and variance σ_m^2. The match-specific skill component remains constant throughout the existence of a match. We also approximate worker-specific skill states x_w by a finite number of states in an ordered set. The smallest (largest) element is x_w^{min} (x_w^{max}), and the immediate predecessor (successor) of x_w is x_w^- (x_w^+). Workers start their life at the lowest skill level and stochastically accumulate skills. Skills accumulate only if a worker stays in the current match. The worker's skill level next period is x_w^+ with age-dependent probability $p_u(a)$, and it remains at x_w with probability $1 - p_u(a)$. The distribution over next period's worker skills x'_w if staying in a match is

$$x'_w = \begin{cases} x_w & \text{with probability } 1 - p_u(a) \\ x_w^+ & \text{with probability } p_u(a), \end{cases}$$

and we set $p_u(a) = 0$ for $x_w = x_w^{max}$. Age dependence follows from a simple recursion $p_u(a) = (1-\delta)p_u(a-1)$ to capture a potential slowdown in skill accumulation with age.

The transferability of worker skills in the labor market is imperfect. A worker of type x_w who takes a new job either from employment or non-employment faces the risk that part of the accumulated skills will not transfer to the new job. If the worker takes a new job, then with probability $1 - p_d$, all of the accumulated skills will transfer to the new job and the worker will remain at skill level x_w. With probability p_d, part of the accumulated skills will not transfer and the skill level next period will be x_w^-. We set $p_d = 0$ for $x_w = x_w^{min}$. The distribution over next period's worker skills x'_w in case of worker mobility is

$$x'_w = \begin{cases} x_w^- & \text{with probability } p_d \\ x_w & \text{with probability } 1 - p_d. \end{cases}$$

A worker who takes up a new job from non-employment faces the same skill transition. In addition, workers in non-employment do not accumulate skills so that skills during non-employment depreciate relative to employment. We discuss a model extension with additional skill depreciation during non-employment in online appendix IV.1.

To ease the exposition, we use $\mathbb{E}_s[\cdot]$ to denote the expectation over future skill states conditional on staying in the match (subscript s for staying) and $\mathbb{E}_m[\cdot]$ to denote the expectation conditional on changing jobs (subscript m for mobility). With this notation in place, we turn to a derivation of endogenous choices.

3.2 Value Functions

A worker-firm match with worker of age a and skill vector $x = \{x_w, x_m\}$ produces output y according to the production function $y = f(x_w, x_m) + \eta_s$, where η_s is an idiosyncratic transitory productivity shock assumed to be logistically distributed with distribution function $H(\eta_s)$ having a mean of zero and variance $\pi^2\psi_s^2/3$. For each match, there exists a cutoff value $\overline{\omega}$ for the productivity shock at which the match separates. Following den Haan et al. (2000a), this cutoff value is determined as part of the bargaining described below. Exploiting the assumption of a logistic distribution, we can write the probability of separating as $\pi_s \equiv H(\overline{\omega}) = (1 + exp(-\overline{\omega}/\psi_s))^{-1}$ and the conditional mean of the realized productivity shocks has a closed form given by $\Psi_s(\pi_s) \equiv \int_{\overline{\omega}}^{\infty} \eta dH(\eta)$.[12] In addition, there is a probability π_f of exogenous separation each period. The exogenous separation shock happens before the endogenous separation decision.

Let $J(x_w, x_m, a)$ denote the value of a firm that is matched at the beginning of the period to a worker of age a with productivity x. The value of the firm is[13]

$$\begin{aligned} J(x_w, x_m, a) &= (1-\pi_f)(1-\pi_s(x_w, x_m, a))\Bigg(f(x_w, x_m) + \frac{\Psi_s(\pi_s)}{1-\pi_s(x_w, x_m, a)} - w(x_w, x_m, a) \\ &+ (1-\pi_{eo}(x_w, x_m, a))\beta\mathbb{E}_s\left[J(x_w', x_m, a')\right]\Bigg). \end{aligned} \tag{1}$$

With probability π_f (π_s), the match separates exogenously (endogenously). Productivity shocks η_s are transitory i.i.d. shocks, and the endogenous separation probability depends on the current state of the match. By contrast, exogenous separations lead to separations irrespective

[12] We derive in online appendix II.2 that $\Psi_s(\pi_s) = -\psi_s(\pi_s \log(\pi_s) + (1-\pi_s)\log(1-\pi_s))$. We suppress arguments of π_s for notational convenience. Note that for endogenous separations there is a one-to-one relationship between the cutoff value and the probability of separating so we refer in a slight abuse of terminology to the bargaining as being over separation rates rather than cutoff values.

[13] When the worker reaches retirement age, the match separates and continuation values are zero.

of the current state of the match. If no separation occurs, the match transits to the production stage. Upon reaching the production stage, the match produces output and pays wages w. Integrating out productivity shocks, output comprises a component $\Psi_s(\pi_s)/(1-\pi_s(x_w, x_m, a))$. The value Ψ_s can be interpreted as an option value from having a choice to separate or not after having received a shock.[14] The fact that an option value arises is not a particular feature of our model but a generic feature of an endogenous mobility choice. The fact that it has an analytic representation results from our distributional assumption on shocks. With probability π_{eo} (described below), the worker makes a job-to-job transition; otherwise the match continues to the next period.

We denote the value function of an employed worker of age a with skill type x_w and matched to a firm of type x_m by $V_e(x_w, x_m, a)$, and $V_n(x_w, a)$ is the corresponding value of a non-employed worker. During non-employment, the worker receives flow utility b. At the search stage, non-employed workers receive job offers with type- and age-dependent probability $p_{ne}(x_w, a)$. Each job offer comes with a stochastic utility component attached to it. We denote the average utility component from job changing by κ_o and the stochastic, idiosyncratic part by η_o. The realization of the idiosyncratic part is independent of the current state. Depending on the match quality of the offer x'_m and the utility component, the worker decides whether to accept the offer or not. A non-employed worker chooses the maximum of $\{V_n(x_w, a'), \mathbb{E}_m\left[V_e(x'_w, x'_m, a')\right] - \kappa_o + \eta_o\}$. As for the productivity shocks η_s, we assume that the idiosyncratic utility component η_o is logistically distributed with mean zero and variance $\pi^2\psi_o^2/3$. The acceptance decision yields an option value $\Psi_{ne}(q_{ne})$ that arises because only job offers with high enough η_o will be accepted. We suppress arguments of q_{ne} for notational convenience. The option value will enter the value functions below. Using standard properties of the logistic distribution, we write the acceptance probability for a job offer of match type x'_m as

$$q_{ne}(x'_m; x_w, a) = \left(1 + \exp\left(\psi_o^{-1}\beta\Big(V_n(x_w, x_m, a') - \left(\mathbb{E}_m\left[V_e(x'_w, x'_m, a')\right] - \kappa_o\right)\Big)\right)\right)^{-1}. \quad (2)$$

Note that we condition the acceptance probability on the offer type x'_m, modeling match quality as an inspection good. The ex-ante value $V_n(x_w, a)$ before the realization of the idiosyncratic

[14] We refer to Ψ_s as the option value because the profile of observed productivity shocks looks like the payoff from a call option. Low productivity shocks will not be realized, and the match separates and high productivity shocks enter output one-for-one.

shock components is given by

$$
\begin{aligned}
V_n(x_w, a) = b + & \overbrace{p_{ne}(x_w, a) \sum_{x'_m} \left(q_{ne}(x'_m; x_w, a) \left(\beta \mathbb{E}_m \left[V_e(x'_w, x'_m, a')\right] - \kappa_o\right) \right) g(x'_m)}^{\text{receiving and accepting offer}} \\
+ & \underbrace{\sum_{x'_m} (1 - p_{ne}(x_w, a) q_{ne}(x'_m; x_w, a)) \beta V_n(x_w, a') g(x'_m)}_{\text{not receiving or not accepting offer}} + \underbrace{p_{ne}(x_w, a) \sum_{x'_m} \Psi_{ne}(q_{ne}) g(x'_m)}_{\text{option value}} \qquad (3)
\end{aligned}
$$

where the first line shows flow value b at the production stage and the case of receiving and accepting an offer at the search stage. The second line shows the case of not receiving or receiving but not accepting an offer and the option value in case an offer is received. The probability of entering employment combines the likelihood of receiving an offer p_{ne} with the probability of accepting an offer q_{ne} and is given by $\pi_{ne}(x_w, a) = \sum_{x'_m} p_{ne}(x_w, a) q_{ne}(x'_m; x_w, a) g(x'_m)$. An employed worker's value function is

$$
\begin{aligned}
V_e(x_w, x_m, a) &= (1 - \pi_f)(1 - \pi_s(x_w, x_m, a)) \left(w(x_w, x_m, a) + V_e^S(x_w, x_m, a) \right) \\
&+ \left((1 - \pi_f) \pi_s(x_w, x_m, a) + \pi_f \right) V_n(x_w, a), \qquad (4)
\end{aligned}
$$

where $V_e^S(x_w, x_m, a)$ denotes the value function for an employed worker at the search stage. With probability $(1 - \pi_f)(1 - \pi_s(x_w, x_m, a))$, the match does not separate and the worker receives wage $w(x_w, x_m, a)$ and enters the search stage providing value $V_e^S(x_w, x_m, a)$. If the match separates, the worker receives the value of non-employment $V_n(x_w, a)$. Note that the separation stage is before the production stage and the search stage, so that a worker who separates at the separation stage receives flow value b during the production stage and searches as non-employed during the search stage of the same period.

The search process on the job is similar to non-employment. The worker receives offers with type-dependent probability $p_{eo}(x_w, x_m, a)$. Each offer comes with the nonpecuniary component as when searching off the job with the stochastic component drawn from the same distribution. The cutoff value above which a competing job offer x'_m is accepted is determined as part of the bargaining. We denote the implied acceptance probability for job offer x'_m by $q_{eo}(x'_m; x_w, x_m, a)$ and the option value from accepting only offers with favorable utility component as $\Psi_{eo}(q_{eo})$.

The search stage value function is

$$
\begin{aligned}
V_e^S(x_w, x_m, a) \quad = \quad & \overbrace{p_{eo}(x,a) \sum_{x'_m} \Big(q_{eo}(x'_m; x, a) \left(\beta \mathbb{E}_m \left[V_e(x'_w, x'_m, a')\right] - \kappa_o\right) \Big) g(x'_m)}^{\text{receiving and accepting offer}} \\
+ \underbrace{\sum_{x'_m} (1 - p_{eo}(x,a) q_{eo}(x'_m; x, a)) \beta \mathbb{E}_s \left[V_e(x'_w, x_m, a')\right] g(x'_m)}_{\text{not receiving or not accepting offer}} & + p_{eo}(x,a) \underbrace{\sum_{x'_m} \Psi_{eo}(q_{eo}) g(x'_m)}_{\text{option value}} . \qquad (5)
\end{aligned}
$$

Note that acceptance probabilities on the job depend on the current match-specific type x_m. The probability of leaving combines acceptance probabilities q_{eo} with the probability of receiving an offer p_{eo}: $\pi_{eo}(x_w, x_m, a) = \sum_{x'_m} p_{eo}(x_w, x_m, a) q_{eo}(x'_m; x_w, x_m, a) g(x'_m)$.

3.3 Bargaining

Every match bargains at the bargaining stage over when to separate to non-employment at the separation stage, the wage that is paid if the match enters the production stage, and when to go to another firm at the search stage. We assume generalized Nash bargaining over the total surplus of the match with the worker's outside option being non-employment. The bargaining conditions on the realization of idiosyncratic shocks but given the risk neutrality of workers and firms only the expected value of the realized shock matters. To ease notation, we suppress state contingency with respect to idiosyncratic shocks and include only expected values in all equations. This bargaining follows den Haan et al. (2000a) or Jung and Kuester (2015) and it leads to an individually efficient outcome in which separations and job-to-job transitions occur only if the joint surplus of the match is too small. The bargaining solution satisfies

$$
\begin{aligned}
[w, \pi_s, q_{eo}(x'_m)] \quad &= \quad \arg\max J(x_w, x_m, a)^{1-\mu} \Delta(x_w, x_m, a)^{\mu} \\
s.t. \quad & \quad a, x_w, x_m \text{ given,}
\end{aligned}
$$

where $\Delta(x,a) = V_e(x,a) - V_n(x,a)$ denotes the worker surplus. We denote by $S(x,a) = \Delta(x,a) + J(x,a)$ the total match surplus at the bargaining stage. Wage payments and mobility decisions happen at different stages within the period. To ease exposition, we therefore define surpluses at the production stage and the search stage. The worker surplus at the search stage is $\Delta^S(x_w, x_m, a) = V_e^S(x_w, x_m, a) - V_n(x_w, a)$ and, in a slight abuse of terminology, we refer to

$S^S(x,a) = \mathbb{E}_s[\beta S(x'_w, x_m, a')] - \mathbb{E}_m[\beta\Delta(x'_w, x'_m, a')]$ as the surplus of staying in the current match relative to an outside offer at the search stage. At the production stage, the worker surplus is $\Delta^P(x,a) = w(x,a) + \Delta^S(x,a)$, and $J^P(x,a) = f(x) - w(x,a) + (1-\pi_{eo}(x,a))\beta\mathbb{E}_s[J(x',a')]$ is the firm's surplus net of idiosyncratic shocks.[15] The total surplus is $S^P(x,a) = \Delta^P(x,a) + J^P(x,a)$. We derive the solution to the bargaining in online appendix II.2. The solutions for $w(x_w, x_m, a)$, $\pi_s(x_w, x_m, a)$, and $q_{eo}(x'_m; x_w, x_m, a)$ are

$$\begin{aligned} \pi_s(x_w, x_m, a) &= \left(1 + \exp\left(\psi_s^{-1} S^P(x,a)\right)\right)^{-1} && (6)\\ w(x_w, x_m, a) &= \mu\left(S^P(x,a) + \frac{\Psi_s(\pi_s)}{1 - \pi_s(x_w, x_m, a)}\right) - \Delta^S(x_w, x_m, a) && (7)\\ q_{eo}(x'_m; x_w, x_m, a) &= \left(1 + \exp\left(\psi_o^{-1}\left(S^S(x,a) + \kappa_o\right)\right)\right)^{-1}. && (8) \end{aligned}$$

Joint bargaining links mobility choices π_s and q_{eo} to wages w. Mobility choices and wages are all functions of the match surplus. In general, the match surplus affects wages positively and mobility decisions negatively. Hence, the joint determination of wages and mobility decisions in our model will lead to high-surplus matches paying high wages and being very stable. This model feature matches the robust empirical correlation between wages and job stability reported in Section 2.3.

The separation probability π_s is proportional to the surplus S^P so that high-surplus matches are less likely to separate because firm and worker agree that they separate only after particularly bad productivity shocks. This is in contrast to exogenous separations that lead to separations independent of the match surplus and therefore let workers fall even from the top of the job ladder.

Wages are a linear function of the worker's share of the total surplus S^P and the option value Ψ_s minus the worker's surplus from searching on the job Δ^S. The fact that Ψ_s enters the wage equation is intuitive because the gains from having a choice to separate are shared between worker and firm. The option value captures the truncated favorable part of the transitory productivity shock distribution.[16] The negative Δ^S term represents a form of a compensating

[15] Note that $J^P(x,a)$ does not include the option value from the value function in eq. (1) but can be interpreted as the permanent component of the surplus.

[16] We assume here that firms provide full insurance against these shocks. Given our assumption of risk neutrality, this is without loss of generality. Alternatively, we demonstrate in online appendix V.2.4 how to interpret these shocks as transitory wage shocks.

differential for differences between on- and off-the-job search. The better on-the-job search is, the lower are wages.

Finally, acceptance decisions for outside offers depend on the match surplus at the search stage and utility component κ_o. A higher surplus of the current match reduces the likelihood of leaving.

3.4 Vacancy posting and matching

To limit computational complexity and to avoid the age structure as an additional aggregate state, we borrow ideas from the literature on directed search (e.g., Menzio and Shi (2011)) and assume that there exist submarkets for all types $\{\varepsilon, a, x\}$. When entering the market, firms direct vacancies to one submarket. To determine the number of vacancies, we impose free entry on each submarket:

$$\kappa = p_{vn}(x_w, a)\beta \sum_{x'_m} q_{ne}(x'_m; x_w, a)\mathbb{E}_m \left[J(x'_w, x'_m, a')\right] g(x'_m) \tag{9}$$

$$\kappa = p_{vo}(x_w, x_m, a)\beta \sum_{x'_m} q_{eo}(x'_m; x_w, x_m, a)\mathbb{E}_m \left[J(x'_w, x'_m, a')\right] g(x'_m), \tag{10}$$

where κ denotes vacancy posting costs, $p_{vn}(x_w, a)$ denotes the contact rate from the firm's perspective with non-employed workers of type x_w and age a, and $p_{vo}(x_w, x_m, a)$ denotes the contact rate from the firm's perspective with employed workers of type x_w in a match of quality x_m and age a. Given the worker's current state, the firm forms expectations about the expected profits, taking into account the worker's acceptance probability for the offer.

Contact rates in each submarket are determined using a Cobb-Douglas matching function $m = \varkappa v^{1-\varrho} u^{\varrho}$ in vacancies v and searching workers u with matching elasticity ϱ and matching efficiency $\varkappa$. We allow for different matching efficiencies between on- and off-the-job search but not across submarkets of skill types or age.[17] The contact rates for non-employed and on-the-job search are

$$p_{vn}(x_w, a) = \varkappa_n \left(\frac{n(x_w, a)}{v_n(x_w, a)}\right)^{\varrho} = \varkappa_n \theta_n(x_w, a)^{-\varrho}, \tag{11}$$

$$p_{vo}(x_w, x_m, a) = \varkappa_o \left(\frac{l(x_w, x_m, a)}{v_o(x_w, x_m, a)}\right)^{\varrho} = \varkappa_o \theta_o(x_w, x_m, a)^{-\varrho}, \tag{12}$$

[17] We provide a model extension to explore the effects of deteriorating search efficiency during non-employment in online appendix IV.1.

where $l(x_w, x_m, a)$ denotes the number of employed workers at the search stage, $v_o(x_w, x_m, a)$ the number of posted vacancies for a particular worker type, and $\theta_o(x, a)$ labor market tightness. The value $n(x_w, a)$ denotes the number of non-employed workers at the search stage, $v_n(x_w, a)$ the number of posted vacancies for a particular worker type, and $\theta_n(x_w, a)$ labor market tightness. Contact rates from the worker's perspective are $p_{eo}(x_w, x_m, a) = \varkappa_o \theta_o(x_w, x_m, a)^{1-\varrho}$ and $p_{ne}(x_w, a) = \varkappa_n \theta_n(x_w, a)^{1-\varrho}$, respectively.

3.5 Parameter identification based on worker transition rates

This section discusses identification of model parameters. The existing literature typically relies on wage data to identify parameters of the skill process (see Bagger et al. (2014) for a recent example). We propose an alternative approach that identifies the parameters of the skill process using the documented worker transition rates from section 2. Our identification approach transforms the ideas of Topel (1991), who also uses wage data, to data on worker transition rates. In our model, wages and worker transition rates are directly linked as bargaining outcomes. In this way, they provide similar information about the evolution of skills over time and across jobs. Here we discuss the identification of the skill process and sketch a general idea about how these data also identify the remaining model parameters. We relegate a detailed discussion on the identification of the remaining parameters and some further discussion on the identification of the skill process parameters to online appendix III. Below, we use wage dynamics from the estimated model to evaluate the model along dimensions not used in the estimation.

Two channels, skill accumulation (experience) and selection (tenure), can explain the declining transition rates by age or tenure. Selection effects are present if idiosyncratic shocks hit matches with heterogeneous quality even if workers are homogeneous. Good matches face a lower probability of separating so that the share of good matches increases with tenure and observed separation rates decline.[18] Hence, selection is an effect associated with tenure accumulation. Skill accumulation instead improves the worker's productivity by age even if match quality is homogeneous. As workers age, they accumulate experience, and become more productive relative to their outside option, and their match-surplus increases so that they separate less. Hence, skill accumulation is an effect associated with experience accumulation. Both channels potentially explain the declining pattern of separations by age. Adopting ideas in Topel (1991),

[18] A related argument can be made for observed job-to-job transitions. Workers in better matches survive, so the likelihood of finding an even better match declines as well.

we use differences between age profiles and newly hired age profiles to disentangle the relative importance of the two effects.

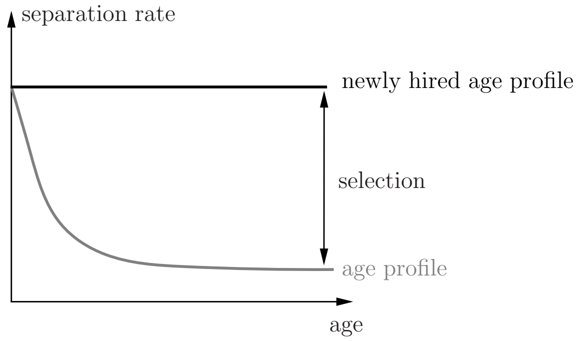
Figure 4 shows separation rates by age and separation rates for newly hired workers for hypothetical economies. Figure 4(a) depicts the case when the decline in the separation rate by age is explained by selection only and skill accumulation is absent. Although age and tenure increase jointly, it is only selection that leads to a declining age profile; the newly hired age profile is flat. In the absence of skill accumulation, a newly hired young worker is identical to a newly hired older worker. Hence, separation rates by age for newly hired workers are independent of age.

Figure 4(b) depicts the case when the decline in separation rates by age is explained by skill accumulation only. Workers accumulate skills with experience, so older workers are on average more skilled and separate less than younger workers. Absent selection effects, skill accumulation by age translates one-to-one into differences in the separation rate by age for newly hired workers. The age profile and the newly hired age profile decrease by the same amount. As discussed in our empirical analysis, the data represent an intermediate case as in figure 4(c), so slope differences in the newly hired age profile and the average age profile identify the relative strength of the two effects.

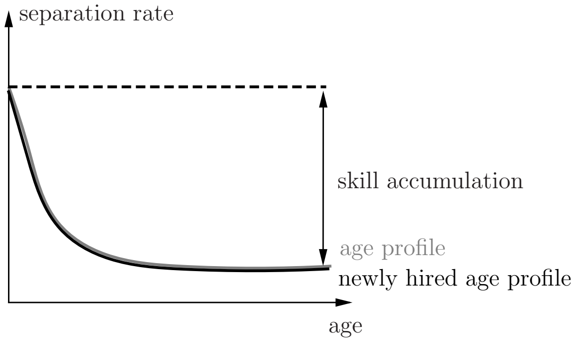
A similar idea applies to the identification of skill transferability across jobs. To disentangle how transferable skills are, we use the newly hired age profile of job-to-job transitions. Workers who accumulate skills face a trade-off between searching for a better match and losing accumulated skills when switching jobs. Consequently, older workers with more accumulated skills are on average more reluctant to accept outside offers than younger workers. As a consequence, older newly hired workers switch jobs less often than younger newly hired workers. If skills were perfectly transferable across jobs, the newly hired age profile would be flat. Hence, the decline in the newly hired age profile for job-to-job transitions identifies how transferable accumulated skills are across jobs (figure 4(d)).

Translating the discussion to model parameters, we explained how the slopes of the newly hired age profiles identify the skill-process parameters p_u and p_d. In online appendix III, we provide a detailed discussion of identification for the remaining model parameters. For this discussion, it is instrumental to recognize that differences between the age profile and the newly

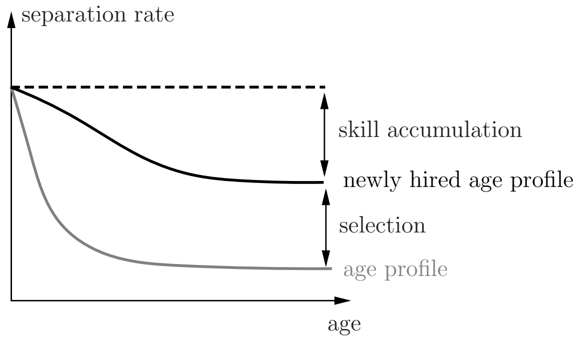
Figure 4: Identification of the skill process



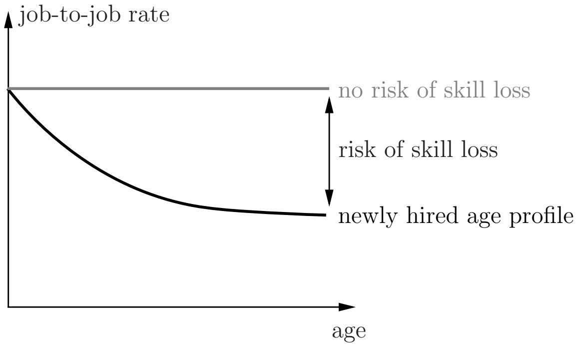
(a) Only selection



(b) Only skill accumulation



(c) Selection and skill accumulation



(d) Transferability of skills

Notes: Panel 4(a) shows stylized age and newly hired age profiles for separation rates in a model with only selection. Panel 4(b) shows stylized age and newly hired age profiles for separation rates in a model with only skill accumulation. Panel 4(c) shows stylized age and newly hired age profiles for separation rates in a model with selection and skill accumulation. Panel 4(d) shows a stylized newly hired age profile for job-to-job transition rates with full and partial transferability of skills. All figures have age on the horizontal axis and transition rates on the vertical axis.

hired age profile also quantify differences in transition rates between low-tenure (newly hired) and high-tenure (average) workers. We now exploit this fact when we summarize the discussion on parameter identification.

The general idea of which dimensions of heterogeneity we exploit for identification already appears in figure 4. The age profiles shown in the figure can be described by three characteristics: their average level, their slope capturing the difference between young and old workers, and their shape describing how quickly the difference between young and old workers materializes.

Concretely, we sketch in section III.1 of the online appendix a stylized model to show that the level of the separation rate, together with separation rate differences between low- and high-tenure workers, and the level of mean tenure identify the outside option b, the dispersion of match-specific skills σ_m, and the dispersion of idiosyncratic productivity costs ψ_s. The discussion surrounding figure 4 already suggests that separation rate differences between low- and high-tenure workers identify σ_m. The outside option b determines the average surplus and, thereby, the level of the separation rate. The dispersion of shocks ψ_s determines differences in separation rates so that it is identified by mean tenure. The speed of skill accumulation δ governs how quickly workers accumulate worker-specific skills and, therefore, how quickly age differences realize. The shape of the separation rate profile identifies this parameter. Exogenous separations limit tenure accumulation of workers by age, so that the slope of the mean tenure profile identifies π_f.

We exploit the level, slope, and shape of the job-to-job transition rate to identify parameters $\varkappa_o$, κ_o, and ψ_o. The matching efficiency $\varkappa_o$ determines the number of job offers for employed workers and is identified by the level of job-to-job transitions. The slope of the job-to-job transition rates depends on the relative importance of nonpecuniary job aspects κ_o. During their working life, workers climb the job ladder so that job-to-job transition rates decline. If nonpecuniary aspects become more important, job-to-job transition rates decline by less; the slope gets smaller. The dispersion of nonpecuniary shocks governed by ψ_o determines the job acceptance elasticity and, thereby, the shape of the job-to-job transition rate profile.

The bargaining power μ is identified by job-to-job transition rate differences between low- and high-tenure workers. A higher bargaining power provides stronger incentives for newly hired workers to climb the job ladder because they will receive a larger fraction of the gains from job switching. The higher the bargaining power, the more newly hired workers want to climb the job ladder. Finally, $\varkappa_n$ and κ are identified by the level and slope of the job finding rate profile. As for job-to-job transitions, $\varkappa_n$ determines the level of the job finding rate. Vacancy posting costs κ, in comparison to the changing surplus due to skill accumulation, determine the slope of the job finding rate.

Compared to existing approaches that mainly focus on heterogeneity in the wage dynamics, such as Bagger et al. (2014), our approach exploits the corresponding heterogeneity in worker mobility over the age-tenure domain for identification. We refer to online appendix III for further details and turn next to a discussion of our estimation procedure and the results.

4 Results

This section starts by discussing our estimation procedure. We then show how the model performs along the mobility dimensions used in the estimation and discuss wage implications as overidentifying restrictions. In Section 5, we then turn to the investigation of earnings losses.

4.1 Estimation Procedure

Before we bring the model to the data, we have to make some assumptions on parameters and functional forms. To align model and data, we set the model period to one month. A worker enters the labor market at age 20 as non-employed, leaves the labor market at age 65, stays retired for a further 15 years, and dies at age 80.[19] The production function is age-independent and log-linear in skills $f(x) = \exp(x_m + x_w)$, as in Bagger et al. (2014).[20] We approximate both skill distributions using five skill states. Mean skill levels are normalized to 1. The match-specific component (x_m) approximates a normal distribution with standard deviation σ_m, and the worker-specific component is constructed such that each increase in skill level leads to a 30% increase in the level of skills ($\sigma_w = 0.3$). In the model, workers and firms care about the expected value of the skill increase ($\sigma_w p_u$), so σ_w constitutes a normalization.[21] In line with the literature, we set a discount factor β to match an annual interest rate of 4% and a matching elasticity of $\varrho = 0.5$ following Petrongolo and Pissarides (2001).

We estimate parameters using a method of moments. We avoid simulation noise and iterate on the cross-sectional distribution from the model. We use age profiles, newly hired age profiles, and mean tenure in the estimation where we weight profiles to focus mostly on ages 20-50. We provide the details on the implementation in appendix B. Table 2 collects the estimated parameters together with the estimated standard errors. Standard errors are computed using the bootstrapped data profiles from Section 2.

All estimated parameters from Table 2 are at economically reasonable magnitudes. The

[19] During retirement, the worker receives entitlements proportionate to the worker-specific skill component in the period before retirement. This retirement scheme makes it less attractive to search on the job in the last few years given that a skill loss has long-lasting effects. In the absence of a retirement value, workers start to increase job-to-job transitions around age 55 only for nonpecuniary reasons. We consider retirement in this stylized form as a convenient abstraction to align model and data along a dimension that is not the focus of this paper.

[20] We provide a discussion of this assumption in Section III of the online appendix.

[21] We also tried other values for σ_w with the most notable change that probabilities of the skill increase adjusted. The only restriction is that σ_w has to be sufficiently large to allow for enough skill increase during the working life.

Table 2: Estimated parameters

Skills		Shocks		Matching and bargaining	
p_u	0.0258	ψ_s	2.8621	μ	0.3097
	(0.0007)		(0.0878)		(0.0299)
p_d	0.0536	κ_o	-0.6933	b	0.3949
	(0.0064)		(0.0942)		(0.0170)
δ	0.0030	ψ_o	1.8503	κ	2.3689
	(0.0001)		(0.1381)		(0.0900)
σ_m	0.0933	π_f	0.0024	$\varkappa_o$	2.3913
	(0.0076)		(0.0001)		(0.1149)
				$\varkappa_n$	0.4591
					(0.0075)

Notes: Estimated parameters and standard errors. Standard errors shown in parentheses. Column Skills shows parameters determining the skill process. The parameter p_u is the probability of worker-specific skill accumulation at age 20, p_d is the probability that worker-specific skills do not transfer at job change, δ governs the declining probability of worker-specific skill accumulation by age, and σ_m denotes the standard deviation of match-specific skills. Column Shocks shows idiosyncratic shock parameters governing worker mobility decisions. The parameter ψ_s determines the dispersion of productivity shocks, κ_o determines the common utility component of all job offers, ψ_o determines the dispersion of the idiosyncratic utility component of job offers, and π_f is the exogenous separation probability. Column Matching and bargaining shows parameters related to the search process. The parameter μ is the bargaining power of the worker, b is the flow utility during non-employment, κ determines vacancy posting costs, and $\varkappa_o$ and $\varkappa_n$ are matching efficiencies for on- and off-the-job search. Standard errors are bootstrapped using 500 repetitions.

parameter p_u refers to age 20 and follows the life-cycle dynamics governed by δ described above. The estimate implies an expected skill increase at age 20 of 0.8 log point per month ($\sigma_w p_u$) or extrapolated to an annual frequency of 9 log points in the first year in the labor market. This skill increase and the decline in its speed governed by δ match the increase and concavity of the empirical log wage profile as shown in figure 6(a). The estimate of p_d implies an expected skill loss from a job change of 1.6 log points ($\sigma_w p_d$). This degree of transferability of skills is consistent with the share of negative wage changes and the average wage gain at job-to-job transitions over the life cycle, as we will demonstrate in Section 4.3.1. Our estimate of σ_m implies a wage difference of roughly 17% (32%) between the average (minimum) match and the best match for the median worker at age 40. This amount of wage dispersion can be compared with empirical estimates of the mean-min ratio of wages, as popularized by Hornstein et al. (2011). As we will discuss in detail below, our model is consistent with empirical estimates of the mean-min ratio in the cross section and over the life cycle. A directly comparable estimate of match-specific wage dispersion is provided in Hagedorn et al. (2017). Their estimate has to be compared to the employment-weighted variance of x_m from our model. Our model delivers a variance of 0.014, close to their reported variance of 0.016.[22]

The size of the parameter estimate for ψ_s is easiest to interpret in relation to transitory wage risk. Imposing some mild additional assumptions on the transmission of these shocks to wages, we quantify the implied transitory wage risk from these shocks in online appendix V.2.4. We find an implied standard deviation of transitory wage shocks of 0.35 that is within the ballpark of the average estimate of 0.29 from Heathcote et al. (2010).

The option value Ψ_o from the acceptance choice of outside offers reflects the nonpecuniary benefits from a new job. The estimates for κ_o and ψ_o imply a modest importance of this nonpecuniary utility component. At age 40, the average utility flow from the nonpecuniary job component to an employed worker corresponds to less than 6% of the average wage. Our estimate for b corresponds to 28% of the average wage of a 40-year-old worker. Our estimate is thereby below the unemployment benefit replacement rate of 40% used in Shimer (2005) but above the effective value estimated in Chodorow-Reich and Karabarbounis (2016). Non-employed workers also receive utility from the acceptance choice of job offers. Their option value is substantially larger than that of employed workers due to higher contact rates. Including the option value from job search, the flow utility in unemployment relative to the average wage is roughly 70% and is above 95% when we compare it to the wages of newly hired workers, thereby moving close to the estimate of Hagedorn and Manovskii (2008). Vacancy posting costs κ correspond to 56% of the quarterly wage of a 40-year-old worker. They therefore capture a broader concept of hiring costs including training costs, as discussed in Silva and Toledo (2009). We estimate a bargaining power of 0.31. The estimate is similar to that in Bagger et al. (2014). It is not directly comparable because their model relies on a different bargaining protocol and data. The estimates for the matching efficiency parameters $\varkappa_o$ and $\varkappa_u$ imply a higher matching efficiency on the job. Despite the higher matching efficiency on the job, employed workers receive fewer job offers than non-employed workers in the model because employers take their lower acceptance rate into account (see eq. (10)). A lower acceptance rate on the job is consistent with the results in Faberman et al. (2016), who report that full-time employed workers have an acceptance rate that is less than half of the acceptance rate of non-employed workers. In the model, a 40-year-old employed worker receives 0.3 job offer per month, whereas a non-employed

[22] Their estimate is based on the National Longitudinal Survey of Youth 1979. Workers in their sample are between ages 18 and 22 in 1979. We consider as a model counterpart the age range from 20 to 43 and report the variance of x_m across employed workers. We report an average of age-specific estimates.

worker receives 0.4 offer. This difference is consistent with estimates in Faberman et al. (2016, Table 3), who report on average 0.2 job offer for employed workers over a four-week period compared to 0.4 job offer for non-employed workers in their data. Furthermore, the mobility pattern on and off the job are in line with the empirical counterparts, as we show below. Finally, our estimate for the exogenous separation rate π_f implies an 8% probability of displacement within three years. This estimate is in line with evidence provided in Farber (2007), who reports a three-year involuntary separation rate of around 10% based on CPS data.

Given that the estimated parameters are at economically reasonable levels, we next show how the estimated model fits the mobility facts used in the estimation. We then present the results on wage dynamics to evaluate the model performance along dimensions that were not part of the estimation.

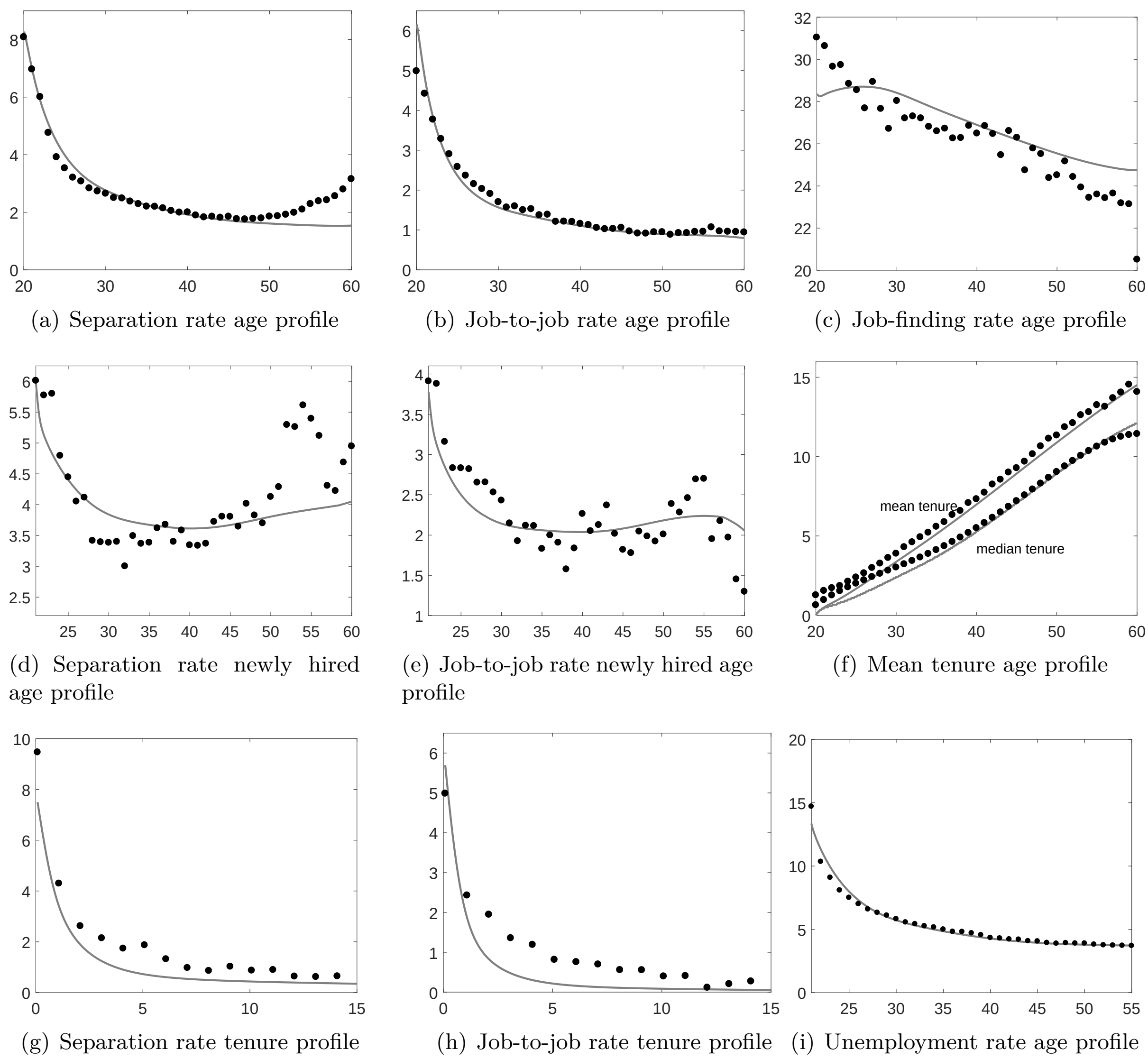
4.2 Labor market mobility

Figure 5 presents, in the upper two rows, the model fit for worker transition rates and mean and median tenure that have been part of the estimation. Figures 5(a), 5(b), and 5(c) show age profiles for separation, job-to-job transition, and job-finding rates. Figures 5(d) and 5(e) show the profiles for separation and job-to-job transition rates by age for newly hired workers. Figure 5(f) shows the age profile of mean and median tenure. All transition rates and mean and median tenure are matched closely.

The bottom row of figure 5 shows transition rates by tenure and unemployment rates by age, both of which have not been directly targeted in the estimation. Figures 5(g) and 5(h) demonstrate the good fit of the model to the transition rates by tenure.[23] The fit of mobility by tenure shows that our model matches the frequency of steps on the job ladder. Importantly, our model matches job stability at the top of the job ladder with very low separation rates for workers with more than 10 years of tenure. In models with high separation rates also at the top of the job ladder, workers fall down the job ladder repeatedly, and differences that result from the job ladder are transitory. Average tenure is low. Matching low separation rates at the top leads to high tenure and to differences in match types that persist over time. Matching the frequency of steps on the job ladder is important for our later analysis because the job ladder

[23] Although not directly targeted, there is a stock-flow relationship in the background that restricts the tenure profile once tenure levels by age are matched. The model profiles have been derived under the assumption of a uniform age distribution. To avoid making any assumptions or requiring an age distribution in the model, we use only age-specific targets in the estimation.

Figure 5: Model prediction and data



(a) Separation rate age profile

(b) Job-to-job rate age profile

(c) Job-finding rate age profile

(d) Separation rate newly hired age profile

(e) Job-to-job rate newly hired age profile

(f) Mean tenure age profile

(g) Separation rate tenure profile

(h) Job-to-job rate tenure profile

(i) Unemployment rate age profile

Notes: Panels 5(a) to 5(c) show age profiles for separation rate, job-to-job transition rate, and job-finding rate from model and data. Panels 5(d) and 5(e) show newly hired age profiles for separation rate and job-to-job transition rate from model and data. Panel 5(f) shows mean and median tenure by age from model and data. The black dots show data, and the gray solid line shows the model. The horizontal axis is age in years, and the vertical axis shows transition rates in percentage points or tenure in years. Newly hired age profiles start at age 21. Panels 5(g) and 5(h) show tenure profiles for separation and job-to-job transition rate from model and data. The black dots show data, and the gray solid line the model. The horizontal axis is tenure in years, and the vertical axis shows transition rates in percentage points. Panel 5(i) shows the age profile of the unemployment rate from model and data. The black dots show the data, and the gray solid line the model. The horizontal axis is age in years, and the vertical axis shows unemployment rates in percentage points. Mean level differences between model and data have been removed. See text for details.

governs the recovery after displacement. We will demonstrate below that our model matches wage gains following job-to-job transitions.

Figure 5(i) shows the unemployment rate by age from the model and CPS data. Non-employment in the model comprises all unemployed workers and some workers who are not classified as unemployed in the CPS but who are attached to the labor market. Recent evidence in Kudlyak and Lange (2014) supports this modeling choice. We discuss this assumption in detail in online appendix II.1.2, and we explain in online appendix V.1 how we construct an adjustment factor to remove the level difference between model and data. Given that all workers start non-employed at age 20 in the model, figure 5(i) shows the age profile of the unemployment rate starting at age 21. The model matches the empirical unemployment rate by age almost exactly.

Finally, note that we focus on the average job-finding rate by age in figure 5(c) because most unemployment spells in the data are short. BLS data show that the share of job losers who are unemployed half a year or more is 18% over our sample period. In our model, the same share at age 40 is 17% with an age variation from 14% at age 25 to 19% at age 55. Hence, our model captures the transitory nature of unemployment spells in the U.S. labor market well. Looking at longer unemployment durations, the model does not generate the empirically observed duration dependence with a decline of only 22% over 24 months. In the data, the decline is slightly more than twice as large. However, very few workers actually face these low job-finding rates because the vast majority of workers finds jobs more quickly. In a model extension described in online appendix IV.1, we match the empirically observed duration dependence. We allow for duration-dependent skill losses during non-employment and deteriorating search efficiency with non-employment duration, capturing two prominent explanations for duration dependence (see Kroft et al. (2013)). The extended model is reestimated and matches the empirically observed duration dependence. We show that accounting for duration dependence of job-finding rates affects our results only marginally, so that we abstract from it for our baseline model.

In sum, the model is consistent with two characteristic features of the U.S. labor market: large average transition rates and a large share of very stable jobs. The coexistence of these facts has so far received little attention in the literature on structural labor market models. Yet, these features are crucial in generating large and persistent earnings losses, as we show below.

Next, we demonstrate that the model is also consistent with a range of other facts on wage dynamics.

4.3 Wage dynamics

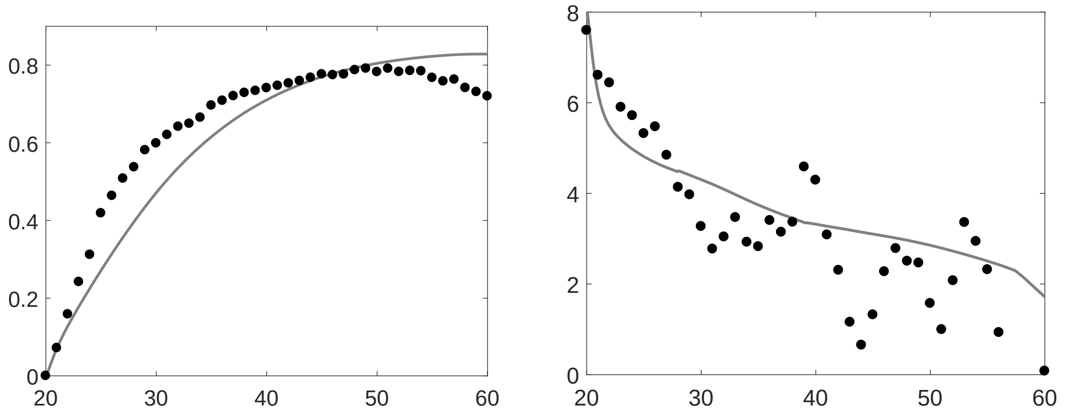
The previous section has shown that the model is consistent with observed worker mobility and job stability pattern. This section demonstrates that the model is also consistent with a range of facts on wage dynamics both on the job and between jobs. For wage dynamics between jobs, we consider average wage gains from job-to-job transitions, the share of negative wage changes following job-to-job transitions, and the share of early career wage growth attributable to job switching. We derive the first two statistics from the SIPP micro data and use the estimate from Topel and Ward (1992) for the decomposition of early career wage growth. For wage dynamics on the job, we consider estimates of the returns to tenure using two alternative identification approaches (Topel (1991) and Altonji and Shakotko (1987)) and the variance of permanent shocks using a permanent-transitory shock decomposition (Storesletten et al. (2004), Guvenen (2009), Heathcote et al. (2010)). Tightly connected to wage dynamics is cross-sectional wage inequality. Therefore, we also discuss the model's ability to match different measures of cross-sectional wage dispersion. Finally, we revisit the correlation between wages and job stability. While the model matches this relationship qualitatively by construction, here we explore the relationship quantitatively. We relegate the details of the estimation procedure using model-simulated data to online appendix V.2.

First, we compare in figure 6(a) the mean (log) wage by age from the model and data. Wage data come from the annual march CPS files. We provide further details on the construction in appendix A.1. Wages from the model are initially not as steep as in the data, but wage growth until age 40 is matched. Generally, the model matches the slope closely but misses some of the concavity of the empirical profile.

4.3.1 Wage gains from job-to-job transitions

Figure 6(b) compares the mean wage gain from a job-to-job transition by age from the model to the data. We derive the empirical profile based on micro data, as in Tjaden and Wellschmied (2014). Online appendix V.2.1 provides details for the construction in the model. The declining age profile of wage gains suggests that the gains from search decline. The model prediction is slightly higher than the empirical estimates but matches a similar decline by age.

Figure 6: Wage profiles



(a) Mean log wage age profile (b) Wage gain at job-to-job transition

Notes: Age profiles of mean log wages and average wage gains following a job-to-job transition from model and data. The gray solid line shows the model, and the black dots show the data. The horizontal axis is age in years, and the vertical axis shows the log-wage change relative to age 20 (left panel) or wage gains relative to the previous job (right panel) in percentage points. Mean log wage profiles come from CPS data, and wage gains are derived using SIPP data, as in Tjaden and Wellschmied (2014).

While figure 6(b) shows that the model generates sizable positive average wage gains following job-to-job transitions, it hides that the model also matches a large fraction (24%) of job-to-job transitions that lead to wage cuts. The fact that a substantial share of job-to-job transitions is associated with wage cuts in the data (32 %) is well known and is, for example, discussed in Tjaden and Wellschmied (2014). Many search models struggle to explain this fact because workers only change jobs if the outside offer is better than the current job. In our model, workers' acceptance decisions depend not only on wages but also on a nonpecuniary utility component. Wage cuts after job-to-job transitions follow naturally in this case.[24]

4.3.2 Early career wage growth

Topel and Ward (1992) document that about one-third of total wage growth in the first 10 years of working life is explained by job-changing activity. In their sample, a typical worker switches jobs frequently and holds on average seven jobs during the first 10 years in the labor market. Similarly, Bagger et al. (2014) find in a structural labor market model that during an initial job-shopping phase, wage growth is strongly driven by job-changing activity. Early career

[24] Alternative explanations for wage cuts at job-to-job transitions are occupation-specific skills, as in Kambourov and Manovskii (2009a) and Kambourov and Manovskii (2009b), or a different bargaining protocol with wage increases over time, as in Postel-Vinay and Robin (2002).

wage growth is an alternative, independent measure for the relative importance of worker- and match-specific skill accumulation. Our model generates on average eight jobs in the first 10 years of working life and a contribution of job-changing activity to wage growth of 30%. Online appendix V.2.2 provides details on the wage growth decomposition in the model.

4.3.3 Returns to tenure

The returns to tenure capture the increase in wages with job duration. So far, no consensus has been reached in the literature on the importance of the returns to tenure relative to the return to general experience. Estimates differ dramatically across studies depending on identification strategies (see e.g., Topel (1991), Altonji and Shakotko (1987), and the survey by Altonji and Williams (2005)).

We implement the estimators by Topel (1991) and Altonji and Shakotko (1987) on simulated data from our model. Online appendix V.2.3 provides details. The model reproduces both estimates very closely. The ordinary least squares (OLS) estimate for the returns to tenure is a common benchmark. Using OLS, Altonji and Shakotko report 26.2% returns from 10 years of tenure for their sample. In the model, we get 24.2%, which is lower than the empirical estimates but still consistent with substantial returns to tenure. Following the instrumental variable approach proposed in Altonji and Shakotko, the model generates 0.0% for returns from 10 years of tenure; this substantial drop is in line with Altonji and Shakotko's estimate of 2.7% (about one-tenth of their OLS estimate).[25] Topel proposes a two-step estimation approach and finds 24.6% for returns from 10 years of tenure, again close to the level of the OLS estimate. Using his approach, the model predicts 29.6% and again matches the empirical pattern of large returns from tenure at the order of the OLS estimate.[26]

4.3.4 Permanent income shocks and wage inequality

We discuss above that in the data and the model, most workers stay on their jobs for several years. We therefore consider the variance of permanent income shocks as an additional measure to describe wage dynamics on the job. As before, we use the empirical estimation approach to capture the statistical properties of the model-generated wage dynamics but do not necessarily take the underlying statistical model as a good description of the model-generated wage process.

[25] Our estimate is within their confidence interval given the standard error of 1.6%.

[26] Exploring the reasons behind the model's ability to match the diverging estimates is beyond the scope of the paper.

We compare our results to findings from Heathcote et al. (2010). Heathcote et al. estimate a standard deviation of 0.084 for the permanent shock. Our model closely matches this number with an estimate of 0.072. We provide the details on the estimation using model data in online appendix V.2.4. There we also discuss how to construct estimates for transitory shocks from the model. When we consider, as in Heathcote et al. (2010), the age range from 25 to 60, we estimate a standard deviation for transitory shocks of 0.35, which is close to the average estimate of 0.29 in Heathcote et al. (2010).

Cross-sectional wage inequality is the result of the described wage dynamics. Hornstein et al. (2011) point out that existing search models struggle to generate substantial wage dispersion. Their preferred measure for wage dispersion is the mean-min ratio of wages (Mm ratio). For a canonical search model calibrated to the U.S. labor market, they find a Mm ratio of 1.046. Tjaden and Wellschmied (2014) use SIPP data to provide empirical estimates of Mm ratios. They report Mm ratios by age that vary between 1.95 and 2.25 over the age range from 25 to 49. At age 36, they report a Mm ratio of 2.12. Our model closely matches this level of wage dispersion and its age variation. The average Mm ratio is 2.53, and it varies from 1.69 at age 25 to 2.93 at age 49 and is 2.50 at age 36. Online appendix V.2.4 provides further details.

Closely related to Hornstein et al. (2011) is the empirical work by Hagedorn et al. (2017). They estimate the contribution of match-specific wage differences to cross-sectional wage inequality. They find that the match-specific variance accounts for 5.7% of the cross-sectional (log) wage variance. We observe match dispersion directly and find that our model aligns well with this estimate. Match dispersion in the model corresponds to 6.4% of the cross-sectional (log) wage variance.[27]

The variance in log wages is another popular measure of wage dispersion. In the data, the variance in log wages increases over the life cycle. Our model matches this increase between ages 20 and 40. The increase is 8 log points in the model in comparison to 10 log points in the CPS data for the same age range. A key challenge in matching the variance of log wages is its sensitivity to the tails of the wage distribution. The parsimony of the worker skill process in our baseline model cannot capture the very right tail of the wage distribution, which limits the increase in the variance of log wages after age 40. In particular, the bounded support for the

[27] We use employment-weighted observations including transitory wage shocks as the correct model counterpart to the empirical approach. We consider workers age 20 to 43 (see footnote 22).

worker-specific skill states leads to a flattening out of the variance age profile. In online appendix IV.2, we provide an extended model where we augment the worker-specific skill process by an additional skill state in the right tail of the skill distribution. We demonstrate that this extension allows us to fit the life-cycle profile of the variance in log wages over the entire working life very closely without sacrificing the fit along other dimensions. We also demonstrate that other results are robust to this model refinement. The caveat is that we have to use the age profile of the variance in log wages to estimate the extended model, so we focus on the parsimonious version in the main text. We relegate further discussion to online appendix V.2.4.

4.3.5 Job stability and wages

Section 2.3 discusses the empirical correlation between wages and job stability. As discussed above, such a link between job stability and wages is a direct implication of the joint bargaining over wages and separation decisions in the model. To show that our model quantitatively accounts for the observed correlation, we redo our empirical analysis on model-generated data using 4-month separation rates. Online appendix V.2.5 provides further details. Our regression coefficient of separation rates on log wages is -0.0368 in the model compared to -0.0392 in the data when looking at all separations, and it is -0.0667 in the model compared to -0.0548 in the data when looking at newly hired workers (see Table 1). We conclude that the wage-stability trade-off from our model is quantitatively consistent with the data.

5 Earnings losses

This section examines implications of the model for estimated earnings losses following displacement. We first provide a model analog of the empirical estimation methodology developed in Jacobson et al. (1993) and show that the model reproduces empirical earnings losses in both size and persistence. We use the structural model to decompose earnings losses into a wage loss effect, an extensive margin effect, and a selection effect. We explore the relative importance of match- and worker-specific skill losses for wage losses and subsequent job stability.

5.1 Group Construction

Jacobson et al. (1993, p.691) define displaced workers' earnings losses as "(...) the difference between their actual and expected earnings had the events that led to their job losses not occurred," and propose an estimation strategy borrowed from the program evaluation literature. The approach is based on the construction of two groups, which we refer to as layoff group and control group. For details on construction of estimates, we follow Couch and Placzek (2010), one of the recent applications of the original estimation strategy. Other recent contributions are von Wachter et al. (2009) and Davis and von Wachter (2011), who apply the same estimation methodology but differ in the construction of the control and the layoff group. We will also compare our model prediction with their results.

The layoff group consists of all workers who separate in a mass-layoff event. The idea of using mass layoffs is that workers are not selected based on their individual characteristics when mass layoffs occur. We associate this event with an exogenous separation in the model. Exogenous separations in the model occur independent of the individual characteristics and are therefore the model analog to a mass layoff event in the data. This mapping is also in line with the discussion in Stevens (1997) and her mapping of separation events in the PSID to displacement.[28] The control group consists of continuously employed workers over the sample period. The empirical analysis covers workers of all ages and controls for age in the regression. In the model, we consider a worker of age 40; this corresponds to the mean age of all workers from the sample used by Couch and Placzek (2010). Online appendix VI.1 reports estimation results for various age groups.[29] The layoff group then consists of all workers who separate as the consequence of an exogenous separation. We provide a discussion of selection effects if separations are endogenous in online appendix VI.2. As in Jacobson et al. (1993) and Couch and Placzek (2010), we initially restrict the sample to workers with at least six years of tenure. For the control group, both studies require a stable job for the next six years because they require continuous employment over their 12-year sample period. We follow the empirical analysis and construct the appropriate model equivalents. In line with all empirical studies, we consider

[28] Couch and Placzek define a separation to be part of a mass layoff if employment in the firm from which the worker separates falls at least by 30% below the maximum level in the year before or after the separation event. Their data covers the period from 1993 to 2004 and the maximum is taken over the period prior to 1999. They restrict attention to firms with 50 or more employees. The empirical literature on earnings losses distinguishes between three separation events termed separation, displacement, and mass layoff and particular selection criteria apply to each event. The general idea behind the selection criteria is that displacement and mass layoff events constitute involuntary separations, while separation events also include voluntary separations like quits to unemployment. See also Stevens (1997) for a discussion. Given that firm size remains undetermined in the model, we cannot impose the size restriction on firms.

[29] In the sample of Couch and Placzek (2010), mean age in the entire sample is 39.7, it is 40.2 in the control group, and 38.9 in the mass layoff group. As we show, earnings losses are almost linear in age, so that the effect at the mean and the mean effect are identical.

non-employment income to be zero. This creates a difference between wage and earnings losses that is quantitatively non-negligible.[30] We also control for worker-specific fixed effects. We reproduce empirical estimates from the model using measures over worker states and transition laws instead of relying on simulation.

5.2 Earnings losses

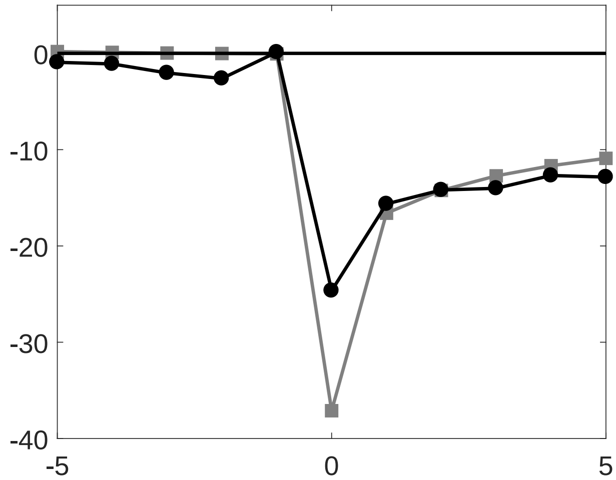
Figure 7 shows earnings losses from the model in comparison to the estimates from Couch and Placzek (2010). The model generates large and persistent earnings losses (gray line with squares). In the first year following the layoff event, earnings losses amount to 37%, and six years after the layoff event, they still amount to 11% of predisplacement earnings. Findings correspond closely with empirical estimates by Couch and Placzek (2010) (black line with circles), which show 25% earnings losses initially and 13% after six years.[31] Standard deviations for estimates from Couch and Placzek are 0.9% to 1.8% of predisplacement earnings so that model predictions are well within the estimated range.

The initial drop in earnings is larger in the model than the empirical estimates. This difference likely results from the fact that the point in time of the layoff event and point in time when the employee is notified in the data can only be determined to be in a certain quarter. The initial earnings losses in the data therefore comprise likely pre- and post-displacement earnings observations, which leads to lower estimated earnings losses than in a case where the exact point in time of the separation can be observed. Pries (2004) makes a similar argument. In online appendix VI.3, we show that small differences in timing of the displacement notification can have a large impact on the initial drop in earnings. We find that one month of advance notification closes the initial difference in estimated earnings losses between model and data by 50%, and two months of advance notification close the gap between the earnings losses from the model and the data completely. In both cases, however, earnings losses after six years remain virtually unaffected.

[30] To get a measure of earnings in the model, we sum the average monthly wages for the layoff and the control group over 12 months for each year. We abstract from the intensive margin for hours worked and refer to wages as salary earned by workers conditional on employment, while earnings is used to refer to total income of a given period including zero income during unemployment.

[31] The earnings losses in Jacobson et al. (1993) are larger, but as Couch and Placzek (2010) argue, they are influenced by the particularly bad economic conditions in Pennsylvania at the time of their study. Davis and von Wachter (2011) also report strong effects on earnings losses from bad economic conditions, but their average estimates for times of good and bad economic conditions are comparable to the estimates by Couch and Placzek (2010) (See also Table 3 below).

Figure 7: Earnings losses following displacement



Notes: Earnings losses after displacement in the model and empirical estimates. The gray line with squares shows earnings losses predicted by the model. The black line with circles shows estimates by Couch and Placzek (2010). The horizontal axis shows years relative to displacement and the vertical axis shows losses in percentage points relative to the control group.

Davis and von Wachter (2011) use the same estimation approach but propose a different construction of the control and layoff group. They require three years of job tenure for both the control and the layoff group prior to their displacement and two years of subsequent job stability following the year of the displacement event for the control group.[32] They consider men aged 50 years and younger. We adjust the average age for displaced workers in the model accordingly to 35 years when comparing the model prediction to their results. Davis and von Wachter (2011) report earnings losses as a present discounted value relative to predisplacement annual earnings, and, alternatively, as a share of the present discounted value of counterfactual earnings. They use an annual discount factor of 5% and extrapolate earnings losses beyond 10 years after the displacement event. We follow them in the implementation. Table 3 reports results from our model in comparison to estimates reported in Davis and von Wachter (2011) for different control and layoff groups and for different age groups.

Our model matches their earnings losses closely except for the oldest group of workers. If we allow for diverging labor force participation trends for workers age 41-50, for example due to early retirement decisions, and match a difference at age 65 of 30%, then the model generates

[32]The classification of mass layoff differs slightly, but given that firm size remains undetermined in our class of models this does not affect the model results. Davis and von Wachter (2011) report that the definition of the mass layoff event does hardly affect the estimated earnings losses.

Table 3: Comparison to earnings loss estimates from Davis and von Wachter (2011)

sample	Davis and von Wachter predisplacement	Davis and von Wachter counterfactual	Model predisplacement	Model counterfactual
all workers	1.7	11.9 %	1.5	10.0 %
age 21-30	1.6	9.8 %	1.7	9.8 %
age 31-40	1.2	7.7 %	1.5	10.0 %
age 41-50	1.9	15.9 %	1.2	8.8 %

Notes: The first column shows the considered sample. All workers in the case of Davis and von Wachter (2011) means men from age 21 to 50. We use midpoints of age intervals to get earnings losses for age groups in the model. See text for further details of sample selection criteria. Column predisplacement reports the discounted sum of earnings losses as a multiple of predisplacement annual earnings. Column counterfactual reports the discounted sum of earnings losses as share of the sum of discounted counterfactual earnings. See text for further details.

earnings losses of 1.8 times predisplacement earnings and 13.8% of the counterfactual present value of earnings; this, again, closely matches the results by Davis and von Wachter (2011).[33] Our model abstracts from early retirement decisions, because they do not have an impact on the mechanism generating large and persistent earnings losses. However, these decisions can potentially become important when looking 20 years ahead after a displacement event for older workers as done in Davis and von Wachter (2011).

5.3 Decomposition

In this section, we decompose earnings losses into three effects: lower wages (wage loss effect), lower employment rates due to higher separation rates in subsequent matches (extensive margin effect), and selection due to restrictions on employment histories of the control group (selection effect). In a second step, we decompose wage loss effect and extensive margin effect in effects due to losses in worker- and match-specific skills. The importance of worker- and match-specific skill losses is the key result for the subsequent policy analysis, because it informs policymakers about the potential effectiveness of re-training and placement support programs.

5.3.1 Selection effect

The control group definition in Jacobson et al. (1993, pp.691) ”compares displacement at date s to an alternative that rules out displacement at date s and at any time in the future”. This construction of the control group leads to a spurious correlation between non-displacement

[33] Chan and Stevens (2001) and Tatsiramos (2010) provide a discussion of the empirical evidence of the effect of displacement on early retirement decisions.

and future employment paths by requiring subsequent stable employment. Viewed through the lens of a structural model, this assumption leads to ex-post selection of employment histories in terms of favorable idiosyncratic shocks and unattractive outside job offers.[34] Ex-post selection applies to workers who are identically ex ante. In addition to ex-post selection, the construction of the control group also leads to selection of workers who differ ex ante. Ex-ante selection occurs because workers who are less likely to separate in the future, either because of higher worker- or match-specific skills, are more likely to be included in the control group today. Ex-ante selection is present if workers and/or matches differ between control and layoff group at displacement.

To obtain an estimate of the importance of this effect, we construct an alternative ideal control group labeled the twin group. For this twin group, we do not impose restrictions on future employment paths, so no ex-post selection arises. Furthermore, we observe the skill distribution and can compare identical workers at age 40 with at least six years of tenure in the control and layoff group. Both groups have the same distribution over skills ex ante and differ only by the fact that one group received the exogenous separation shock while the other group did not. Hence, using our model, we can do the counterfactual experiment that must remain unobserved in the data of what would have happened had the worker not been displaced. We track the average earnings paths of these two groups.

If we compare the earnings losses to the benchmark case where the control group is employed continuously, we find that initial earnings losses are nearly identical and driven largely by the length of the initial non-employment period. However, earnings losses after six years are substantially different. The difference is solely due to the selection of the control group as the layoff group is identical in the twin experiment and in the benchmark. The resulting selection effect is sizable, accounting for 31% of the total earnings losses after six years. In online appendix VII, we provide a graphic illustration of the decomposition.

Couch and Placzek (2010) report results using an estimation approach that involves matching workers based on propensity scores. The idea is to compare workers who have identical

[34] Jacobson et al. (1993) discuss a potential bias in their estimation approach if error terms are correlated over time. They argue that the effect will disappear as long as the error term is mean stationary but that their estimates will be biased if the error term conditional on displacement is not zero. In their discussion, they focus on the group of workers that is displaced. However, focusing on workers that do not get displaced it becomes apparent that these workers stay continuously employed because of a particularly good history of shock realizations. In this case, the conditional error term is generally not zero and the bias can become substantial.

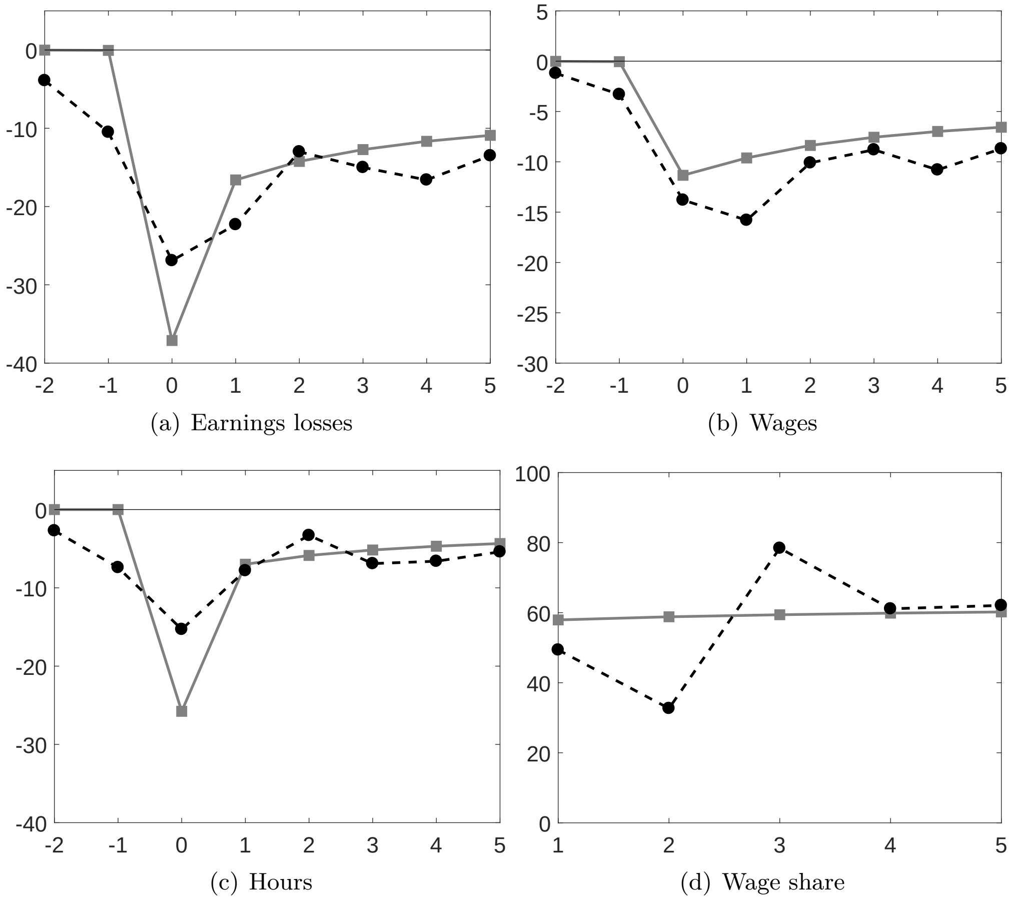
probabilities for being laid-off to control for individual heterogeneity. Still, they require continuous employment for the control group, so ex-post selection arises. They find that accounting for ex-ante selection in this way can account for 20% of the estimated earnings losses at the maximum. Davis and von Wachter (2011) reduce the non-displacement period for the control group after the displacement event. If we decompose earnings losses using their control group, we find that after six years, the selection effect is roughly cut by half and accounts for 14% of estimated earnings losses. Regarding ex-post selection, Davis and von Wachter (2011) discuss results for a case when non-mass layoff separators are included in the control group, in which case workers with less favorable employment histories are also part of the control group. In this case, they find that estimated earnings losses are up to 25% lower. This result and the result from the matching estimator by Couch and Placzek (2010) already indicate that both ex-ante and ex-post selection might be substantial in the empirical studies.

5.3.2 Extensive margin and wage loss effect

The literature does not always make a clear distinction between wage and earnings losses when interpreting empirical estimates. A notable exception is Stevens (1997). She empirically decomposes earnings losses into wage losses and an effect due to lower job stability. When we decompose earnings losses based on our model, we control for the selection effect using the twin group as our control group. The wage loss effect of our decomposition captures wage differences of employed workers between the control group and the layoff group. The extensive margin effect accounts for the remainder of earnings losses resulting from differences in employment rates between control and layoff group. Based on this decomposition, we find that the wage loss effect accounts for 48% of total earnings losses after six years and the remaining 21% are due to the extensive margin effect. Looking at the evolution of the decomposition over time, we find the extensive margin effect to be largest on impact, but even after six years, the layoff group is more often non-employed than the control group. We show the decomposition over time in online appendix VII.

To validate this decomposition, we compare the model-based decomposition of earnings losses to data from the Panel Study of Income Dynamics (PSID) closely following the analysis in Stevens (1997). For this comparison, we neither in the model nor in the data control for the selection effect to make results directly comparable. Stevens (1997) uses PSID data spanning

Figure 8: Empirical decomposition of earnings losses



(a) Earnings losses

(b) Wages

(c) Hours

(d) Wage share

Notes: Earnings losses and decomposition of earnings losses from model and PSID data. Top left panel: Earnings losses from model and estimates based on PSID data. Top right panel: Wage losses from model and estimates based on PSID data. Bottom left panel: Extensive margin from the model and hours losses estimated in PSID data. Bottom right panel: Share of wage losses in earnings losses from the model and based on empirical estimates. Horizontal axes show time relative to the displacement event in years. Vertical axes in the first three panels show losses in percentage points relative to the control group for earnings, wages, and extensive margin. Vertical axis in the bottom-right panel shows wage losses as share of earnings losses in percentage points. The gray solid lines with squares shows model results. The black dashed line with circles show empirical estimates. See text for further details.

the years between 1968 to 1988 to estimate earnings losses from job displacement. Unlike the administrative data, as used in Jacobson et al. (1993), Couch and Placzek (2010), or Davis and von Wachter (2011), PSID data provides information on earnings and hours worked that allow estimating extensive margin and wage loss effect directly. We follow Stevens (1997) in terms of sample selection and definition of worker displacement. We adopt her empirical specification and focus on first displacements consistent with the implementation in the model and the empirical approach in Couch and Placzek (2010). We provide further details about PSID data and the implementation in appendix A.3. Figure 8(a) shows the estimated earnings losses based on the specification in Stevens (1997) in comparison to the model. One caveat of the PSID data is its small sample size compared to administrative sources so that point estimates are less precise. Differences between empirical estimates and model counterparts are therefore typically not statistically significant. For example, the estimated earnings losses from Figure 8(a) are slightly larger than their model counterpart, but these differences are not statistically significant. Estimated earnings losses show the same dynamic evolution with large and persistent losses after six years.

In a second step, we make use of that the PSID provides information about annual hours worked. Annual working hours are affected by periods of non-employment because non-employment periods imply lost working hours. We use the information on working hours to decompose earnings losses into contributions from lower wages and lower employment. We proceed with the same estimation approach as for earnings losses but replace earnings on the left-hand side of the regression by wages and hours worked. For wages, we use annual earnings divided by annual hours worked. In figures 8(b) and 8(c), we compare the reduction in hours and wages from the data to wages and the extensive margin from the model. The model matches the reduction in wages and working time closely. The reduction in working time is matched almost exactly while the wage loss is slightly larger in the data.

Earnings losses between model and data in figure 8(a) differ slightly in size. To control for this level difference in the decomposition, we consider the share of earnings losses accounted for by wage losses from year 2 to 6 after displacement in figure 8(d). The model predicts a relatively constant share of 60%. This number differs from the decomposition above because earnings losses still comprise the selection effect. The decomposition from the data varies over time but stays always around 60%. We conclude that the model aligns well with the empirical evidence regarding the decomposition of earnings losses.

5.3.3 Decomposition in worker- and match-specific effects

The literature has proposed both match- and worker-specific skill losses as explanation for the observed earnings losses.[35] The distinction is important to inform policymakers if retraining

[35] Ljungqvist and Sargent (2008) and Rogerson and Schindler (2002) model earnings losses as loss of worker-specific skills. Earnings losses in this case are large and persistent by construction. These models do not consider the effects of this assumption on worker mobility. Low et al. (2010) and Davis and von Wachter (2011) propose match-specific skill losses in models that match average worker mobility. In this case, earnings losses are small and transitory.

in case of worker-specific skill losses or placement support in case of match-specific skill losses should be at the heart of labor market policies targeted at displaced workers. We use counterfactual employment paths from our structural model to inform the debate about the relative importance of the two explanations. We construct three counterfactual groups of workers for whom we track the evolution of earnings and wage losses after an initial skill loss. All losses are expressed relative to a benchmark group that corresponds to the control group from the twin experiment so that no selection effect will be present in the decomposition. The first group loses worker-specific skills as in the case of a single job change, but keeps the match-specific component. A second group keeps the worker-specific component, but loses the match-specific component. This group draws a new match-specific component from $g(x_m)$. A third group loses both their worker- and match-specific components. Earnings and wage losses of this third group correspond closely in size to the earnings and wage losses from the original estimation in the twin experiment.[36] We again provide a graphic illustration of the decomposition in online appendix VII.

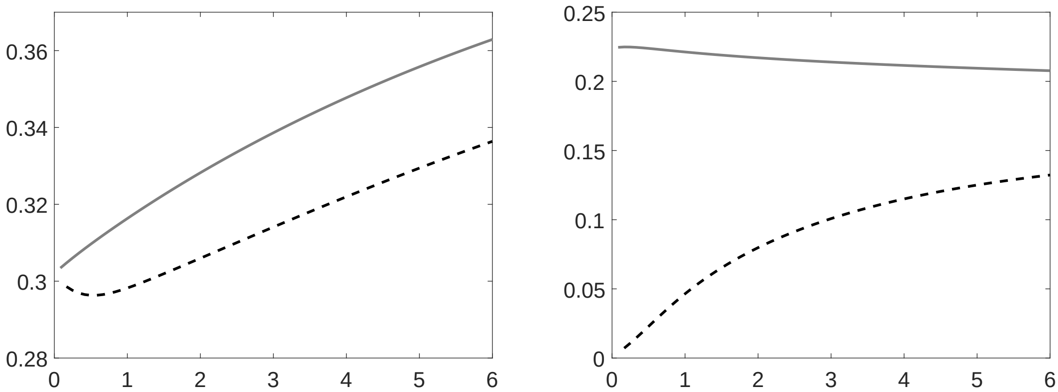
For the group with the worker-specific skill loss, we find wage losses that are small but highly persistent. After six years, their wage loss corresponds to 14.7% of the wage loss for the group that loses worker- and match-specific skills. The group with the match-specific skill loss experiences a significant recovery in wages from an initial drop of roughly 12% to 4% after six years. However, the wage loss is persistent. The wage loss after six years of this group corresponds to 85.8% of the wage loss of the group that loses both match- and worker-specific skills. The decomposition has a negative residual of -0.4 %. Turning to earnings losses, we find that the group with the match-specific skill loss experiences a strong divergence of wages and earnings initially due to increasing job instability. The difference between wages and earnings reduces over time but remains significant and persistent. If we decompose the difference between wage and earnings losses (the extensive margin effect), we find that 94.2% is due to match-specific skill loss and 4.5% due to worker-specific skill loss. The remaining 1.3% are a residual of the decomposition. Hence, match-specific skill losses are the dominant driver of wage and earnings losses.

[36] The fact that they do not match exactly results from the fact that we do not put workers into non-employment initially. We do this because otherwise we cannot keep the match-specific skills of the second group initially fixed.

5.3.4 Discussion

The loss of a particularly good job, meaning a job with high match-specific skills, accounts for most of the large and persistent earnings losses. To generate large and persistent skill differences in the match type, it is important that good jobs at the top of the job ladder are very stable. Workers who have lost their good jobs due to displacement search the market and recover to the average job in the economy, so there is mean reversion from below. If good jobs are very stable, there is no mean reversion from above leading to large and persistent differences. Figure 9 visualizes the skill dynamics for the worker- and the match-specific skills following the initial displacement event.

Figure 9: Skill dynamics following displacement



(a) Worker-specific skills

(b) Match-specific skills

Notes: Left panel: Average worker-specific skill level in control group (gray solid line) and layoff group (black dashed line) after displacement event. Right panel: Average match-specific skill level in control group (gray solid line) and layoff group (black dashed line) after displacement event. Vertical axes show mean skill levels (x_w and x_m). Horizontal axes show time in years relative to the displacement event.

Looking at worker-specific skills from our twin experiment in figure 9(a), we see that there is an initial drop followed by diverging paths due to job instability and high worker mobility in the layoff group (black dashed line). Looking at match-specific skills from our twin experiment in Figure 9(b), we find that the initial drop is followed by a recovery of the layoff group towards the mean (black dashed line). There is little mean reversion from above due to very stable jobs at the top of the job ladder (gray solid line). Although the job ladder allows for mean-reversion from below, the low mean-reversion from above leads to persistent differences in match-specific skills.

The good jobs at the top of the job ladder are the result of search rather than of accumulated worker-specific skills, and might therefore be considered as a source of transitory differences across workers. The fact that persistent earnings losses are driven by this skill component might hence be surprising. Our skill process is not confined to provide this explanation. While different explanations which we encompass in our model could potentially generate large and persistent earnings losses, it is worker mobility that pins down the skill process in our model. An explanation that focuses on the deterioration of worker-specific skills during unemployment or upon transition as the key driver of earnings losses faces the challenge of having to match the empirical mobility pattern (Ljungqvist and Sargent (1998)). Such an explanation might generate large earnings losses at least initially as it affects workers' persistent skill component but is at odds with observed worker mobility (see den Haan et al. (2000b) for a related point). If worker-specific skills were the main source of earnings losses, this would imply that expected losses from mobility are high and workers who have a mobility choice will be very reluctant to engage in mobility. As a result, average worker mobility would be low, both because expected losses of mobility are high due to low transferability of skills and because gains from mobility are little because of little persistent job heterogeneity.[37]

To explain high average worker mobility, we need a skill process that features a high degree of transferability of accumulated skills and sufficiently large gains from mobility. Our skill process has these features with gains from mobility being large because jobs further up on the job ladder are more stable and pay higher wages. As a consequence, earnings losses are driven by the loss of a particularly good job rather than by the deterioration of accumulated worker-specific skills.

5.4 Sensitivity

We provide a detailed discussion of the sensitivity of our results for earnings losses in online appendix VI. Here, we highlight the most important findings. We demonstrate that our model closely reproduces the earnings losses for the non-mass layoff sample in Couch and Placzek (2010). We do this by including all separators, i.e. endogenous separations and job-to-job transitions, in the layoff group. Including endogenous separations and job-to-job transitions implies that we include workers that are negatively selected based on their worker- and match-

[37] In online appendix IV.1, we discuss a model extension with additional skill depreciation during unemployment. We estimate the model using data on unemployment duration dependence. We find that the estimated skill depreciation is small. On average less than 3 % of workers lose skills due to one additional month of unemployment.

specific skill type. Even in this case, we get large and persistent earnings losses, although they are slightly lower in line with the empirical evidence. We also show that earnings losses change little with age in line with Jacobson et al. (1993). We also report the profile of long-run earnings losses underlying our comparison to the results by Davis and von Wachter (2011). We show that earnings losses are still significant 20 years after the initial displacement event. We discuss in detail the effects of varying selection criteria for the control group that is the key difference between Davis and von Wachter (2011) and Couch and Placzek (2010). We also demonstrate that when we select separators with good labor market prospects, then earnings losses vanish in line with the empirical findings for separators who do not claim unemployment benefits. Finally, we use age-specific job stability thresholds to account for the fact that tenure increases linearly with age. We still find earnings losses to be large and persistent.

Regarding the decomposition of earnings losses, we discuss in Section IV of the online appendix results from the two model extensions with skill depreciation during non-employment (Section IV.1) and additional skill accumulation on the job to match the tail of the wage distribution (Section IV.2). We find for both extensions large and persistent earnings losses in line with the baseline model. As in the baseline model, the decomposition of earnings losses attributes the largest contribution to the wage loss effect, followed by the selection effect and the extensive margin effect. The wage loss effect is largest in the extension with skill depreciation during non-employment (59%), compared to the extension with additional skill accumulation on the job (55%) and the benchmark (49%). As we discuss above, the mobility dynamics that identify the parameters of the skill process put strong discipline on skill dynamics associated with worker mobility and job loss. The case with skill depreciation during non-employment assumes a skill process that in principle has the most adverse consequences for a displaced worker who has accumulated a lot of worker-specific skills. Even under the assumption of duration dependent skill losses, the wage loss effect after displacement increases only modestly suggesting that our baseline skill dynamics capture well the main sources of earnings losses after job displacement.

6 Policy analysis

Understanding the sources of earnings losses is vital for designing labor market policies. Viewed through the lens of our structural model, active labor market policy can potentially help displaced workers along two margins: First, it can help to avoid the loss of worker-specific skills by providing retraining services. Second, it can help to regain match-specific skills by providing placement support to foster better matches between jobs and workers.

In practice, placement support and retraining are the two pillars of the Dislocated Worker Program (DWP) of the Workforce Investment Act. The DWP "is designed to provide quality employment and training services to assist eligible individuals in finding and qualifying for meaningful employment, and to help employers find the skilled workers they need to compete and succeed in business."[38] The DWP is targeted explicitly at displaced workers who lost their jobs due to layoff, plant closures, or downsizing.[39] The targeted group, therefore, corresponds in principle to the group of displaced workers in our model.

We examine the effectiveness of the DWP in reducing earnings losses within our model. Leaving aside costs to run the program, we consider retraining and placement support for 40-year-old displaced workers. It is important to bear in mind that, using our structural model we take into account all endogenous responses on wages, mobility, and vacancy posting decisions when evaluating the effects of the program. As measures for policy evaluation, we report changes in persistent earnings losses, changes in job stability, and the associated welfare changes in terms of the equivalent variation in monthly earnings.[40]

Concretely, we implement retraining by reducing the probability of skill loss for displaced workers to zero ($p_d = 0$). We keep the probability of skill loss for all job-to-job transitions and transitions from non-employment to employment if workers did not separate in a displacement event. Displaced workers receive the policy on their initial non-employment spell after displacement but not in case of future separations. We assume that retraining takes place as intensive

[38]http://www.doleta.gov/programs/general_info.cfm (retrieved September 14, 2015).

[39]The program also comprises special funds that can be channeled to areas that suffer from plant closings, mass layoffs, or job losses due to natural disasters or military base realignment and closures. The median worker in the program is between age 30 and 44, has a high school education, and earns about median earnings before displacement. Males and females are equally likely to be part of the program. See http://www.doleta.gov/programs/dislocated.cfm for a more detailed description of the program.

[40]The latter measure accurately reflects welfare in our model as it takes utility flows from non-employment and utility flows from the nonpecuniary utility component of job search into account.

class-room training so that there are opportunity costs for workers who cannot, by assumption, search for jobs during the program. We denote the duration of the program by t and report results for varying program durations including $t = 0$ and discuss the trade-off between skill recovery and lost search time.

We implement placement support by replacing the unconditional offer distribution $g(x_m)$ by a distribution of match-specific skills of workers who were displaced τ months ago but had not received the policy. These workers have already searched τ months on and off the job. We call τ the "leapfrogged" search time that is offered by the policy to currently displaced workers. Receiving a "leapfrogged offer distribution" of τ months each period makes searching a new job much more efficient for displaced workers, and results in a better match between jobs and workers. One interpretation of τ is that it measures the effectiveness of the employment agency to deal with search frictions when generating job offers. A non-employed worker generates π_{ne} offers per month. After τ months of search, a non-employed will have generated $\pi_{ne}\tau$ offers. The employment agency leapfrogging τ months of search therefore generates τ times as many offers. Selection on these offers during the search process shifts the distribution so that it first-order stochastically dominates the offer distribution $g(x_m)$ without policy. Displaced workers receive this shifted offer distribution each period during their initial non-employment spell after the displacement event. Hence, each period's offer distribution is equivalent to a distribution that comprises τ months of search.

Table 4 reports results for retraining of different program durations t in the first four columns. The last four columns report results for placement support as a function of leapfrogged search time τ. Looking at retraining, the best potential outcome of the program is being immediately effective and the duration being zero ($t = 0$), the welfare gain of the worker amounts to 0.7% of earnings. Earnings losses reduce by 11% and job stability measured as the change in unemployment rates six years after displacement increases so that the unemployment rate decreases by 5%. The worker is indifferent between participating in the policy or not at a program duration of 3.2 weeks (0.74 months). Earnings losses reduce by 9.1% and job stability decreases slightly, which in turn increases the unemployment rate by 1%. The gradient over the program duration is very steep. If the program lasts for 3 months, the worker will not like to participate and would be even willing to give up 1.8% of earnings to avoid participating in the program. Earnings losses are 3.2% lower than in the case without policy intervention, although welfare effects are negative. Job stability decreases substantially, which raises the unemployment rate by 20% and, thereby, increases earnings losses from the extensive margin effect. If the program lasted for 6 (12) months, the lost search time increased the earnings losses and workers would experience 7.5% (60.1%) higher earnings losses and higher job instability. Hence, the policy must quickly be effective in order to avoid outcomes that are worse than without policy intervention.

Table 4: Effects of placement support and retraining on welfare, earnings losses, and job stability

retraining				placement support			
t	ΔV	Δw	Δu	τ	ΔV	Δw	Δu
0	0.7 %	-11.5 %	-5.0 %	3	0.2 %	-5.4 %	-4.6 %
0.74	0.0 %	-9.1 %	0.9 %	6	0.4 %	-10.1 %	-8.3 %
3	-1.8 %	-3.2 %	19.9 %	12	0.7 %	-20.9 %	-15.9 %
6	-4.0 %	7.5 %	51.0 %	24	1.2 %	-42.5%	-29.2 %
12	-8.3 %	60.1 %	158.0 %	$\bar{\tau}$	0.6 %	-15.2%	-6.8 %
earnings loss without policy					7.5%		
unemployment rate without policy					4.2%		

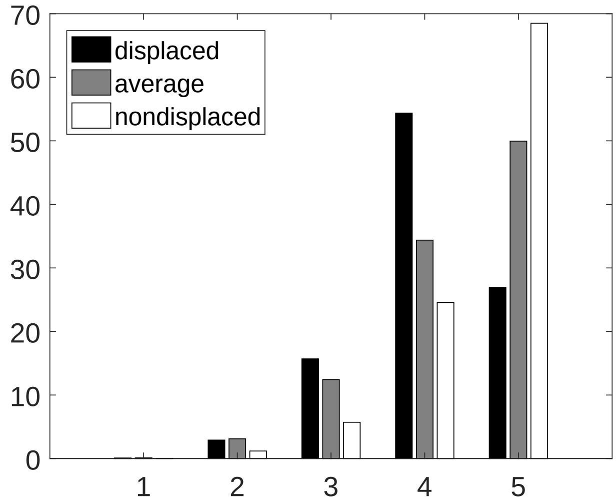
Notes: Effects of placement support and training policies on welfare, earnings losses, and job stability. The term ΔV denotes the average welfare effect expressed as a multiple of median earnings. The term Δw denotes the reduction in earnings losses from the twin experiment in the sixth year after the displacement event relative to earnings losses without policy intervention (positive numbers indicate an increase of earnings losses). The term Δu denotes the percentage change in the unemployment rate in comparison to the unemployment rate without policy intervention in the sixth year after the displacement event (positive numbers indicate an increase of the unemployment rate). The welfare effect is the present discounted value of the consumption equivalent variation over the life cycle of a worker entering the labor market. The parameter t denotes the duration of the worker training program that avoids skill loss but prevents job search. The parameter τ denotes the shift of the offer distribution to τ periods ahead in the search process. The parameter $\bar{\tau}$ denotes the case of the offer distribution to match the average distribution six years after the displacement event. Bottom rows show earnings losses and unemployment rate without policy intervention in the sixth year after the displacement event from the twin experiment. See text for further details.

A placement support program that is equivalent to the retraining program in terms of its welfare effect, the duration of which is $t = 0$, has to offer the equivalent of 12 months of search ($\tau = 12$). Given that a displaced worker in the model manages to obtain on average about 0.5 offers a month, leapfrogging 12 months of search implies that the agency would need to generate roughly 6 offers each month decreasing the time between job offers from 60 days to 5 days. This constitutes a substantial increase in efficiency regarding job search from placement support. However, even if the agency managed to do so, earnings losses would still be large and would reduce by a mere 21%; job stability would increase reducing the unemployment rate by 16%. To see that this is a substantial policy intervention, we compare it to a policy where workers receive full mean reversion and get back to the average match distribution of their cohort ($\bar{\tau}$). In this case, the welfare gain is 0.6% and earnings losses are 15.2% lower. Job stability increases and reduces the unemployment rate by 6.8%. The effect is smaller than that from leapfrogging 12 months of search. Leapfrogging 12 months therefore constitutes a substantial policy intervention that overcomes search frictions to an extent that workers will have even better matches than the average worker. It is important to keep in mind that the policy increases the search efficiency of displaced workers permanently during their initial search period because each period, they receive offers from a distribution that contains τ months of search. Hence, as an example, receiving three offers with the policy corresponds to 36 months of search off the job without the policy.

Combining placement support ($\bar{\tau}$) and retraining ($t = 0$) yields complete mean reversion for displaced workers from below, in the sense that the workers receive the average match-type distribution of their cohort and experience no worker-specific skill loss. This policy yields a welfare gain of 1.3% and reduces earnings losses by 26.6%. Still, earnings losses are large and persistent with 5.5% after six years compared to 7.5% without policy intervention. We find that the effects on earnings losses from the two policies are approximately additive in the combined program that reduces earning losses by 26.6%; the programs in isolation yield a reduction in earnings losses of 15.2% from placement support and 11.5% from retraining ($15.2\% + 11.5\% = 26.7\% \approx 26.6\%$). This reduction of earnings losses is modest compared to the substantial and very effective policy intervention.

To investigate the reason behind this ineffectiveness, Figure 10 shows the distribution of match-specific skill types six years after displacement for displaced workers (without policy intervention), the average worker, and nondisplaced workers (Figure 9 shows the corresponding mean skill levels for displaced and nondisplaced workers). First, when comparing displaced workers to the average worker, we see that without policy intervention, there is modest mean reversion and search frictions contribute to earnings losses. Second, when comparing the average worker to the group of nondisplaced workers, we see that displaced workers come from very good and stable jobs. Job stability of nondisplaced workers leads to the persistent differences between them and the average worker. Hence, even if a policy manages to bring displaced workers back to the average as does our placement support policy with retraining, these workers still suffer

Figure 10: Distribution across match-types following displacement



Notes: Distribution over match-types x_m for displaced workers, the average worker, and workers in the control group of the twin experiment (nondisplaced) six years after the displacement event. Vertical axis shows five discretized match states (1: lowest, 5: highest), vertical axis shows share of employed workers in each of the skill states in percentage points.

substantial earnings losses despite full mean-reversion from below.

Our policy analysis offers a structural interpretation to several empirical studies evaluating the DWP (see Card et al. (2010) for a survey). These studies estimate that the effectiveness of the DWP is moderate at best and counterproductive at worst. The studies on the DWP surveyed in Heckman et al. (1999) typically conclude that wage effects of active labor market policies are small or have no impact on displaced workers. More recently, Heinrich et al. (2013) even find a negative lock-in effect in the first two years after exiting the program and a zero impact thereafter for men.

Our model suggests that even if more money is invested into active labor market policies to help displaced workers, it is unlikely that these policies will significantly help to reduce earnings losses. Both retraining and placement support will likely affect only a small fraction of the total earnings losses. Of course, any program that increases worker-specific skills beyond the predisplacement skill level would be beneficial and would decrease earnings losses further. Such a policy constitutes general education and would equally apply for workers on the job, who would benefit similarly. Any type of placement support that implicitly or explicitly helps to improve the match distribution would be welcome but it is hard to envision a governmental program that overcomes search frictions to an extent that leads to matches that are even better than the average of the market. Our negative perception of the effectiveness of active labor market policy is rooted in our view on the sources of the earnings losses. An active policy can help to remove frictions and foster mean reversion making displaced workers recover to the average worker. However, it cannot affect the downward force so that nondisplaced workers have persistently better jobs than the average worker.

7 Conclusions

Large and persistent earnings losses of displaced workers are a prime source of income risk in macroeconomic models with adverse individual and macroeconomic consequences. Understanding the size and sources of earnings losses poses a considerable challenge to existing labor market models predicting small and transitory losses. We provide a novel explanation and study the size and sources of earnings losses from a structural labor market perspective.

In our model, good jobs at the top of the job ladder do not only pay high wages but are also very stable. We support this argument by providing new empirical evidence on job stability, heterogeneity in worker mobility, and the correlation of wages and job stability for the United States. While wage heterogeneity has been studied extensively, we show that accounting for heterogeneity in job stability is important to explain the observed earnings dynamics. Our results highlight the tight link between job stability and earnings dynamics. After accounting for the empirically observed job stability at the top of the job ladder, our model generates large and persistent earnings losses consistent with the empirical evidence.

Once a worker has lost a job at the top of the job ladder due to displacement, the job ladder provides the opportunity for mean reversion from below but the counterfactual employment path —a stable job at the top of the job ladder— prevents mean reversion from above, so that large and persistent differences between displaced and nondisplaced workers arise. We explore the effectiveness of active labor market policies like the Dislocated Worker Program to help displaced workers. Our findings suggest that job stability for nondisplaced workers is key to understand the empirically documented ineffectiveness of these programs because they only affect mean reversion from below.

On the theoretical side, we provide a life-cycle labor market model with search frictions together with an identification approach for model parameters based on heterogeneity in worker mobility. Our model provides a unified framework to jointly study worker mobility, job stability, and earnings dynamics and can serve as a starting point for several avenues of future research. The life-cycle dimension and skill process make the model broadly applicable to important policy questions we have not considered here. For example, one can study the long-term effects of the increase in youth unemployment on skill accumulation and earnings, a problem many European countries currently face. More generally, the model can be used to study the impact of policy interventions on different demographic groups, the effect of taxation on worker reallocation or the effect of changes in the unemployment insurance system on earnings and mobility dynamics. Because of its tractability, the most obvious extension though is to incorporate business cycle shocks. Davis and von Wachter (2011) find that estimated earnings losses increase substantially in recessions. In light of the recent crises, a better understanding of the underlying causes is urgent. We leave these applications to future research.

A Data

A.1 Current Population Survey (CPS)

We use data from the basic monthly files of the Current Population Survey (CPS) between January 1980 and December 2007 and the Occupational Mobility and Job Tenure supplements for 1983, 1987, 1991, 1996, 1998, 2000, 2002, 2004, and 2006. The CPS is a monthly household survey representative of the U.S. noninstitutionalized population and constitutes the main data source for labor market statistics. Every household is interviewed for four consecutive months, not interviewed for the following eight months, and then interviewed for four consecutive months again before leaving the survey permanently. The survey provides information on approximately 60,000 households with 110,000 individuals each month. We link data from the monthly files and the supplements using the matching algorithm as in Madrian and Lefgren (1999). From the matched files we construct worker flows as in Shimer (2012) or Fallick and Fleischman (2004). In particular, we use the approach proposed in Fallick and Fleischman (2004) to construct job-to-job worker flows.[41] Worker flows are derived using adjusted observation weights to account for attrition in matching as in Feng (2013). Worker flows are furthermore adjusted for misclassification. Misclassification of the labor force status is a well-known problem in the CPS already since the early work of Poterba and Summers (1986) and Abowd and Zellner (1985) and has recently received renewed attention in the literature (see Feng (2013)). We adjust flows using the approach in Hausman et al. (1998) with data from the supplement files where information on age and tenure is available and run separate logit regressions for separation and job-to-job rates for each year.[42] We use the average estimated error across regressions to adjust transition rates.[43] The estimated misclassification probabilities are 0.0074 for separations and 0.0094 for job-to-job transitions. When compared to the misclassification adjustments surveyed in Feng (2013), the adjustment appears modest for separation rates. For job-to-job rates, our estimated misclassification probabilities are the first attempt to adjust job-to-job flows for misclassifica-

[41] Given that the approach in Fallick and Fleischman (2004) uses dependent interviewing these flows can only be constructed from 1994 onwards.

[42] As controls, we include age and tenure terms up to order three, age and tenure interactions up to total degree three, education dummies grouping workers into four education groups (high school dropouts, high school, some college, and college), as well as interactions between education and age, education and tenure.

[43] The results are similar when we use the median error instead of the mean. The adjusted transition rates are $\pi_{adj} = \frac{\pi - \alpha}{1 - 2\alpha}$ where α denotes the misclassification error and π the measured transition rate.

tion, to the best of our knowledge. However, our model provides some indication regarding the validity of the adjustment because it shows that the adjusted rates match the observed level of job stability (mean tenure) as it has to be the case in a consistent stock-flow relationship.

To derive transition rate profiles by age and tenure, we construct worker flows for cells that share the same characteristics for each pair of linked cross sections where this information is available. Specifically, we construct age profiles of newly hired workers by considering those workers that have one year of tenure. To increase the number of observations at each age, we use moving age windows centered at each age with a range of plus and minus two years. We average all transition rate profiles across surveys to remove business cycle variation from transition rate profiles. We made sure that age profiles from the basic monthly CPS files and the average age profiles from the irregular supplement files are consistent by adjusting mean age profiles using age-specific adjustment factors. The reported confidence bands are calculated using bootstrapping with 10,000 repetitions from the pooled sample stratified by age. We always report $-/+2$ standard deviations around the mean.

To bring the model to the data, we have to derive worker flows from the model. For the model, we assume that the production stage includes the reference week for which the CPS labor force status is reported.

Wage data comes from the CEPR March CPS uniform extracts for the period from 1980 to 2008. We use hourly wages constructed by dividing total wage and salary income by total (usual) hours worked. We winsorize the top and bottom 1% of the log wage distribution and regress log wages on age and year dummies. To construct mean log wage profiles by age, we run a regression of log wages on age and time dummies and use the estimated age coefficients as in Heathcote et al. (2010). To construct the age profile for the variance of log wages, we use variance of residuals by age from the regression.

A.2 Survey of Income and Program Participation (SIPP)

We use data from the 2004 Panel of the Survey of Income and Program Participation (SIPP) conducted by the Census Bureau.[44] The 2004 Panel provides data on roughly 68,600 individuals representative of the U.S. noninstitutionalized population. It provides information on demographic characteristics and labor market histories including wages. The SIPP is a household

[44] We use the 2014 SIPP Uniform Extracts data provided by the Center for Economic and Policy Research.

survey where each household in the panel is interviewed every four months and each household has nine interviews in total over the survey period. At each interview, information for the four months preceding the interview is collected so that there are in total 36 observations per individual. We restrict the sample to workers age 20 to 55. Workers can report information on more than one job. We only keep primary jobs and drop contingent workers.

We code individuals as employed if they had a job and worked on this job the entire month. Given that we are interested in job stability for all workers including those with more stable jobs, we code all other employment states as non-employed. We code a separation if a worker is employed in month t and non-employed in $t+1$. If hourly wages are not reported, we use the calculated wage derived by dividing earnings by hours worked.

The regression coefficients reported in the main text are the coefficients from a regression of an indicator variable $I^T_{i,t}$ on *log* wage $w_{i,t}$, and a full set of dummy variables for age and year effects. The indicator variable $I^T_{i,t}$ is 1 if individual i observed in period t separates within the next T months from her or his employer. Wages, age, and year information are used for individual i at time t in the regression

$$I^T_{i,t} = \beta w_{i,t} + \sum_{a=20}^{55} \alpha_a I_a + \sum_{y=2004}^{2007} \gamma_y I_y + u_{i,t}.$$

where $u_{i,t}$ denotes the error term and I_a and I_y denote age and year dummies. Table 1 in the main text reports the coefficient β from this regression. For the regression with newly hired workers, we restrict the regression sample to workers with zero tenure at t.

A.3 Panel Study of Income Dynamics (PSID)

The Panel Study of Income Dynamics (PSID) is a longitudinal panel survey of individuals that is representative of the U.S. population. It collects demographic information including age, sex, education, and marital status together with detailed information on employment and income histories. We follow Stevens (1997) regarding sample selection, definition of displacement events, and econometric approach. We use data from 1968 to 1988. We keep individuals who are present in 1968 and are the head of household. We require that they have positive annual earnings in each year they are in the survey and that they did not get displaced in the 10 years before 1968. This leaves us with 1,615 individuals in 1969. In total, 445 workers get displaced between 1969 and 1987.[45]

A displacement event is coded when a worker lost the job due to plant closing, the worker reports having been laid off or fired. The information is reported by individuals in January of each survey year when they are asked about what happened to their previous job. Stevens discusses potential caveats of this definition in detail.

We follow Stevens' econometric implementation of the original Jacobson et al. (1993) approach. We regress the outcome variable of interest y_{it} (*log* earnings, *log* wages, *log* hours) for individual i in year t on gender-specific age profiles collected in X_{it}, individual fixed effects α_i, year fixed effects γ_t, and dummy variables D_{it}^s where s indicates the distance to the displacement event in years. We estimate the following regression

$$y = \beta X_{it} + \alpha_i + \gamma_t + \sum_{s=-2}^{10} D_{it}^s + u_{it}$$

where u_{it} denotes the error term. We use dummies from two years before ($s = -2$) to 10 years or more after the displacement event ($s = 10$) as reported in Table 4 of Stevens (1997). We only report results for the first displacement event. The first displacement event corresponds to the displacement event of high-tenure workers considered in Couch and Placzek (2010) and in the model.

B Model estimation

We estimate model parameters using a method of moments. We use as objective function the sum of squared percentage deviations of the model implied age profiles, newly hired age profiles, and the age profile of mean tenure from the empirical counterparts. We avoid simulation noise from the model and iterate instead on the cross-sectional distributions over age, tenure, and skill types to determine model moments. If we denote the parameter vector by θ, then the objective is

$$\min_{\theta} \sum_{a=20}^{50} \left(\frac{\pi_s(a,\theta) - \hat{\pi}_s(a)}{\hat{\pi}_s(a)} \right)^2 + \sum_{a=20}^{50} \left(\frac{\pi_{eo}(a,\theta) - \hat{\pi}_{eo}(a)}{\hat{\pi}_{eo}(a)} \right)^2 + \sum_{a=20}^{50} \left(\frac{\pi_{ne}(a,\theta) - \hat{\pi}_{ne}(a)}{\hat{\pi}_{ne}(a)} \right)^2$$
$$+ \sum_{a=21}^{50} \left(\frac{\pi_s^{NH}(a,\theta) - \hat{\pi}_s^{NH}(a)}{\hat{\pi}_s^{NH}(a)} \right)^2 + \sum_{a=21}^{50} \left(\frac{\pi_{eo}^{NH}(a,\theta) - \hat{\pi}_{eo}^{NH}(a)}{\hat{\pi}_{eo}(a)} \right)^2 + \sum_{a=25}^{60} \left(\frac{t(a,\theta) - \hat{t}(a)}{\hat{t}_s(a)} \right)^2$$

[45]These numbers differ marginally from the numbers reported in Stevens (1997).

with $\pi_s(a,\theta)$ denoting the average separation rate from the model using parameter vector θ. π_{eo} and π_{ne} denote the job-to-job and job finding rate, accordingly. $t(a,\theta)$ denotes mean tenure at age a under parameter vector θ from the model. The newly hired age profiles are denoted by a superscript NH. Data profiles are indicated using a hat. For separations, job-to-job transitions, and job-finding rates we use the age profile from age 20 to 50; we use the newly hired age profiles for separations and job-to-job transitions from age 21 to 50; we use the mean tenure profile from age 25 to 60. We only use information up to age 50 for transition rates to abstract from early retirement, which becomes particularly strong for the separation rate. We use tenure information from age 25 onwards to abstract from the initial differences between data and model in tenure at age 20. We use information until age 60 to put additional emphasis on job stability in the estimation. Initial differences in tenure arise because the model is restricted to generate a tenure level of zero at the beginning of working life, so that we can target the newly hired age profile only from age 21 onwards. We use a standard Newton-type solver for the optimization. We experimented with different starting values and solvers for global optimization.

References

Abowd, J. M. and A. Zellner (1985): "Estimating Gross Labor-Force Flows," Journal of Business & Economic Statistics, 3, 254–83.

Altonji, J. G. and R. A. Shakotko (1987): "Do Wages Rise with Job Seniority?" Review of Economic Studies, 54, 437–59.

Altonji, J. G. and N. Williams (2005): "Do Wages Rise With Job Seniority? A Reassessment," Industrial and Labor Relations Review, 58, 370–397.

Bagger, J., F. Postel-Vinay, F. Fontaine, and J. Robin (2014): "Tenure, Experience, Human Capital and Wages: A Tractable Equilibrium Search Model of Wage Dynamics," American Economic Review, 1–51.

Becker, G. S. (1962): "Investment in human capital: A theoretical analysis," The Journal of Political Economy, 9–49.

Bonhomme, S. and G. Jolivet (2009): "The pervasive absence of compensating differentials," Journal of Applied Econometrics, 24, 763–795.

Card, D., J. Kluve, and A. Weber (2010): "Active Labour Market Policy Evaluations: A Meta-Analysis," The Economic Journal, 120, F452–F477.

Chan, S. and A. H. Stevens (2001): "Job Loss and Employment Patterns of Older Workers," Journal of Labor Economics, 19, pp. 484–521.

Cheron, A., J.-O. Hairault, and F. Langot (2013): "Life-Cycle Equilibrium Unemployment," Journal of Labor Economics, 31, 843–882.

Chodorow-Reich, G. and L. Karabarbounis (2016): "The Cyclicality of the Opportunity Cost of Employment," Journal of Political Economy, 124, 1563–1618.

Couch, K. A. and D. W. Placzek (2010): "Earnings Losses of Displaced Workers Revisited," American Economic Review, 100, 572–89.

Daly, M., , D. Hryshko, and I. Manovskii (2016): "Improving the Measurement of Earnings Dynamics," Working Paper 22938, National Bureau of Economic Research.

Davis, S. J. and T. von Wachter (2011): "Recessions and the Costs of Job Loss," Brookings Papers on Economic Activity, 43, 1–72.

den Haan, W., C. Haefke, and G. Ramey (2005): "Turbulence and Unemployment in a Job Matching Model," Journal of the European Economic Association, 3, 1360–1385.

den Haan, W., G. Ramey, and J. Watson (2000a): "Job Destruction and Propagation of Shocks," American Economic Review, 90, 482–498.

den Haan, W. J., G. Ramey, and J. Watson (2000b): "Job destruction and the experiences of displaced workers," Carnegie-Rochester Conference Series on Public Policy, 52, 87–128.

Diebold, F. X., D. Neumark, and D. Polsky (1997): "Job Stability in the United States," Journal of Labor Economics, 15, 206–33.

Dustmann, C. and C. Meghir (2005): "Wages, Experience and Seniority," Review of Economic Studies, 72, 77–108.

Esteban-Pretel, J. and J. Fujimoto (2014): "Life-Cycle Labor Search With Stochastic Match Quality," International Economic Review, 55, 575–599.

Faberman, R. J., A. Mueller, A. Sahin, and G. Topa (2016): "Job Search Behavior among the Employed and Non-Employed," Federal Reserve Bank of New York, mimeo.

Fallick, B. and C. A. Fleischman (2004): "Employer-to-employer flows in the U.S. labor market: the complete picture of gross worker flows," Finance and Economics Discussion Series 2004-34, Board of Governors of the Federal Reserve System (U.S.).

Farber, H. S. (1995): "Are Lifetime Jobs Disappearing? Job Duration in the United States: 1973-1993," NBER Working Papers 5014, National Bureau of Economic Research, Inc.

——— (1999): "Mobility and stability: The dynamics of job change in labor markets," in Handbook of Labor Economics, ed. by O. Ashenfelter and D. Card, Elsevier, vol. 3 of Handbook of Labor Economics, chap. 37, 2439–2483.

——— (2007): "Job Loss and the Decline in Job Security in the United States," Working papers, Princeton University, Department of Economics, Center for Economic Policy Studies.

——— (2008): "Employment Insecurity: The Decline in Worker-Firm Attachment in the United States," Working Papers 1056, Princeton University, Department of Economics, Center for Economic Policy Studies.

Feng, S. (2013): "Misclassification errors and the underestimation of the US unemployment rate," The American Economic Review, 103, 1054–1070.

Fujita, S. (2011): "Reality of on-the-job search," Federal Reserve Bank of Philadelp working paper NO. 10-34/R.

Guvenen, F. (2009): "An Empirical Investigation of Labor Income Processes," Review of Economic Dynamics, 12, 58–79.

Hagedorn, M. and I. Manovskii (2008): "The Cyclical Behavior of Equilibrium Unemployment and Vacancies Revisited," American Economic Review, 98, 1692–1706.

Hagedorn, M., I. Manovskii, and Y. Wang (2017): "Search Frictions and Wage Dispersion," Tech. rep.

Hall, R. E. (1982): "The Importance of Lifetime Jobs in the U.S. Economy," American Economic Review, 72, 716–24.

Hausman, J., J. Abrevaya, and F. Scott-Morton (1998): "Misclassification of the dependent variable in a discrete-response setting," Journal of Econometrics, 87, 239 – 269.

Heathcote, J., F. Perri, and G. L. Violante (2010): "Unequal We Stand: An Empirical Analysis of Economic Inequality in the United States: 1967-2006," Review of Economic Dynamics, 13, 15–51.

Heckman, J. J., R. J. Lalonde, and J. A. Smith (1999): "The economics and econometrics of active labor market programs," in Handbook of Labor Economics, ed. by O. Ashenfelter and D. Card, Elsevier, vol. 3 of Handbook of Labor Economics, chap. 31, 1865–2097.

Heinrich, C. J., P. R. Mueser, K. R. Troske, K.-S. Jeon, and D. C. Kahvecioglu (2013): "Do public employment and training programs work?" IZA Journal of Labor Economics, 2, 1–23.

Hornstein, A., P. Krusell, and G. L. Violante (2011): "Frictional Wage Dispersion in Search Models: A Quantitative Assessment," American Economic Review, 101, 2873–98.

Jacobson, L. S., R. J. LaLonde, and D. G. Sullivan (1993): "Earnings Losses of Displaced Workers," American Economic Review, 83, 685–709.

Jarosch, G. (2014): "Searching for Job Security and the Consequences of Job Loss," Tech. rep.

Jolivet, G., F. Postel-Vinay, and J.-M. Robin (2006): "The empirical content of the job search model: Labor mobility and wage distributions in Europe and the US," European Economic Review, 50, 877–907.

Jung, P. and K. Kuester (2011): "The (un)importance of unemployment fluctuations for the welfare cost of business cycles," Journal of Economic Dynamics and Control, 35, 1744 1768.

——— (2015): "Optimal Labor-Market Policy in Recessions," AEJ: Macroeconomics, 7, 124–156.

Jung, P. and M. Kuhn (2014): "Labour market institutions and worker flows: comparing Germany and the US," The Economic Journal, 124, 1317–1342.

Kambourov, G. and I. Manovskii (2009a): "Occupational Mobility and Wage Inequality," Review of Economic Studies, 76, 731–759.

——— (2009b): "Occupational Specificity of Human Capital," International Economic Review, 50, 63–115.

Krebs, T. (2007): "Job Displacement Risk and the Cost of Business Cycles," American Economic Review, 97, 664–686.

Kroft, K., F. Lange, and M. J. Notowidigdo (2013): "Duration dependence and labor market conditions: Evidence from a field experiment," The Quarterly Journal of Economics, 128, 1123–1167.

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

Krusell, P. and A. A. Smith (1999): "On the Welfare Effects of Eliminating Business Cycles," Review of Economic Dynamics, 2, 245–272.

Kudlyak, M. and F. Lange (2014): "Measuring heterogeneity in job finding rates among the nonemployed using labor force status histories," Tech. rep., FRB Richmond Working Paper.

Ljungqvist, L. and T. Sargent (1998): "The European Unemployment Dilemma," Journal of Political Economy, 106, 514–550.

Ljungqvist, L. and T. J. Sargent (2008): "Two Questions about European Unemployment," Econometrica, 76, 1–29.

Low, H., C. Meghir, and L. Pistaferri (2010): "Wage Risk and Employment Risk over the Life Cycle," American Economic Review, 100, 1432–67.

Madrian, B. C. and L. J. Lefgren (1999): "A Note on Longitudinally Matching Current Population Survey (CPS) Respondents," Working Paper 247, National Bureau of Economic Research.

Marinescu, I. and R. Wolthoff (2016): "Opening the black box of the matching function: The power of words," Tech. rep., National Bureau of Economic Research.

Menzio, G. and S. Shi (2011): "Efficient Search on the Job and the Business Cycle," Journal of Political Economy, 119, 468 – 510.

Menzio, G., I. A. Telyukova, and L. Visschers (2016): "Directed search over the life cycle," Review of Economic Dynamics, 19, 38–62.

Moscarini, G. (2005): "Job Matching and the Wage Distribution," Econometrica, 73, 481–516.

Parent, D. (2000): "Industry-specific capital and the wage profile: Evidence from the national longitudinal survey of youth and the panel study of income dynamics," Journal of Labor Economics, 18, 306–323.

Petrongolo, B. and C. A. Pissarides (2001): "Looking into the Black Box: A Survey of the Matching Function," Journal of Economic Literature, 2001, 390–431.

Postel-Vinay, F. and J.-M. Robin (2002): "Equilibrium wage dispersion with worker and employer heterogeneity," Econometrica, 70, 2295–2350.

Poterba, J. M. and L. H. Summers (1986): "Reporting Errors and Labor Market Dynamics," Econometrica, 54, 1319–1338.

Pries, M. J. (2004): "Persistence of employment fluctuations: a model of recurring job loss," The Review of Economic Studies, 71, 193–215.

Rogerson, R. and M. Schindler (2002): "The welfare costs of worker displacement," Journal of Monetary Economics, 49, 1213–1234.

Ruhm, C. J. (1991): "Are Workers Permanently Scarred by Job Displacements?" American Economic Review, 81, 319–24.

Rupert, P. et al. (2004): "Wage and employer changes over the life cycle," Economic Commentary.

Rust, J. (1987): "Optimal replacement of GMC bus engines: An empirical model of Harold Zurcher," Econometrica, 999–1033.

Shimer, R. (2005): "The Cyclical Behavior of Equilibrium Unemployment, Vacancies, and Wages: Evidence and Theory," American Economic Review, 95, 25–49.

——— (2012): "Reassessing the ins and outs of unemployment," Review of Economic Dynamics, 15, 127 – 148.

Silva, J. I. and M. Toledo (2009): "Labor Turnover Costs and the Cyclical Behavior of Vacancies and Unemployment," Macroeconomic Dynamics, 13, 76–96.

Stevens, A. H. (1997): "Persistent Effects of Job Displacement: The Importance of Multiple Job Losses," Journal of Labor Economics, 1, 165–188.

Storesletten, K., C. I. Telmer, and A. Yaron (2004): "Cyclical Dynamics in Idiosyncratic Labor Market Risk," Journal of Political Economy, 112, 695–717.

Tatsiramos, K. (2010): "Job displacement and the transitions to re-employment and early retirement for non-employed older workers," European Economic Review, 54, 517–535.

Tjaden, V. and F. Wellschmied (2014): "Quantifying the Contribution of Search to Wage Inequality," American Economic Journal: Macroeconomics, 6, 134–61.

Topel, R. H. (1991): "Specific Capital, Mobility, and Wages: Wages Rise with Job Seniority," Journal of Political Economy, 99, 145–76.

Topel, R. H. and M. P. Ward (1992): “Job Mobility and the Careers of Young Men,” The Quarterly Journal of Economics, 107, pp. 439–479.

Violante, G. L. (2002): “Technological Acceleration, Skill Transferability, and the Rise in Residual Inequality,” The Quarterly Journal of Economics, 117, pp. 297–338.

von Wachter, T., J. Song, and J. Manchester (2009): “Long-Term Earnings Losses due to Mass-Layoffs During the 1982 Recession: An Analysis Using Longitudinal Administrative Data from 1974 to 2004,” Tech. rep., Columbia University.

Online Appendix
Not for Publication

This online appendix accompanies the paper 'Earnings losses and labor mobility over the life cycle'. Section I offers a simple statistical model that illustrates the close connection between earnings losses and heterogeneity in job stability. Section II discusses in depth some of the modeling assumptions made in section 3 of the main part of the paper and derives the bargaining outcomes in detail. Section III provides the details on the identification of model parameters not discussed in the main text and some supplementary discussion. Section IV presents two model extensions of the baseline model. The first extension demonstrates how to account for duration dependence in job finding rates. It also allows for additional skill depreciation during unemployment. The second extension demonstrates how to account for the life-cycle increase in wage inequality. We demonstrate in both cases that our main results are not affected when we match these additional data features. Section V explains the mapping of non-employment rates from model to data and how we implement our estimation of the wage dynamics in the model. It also provides some additional discussion of empirical estimates. Section VI offers an extensive sensitivity analysis on earnings losses from the model. Finally, section VII provides the graphic illustration of the decomposition of earnings losses discussed in section 5.3 of the main part of the paper.

I Simple model

This section develops a simple statistical model to demonstrate the tight link between job stability and large and persistent earnings losses. While job stability at the top of the job ladder is important to generate large and persistent earnings losses, heterogeneity in job stability is important to still match high average worker mobility.

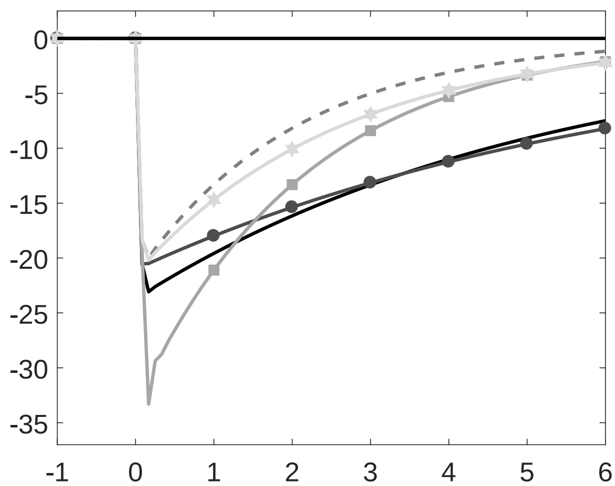
There are two types of jobs: good and bad.[46] Unemployment spells last for only one period and at reemployment all jobs are bad.[47] Good (bad) jobs separate with probability $\pi_g = 0.003$ ($\pi_b = 0.04$) and pay $w_g > w_b$, so that good jobs are more stable and yield higher earnings. Bad

[46]Here, we are agnostic about whether it is the worker or the match that makes a job good or bad. We discuss this identification problem for the full model in section 3.5.

[47]Results remain unaffected if we allow for a 10% probability of starting a good job, for example.

jobs turn into good jobs with probability γ every period either through job-to-job mobility or experience accumulation. We set $\gamma = 0.01$ so that the upward friction is considerable and the duration of bad jobs is more than eight years. We set the wage differences across good and bad jobs to match earnings losses of 7.5% after six years in line with our results from the full model (implying a wage difference of 30%). Figure A shows the resulting earnings losses. We measure earnings losses in this simple case as earnings difference between a group of workers displaced from good jobs and a group of workers not displaced from such jobs. Workers that were not displaced initially will still separate in the future according to their separation rate π_g. Our discussion in section 5 in the main part of the paper shows that this provides a good approximation to the more sophisticated approach from the empirical literature. Note, however, that we condition on job quality here, and job quality remains unobserved in the data. In the baseline case (black solid line), earnings losses are large and persistent and amount to 7.5% after six years, reproducing the empirical estimates due to calibration.

Figure A: Earnings losses in simple model



Notes: Earnings losses from simple model. The horizontal axis shows years since the displacement event. The vertical axis shows earnings losses in percentage points. The black solid line shows benchmark with large share of stable jobs and heterogeneity in mobility rates. The dark gray dashed line shows first counterfactual without stable jobs and heterogeneity. The dark gray line with circles shows second counterfactual with large share of stable jobs but no heterogeneity in mobility rates. The light gray line with squares shows third counterfactual without share of stable jobs but with heterogeneity in mobility rates. The light gray line with stars shows fourth counterfactual without stability and heterogeneity and without job upgrading.

We look at four experiments to demonstrate that job stability generates persistent earnings losses while heterogeneity in worker mobility is necessary to account for high average worker mobility. In the first experiment (gray dashed line), we set separation rates uniformly to $\pi_g = \pi_b = 0.03$, neither accounting for stable jobs nor for heterogeneity in separations rates. Earnings losses are now small and transitory at 1.2% after six years. In the second case (dark gray line with circles), we set $\pi_g = \pi_b = 0.003$, removing heterogeneity but keeping job stability. Earnings losses remain large and persistent at 8.2% after six years. However, the model fails to account for high average worker mobility, a key feature of the data. In the third case (light gray line with squares), we keep heterogeneity in mobility rates but remove job stability of good jobs. We set $\pi_g = 0.015$ but keep the ratio $\frac{\pi_b}{\pi_g}$ as in the benchmark model ($\pi_b = 0.2$). In this case, the job ladder is initially quite slippery and it takes a long time to climb it up. However, earnings losses are again small and transitory at 2.2% after six years. Heterogeneity in transition rates alone is therefore not sufficient to get large and persistent earnings losses. Finally, in the fourth experiment (light gray line with stars), we set $\gamma = 0.001$, preventing the worker to climb up the ladder for an expected 83 years but let separation rates stay uniformly at $\pi_g = \pi_b = 0.03$. We can now investigate whether it is the upward friction and the resulting persistence of bad jobs that leads to large and persistent earnings losses, as often claimed in the literature. As the figure shows, earnings losses in this case are again small and transitory at 2.2% after six years.

The simple model demonstrates that a model of worker mobility that aims at explaining large and persistent earnings losses must also explain a large share of stable jobs; at the same time it also has to account for heterogeneity in worker mobility rates to match the observed high average worker mobility. A considerable upward friction is necessary to prevent workers to immediately regain their skills. However, it needs very stable jobs as well, preventing nondisplaced workers to become similar to displaced workers too quickly (little mean reversion from above). In section 3 of the main part of the paper, we present a micro-founded model of labor market behavior that accounts for all of these facts jointly.

II Model details and discussion

II.1 Discussion

The building blocks of our model follow in most part a large strand of the literature. This section discusses some of our modeling choices in more detail.

II.1.1 Finite life cycle

We depart from an infinite-horizon benchmark and explicitly account for age and a finite working life for three reasons. First, our empirical analysis highlights age as a driver of heterogeneity in worker mobility. Second, our empirical analysis documents that mean and median tenure increase almost linearly with age. A linear increase with age indicates that the data are inherently non-stationarity. We consider a finite working life as the most appealing and natural way to deal with this non-stationarity. Otherwise, combining heterogeneity in mobility rates and a large share of stable jobs in an infinite horizon model needs some other way to deal with the concentration of workers in the best jobs over time. Third, the life cycle naturally allows for a distinction between the accumulation of labor market experience and tenure on the job. We discuss in section 3.5 of the paper that this distinction contains information to determine the relative importance of worker- and match-specific skills.

II.1.2 Non-employment

We assume that the non-employment state comprises workers either in unemployment or not in the labor force (NILF) who are attached to the labor market. We consider this a convenient modeling tool that allows us to abstract from an additional job search decision in the model that distinguishes states of unemployment and NILF in the data. Two pieces of empirical evidence support this modeling choice. First, Kudlyak and Lange (2014) provide evidence that job finding rates of unemployed and NILF workers are almost identical if they have recent employment spells. Hence, for workers attached to the labor market, the abstraction from NILF is irrelevant. We discuss in section V.1 of this online appendix that the estimated model closely matches the evolution of unemployment rates over the life cycle once we remove the level difference resulting from the broader unemployment concept. Second, a large fraction of inflows to employment from NILF are labor market entrants, and therefore, flows that are exogenous to our model. Over the time period from 1980 to 2005, 23% of all inflows from out of the labor force to employment come from workers 20 and younger, the number rises to 39% if we consider workers 25 and younger. This suggests that a large fraction of these flows are labor market entrants that our model accounts for directly through its life-cycle structure.

II.1.3 Skill process

It is common at least since Becker (1962) to distinguish between worker- and match-specific skills.[48] Examples of worker-specific skills include the ability for general problem solving, social interaction with clients and colleagues, dealing with requests by foremen and clients, or a more efficient organization of the work flow. Examples of match-specific skills include working with technology, software, or products of the firm, the particular combination of tasks at a job, or leadership by foremen or senior colleagues.

One way to distinguish the two skill components through the lens of our model is by their accumulation process. Worker-specific skills are skills that are acquired by training or labor market experience; once they are lost they can be retrained. Match-specific skills are an inherent feature of a worker-firm relationship. They are lost once the worker changes jobs and require search to be regained. We refer to this distinction in the policy analysis of section 6.

In addition to the components of our skill process, some scholars allow for worker-specific skill depreciation during non-employment. Our skill process captures this effect, too, because only employed workers accumulate skills but non-employed workers do not. Hence, there is skill depreciation during non-employment as the average skill difference between employed and non-employed workers widens with non-employment duration. Explicitly allowing for skill depreciation during non-employment creates an asymmetry between on- and off-the-job search that makes it very attractive to accept any job and then keep on searching while employed. This would put the search technology on and off the job on different footings with respect to the skill process rather than with respect to the average contact rate, which we focus on and which has been shown to be empirically different (Faberman et al. (2016)). We discuss the effects of introducing additional skill depreciation when out of work in the model extension of section IV.1. For the estimated model, we find that additional skill depreciation during non-employment is rather small and affects relatively few workers due to the transitory nature of unemployment in the U.S.

Match-specific skills do not transfer to other jobs but mobility choices in the model will lead to skill dynamics where match-specific skills typically increase and only infrequently decrease. This pattern is similar in terms of outcomes to a model with industry- or occupation-specific skills

[48] Becker refers to those skills as general and specific human capital.

where workers typically stay within their industry or occupation to avoid skill losses (Parent (2000), Kambourov and Manovskii (2009b)). The match-specific component could also have a broader interpretation and capture characteristics of the match that increase the joint surplus relative to a fixed outside option. For example, it could include effects from monopoly rents or government subsidies. If rents are part of earnings losses following job displacement, they will show up as match-specific skill losses.

The match-specific skill component is the outcome of search in a frictional market. Over time, workers receive job offers and climb up the job ladder. High match-specific skill components characterize good jobs and constitute the top of the job ladder in our model. The improvement of job quality through labor market search and mobility has been found to be important for early career wage growth and high mobility rates at the beginning of workers' careers (Topel and Ward (1992)). We demonstrate that our model fits this evidence in section 4.3.2 in the main part of the paper.

Finally, regarding the distinction between worker- and match-specific skills Becker himself already acknowledges that it might not always be possible to clearly distinguish between the two.[49] It is easy to criticize some of the above examples as being not fully worker- or match-specific. In fact, our skill process captures this inherent uncertainty by making the transferability of accumulated skills risky. When switching jobs, workers do not know if skills transfer to the new job. Switching jobs entails the risk that some skills that have been thought of as being productive in all jobs are not. We are not the first to assume partial transferability of skills. Similar skill processes have been used in the literature using various headings, for example, in Ljungqvist and Sargent (1998) (turbulence), Jolivet et al. (2006) (reallocation shocks), or Violante (2002) (vintage-human capital).

II.1.4 Productivity and utility shocks

Our approach to model endogenous separations using productivity shocks is closely related to the endogenous separations model of den Haan et al. (2000a) who use transitory log-normally distributed shocks. With additive logistically distributed shocks, their basic mechanism remains unaffected and the outcomes of the two modeling approaches can be made very similar by

[49] Becker (1962) "Much on-the-job training is neither completely specific nor completely general [...] "(p.17).

recalibrating the underlying variances (see Jung and Kuhn (2014)).[50] The advantage of our distributional assumption is that it saves on the maximization step in the numerical solution routine because optimal choices have analytic expressions. The fact that productivity shocks are in some cases negative under this formulation is equivalent to negative productivity shocks in the setting with log-normally distributed shocks. In both cases, the realized output is smaller than the expected output given skills of the match. Exogenous separations happen independent of the skill state of the current match. One interpretation is that the underlying shock renders the match permanently unproductive so that all workers separate from the employer independent of the previous skill state. We discuss productivity shocks as source of transitory wage risk in section V.2.4 of this online appendix. We find that the estimated variance of transitory wage shocks aligns well with the empirical evidence.

The utility shocks to outside offers capture the possibility that job characteristics other than wages affect job mobility decisions in a tractable way. It captures, for example, job characteristics like distance from home, working arrangements, workplace atmosphere, or other amenities of the new job that in practice might affect job mobility decisions. In the limit as ψ_o approaches zero, the model nests the case without additional idiosyncratic job characteristics. The alternative limit as ψ_o approaches infinity considers the other extreme when wages play no role and idiosyncratic utility components alone govern acceptance. An intermediate value of ψ_o together with utility component κ_o parameterizes the relative importance of having a choice along a second dimension that captures the attractiveness of a job offer to an individual. We also discuss this choice and the change in the relative importance over the life cycle in section III on identification. There, we also discuss and exploit that utility shocks bound the elasticity of switching jobs conditional on having received a job offer. Without this second dimension, workers leave the current job whenever the outside offer is only slightly better. Utility shocks smooth this discontinuity. The assumption that the shock is a one-time shock and is i.i.d. is restrictive. The obvious alternative would be to replace it by a persistent nonpecuniary utility component. This adds an additional state variable and further complicates the model.

The fact that other dimensions than the wage govern mobility decisions has been suggested in a growing body of literature that documents the importance of nonpecuniary job characteristics

[50] Similar distributional assumptions are widely used in the literature that deals with discrete choice problems (cp. Rust (1987)) because they allow for a convenient closed form solution of the maximization choice, see Jung and Kuester (2011) for an application.

for mobility decisions, for example, Bonhomme and Jolivet (2009), Rupert et al. (2004), and Fujita (2011). These utility shocks help explain why many observed job-to-job switches involve wage cuts (Tjaden and Wellschmied (2014)). Alternative explanations for wage cuts at job-to-job transitions are occupation-specific skills as in Kambourov and Manovskii (2009a) and Kambourov and Manovskii (2009b) or a different bargaining protocol with wage increases over time as in Postel-Vinay and Robin (2002).

II.1.5 Directed search

On economic grounds the assumption implies that firms direct vacancies towards a particular worker type, for example, firms post vacancies for ”junior” or ”senior” positions, a pattern strongly supported by the data (Marinescu and Wolthoff (2016)). Experience can typically be observed during interviews and resumes can serve as indicators as well. Our setup can be interpreted as one where a position has zero productivity if a firm hires a worker of a different type than the one it is looking for so that there are no incentives for workers to search in other sub-markets. Sub-markets for workers on the job with a particular match type imply that in the data we should see that workers with the same experience level but lower wages receive more offers because they are in jobs of lower match quality. We are not aware of any evidence regarding this pattern but consider it a reasonable model prediction.

On technical grounds, the assumption of directed search makes the model computationally simpler because the cross-sectional distribution across worker- and match-types does not enter individual decisions and the age distribution does not enter as an additional aggregate state. A single search market would add a layer of complexity because the cross-sectional distribution would enter the vacancy posting decision.

We allow for different matching efficiencies on and off the job. We do not impose that either on- or off-the-job search is more efficient when we bring the model to the data. Examples that can cause differences in matching efficiency are potential network effects for job seekers, be it through colleagues, business contacts, or access to information on open positions at competitors, suppliers, or clients.

II.2 Details on derivation

In this section, we provide additional details and derivations for our model presented in section 3 of the main part of the paper.

II.2.1 Truncated expectation for the logistic distribution

We will repeatedly use properties of the logistic distribution. Here, we derive these properties for reference. Let H be a logistic distribution with mean μ_η and variance $\frac{\pi^2}{3}\psi_\eta^2$. Let ϖ be the cut-off value, so we can solve the truncated expectation as

$$\begin{aligned}
\int_{-\infty}^{\varpi} \eta\ h(\eta)\ \mathrm{d}\eta &= \Big[\eta\ H(\eta)\Big]_{-\infty}^{\varpi} - \int_{-\infty}^{\varpi} H(\eta)\ \mathrm{d}\eta \\
&= \Big[\eta\ H(\eta)\Big]_{-\infty}^{\varpi} - \left(1+\exp\left(-\frac{\eta-\mu_\eta}{\psi_\eta}\right)\right)^{-1}
\end{aligned}$$

Applying de l'Hôpital's rule, the first term simplifies to $\bar{\omega}H(\bar{\omega})$. For the integral, multiply the numerator and the denominator by $\exp\left(\psi_\eta^{-1}(\eta-\mu_\eta)\right)$. Define $y=\exp\left(\psi_\eta^{-1}(\eta-\mu_\eta)\right)$ which implies $\mathrm{d}\eta=\psi_\eta y^{-1}\ \mathrm{d}y$. Using this definition, the equation simplifies to

$$\begin{aligned}
\int_{-\infty}^{\bar{\omega}} \eta\ h(\eta)\ \mathrm{d}\eta &= \varpi\ H(\bar{\omega}) - \psi_\eta \int_{-\infty}^{\bar{\omega}} \frac{1}{1+y}\mathrm{d}y \\
&= \bar{\omega}\ H(\bar{\omega}) - \psi_\eta\Big[\log(1+y)\Big]_{-\infty}^{\bar{\omega}}
\end{aligned}$$

Re-substitution yields

$$\begin{aligned}
\int_{-\infty}^{\bar{\omega}} \eta\ h(\eta)\ \mathrm{d}\eta &= \bar{\omega}H(\bar{\omega}) - \psi_\eta\left[\log\left(1+\exp\left(\frac{\eta-\mu_\eta}{\psi_\eta}\right)\right)\right]_{-\infty}^{\bar{\omega}} \\
&= \bar{\omega}\ H(\bar{\omega}) + \psi_\eta \log(1-H(\bar{\omega}))
\end{aligned}$$

where the last step uses the fact that $\exp\left(\frac{\eta-\mu_\eta}{\psi_\eta}\right)=\frac{H(\eta)}{1-H(\eta)}$, which, evaluated at $\bar{\omega}$, can be solved for $\bar{\omega}=\psi_\eta\ (\log H(\bar{\omega})-\log(1-H(\bar{\omega})))+\mu_\eta$. Plugging the solution for $\bar{\omega}$ back into the solution of the integral, we finally arrive at

$$\int_{-\infty}^{\bar{\omega}} \eta \; h(\eta) \; \mathrm{d}\eta = \psi_\eta \; \Big(H(\bar{\omega}) \log(H(\bar{\omega})) + (1 - H(\bar{\omega})) \log(1 - H(\bar{\omega})) \Big) + H(\bar{\omega}) \; \mu_\eta$$

II.2.2 Bargaining Details

The value functions have been derived as

$$\begin{aligned}
J(x_w, x_m, a) \;\; &= \;\; (1-\pi_f)(1-\pi_s(x_w, x_m, a)) \Bigg(f(x_w, x_m) + \frac{\Psi_s(\pi_s)}{(1-\pi_s(x_w, x_m, a))} - w(x_w, x_m, a) \\
&+ \;\; (1-\pi_{eo}(x_w, x_m, a)) \beta \mathbb{E}_s \left[J(x_w', x_m, a') \right] \Bigg) && (13) \\
V_n(x_w, a) \;\; &= \;\; b + p_{ne}(x_w, a) \sum_{x_m'} \left(q_{ne}(x_w, x_m', a) \left(\beta \mathbb{E}_m \left[V_e(x_w', x_m', a') \right] - \kappa \right) \right) g(x_m') \\
&\quad + \sum_{x_m'} (1 - p_{ne}(x_w, a) q_{ne}(x_w, x_m', a)) \beta V_n(x_w, a') g(x_m') \\
&\quad + p_{ne}(x_w, a) \sum_{x_m'} \Psi_{ne}(q_{ne}) g(x_m') && (14) \\
V_e(x_w, x_m, a) \;\; &= \;\; (1-\pi_f)(1-\pi_s(x_w, x_m, a)) \left(w(x_w, x_m, a) + V_e^S(x_w, x_m, a) \right) \\
&\quad + ((1-\pi_f)\pi_s(x_w, x_m, a) + \pi_f) \, V_n(x_w, a) && (15)
\end{aligned}$$

with the value function at the search stage defined as

$$\begin{aligned}
V_e^S(x_w, x_m, a) \;\; &= \;\; \overbrace{p_{eo}(x, a) \sum_{x_m'} \Bigg(q_{eo}(x_m'; x, a) \left(\beta \mathbb{E}_m \left[V_e(x_w', x_m', a') \right] - \kappa_o \right) \Bigg) g(x_m')}^{\text{receiving and accepting offer}} \\
+ \underbrace{\sum_{x_m'} (1 - p_{eo}(x, a) q_{eo}(x_m'; x, a)) \beta \mathbb{E}_s \left[V_e(x_w', x_m, a') \right] g(x_m')}_{\text{not receiving or not accepting offer}} &+ p_{eo}(x, a) \underbrace{\sum_{x_m'} \Psi_{eo}(q_{eo}) g(x_m')}_{\text{option value}} . \qquad (16)
\end{aligned}$$

For later reference, we provide expressions of surplus definitions from the main text

$$\begin{aligned}
\Delta(x_w, x_m, a) &= V_e(x_w, x_m, a) - V_n(x_w, a) \\
&= (1-\pi_f)(1-\pi_s(x_w, x_m, a))\left(w(x_w, x_m, a) + V_e^S(x_w, x_m, a) - V_n(x_w, a)\right) \\
S(x_w, x_m, a) &= (V_e(x_w, x_m, a) - V_n(x_w, a) + J(x_w, x_m, a)) \\
&= (1-\pi_f)(1-\pi_s(x_w, x_m, a))\left(f(x_w, x_m) + \frac{\Psi_s(\pi_s)}{(1-\pi_s(x_w, x_m, a))} + \Delta^S(x_w, x_m, a)\right) \\
&\quad +(1-\pi_f)(1-\pi_s(x_w, x_m, a))\left((1-\pi_{eo}(x_w, x_m, a))\beta\mathbb{E}_s\left[J(x_w', x_m, a')\right]\right) \\
S^P(x, a) &= \Delta^P(x, a) + J^P(x, a) = f(x) + \Delta^S(x, a) + (1-\pi_{eo}(x, a))\beta\mathbb{E}_s[J(x', a')].
\end{aligned}$$

Next, we provide the details on the derivation. First, note that maximization with respect to wages delivers the classical formula that $\Delta(x_w, x_m, a) = \mu(J(x_w, x_m, a) + \Delta(x_w, x_m, a))$ and after rearranging

$$\begin{aligned}
&\mu(1-\pi_f)(1-\pi_s(x_w, x_m, a))\Big(f(x_w, x_m) - w(x_w, x_m, a)\Big) + \\
&\mu(1-\pi_f)(1-\pi_s(x_w, x_m, a))(1-\pi_{eo}(x_w, x_m, a))\beta\mathbb{E}_s\left[J(x_w', x_m, a')\right] + \\
&\mu(1-\pi_f)\Psi_s(x_w, x_m, a) \\
= \;&(1-\mu)(1-\pi_f)(1-\pi_s(x_w, x_m, a))\left(w(x_w, x_m, a) + V_e^S(x_w, x_m, a) - V_n(x_w, a)\right)
\end{aligned}$$

Hence, we get

$$\begin{aligned}
w(x_w, x_m, a) &= \mu\Big(f(x_w, x_m) + (1-\pi_{eo}(x_w, x_m, a))\beta\mathbb{E}_s\left[J(x_w', x_m, a')\right]\Big) \\
&\quad + \frac{\mu}{1-\pi_s(x_w, x_m, a)}\Psi_s(x_w, x_m, a) - (1-\mu)\left(V_e^S(x_w, x_m, a) - V_n(x_w, a)\right).
\end{aligned}$$

Using $S^P(x, a) - \Delta^S(x, a) = f(x_w, x_m) + (1-\pi_{eo}(x_w, x_m, a))\beta\mathbb{E}_s\left[J(x_w', x_m, a')\right]$ and

$$w(x_w, x_m, a) = \mu\Big(S^P(x, a) - \Delta^S(x, a)\Big) + \frac{\mu}{(1-\pi_s(x_w, x_m, a))}\Psi_s(x_w, x_m, a) - (1-\mu)\Delta^S(x, a)$$

we obtain

$$w(x_w, x_m, a) = \mu\left(S^P(x, a) + \frac{\Psi_s(\pi_s)}{1-\pi_s(x_w, x_m, a)}\right) - \Delta^S(x_w, x_m, a)$$

as claimed. Next, we use that maximization with respect to $\pi_s(x_w, x_m, a)$ delivers

$$
\begin{aligned}
&(1-\pi_f)\Big(f(x_w,x_m) - w(x_w,x_m,a) + (1-\pi_{eo}(x_w,x_m,a))\beta\mathbb{E}_s\left[J(x'_w,x_m,a')\right]\Big) \\
&-(1-\pi_f)\psi_s \log\left(\frac{(1-\pi_s(x_w,x_m,a))}{\pi_s(x_w,x_m,a)}\right) + (1-\pi_f)\left(w(x_w,x_m,a) + V_e^S(x_w,x_m,a) - V_n(x_w,a)\right) \\
&= \quad 0
\end{aligned}
$$

and, after rearranging and using that $S^P(x,a) = f(x) + \Delta^S(x,a) + (1-\pi_{eo}(x,a))\beta\mathbb{E}_s[J(x',a')]$, we obtain

$$
\pi_s(x_w,x_m,a) = \left(1 + \exp\left(\psi_s^{-1} S^P(x,a)\right)\right)^{-1}
$$

as claimed. Finally, maximization with respect to the acceptance probability for the different outside offers x'_m that we denote by $q_{eo}(x'_m; x_w, x_m, a)$ delivers

$$
\beta\mathbb{E}_s\left[J(x'_w,x_m,a')\right] + \beta\mathbb{E}_s\left[V_e(x'_w,x_m,a')\right] - \beta\mathbb{E}_m\left[V_e(x'_w,x'_m,a')\right] + \kappa_o = \psi_{eo}\log\left(\frac{1-q_{eo}(x'_m;x_w,x_m,a)}{q_{eo}(x'_m;x_w,x_m,a)}\right)
$$

and after rearranging

$$
\begin{aligned}
&q_{eo}(x'_m;x_w,x_m,a)) \quad = \\
&\left(1+\exp\left(\psi_o^{-1}\Big(\beta\mathbb{E}_s\left[J(x'_w,x_m,a')\right] + \beta\mathbb{E}_s\left[V_e(x'_w,x_m,a')\right] - \beta\mathbb{E}_m\left[V_e(x'_w,x'_m,a')\right] + \kappa_o\Big)\right)\right)^{-1}
\end{aligned}
$$

Recalling that

$$
\begin{aligned}
&\beta\mathbb{E}_s\left[J(x'_w,x_m,a')\right] + \beta\mathbb{E}_s\left[V_e(x'_w,x_m,a')\right] - \beta\mathbb{E}_m\left[V_e(x'_w,x'_m,a')\right] + \kappa_o \\
&\quad = \beta\mathbb{E}_s\left[J(x'_w,x_m,a')\right] + \beta\mathbb{E}_s\left[V_e(x'_w,x_m,a') - \beta V_n(x'_w,a')\right] - \beta\mathbb{E}_m\left[V_e(x'_w,x'_m,a') - V_n(x'_w,a')\right] + \kappa_o \\
&\quad = \beta\mathbb{E}_s S(x'_w,x_m,a') - \mathbb{E}_m\Delta(x'_w,x_m,a') + \kappa_o
\end{aligned}
$$

Using $S^S(x,a) = \mathbb{E}_s[S(x'_w,x_m,a')] - \mathbb{E}_m[\Delta(x'_w,x'_m,a')]$ we obtain

$$
q_{eo}(x'_m;x_w,x_m,a)) = \left(1+\exp\left(\psi_o^{-1}\Big(S^S(x,a)+\kappa_o\Big)\right)\right)^{-1}
$$

as claimed.

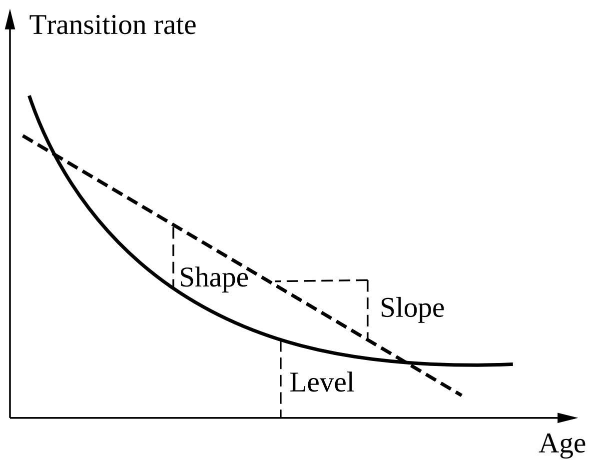
III Identification

This section discusses first the identification of the remaining model parameters (section III.1). In addition, we provide some additional discussion on the identification of the skill process described in the main part of the paper (section III.2). Finally, we provide some short discussion on the functional form of the production function and how it affects estimated parameters (section III.3).

III.1 Identification of remaining model parameters

This section provides a discussion of identification of the remaining model parameters. We abstain from a formal proof of identification; instead, we provide a combination of formal and intuitive arguments for parameter identification. To ease this discussion, we show in figure B a stylized profile of transition rates and highlight level, slope, and shape as three characteristics of the transition rate profile that we will exploit for our identification arguments. The meaning of the level is obvious. The slope captures the age differences, while the shape captures how quickly the age differences materialize. We will also refer to differences in transition rates of low- and high-tenure workers in our discussion, although tenure profiles are not explicitly used in the estimation. For this, it is important to note that differences between newly hired workers and the average worker at each age reflect differences between low- and high-tenure workers conditional on age. Hence, information about tenure profiles can be inferred from differences of the two age profiles. A key advantage of only using data moments conditional on age is that it makes our estimation independent of the age distribution and its changes over time.

Figure B: Identification



In the main text, we discuss how the slopes of newly hired age profiles identify the parameters of the worker skill process, p_u and p_d, given the dispersion of match-specific skills σ_m. Next, we explain how the dispersion of match-specific skills σ_m, the dispersion of idiosyncratic productivity costs ψ_s, and the outside option b are identified exploiting level differences in separation rates between low- and high-tenure workers, the level of the separation rate, and mean tenure. To map these targets back onto the age profiles used in the estimation, note that the slope of the separation rate age profile in combination with the slope of the newly hired age profile provide the differences in separation rates between low- and high-tenure workers. The level of the separation rate age profile gives the average separation rate and the age profile of mean tenure provides the mean tenure level. Note that despite the stock-flow relationship of tenure and transition rates, the level of mean tenure remains an independent moment as long as we do not target the profile of separation rates by tenure.

For our argument, we abstract from age differences and on-the-job search so that jobs only differ in their match component x_m. Contact rates are exogenous and non-employed workers accept all offered jobs so that the distribution across newly hired workers coincides with the offer distribution. We maintain the distributional assumptions from the main part of the paper for match-specific skills x_m and productivity shocks η_s. We consider a case with three skill types x_m^L, x_m^M, and x_m^H. We set skill types $x_m^L = x_m^M - \sigma_m$ and $x_m^H = x_m^M + \sigma_m$ to approximate the normal distribution and normalize $x_m^M = 1$.[51] Denote the separation rates of the different skill types by $\pi_L = \pi_s(x_m^L)$, $\pi_M = \pi_s(x_m^M)$, and $\pi_H = \pi_s(x_m^H)$. It can be shown that the three separation rates together with the three parameters σ_m, ψ_s, and b constitute a non-linear system with three equations and three unknowns that has a unique solution if $\pi_L > \pi_M > \pi_H > 0$.[52] What remains to be shown is that mean separation rate, mean tenure, and separation rate difference between a newly hired (low tenure) and the average worker (high tenure) provide the

[51] There always exist probabilities so that mean and variance are matched. The three skill types are sufficient given the assumption of a normal distribution because the first two moments characterize the distribution.

[52] The non-linear system of equations is

$$\pi_s(x_m^j) = \left(1 + \exp\left(\psi_s^{-1} S(x_m^j)\right)\right)^{-1} \qquad j = \{L, M, H\}$$

with

$$\begin{aligned} S(x_m^j) &= x_m^j - b + \beta(1 - \pi_s(x_m^j))S(x_m^j) + \Psi(\pi_s(x_m^j)) - \beta\pi_{ne}\mu\mathbb{E}_0[S(x_m)] \\ \Psi(\pi_s(x_m^j)) &= -\psi_s\Big(\pi_s(x_m^j)\log(\pi_s(x_m^j)) + (1 - \pi_s(x_m^j))\log(1 - \pi_s(x_m^j))\Big) \end{aligned}$$

with $\mathbb{E}_0[S(x_m)]$ being the expected value of employment for a newly hired worker. All other parameters are given.

necessary information to pin down the underlying separation rates π_L, π_M, and π_H. Denote the distribution of separation rates (skill types) of newly hired workers by $g(\pi)$ and the distribution over all separation rates (skill types) by $h(\pi)$, then it holds that $h(\pi_j) = \frac{N}{\pi_j} g(\pi_j)$ for $j = \{L, M, H\}$ with $N = \left(\frac{g(\pi_L)}{\pi_L} + \frac{g(\pi_M)}{\pi_M} + \frac{g(\pi_H)}{\pi_H}\right)^{-1}$. The mean separation rate $\bar{\pi}_s$ is

$$\bar{\pi}_s = \pi_L h(\pi_L) + \pi_M h(\pi_M) + \pi_H h(\pi_H) = N \left(g(\pi_L) + g(\pi_M) + g(\pi_H)\right) = N.$$

Mean tenure $\bar{T}$ is

$$\begin{aligned} \bar{T} &= \pi_L^{-1} h(\pi_L) + \pi_M^{-1} h(\pi_M) + \pi_H^{-1} h(\pi_H) = N \left(\pi_L^{-2} g(\pi_L) + \pi_M^{-2} g(\pi_M) + \pi_H^{-2} g(\pi_H)\right) \\ &= \bar{\pi}_s \left(\pi_L^{-2} g(\pi_L) + \pi_M^{-2} g(\pi_M) + \pi_H^{-2} g(\pi_H)\right). \end{aligned}$$

The separation rate of newly hired workers is $\pi_s^{NH} = \pi_L g(\pi_L) + \pi_M g(\pi_M) + \pi_H g(\pi_H)$. Denote the vector of the stacked three moments by $m = [\bar{\pi}_s, \bar{T}, \pi_s^{NH}]$. We get

$$\frac{\partial m}{\partial \pi_j} = -g(\pi_j) \left[(N\pi_j)^{-2}, \pi_j^{-4}\left(N^{-2} g(\pi_j) + \pi_j N\right), -1\right] \qquad j = \{L, M, H\}.$$

The matrix of derivatives $\left[\frac{\partial m}{\partial \pi_L}, \frac{\partial m}{\partial \pi_M}, \frac{\partial m}{\partial \pi_H}\right]$ has full rank so that there is a unique solution for $[\pi_L, \pi_M, \pi_H]$ given a vector of moments $m = [\bar{\pi}_s, \bar{T}, \pi_s^{NH}]$. Given the unique mapping from separation rates to parameters, it follows that there is also a mapping from these data moments to parameters $[\sigma_m, \psi_s, b]$. This completes the argument.

The last parameter of the skill process that needs to be discussed is δ. The parameter governs how quickly workers accumulate worker-specific skills by age, and therefore, how quickly age differences materialize. Hence, the shape of the separation rate profile identifies the speed of skill accumulation δ.

Next, we turn to the three parameters κ_o ,ψ_o, and $\varkappa_o$ governing on-the-job search. The likelihood of receiving a competing offer $p_{eo}(x_w, x_m, a)$ is, on average, determined by the efficiency of the matching function $\varkappa_o$ so that the level of the job-to-job transition rate identifies $\varkappa_o$. Recall that a worker of age a with worker-specific skill level x_w who is currently in a match of type x_m accepts a competing job offer with match-specific component x'_m with probability

$q_{eo}(x'_m; x_w, x_m, a)$ given by

$$q_{eo}(x'_m; x_w, x_m, a) = \left(1 + \exp\left(\psi_o^{-1}(\mathbb{E}_s[S(x'_w, x_m, a')] - \mu\mathbb{E}_m[S(x'_w, x'_m, a')] + \kappa_o)\right)\right)^{-1}.$$

This acceptance decision depends on the surplus difference $\mathbb{E}_s[S(x'_w, x_m, a')] - \mu\mathbb{E}_m S[(x'_w, x'_m, a')]$ and the utility component κ_o. The surplus part captures productivity-related aspects of the acceptance decision and the utility component κ_o captures the non-pecuniary aspects of the acceptance decision. Over the life cycle, workers move to better and better matches ("climbing the job ladder") so that older workers are on average in more productive matches than younger workers. This leads to lower acceptance rates from a smaller surplus component and to a declining job-to-job transition rate by age. The size of the decline, the slope of the job-to-job transition rate profile, depends on the importance of nonpecuniary aspects of job search. If nonpecuniary aspects are important, the relative importance of job quality declines and job-to-job transition rates decline less by age; the slope of the job-to-job transition rate profile becomes smaller. Hence, for a given skill process, the slope of the job-to-job profile identifies the importance of nonpecuniary aspects κ_o. It is important to note that κ_o enters relative to ψ_o so that only the ratio is identified by this argument. To disentangle the levels of κ_o and ψ_o, it is informative to look at the elasticity of the acceptance decision with respect to ψ_o

$$\frac{\partial q_{eo}}{\partial \psi_o}\frac{\psi_o}{q_{eo}} = \frac{(1-q_{eo})}{\psi_o}\bigg(\mathbb{E}_s[S(x'_w, x_m, a')] - \mu\mathbb{E}_m S[(x'_w, x'_m, a')] + \kappa_o\bigg).$$

A large variance of utility shocks (large ψ_o) implies a small elasticity of the acceptance decision to surplus differences so that workers do not react much to competing outside offers. In the limiting case of $\psi_o \to \infty$, acceptance probabilities converge to 0.5 and are independent of surplus differences. By contrast, a small ψ_o implies that a small surplus difference offered by a competing firm makes it very likely that the worker leaves the current employer. Hence after fixing the skill process and the importance of nonpecuniary aspects of job search, ψ_o governs the speed at which workers climb the job ladder and it is identified by the shape of the job-to-job transition rate profile.

Looking at the expression for the acceptance rate q_{eo} again, we also see that the worker's bargaining power μ enters this decision multiplying the surplus of the outside offer. This distinguishes job-to-job mobility decisions from separation decisions. Separation decisions only depend on the total surplus in the current match and are independent of the surplus split governed by μ (see eq. (6)). The reason the bargaining power is part of the job-to-job mobility decision is that it determines the share of the surplus of the new match that the worker will receive. If a large part of the surplus of the new job accrues to the worker (high μ), surplus differences weight strongly in the acceptance decision. This implies that a high bargaining power leads to strong incentives for newly hired workers to climb the job ladder. The higher the bargaining power, the larger the participation in the productivity increase and the more attractive becomes job switching, and hence, the more quickly newly hired workers want to climb the job ladder. Given the skill process and the distribution of nonpecuniary shocks, the bargaining power is identified by the shape of the job-to-job age profile of newly hired workers relative to the shape of the job-to-job age profile of the average worker. Put differently, the shape of the tenure profile of job-to-job transitions identifies the bargaining power μ. This information on the shape of the tenure profile of job-to-job transitions is traced out by the different shapes of the age and newly hired age profiles of job-to-job transitions. The fact that job-to-job mobility identifies the worker's bargaining power is also exploited in Bagger et al. (2014) using wage changes.

Parameters related to job search in non-employment, κ and $\varkappa_n$, are identified by level and slope of the job finding rate profile. As for on-the-job search, the efficiency of the matching function $\varkappa_n$ determines the level of the job finding rate. The slope of the job finding rate identifies vacancy posting costs κ. When workers accumulate worker-specific skills by age, the match surplus varies with age so that the costs to post a vacancy κ as share of the surplus vary with age. The age differences in job finding rates, the slope of the job-finding rate, therefore identify vacancy posting costs κ.

Finally, the exogenous separation rate π_f constitutes a counteracting force to tenure accumulation. It governs the right tail of the tenure distribution and thereby limits the slope of the mean tenure profile. As a consequence, the slope of the mean tenure profile identifies the exogenous separation rate π_f.

III.2 Further discussion for identification of skill process parameters

The main part of the paper discusses identification of the parameters of the skill process. These identification arguments assume that all newly hired workers come from non-employment but this was for illustration purpose only. It is important for our identification that some newly hired workers have been in non-employment before. If not all newly hired workers come from non-employment, the argument still applies in relative terms and the decline in transition rates of newly hired workers is a convex combination of skill accumulation and a selection effect due to a fraction of newly hired workers from other employers. newly hired workers coming from other employers will on average be in better matches than workers coming from non-employment. The selection effect would be weaker but still present, so that in relative terms the newly hired age profile is less affected by selection than the average age profile. In the data about 60% of newly hired workers come from non-employment so that we expect the effect to be strong enough for our identification argument to be valid. Importantly, our argument does not rely on the fact that the newly hired age profile captures a pure experience effect, as for example in Topel (1991), but only on the fact that the experience effect is stronger for transition rates of newly hired workers. In the model, transition rates for newly hired workers will also be composed of an experience effect and a selection effect due to job-to-job transitions. This is not the case in Topel's (1991) two-step estimation approach. Topel uses the point estimate from the first-step as an estimate of accumulated worker-specific skills. He discusses that if there is an increasing correlation between worker- and match-specific skills with age, then his results provide a lower bound on the returns to tenure. Dustmann and Meghir (2005) discuss this problem and only use workers from displaced firms when estimating the returns to tenure to avoid a correlation between worker and match types.

III.3 Functional form of the production function and identification

Finally, we discuss the importance of the functional form of the production function for identification. For our estimation, we assume that the production function is age-independent and log-linear in skills $f(x) = \exp(x_m + x_w)$ as in Bagger et al. (2014). We do not identify the shape of the production function. The assumed production function has strictly positive cross-partial derivatives which induces positive assortative matching. As discussed in section 3, mobility and wage dynamics in the model are surplus-driven. A positive cross-partial derivative adjusts the distance between different jobs for workers of different types. As long as the dispersion of skills and of productivity and utility shocks can adjust during the estimation process, the cross-partial derivative will mainly adjust parameter estimates but will not affect the general mechanism that we concentrate on in this paper. The general mechanism only relies on endogenous mobility decisions and that wages and job stability are inversely related.

IV Model extensions

This section describes two extensions of the baseline model from section 3 in the main part of the paper. The first extension accounts for duration dependence in job finding rates by introducing skill depreciation during unemployment and a deterioration in search efficiency. The second extension accounts for the right tail of the wage distribution to match the life-cycle increase in wage inequality. In both cases, we first describe the changes of the model for the extended version and demonstrate afterwards how results on earnings losses and their decomposition are affected by these extensions.

IV.1 Duration dependence

The extension described in this section is designed to explain the extent of duration dependence in job finding rates observed in the data. We introduce two additional channels to generate declining job finding rates with non-employment duration. The first channel is skill depreciation during non-employment. The second is a depreciation of search efficiency with non-employment duration. The literature on negative duration dependence did not yet settle on the relative importance of different channels through which negative duration dependence might arise. By allowing for two different sources, we remain agnostic about the importance of skill loss as a source of duration dependence.

In the baseline model, worker-specific skills x_w stay constant during non-employment.[53] In the extended model, we introduce stochastic skill depreciation of worker-specific skills if the worker stays non-employment from one period to the next. The distribution over worker skills

[53]See section II.1.3 of this online appendix for a detailed discussion.

x'_w in the next period if staying non-employed during this period is

$$x'_w = \begin{cases} x_w^- & \text{with probability } p_\delta \\ x_w & \text{with probability } 1 - p_\delta \end{cases}$$

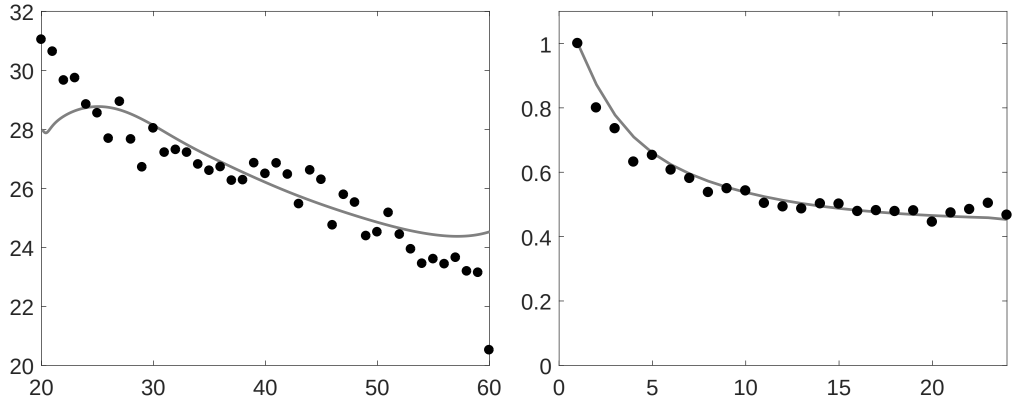
and we set $p_\delta = 0$ if $x_w = x_w^{min}$. In addition, we allow for a depreciation of search efficiency for long-term non-employed workers. This depreciation of search efficiency captures the loss of professional contacts used for job search but also discrimination of employers (Kroft et al. (2013)). We model this depreciation as occurring with probability p_s and resulting in a search efficiency $s < 1$. The lower search efficiency implies that the contact rate p_{ne} is only a fraction s of the short-term non-employed. To preserve tractability of the model, we model long-term non-employment as an additional state that employers can condition upon when posting vacancies.

When estimating the model, we use the negative duration dependence of job finding rates as additional target. We associate long-term non-employment with 52 weeks and set p_s to $\frac{1}{12}$. This choice is motivated by the empirical observation of disappearing duration dependence at longer durations of unemployment as discussed below. We include the parameters p_δ and s as additional parameters in the estimation. We add the profile of duration dependence during non-employment as additional target in the estimation.[54] We estimate $p_\delta = 0.0285$ and $s = 0.4766$. Figure C shows the age profile of job finding rates and the negative duration dependence in the extended model. The extended model with its additional parameters matches the negative duration dependence almost exactly. The data shows that job finding rates decline by roughly 50% over the first year in unemployment and that they stay relatively constant afterwards. This pattern matches typical patterns found in the literature; see for example, Kroft et al. (2013). The age profile is matched equally well in comparison to the baseline model. We provide the model fit to the other mobility profiles in Figure D. The figure shows that the extended model fits the life-cycle mobility profiles and the mean and median tenure profile equally well as our baseline model.

Figure E shows the estimated earnings losses and their decomposition in the extended model. Both figures are qualitatively and quantitatively very similar to the results reported for the baseline model in section 5. The decomposition of long-run earnings losses attributes a slightly

[54] We put additional weight on matching this profile in the estimation.

Figure C: Job-finding rates by age and duration dependence of job finding rates for extended model (section IV.1)

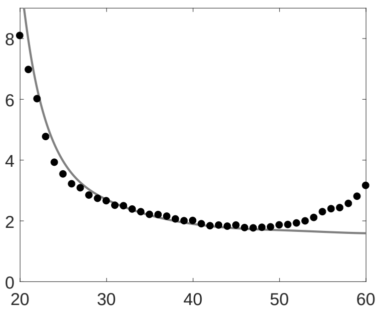


(a) Age profile (b) Duration profile

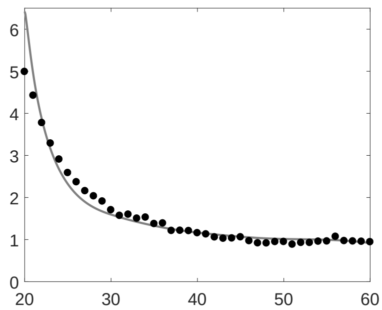
Notes: Job-finding rates by age and duration dependence of job-finding rates in the extended model. Left panel: Job-finding rates by age. The black dots show the data and the gray solid line shows the model. The horizontal axis shows age in years and the vertical axis shows job-finding rates in percentage points. Right panel: Duration dependence of job-finding rates. The black dots show the data and the gray solid line shows the model. Vertical axis shows duration job-finding rates relative to job-finding rates in the first month of unemployment. The horizontal axis shows unemployment duration in months.

larger component to the wage loss effect (59%) and slightly lower components to the extensive margin effect (16%) and the selection effect (24%). In the baseline model, the corresponding contribution shares are 48%, 21%, and 31%.

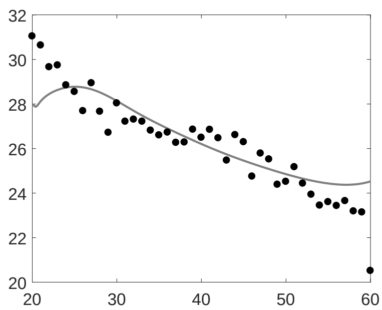
Figure D: Model prediction and data



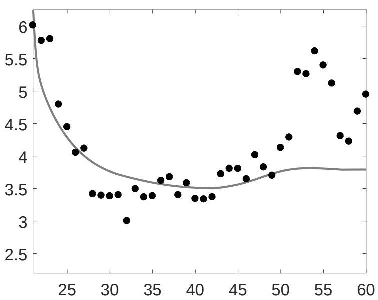
(a) Separation rate age profile



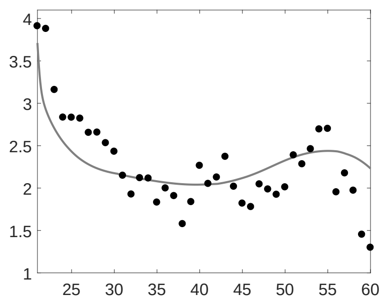
(b) Job-to-job rate age profile



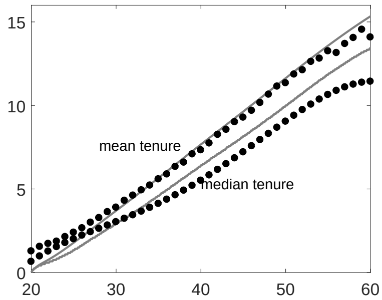
(c) Job finding rate age profile



(d) Separation rate age-tenure profile



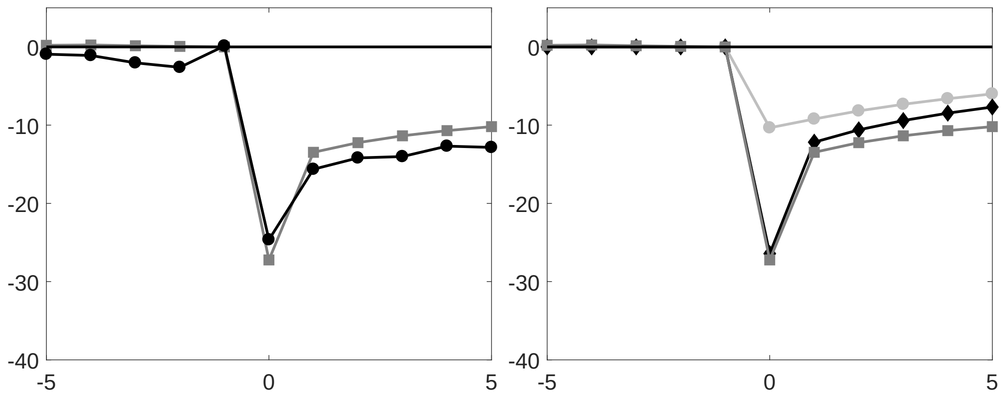
(e) Job-to-job rate age-tenure profile



(f) Mean tenure age profile

Notes: Age, age-tenure, and tenure profiles from the extended model and the data. The black dots show the data and the gray solid line the model. The horizontal axis is age in years and the vertical axis shows transition rates in percentage points or tenure in years.

Figure E: Earnings losses following displacement and decomposition of earnings losses for extended model (section IV.1)



(a) Earnings losses (b) Decomposition of earnings losses

Notes: Left panel: Earnings losses after displacement in the extended model and empirical estimates. Gray line with squares shows earnings losses predicted by the model and the black line with circles shoes the estimates by Couch and Placzek (2010). The horizontal line shows years relative to the displacement event and the vertical axis shows losses in percentage points relative to the control group. Right panel: The gray line with squares show earnings losses relative to the control group. The black line with diamonds show earnings losses relative to a control group without additional selection criteria (twin group). The light gray line line with circles show wage losses for employed workers relative to a control group without additional selection criteria (twin group). The horizontal line shows years relative to the displacement event and the vertical axis shows losses in percentage points relative to the control group.

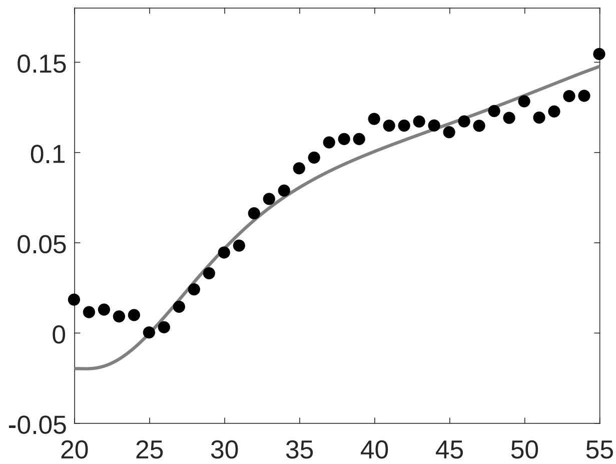
IV.2 Life-cycle inequality

The extension described in this section is designed to match the entire life-cycle profile of the variance of log wages. For the extended model, we augment the worker-specific skill set by an additional state in the right tail of the skill distribution $x_w^{new} > x_w^{max}$ where x_w^{new} denotes the additional state and x_w^{max} the maximum state of the skill set from the baseline model.[55] The probability that the new skill state is reached through stochastic skill accumulation p_u^{new} differs from p_u. We set p_u^{new} so that at age 20 the probability of reaching state x_w^{new} from the next lower state is 1% per year. The probability p_u^{new} changes with age following the same recursion as p_u. We estimate the extended model and include the new skill state as an additional parameter to be estimated. We add the change of the life-cycle profile of the variance of log wages relative to age 25 as an additional target to the estimation. We find that the additional state x_w^{new} has a productivity 144 log points above the highest skill state that we have in the baseline model. The difference between the highest skill state from the baseline model and the new skill state

[55]The skill process has to be changed accordingly, with x_w^{new} being the maximum element of the worker-specific skill set of the extended model.

exceed the difference between the remaining skill states by roughly 100 log points capturing the very long right tail of the wage distribution. We find that in the cross-sectional distribution of the estimated model, 7% of workers have reached this skill state by age 40. Figure F shows the variance of log wages of the extended model in comparison to the empirical profile based on CPS data. We provide details on the construction of the empirical variance profile in section A of the appendix to the main part of the paper.

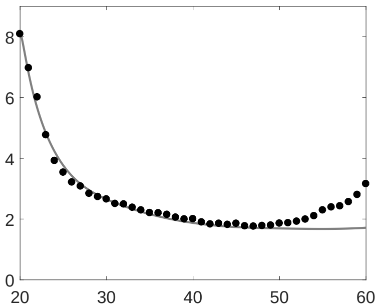
Figure F: Model prediction and data for variance of log wages from extended model (section IV.2)



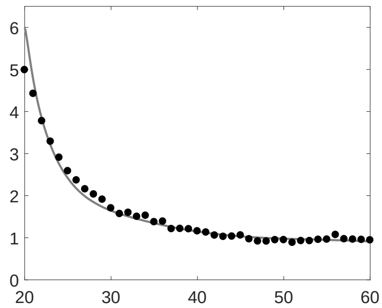
Notes: Life-cycle profile of variance of log wages from CPS data (black dots) and model prediction for extended model (gray solid line). Both profiles have been normalized to zero at age 25. See text for details on construction of life-cycle variance profiles.

Figure G shows that the extended model fits the life-cycle mobility profiles and the mean and median tenure profile equally well as our baseline model. Figure H shows earnings losses and decomposition of earnings losses for the extended model. The estimated earnings losses and decomposition of earnings losses show only small differences to the results reported for the baseline model in section 5. Earnings losses in the sixth year after displacement are 11.1% relative to 10.8% in the baseline model and the decomposition of the earnings losses is 55%, 17%, and 28% for the wage loss effect, extensive margin effect, and the selection effect. We get a decomposition with 48%, 21%, and 31% in the baseline model.

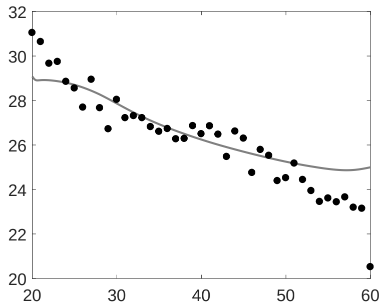
Figure G: Model prediction and data



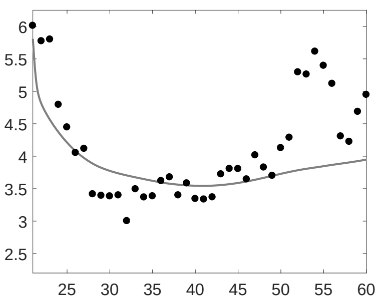
(a) Separation rate age profile



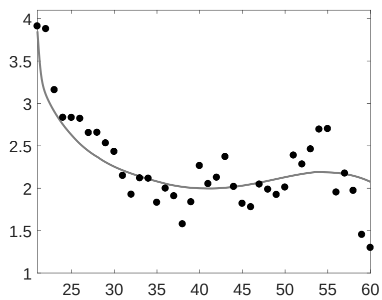
(b) Job-to-job rate age profile



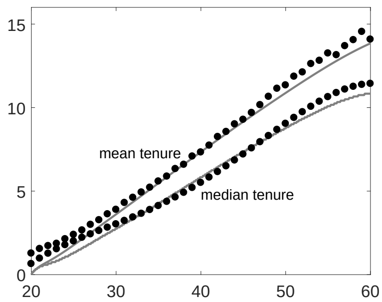
(c) Job finding rate age profile



(d) Separation rate age-tenure profile



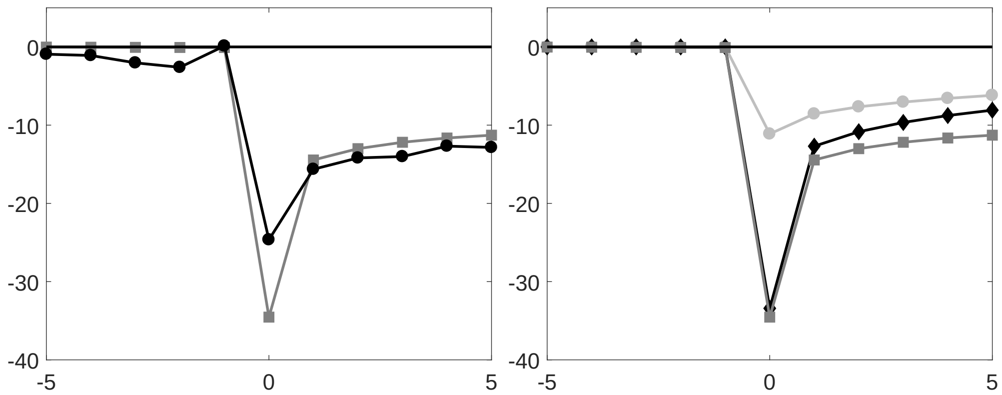
(e) Job-to-job rate age-tenure profile



(f) Mean tenure age profile

Notes: Age, age-tenure, and tenure profiles from the extended model and the data. The black dots show the data and the gray solid line the model. The horizontal axis is age in years and the vertical axis shows transition rates in percentage points or tenure in years.

Figure H: Earnings losses following displacement and decomposition of earnings losses from extended model (section IV.2)



(a) Earnings losses (b) Decomposition of earnings losses

Notes: Left panel: Earnings losses after displacement in the extended model and empirical estimates. The gray line with squares shows earnings losses predicted by the model and the black line with circles are estimates by Couch and Placzek (2010). The horizontal axis shows years relative to the displacement event and the vertical axis shows earnings losses in percentage points relative to the control group. Right panel: The gray line with squares are earnings losses relative to the control group. The black line with diamonds are earnings relative to a control group without additional selection criteria (twin group). The light gray line with circles are wage losses for employed workers relative to a control group without additional selection criteria (twin group). The horizontal axis shows years relative to the displacement event and the vertical axis shows losses in percentage points relative to the control group.

V Non-employment profiles and wage dynamics

This section provides details of the mapping of non-employment from the model to the data and the construction of estimators of wage dynamics in the model.

V.1 Non-employment profiles

In the main part of the paper, we compare the age profile of the non-employment rate from the model to the age profile of the unemployment rate from CPS data. Non-employment in the model comprises all unemployed workers and some workers who are not classified as unemployed in the CPS but who are attached to the labor market. Recent evidence in Kudlyak and Lange (2014) supports this modeling choice that we discuss in detail in section II.1.2 of this online appendix. Including some non-employed workers who are attached to the labor market leads to a level difference between the unemployment rate in the data and the non-employment rate in the model. We remove this level difference by multiplying the model non-employment rate by a constant so that the mean non-employment rate from the model matches the mean unemployment rate from the data for the age range from 21 to 55. We do not adjust slope or shape of the model profile. The resulting adjustment factor implies that there is roughly one additional non-employed worker for every two unemployed workers. The adjustment almost exactly matches the difference between the Hornstein-Kudlyak-Lange non-employment index and the official unemployment rate.[56] The average ratio of the Hornstein-Kudlyak-Lange non-employment index to the BLS unemployment rate from 1994 to today is 1.57. The average ratio of the non-employment rate in the model to the empirical unemployment rate is also 1.57. We consider the coinciding level differences as further evidence that our model closely matches the concept of unemployment based on observed worker flow rates and labor force attachment of the Hornstein-Kudlyak-Lange non-employment index.

V.2 Wage dynamics

In the main part of the paper, we discuss wage dynamics from the model and compare them to the empirical evidence. Here, we provide details on how we derive the wage dynamics using model data. Readers are referred to the literature for further details of the estimation and discussion on the estimation methods.

V.2.1 Wage gains from job-to-job transitions

This section explains how we compute wage gains from job-to-job transitions discussed in section 4.3.1. We compute wage gains from job-to-job transitions using the conditional distribution functions from the model. For each job-to-job transition, we compute the expected wage conditional on the current state x by taking into account offer probabilities $g(x_m)$, acceptance probabilities $q_{eo}(x'_m; x_w, x_m, a)$, and skill transitions $\mathbb{E}_m[\cdot]$. This yields $\mathbb{E}_{j2j}[w|x_w, x_m, a]$, where we use subscript $j2j$ to indicate that we condition on a job-to-job transition taking place. We compute wage growth $\frac{\mathbb{E}_{j2j}[w|x_w,x_m,a]}{w(x_w,x_m,a)}$ as expected wage after realized job-to-job transitions relative to current wage $w(x_w, x_m, a)$. We average across worker types by age using weights implied by the transition probabilities $\pi_{eo}(x_w, x_m, a)$. Recall, that transition probabilities $\pi_{eo}(x_w, x_m, a)$ also depend on the probability of receiving an offer $p_{eo}(x_w, x_m, a)$.

[56] See https://www.richmondfed.org/research/national_economy/non_employment_index for further details on the Hornstein-Kudlyak-Lange non-employment index.

V.2.2 Early career wage growth

This section explains how we derive the contribution of job changing to early career wage growth from the model as discussed in section 4.3.2. The estimation of the contribution of job changing to early career wage growth requires path dependent information over long time intervals so that we resort to model simulation. We simulate a cross-section of 10,000 workers from the model and track their employment and wage histories. We aggregate data to quarterly frequency to be consistent with the data used in Topel and Ward (1992). We compute wage growth in the first 10 years in the labor market as the log difference in wages. We compute the wage growth due to job changing activity as the sum of wage gains due to job changes over the same period. We follow Topel and Ward (1992) and determine the wage gain from a job change as

$$log(w_{a+1}) - log(w_{a-2}) - d\hat{w}_{a+1} - d\hat{w}_{a-1}$$

where a denotes age in quarters and $d\hat{w}_a$ denotes the predicted quarterly wage growth from age a to $a+1$ based on an independent regression of job stayers. For the wage growth regression for job stayers, we follow Topel and Ward (1992) in the choice of controls and include potential experience, tenure, completed tenure of the job spell, and a job change indicator that is 1 for the last year on a job. We include higher order terms for tenure and experience as in Topel and Ward (1992) (Table VI, row 5). We restrict the sample to be consistent with the data used in Topel and Ward (1992). We only use observations of job stayers who are age 33 and younger with at least two quarters of tenure at the first wage observation. For further details on the estimation or on the derivation of wage gains see Topel and Ward (1992).

V.2.3 Returns to tenure

This section explains how we estimate the returns to tenure in the model discussed in section 4.3.3. We use a simulated sample of 10,000 workers from the model. We follow the instrumental variable approach in Altonji and Shakotko (1987) and the two-step approach in Topel (1991) to estimate returns to tenure. To make the data consistent, we drop unemployment spells from the sample, employment spells that last less than 3 months, and all workers with more than 45 jobs. We choose the 45 job threshold to match average tenure of 7.7 years in Altonji and

Shakotko (Table 1) within our sample.[57] The data aligns closely with the other unconditional means reported in Altonji and Shakotko (Table 1). We aggregate employment histories to annual frequency and use average wages as a measure for the annual wage. This approach is equivalent to keeping unemployment spells in the sample but average wages over employment spells only. Both approaches correspond to the empirical approach of dividing annual income by hours worked. We construct instrumental control variables as in Altonji and Shakotko by constructing within-spell deviations. We also include an indicator variable for the first year on the job. When we run the OLS regression, we use the indicator variable for the first year on the job, experience, and tenure terms as in Altonji and Shakotko (1987). We follow their assignment of wage observations to controls and use tenure lagged by one period.

For the two-step estimator in Topel (1991), we use the same simulated data from the model. We run a first-stage regression on wage growth using the same experience and tenure controls as in Topel (1991). We follow Topel and assign wage observations to controls from the current period. Accordingly, we restrict the sample to spells with more than one year of tenure. We construct initial wages on the job spell by subtracting predicted wage growth and construct initial experience by subtracting accumulated tenure. We run the linear regression of initial wages on initial experience to derive the linear experience effect (β_1) as in Topel (Table 3).

In both cases, we construct the returns to 10 years of accumulated tenure using the point estimates from the regressions on our sample and compare it to the predictions using the reported point estimates from Altonji and Shakotko (1987) (Table 1 columns 2 for OLS and 4 for IV) and Topel (1991) (Table 2 model 3 for "experience effect", Table 3 "tenure effect").

V.2.4 Permanent income shocks and wage inequality

This section is divided into three parts; the first part explains how we estimate the variance of permanent income shocks discussed in section 4.3.4 from the model. The second part discusses how we derive an estimate for transitory wage shocks from the model, and the third part discusses wage inequality in the model and the measurement in the data.

We derive wage residuals in the model by subtracting age-specific mean (log) wages. We denote the residual for a worker of type x at age a by $\hat{w}(x, a)$. As discussed in the main part of the paper, we use the estimation to describe the statistical properties of the model relative to

[57] Average tenure in our sample is 7.5 years.

the data and we assume that these residuals follow a random walk.

$$\hat{w}(x,a) = \zeta(x,a) + \iota \qquad \text{and} \qquad \zeta(x,a) = \zeta(x,a-1) + \nu$$

We denote the standard deviation of permanent shocks by σ_ν. We discuss the estimation of the standard deviation of transitory shocks denoted by σ_ι below. We follow the macroeconomic literature (Storesletten et al. (2004), Guvenen (2009), Heathcote et al. (2010)) and use identification in levels. This choice is supported by recent results from Daly et al. (2016), who show theoretically that if identification in levels is used, the estimate of the variance of permanent shocks is typically unbiased in the data. The identification requires variances and covariances of wage residuals.

$$\begin{aligned}
\sigma^2_{\nu,a} &= cov(\hat{w}(x,a), \hat{w}(x,a+1)) - cov(\hat{w}(x,a-1), \hat{w}(x,a+1)) \\
\sigma^2_{\iota,a} &= var(\hat{w}(x,a)) - cov(\hat{w}(x,a), \hat{w}(x,a-1)) - \sigma^2_{\nu,a}
\end{aligned}$$

These moments can be derived using model distributions so that we do not have to resort to simulation. We restrict the sample to workers age 20 to 50. We estimate age-specific variances and report the average in the main part of the paper. The estimated standard deviation is 0.072. Heathcote et al. (2010) report an estimate of 0.084.

In the baseline model, the estimated standard deviation of transitory wage shocks σ_ι is low. The reason is that we abstract from transitory wage risk by integrating out transitory productivity shocks η_s. Given that workers and firms are risk-neutral, this is without loss of generality. Alternatively, we could add transitory wage fluctuations from productivity shocks. In this case, the option value Ψ_s that captures the conditional expected value of shocks has to be subtracted from wages and the shock realization from the logistic distribution has to included in the wage. This can be done as follows. Denote by $w^n(x_w, x_m, a)$ the wage net of the option value

$$w^n(x_w, x_m, a) = w(x_w, x_m, a) - \mu \frac{\Psi_s(\pi_s)}{1-\pi_s}.$$

We add productivity shocks by using inverse transform sampling to sample from the logistic distribution. The sampling has to take into account separation decisions, so that draws of the uniform distribution are restricted to support $[\pi_s(x_w, x_m, a), 1]$. We denote the stochastic realization of the wage including the transitory wage component by $\tilde{w}(x_w, x_m, a)$. Given a productivity shock η_s, the wage becomes

$$\tilde{w}(x_w, x_m, a) = w^n(x_w, x_m, a) + \mu\eta_s.$$

We resort to simulation of the model in this case and simulate a cross section of 100,000 individuals. In very few cases it happens that due to the additive structure of shocks wages take negative values, we therefore approximate log wages including transitory shocks by $\frac{\tilde{w}}{\bar{w}}$ where $\bar{w}$ denotes the average wage conditional on age. The model is at monthly frequency and we aggregate the simulation paths to annual frequency. We derive average annual wages by summing all wages of a worker of a certain age and divide the sum by the number of employment periods so that there are no transitory fluctuations due to intervening spells of unemployment. This construction of annual wage observation corresponds to the empirical approach of constructing wages as ratio of annual earnings to annual hours worked. The estimated standard deviation σ_ι averaged over age groups is 0.41 compared to the average estimate of 0.29 from Heathcote et al. (2010). The estimate of the standard deviation of permanent shocks remains largely unaffected at 0.071 compared to 0.072 in case of full insurance of productivity shocks. Transitory variances are particularly high during the initial years in the labor market. If we consider the age range from 25 to 60 as in Heathcote et al. (2010) the standard deviation of transitory shocks goes down to 0.351. We conclude that our model is also consistent with substantial transitory wage fluctuations.[58]

Regarding wage inequality, we compare our model and the data using two measures of wage dispersion. The two measures we consider are the mean-min ratio (Mm ratio) as popularized in the theoretical literature on job search by Hornstein et al. (2011) and the variance of log wages as a popular measure from the empirical labor literature. Measuring the mean-min ratio in the data requires taking a stand on the minimum wage in the labor market. Tjaden and Wellschmied (2014) report Mm ratios between 1.83 and 3.02 depending on which bottom

[58]Studying the reasons behind diverging estimates for identification in levels and growth rates is beyond the scope of this paper. This question has been addressed in a recent paper by Daly et al. (2016). In our model, differences of estimates are rather modest. This is consistent with the result of Daly et al. (2016) that for balanced panels the difference in estimated variances based on the two methods largely disappears.

percentile of the wage distribution they use as representing the minimum wage. Using the 5th percentile as representing the minimum wage, they report Mm ratios by age that vary from 1.95 to 2.25 between age 25 and 49. At age 36, they report a ratio of 2.12. The average Mm ratio is 2.14. Using the 1st percentile, they report an average Mm ratio of 3.02 and using the 10th percentile they report an average Mm ratio of 1.83. Our model has more than 5% of workers at the lowest grid point of the wage distribution, so it is less sensitive to measurement problems regarding the minimum wage.

We explain in section A.1 of the appendix to the main part of the paper how we construct an age profile of the variance of log wages from CPS data. We discuss how the model fits to the variance age profile as part of the model extension presented in section IV.2 of this online appendix.

V.2.5 Job stability and wages

This section explains how we simulate data from the model to compare the trade-off between job stability and wages in the model to the evidence based on SIPP data presented in section 2.3. We discuss results of this comparison in section 4.3.5 of the main part of the paper.

At each age, we construct the 4-month separation rate. Consistent with the data, we allow for multiple separations over the 4-month period. We include transitory shocks to wages from the model. We discuss the construction of wages including transitory shocks in section V.2.4 of this online appendix. We denote the stochastic realization of the wage including the transitory wage component by $\tilde{w}(x_w, x_m, a)$. We average wages over the same period for which the separation rate is constructed (4 months). In very few cases it happens that due to the additive structure of shocks wages take negative values, we therefore approximate log wages including transitory shocks by $\frac{\tilde{w}}{\bar{w}}$ where $\bar{w}$ denotes the average wage conditional on age. We round age from the model to the next lower integer and implement all remaining steps following the empirical analysis on SIPP data described in section 2.3 of the main part and in section A.2 of the appendix to the main part of the paper.

VI Details on earnings losses

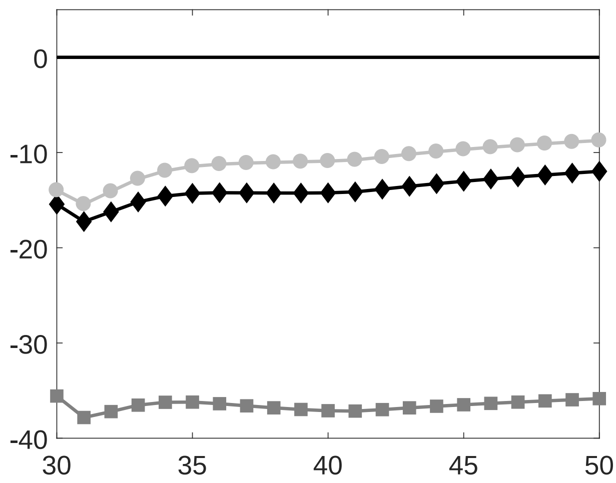
This section provides further details on earnings losses from the model. Section VI.1 discusses how earnings losses vary with age at displacement. Section VI.2 provides earnings losses for a sample including endogenous separations and job-to-job transitions and compares them to the non-mass layoff sample of Couch and Placzek (2010). Section VI.3 discusses the effect of advance notification of displacement events on estimated earnings losses in the first year after displacement. Section VI.4 discusses the evolution of earnings losses over a span of 20 years and shows the time path of earnings losses underlying the comparison to results by Davis and von Wachter (2011). Section VI.5 provides earnings losses for workers with good labor market prospects and compares them to earnings losses in a sample of non-benefit claimants from Couch and Placzek (2010). Finally, section VI.6 explores the effect of varying the job stability criterion before displacement with age to account for the increase in mean tenure by age.

VI.1 Earnings losses by age

In figure I, we show short, medium, and long-run earnings losses from displacement by age. The selection criteria and the construction of the control and layoff group follow the construction underlying figure 7 in section 5 in the main part of the paper, except that we vary the age at displacement. The gray line with squares shows earnings losses in the first year following displacement, the black line with diamonds shows earnings losses in the third year following displacement, and the light gray line with circles shows earnings losses in the sixth year following displacement. Age on the horizontal axis shows the age at displacement.

We report earnings losses for workers being aged between 30 and 50 at the time of the job loss. We see that the losses vary only little with age and that losses are almost linear in age so that the loss at average age is equivalent to the average loss over all ages for a symmetric age distribution. This shows that as long as the distribution in the samples of the empirical studies is not heavily skewed, considering losses at mean age will be nearly identical to mean losses across different ages. Indeed, this age range covers the relevant age range of the empirical studies. In the sample by Couch and Placzek (2010), mean age of the entire sample/separators/continuously employed workers is 39.7/38.9/40.2 years, the median is at 40/39/41 years and the 10^{th} percentile is always 9 years below the median and the 90^{th} is 8/8/7 years above the median. This shows that the distribution is highly symmetric around age 40 and mainly concentrated between ages 30 and 50. This justifies our focus on the mean/median worker in the main part of the paper.

Figure I: Earnings losses following displacement for different age groups



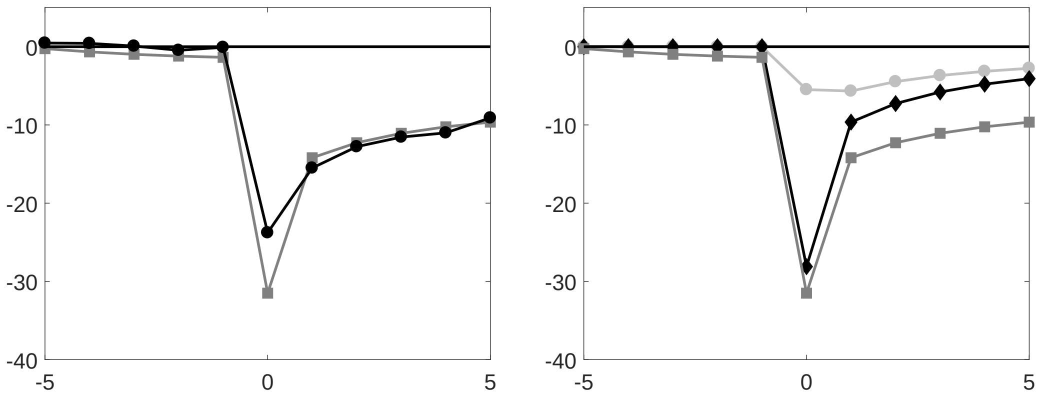
Notes: Earnings losses following displacement for different age groups. Construction and sample selection are as described in the main text. The gray line with squares shows earnings losses relative to the control group in the year of the displacement event. The black line with diamonds shows earnings losses three years after the displacement event. The light gray line with circles shows earnings losses six years after the displacement event. The horizontal axis shows age at the displacement event and the vertical axis shows earnings losses in percentage points.

VI.2 Earnings losses following separations

In figure J, we consider the earnings losses following a separation event. In this case, a separation comprises all workers that separate from their firm at the separation stage or do a job-to-job transition. The control group remains the same as in the case of displacement but the layoff group now comprises a particular selection of workers who, on average, have worse match- and/or worker-specific skills. We consider this the analog to non-mass layoff separators in Couch and Placzek (2010). We use the same methodology to derive earnings losses from the model as in the case of displacement and compare earnings losses from the model to the empirical estimates reported for separators in the non-mass layoff sample in Couch and Placzek (2010). Figure 11(a) shows earnings losses. We find that the model matches the empirical estimates of earnings losses very closely, both in the short and in the longer run. Figure 11(b) provides the decomposition in selection effect, extensive margin effect, and wage loss effect as before. For the twin experiment, we construct the control group to have the same skill composition as the layoff group at six years of tenure just before the separation event, in both match and worker type. The remainder of the decomposition is exactly as in the main text.

Selection now becomes significantly more important. Our decomposition assigns 57.7% of the

Figure J: Earnings losses following separation



(a) Earnings losses (b) Decomposition

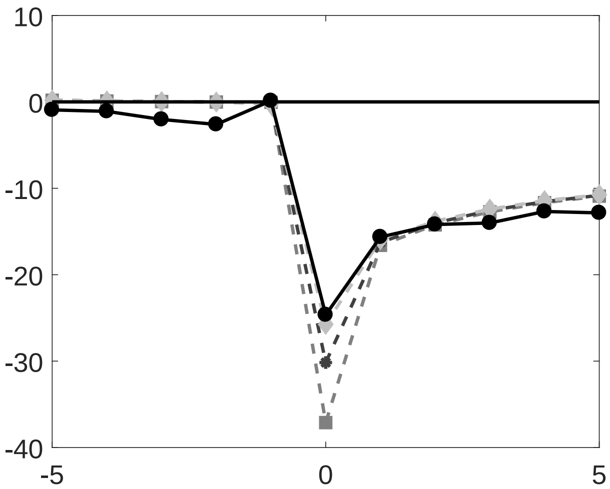
Notes: Left panel: Earnings losses after separation from model and empirical estimates. The gray line with squares shows earnings losses predicted by the model. The black line with circles shows the estimates by Couch and Placzek (2010). Separations include all separations and job-to-job transitions. Right panel: The gray line with squares shows earnings losses relative to the control group. The black line with diamonds shows the earnings relative to a control group without additional selection criteria (twin group). The light gray line with circles shows the wage losses for employed workers relative to a control group without additional selection criteria (twin group). The horizontal axis shows years relative to the separation and the vertical axis shows losses in percentage points relative to the control group.

earnings losses to selection, 13.7% to the extensive margin, and only 28.6% to wage losses. The reason for the increased importance of the selection effect is that the layoff group comprises workers that want to change jobs or who separate endogenously from their employer. These workers are a negative selection in terms of skills of workers with six or more years of tenure. This makes the control group even more selective than in the case of exogenous separations.

VI.3 Advance notification

Figure K demonstrates the effect of advance notification on first-year earnings losses from the model. In the data, displacement events can only be determined to lie in a certain quarter and no information is available when the worker is notified about the upcoming displacement event. In the model, the displacement happens in the moment it is announced. Here, we provide two alternative scenarios. The first scenario is one in which the worker is notified at the beginning of the month and has one additional month to search for a new employment before the displacement event. We assume that the search technology is the same as when the worker has already been displaced. This case is shown as dark gray dashed line with stars in figure K. We see already that the difference between the model prediction and the data reduces by roughly 50% in the initial year after displacement. There is no notable effect on earnings losses after six years. The second scenario provides workers with the opportunity to search for two months before the displacement occurs (light gray dashed line with diamonds). In this case, the difference in earnings losses in the initial year after displacement disappears almost completely and earnings losses after six years show no notable effect. We consider advance notification about the displacement event the likely explanation for the difference between the earnings losses predicted by the model and the empirical estimates in the first year after displacement.

Figure K: Earnings losses following displacement with advance notification

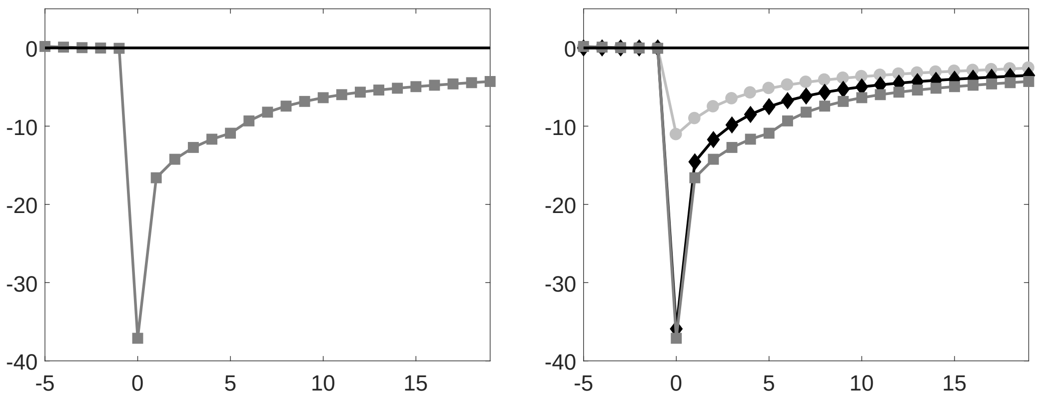


Notes: Earnings losses after displacement from model and empirical estimates. The gray line with squares shows earnings losses predicted by the model without advance notification. The black line with circles shows estimates by Couch and Placzek (2010). The dark gray dashed line with stars shows earnings losses predicted by the model when workers can search one months before the displacement event on the old job. The light gray dashed line with diamonds shows earnings losses predicted by the model when workers can search two months before the displacement event on the old job. The horizontal axis shows years relative to the displacement event and the vertical axis shows losses in percentage points relative to the control group.

VI.4 Long-run earnings losses following displacement

The left panel of figure L reproduces figure 7 from the main part of the paper for up to 20 years after the initial displacement. The right panel of figure L reproduces figure P from this online appendix again for up to 20 years after the initial displacement. In the main part of the paper, we restrict the analysis to a six-year time span after the initial displacement as in most empirical studies. Our structural model has been shown to reproduce earnings losses over this time span very closely. We can use the model to provide predictions for earnings losses for a longer time span (20 years following displacement).

Figure L: Long-run earnings losses following displacement



(a) Earnings losses

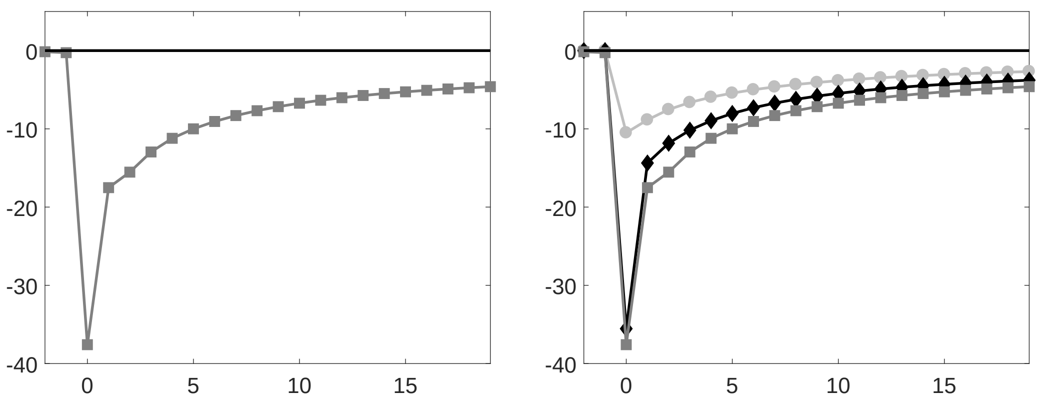
(b) Decomposition

Notes: Left panel: Long-run earnings losses after displacement from model. Right panel: The gray line with squares shows earnings losses relative to the control group. The black line with diamonds shows the earnings relative to a control group without additional selection criteria (twin group). The light gray line with circles shows the wage losses for employed workers relative to a control group without additional selection criteria (twin group). The horizontal axis shows years relative to the displacement event and the vertical axis shows losses in percentage points relative to the control group.

The left panel shows the earnings losses following displacement. The losses up to six years following displacement are as in the main part of the paper. After six years there is a small kink in earnings losses. This kink results from the selection criteria imposed on the control group. Following the sixth year after displacement the control group is no longer restricted to be continuously employed. This leads to non-employment in the control group from this point onwards. This reduces the selection effect instantaneously and causes a kink in the earnings losses. Below, we provide a further sensitivity analysis with respect to the construction of the control and the layoff group. Still, 20 years after the displacement event, the group of displaced workers suffers sizable earnings losses compared to the control group of roughly 5%. Looking at the right panel of figure L, we see the decomposition into selection, extensive margin, and wage loss effect as described in section 5.3 of the main text. We see that while the extensive margin effect reduces over time, the selection effect remains fairly constant in size and therefore gains in relative importance. The wage loss effect reduces but remains sizable even 20 years following the displacement event.

For figure L, we follow the selection criteria from Couch and Placzek (2010) that originate from Jacobson et al. (1993). Jacobson et al. (1993) argue that this choice of the control and layoff group simplifies the interpretation of their estimates. However, other group selection criteria have been proposed in the literature. For example, Davis and von Wachter (2011) look at workers with three years of prior job tenure and restrict the control group to workers who do not separate for two years following the displacement event rather than requiring continuous employment over the sample period. We provide a comparison to their results based on discounted earnings losses in table 3, section 5 of the main text, and find that our model matches their estimates closely. Figure M shows the time path of earnings losses underlying the discounted earnings losses from table 3.

Figure M: Long-run earnings losses following displacement with different group construction



(a) Earnings losses

(b) Decomposition

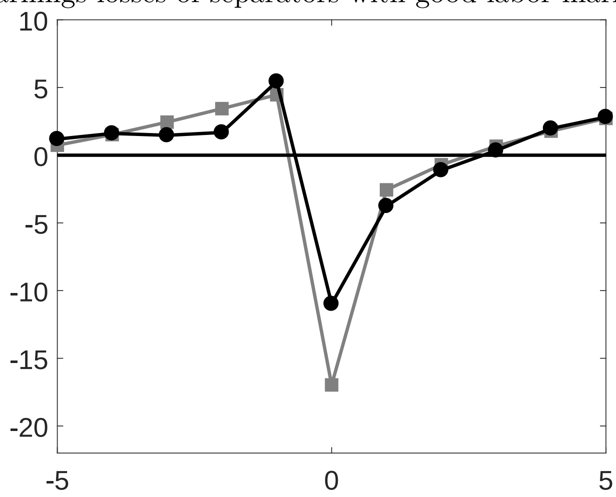
Notes: Left panel: Long-run earnings losses after displacement from model. Control and layoff group constructed as in Davis and von Wachter (2011). See text for details. Right panel: The gray line with squares shows earnings losses relative to a control group constructed as in Davis and von Wachter (2011). The black line with diamonds shows the earnings relative to a control group without additional selection criteria (twin group). The light gray line with circles shows the wage losses for employed workers relative to a control group without additional selection criteria (twin group). The horizontal axis shows years relative to displacement and the vertical axis shows losses in percentage points relative to the control group.

Qualitatively, the earnings losses in the left panel as well as the decomposition in the right panel look very similar to earnings losses and the decomposition in figure L. However, two points are noteworthy. First, the earnings losses uniformly decrease in size. Second, the selection effect in the decomposition effect of earnings losses decrease because the shorter non-separation period for the control group reduces the imposed correlation on the employment history of these workers. Quantitatively, we still find sizable earnings losses six years after displacement of roughly 8.3%. Selection becomes less important. Our decomposition assigns 13.8% of the earnings losses to selection, 26.8% to the extensive margin, and 59.4% to wage losses.

VI.5 Earnings losses for workers with good labor market prospects

Couch and Placzek (2010) examine the variation in earnings losses among different subgroups of separators. In line with Jacobson et al. (1993), they find strong variation in earnings losses across groups of workers who claim unemployment benefits and those who do not. In particular, Couch and Placzek (2010) find that workers who do not claim unemployment benefits do not experience large and persistent earnings losses. Arguably, these are workers who have good labor market prospects and therefore avoid the burdensome process of applying for unemployment benefits. Our model generates large and persistent earnings losses for displaced workers (section 5 of main part) and for all separators (section VI.2 of online appendix) in line with the empirical evidence.

Figure N: Earnings losses of separators with good labor market prospects



Notes: Earnings losses after separation for workers with good labor market prospects from model and empirical estimates for non-claimants from Couch and Placzek (2010). See text for details on group construction in the model. The gray line with squares shows the earnings losses predicted by the model. The black line with circles shows the estimates by Couch and Placzek (2010). The horizontal axis shows years relative to the separation and the vertical axis shows losses in percentage points relative to the control group.

Next, we show that such losses are not hard-wired in the model but that the model is also consistent with zero long-run losses for workers that have good labor market prospects. Given that there is no direct counterpart to workers who do not claim unemployment benefits in the model, we look at separating high-tenure workers who can improve their labor market situation after leaving the current employer. We refer to them as workers with good labor market prospects. We leave all other sample selection criteria as before but select only the bottom half of the earnings distribution out of workers with at least six years of job tenure. We consider the separators from this group as the corresponding group to the sample of workers who do not claim benefits from Couch and Placzek (2010) (non-claimant sample) shown in their Figure IV. The control group are continuously-employed workers. Figure N shows the resulting earnings losses for these workers with good labor market prospects against the estimates for the non-claimant sample from Couch and Placzek (2010). Figure N shows that our model matches the empirical results very closely. Like in the data, our model generates for workers with good labor market prospects earnings losses that are small in the long-run.

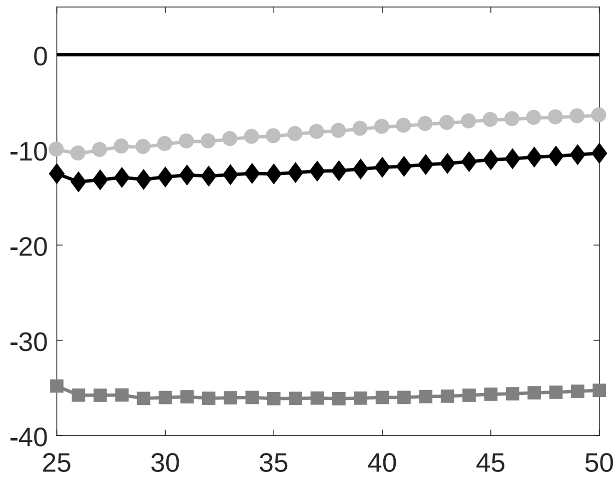
VI.6 Earnings loss with age-specific stability threshold

The empirical studies use six or three years of prior job tenure as a threshold to identify stable jobs. In our empirical analysis in section 2 of the main text, we show that tenure increases almost linearly with age. An important reason for this increase in job stability is that workers find better jobs over time. This implies, however, that three years of prior job tenure selects a very different group of workers among the 25-year-old workers than among the 40-year-old workers. While a 25-year-old worker with three years of tenure is at the mean of the age-specific tenure distribution, a 40-year-old worker with three years of tenure is in the lower part of the age-specific tenure distribution. Hence, the 25-year-old worker has found a stable job relative to his cohort, while the 40-year-old is compared to his cohort on a rather unstable employment path. To account for this effect, we compute earnings losses for workers in stable jobs using age-specific mean tenure as stability threshold according to the age-specific means from figure 2(a).[59] We focus on the twin experiment, i.e. we do not impose any future job stability requirements for the control group after the displacement event.

Figure O shows the short-run, medium-run, and long-run earnings losses by age at displacement using the age-specific job stability criterion. The earnings losses are large and vary only little with age although job stability thresholds vary. The reason is that in all cases, workers in the control group hold the best jobs and face very stable employment relationships. Their jobs are by construction above the average in terms of job stability of their age group and will persistently remain better than the average worker of the cohort. Even if displaced workers manage to recover to the average of the age cohort, there will be large and persistent earnings losses among these workers. Hence, this shows that the observed earnings losses result in large

[59] For example, we use two years of job tenure for a 25-year-old worker to classify stable jobs but seven years of job tenure to classify a stable job for a 40-year-old worker.

Figure O: Age-specific earnings losses with age-specific stability threshold



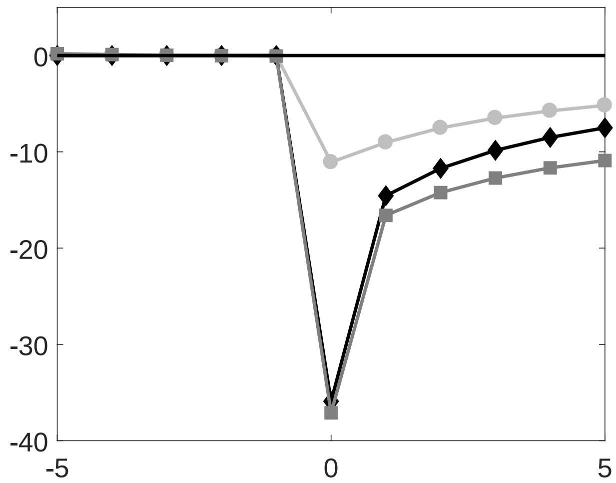
Notes: Earnings losses following displacement for different age groups and with age-specific job stability threshold. Stable jobs are defined by age-specific mean tenure. Construction and sample selection is otherwise as described in the main text. The gray line with squares shows earnings losses relative to the control group in the year of the displacement event. The black line with diamonds shows earnings losses three years after the displacement event. The light gray line with circles shows earnings losses six years after the displacement event. The horizontal axis shows age at the displacement event and the vertical axis shows earnings losses in percentage points.

part from a mean-reversion of workers from very good jobs to the average.

VII Decomposition of earnings losses

In section 5.3, we decompose earnings losses into a selection effect, an extensive margin effect, and a wage loss effect. Figure P documents the quantitative importance of each factor over time after the initial displacement event. The light gray line with circles gives the wage loss effect, the difference between the light gray line with circles and the black line with diamonds gives the extensive margin effect, and the difference between the black line with diamonds and the gray line with squares gives the selection effect.

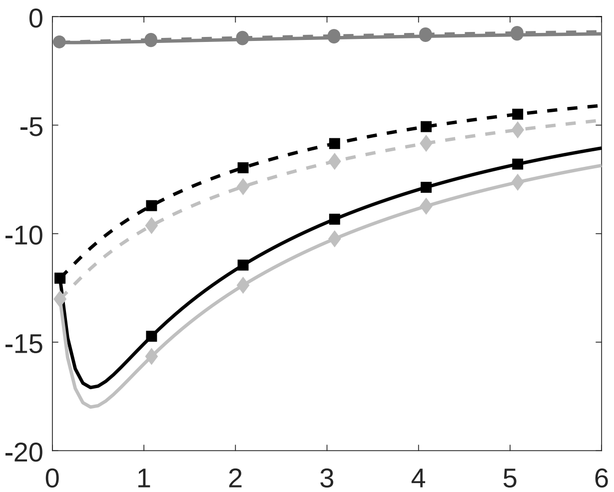
Figure P: Decomposition of earnings losses



Notes: The gray line with squares shows earnings losses relative to the control group. The black line with diamonds shows earnings losses relative to a control group without additional selection criteria (twin group). The light gray line with circles shows wage losses for employed workers relative to a control group without additional selection criteria (twin group). The horizontal axis shows years relative to displacement and the vertical axis shows losses in percentage points relative to the control group.

In section 5.3 of the main part, we also decompose the wage loss effect and the extensive margin effect in effects coming from a loss of worker- and match-specific skill losses. Figure Q provides the graphical decomposition. We consider three cases: The gray lines with circles show the case with worker-specific skill loss, while the black lines with squares show the loss of match-specific skills. The light gray lines with diamonds illustrate the loss of worker- and match-specific skills. The wage effect is decomposed by the dashed lines, earnings losses are shown as solid lines and the difference between earnings and wage losses constitutes the extensive margin effect.

Figure Q: Decomposition of wage loss and extensive margin effect



Notes: Wage and earnings losses of counterfactual experiment. Wage losses are indicated by dashed lines and earnings losses by solid lines. The lines with gray circles correspond to a group with only worker-specific skill losses, the lines with black squares to a group with only match-specific skill losses, and the lines with light gray diamonds to a group with worker- and match-specific skill losses. All losses are relative to a control group without any skill losses. Details of group construction can be found in the main text. Horizontal axis shows years since skill loss. Vertical axis shows losses in percentage points relative to the control group.