Mismatch Cycles

Isaac Baley
Ana Figueiredo
Robert Ulbricht

This version: August 2021
(January 2020)

Barcelona GSE Working Paper Series

Working Paper nº 1151

Mismatch Cycles

Isaac Baley* Ana Figueiredo† Robert Ulbricht‡

August 3, 2021

Abstract

This paper studies the cyclical dynamics of skill mismatch and quantifies its impact on labor productivity. We build a tractable directed search model, in which workers differ in skills along multiple dimensions and sort into jobs with heterogeneous skill requirements. Skill mismatch arises due to information frictions and is prolonged by search frictions. Estimated to the U.S., the model replicates salient business cycle properties of mismatch. Job transitions in and out of bottom job rungs, combined with career mobility, are key to account for the empirical fit. The model provides a novel narrative for the scarring effect of unemployment.

Keywords: Business cycles, cleansing, learning about skills, multidimensional sorting, scarring effect of unemployment, search-and-matching, skill mismatch, sullying.

JEL Classification: E24, E32, J24, J64.

*Universitat Pompeu Fabra, CREi, Barcelona GSE, and CEPR, isaac.baley@upf.edu.
†Erasmus School of Economics, figueiredo@ese.eur.nl
‡Boston College, ulbricht@bc.edu

Research support is gratefully acknowledged by Baley from the Spanish Ministry of the Economy and Competitiveness through the Severo Ochoa Programme for Centres of Excellence in R&D (CEX2019-000915-S) and by Figueiredo from Fundação para a Ciência e Tecnologia, Grant SFRH/BD/94516/2013. The research leading to these results has received funding from "la Caixa" Foundation Project Grant CG-2019_01: Worker Reallocation, Occupational Mismatch, and Business Cycles.

> "In a regime of ignorance, Enrico Fermi would have been a gardener, Von Neumann a checkout clerk at a drugstore." (Stigler, 1962)

1 Introduction

Over the business cycle, labor markets face a large amount of reallocation: firms create and destroy vacancies, work-relationships are formed and resolved, and workers change jobs and careers. In this paper, we investigate—theoretically and empirically—how business cycles affect the skill allocation of workers to jobs.

Our theoretical framework is a version of the directed search model of Menzio and Shi (2010, 2011), in which we incorporate two key features. First, workers differ along multiple skill dimensions and sort into jobs with heterogeneous skill requirements along those dimensions. The job search of workers encompasses a career choice, determining the *type of skill* that workers seek to employ, and a vertical choice of *task complexity*, which entails varying ability requirements on the employed skill. Second, workers and firms have incomplete information about worker skills, which generates skill mismatch in equilibrium. Workers and firms revise their beliefs about worker skills based on a noisy learning technology, with the important assumption that learning is more accurate regarding skills currently used in production. In equilibrium, workers reallocate both up and down job ladders within a given career path (utilizing the same skill at varying complexities) and across different career paths (utilizing different skills).

We estimate the framework using a combination of worker-level data from the NLSY79 and occupation-level descriptors of job requirements (O*NET).[1] The estimation builds on a novel skill-based strategy to identify career switches in the data. We find that the business cyclicality of mismatch is determined by two opposing forces. On the one hand, we find that in recessions underqualified workers are fired, specifically those that are occupied at the bottom rungs of the job ladder. This *cleansing* effect reduces mismatch among ongoing work-relations, raising the average labor productivity of workers that have been continuously employed for two years by 1.3 percent. On the other hand, we find that mismatch among new hires goes up in recessions, which is primarily caused by an increase in overqualification among workers hired for low-complexity jobs. This *sullying* effect reduces labor productivity of new hires by 0.9 percent. Both the cleansing and sullying effect are consistent with direct evidence on the cyclicality of mismatch, which we document among workers in the NLSY79.

[1] See Yamaguchi (2012), Lindenlaub (2017), and Lise and Postel-Vinay (2020) for related calibration strategies using the same combination of NLSY79 and O*NET.

Our theoretical findings are explained by a non-trivial interaction between job mobility and mismatch: Whereas transitions within a given career path (to jobs that employ similar skills) tend to reduce mismatch as workers re-sort across job rungs in response to belief revisions, transitions into new career paths (to jobs that employ previously untried skills) tend to increase mismatch as a consequence of higher uncertainty. Accordingly, the cyclicality of mismatch is closely entangled with the business cycle dynamics of career mobility. Specifically, our model predicts that career mobility is countercyclical (which we confirm in the data). This is because workers that are fired from the bottom rungs of a given career path will optimally seek to find jobs that utilize a different skill set rather than re-applying to jobs, for which they are underqualified. In that sense, the two opposing forces shaping the cyclicality of mismatch are in fact both manifestations of the *cleansing* of underqualified workers, which increases career mobility in recessions and in turn heightens mismatch among new hires.

At the worker-level, our framework gives rise to considerable inertia in mismatch and earnings, reflecting, on the one hand, the time needed to learn about any subsisting mismatch and, on the other hand, its slow dissolution due to search frictions. The inertia provides a novel narrative for the "scarring effect of unemployment", which complements recent explanations by Jung and Kuhn (2019), Jarosch (2021), and Huckfeldt (2021). In line with empirical evidence, workers that are displaced from their careers suffer large and persistent earnings losses, even after they have been re-employed. In the calibrated model, these earnings losses amount to 19 percent five years after displacement, and to about 10 percent ten years after displacement.

We conclude the paper with direct evidence for workers having imperfect information about their skills. Using workers' forecasts about their own future occupation, we document that the forecast errors entailed in these forecasts can be systematically predicted by a measure of worker ability that has been realized at the time the forecasts are formed. The evidence complements recent work by Conlon et al. (2018) who document substantial forecast errors in workers expectations regarding future labor market outcomes using the Survey of Consumer Expectations of the NY Fed.[2] In addition, we provide indirect evidence towards the model's mechanism. First, career mobility is predicted by the suitability of workers' skills for their current career. Second, mismatch among workers starting a new career is on average larger and more dispersed compared to workers switching jobs within careers.

Related literature Our model combines ingredients from several strands of the literature. Our formulation of the labor market is based on the directed search models of Menzio and

[2] Fredriksson, Hensvik and Skans (2018) also provide indirect evidence pointing to information frictions using Swedish administrative data.

Shi (2010, 2011), Menzio, Telyukova and Visschers (2016) and Schaal (2017), which provide us with the analytical framework to explore out of steady state dynamics in a model with many degrees of heterogeneity.

The multidimensional modeling of skills is closely related to recent theoretical works by Lise and Postel-Vinay (2020) and Lindenlaub and Postel-Vinay (2017) that also emphasize the irreducibility of worker heterogeneity into a single unidimensional index.[3] There are two important differences with respect to our paper. First, both papers consider a random search model of the labor market, effectively accounting for skill mismatch by an exogenous friction that prevents workers from applying to the best-fitting jobs. In contrast, our approach abstracts from such frictions by allowing search to be directed, and instead motivates skill mismatch using incomplete information.[4] Second, both papers focus on steady states, whereas our framework allows for aggregate shocks and is tractable enough to explore out of steady state dynamics, which is at the core of our exploration.

Finally, our model incorporates learning à la Jovanovic (1979, 1984). Our paper particularly relates to more recent works, in which learning is about worker skills, rather than a match-specific productivity term (e.g., Groes, Kircher and Manovskii, 2013, Papageorgiou, 2014, and Wee, 2016). In our model, this implies that the assessment of future match qualities varies with the prior work experience of workers and, in particular, leads to countercyclical fluctuations in uncertainty. Relatedly, Acharya and Wee (2020) explore a complementary mechanism that similarly gives rise to countercyclical uncertainty that reduces matching efficiency in recessions.

Our paper also contributes to an old debate on the cyclicality of worker–occupation mismatch.[5] On the one hand, matching models with endogenous separations suggest that mismatch is procyclical due to a *cleansing* of unproductive matches (e.g., Mortensen and Pissarides, 1994; see also, Lise and Robin, 2017 for a variant with ex ante heterogeneous workers). On the other hand, others have argued that mismatch is countercyclical due to various *sullying* forces (e.g., Barlevy, 2002; Moscarini, 2001; Barnichon and Zylberberg, 2019). Our analysis provides a more nuanced view, suggesting that in fact both forces are present among different sets of workers, although the cleansing effect unambiguously dominates at the aggregate. Our evidence complements Crane, Hyatt and Murray (2018) who provide direct evidence that overall sorting is countercyclical, Bowlus (1995) who provides indirect evidence

[3] Neal (1999) also studies an environment that distinguishes between career and firm matches.

[4] While labor market frictions by themselves do not cause mismatch to arise in our framework, they do contribute to its persistence as they make reallocation costly. Related to the role of imperfect information in our model, Guvenen et al. (2020) use a similar narrative to motivate their empirical exploration of multidimensional skill mismatch.

[5] Şahin et al. (2014) explore an alternative notion of mismatch between vacancies and job seekers.

that match quality of new hires is procyclical, and Haltiwanger et al., 2021 who find evidence of both sullying and cleansing during recessions.

Layout The paper is organized as follows. In Section 2, we set up the model and characterize equilibrium. In Section 3, we describe the calibration strategy used to quantify the model. In Section 4, we explore implications of mismatch at the worker-level. In Section 5, we describe the predicted business cycle dynamics of mismatch and contrast them with the data. In Section 6 we present suggestive evidence towards the learning friction at the core of the model and towards its implications for career mobility and mismatch. Section 7 concludes.

2 Model

We develop a directed search model of the labor market with endogenous sorting and aggregate fluctuations in productivity. There are two key features. First, workers are characterized by a high-dimensional vector of different *skill types*. Given their skills, workers sort into jobs that are characterized by the type of skill they employ and are further differentiated by the *intensity* they make use of this skill ("task complexity"). Second, information about worker skills is imperfect and needs to be inferred from noisy signals.

2.1 Environment

Population and technology Time is continuous and extends forever. There is a unit mass of workers, indexed by $i \in [0, 1]$, and an endogenous measure of one-vacancy firms with free entry. Firms and workers are risk neutral and share the same discount rate ρ. Each worker is characterized by a continuum of time-invariant abilities, $\{a_{i,k}\}_{k\in[0,1]}$, where $a_{i,k}$ are Normally distributed with mean a_0 and variance S_0 and are i.i.d. across skill types k and across workers i. Abilities are not observed (directly), but their distribution is public information.

Jobs are characterized by a unique skill type $k \in [0, 1]$ utilized in production, and a skill requirement or "task complexity" $r \in \mathcal{R}$ where $\mathcal{R} \subset \mathbb{R}$ is compact. Henceforth, we label jobs sharing the same skill type k as "career", and refer to distinct levels of r within a given career as "job ladder". The log-output flow of worker i in job (k, r) is given by

$$\log y_{i,k,r}(t) = z(t) + \eta r - \max\{r - a_{i,k}, 0\}. \tag{1}$$

Here, $z(t)$ is an aggregate productivity component, which follows a Poisson process that takes two values, $z(t) \in \{z_L, z_H\}$, with switching intensities λ_{z_L} and λ_{z_H}; we normalize $z_L \leq z_H$ and identify the first state with a recession. The second term in (1), ηr, defines the gains in (potential) output associated with more complex tasks, whereas the third term captures losses due to underqualification. We assume $\eta \in (0, 1)$, so that the net return on raising the skill requirement is positive if and only if the worker is skilled enough to operate the more complex technology ($a_{i,k} > r$).

Unemployed workers receive a constant utility flow b from home production.

Evolution of beliefs Agents learn about workers' skills while producing. Specifically, in each instant that a worker is employed, workers and firms update their beliefs about the utilized skill, $a_{i,k}$, based on the noisy signal

$$\mathrm{d}s_{i,k}(t) = a_{i,k}\mathrm{d}t + \sigma \mathrm{d}W_{i,k}(t),$$

where $\sigma > 0$ parametrizes the noisiness of the signal and $W_{i,k}$ follows a standard Brownian motion that is independent across all i and k. We assume that all learning is common knowledge and no direct inference is made from $y_{i,k,r}$ (we view the signal $s_{i,k}$ as an approximation to the information that could be inferred if agents were to observe a noisy version of output[6]).

Specifically, the assumed process for $s_{i,k}$ implies that for all i and k the posterior distribution entertained about $a_{i,k}$ is Gaussian at all times. Let $\hat{a}_{i,k}(t)$ and $\Sigma_{i,k}(t)$ denote the first two moments of this posterior. While employed in a job utilizing skill k, the posterior moments follow a diffusion given by the usual Kalman-Bucy filter,

$$\begin{aligned}
\mathrm{d}\hat{a}_{i,k}(t) &= \frac{\Sigma_{i,k}}{\sigma^2}\left(\mathrm{d}s_{i,k}(t) - \hat{a}_{i,k}\mathrm{d}t\right) \\
\mathrm{d}\Sigma_{i,k}(t) &= -\left(\frac{\Sigma_{i,k}}{\sigma}\right)^2 \mathrm{d}t.
\end{aligned}$$

Upon switching to a previously untried skill type k, the belief is initialized at the objective prior distribution, $(\hat{a}_{i,k}, \Sigma_{i,k}) = (a_0, S_0)$.

Labor markets, vacancy creation, and separations The labor market is organized in a continuum of submarkets indexed by the job characteristics (k, r), the relevant worker type

[6] In fact, this interpretation could be made exact with two slight changes to the environment: (i) time is discrete, (ii) the penalty on underqualification is given by $g(r - a_{i,k} - \sigma\epsilon_{i,t})$ where $\epsilon_{i,t} \sim \mathcal{N}(0, 1)$ is i.i.d. across i and t. Here g can be any monotonic approximation to $\max\{r - a_{i,k}, 0\}$ which sustains some arbitrary small return on skills when $a_{i,k} > r$. E.g., one could set $g(x) = \max\{x, 0\} + \beta x$ with $\beta > 0$ small. As long as g is strictly increasing in x, it holds that observing $y_{i,k,r}$ is informationally equivalent to observing a noisy signal $a_{i,k} + \sigma\epsilon_{i,t}$, demonstrating our claim.

$(\hat{a}_{i,k}, \Sigma_{i,k})$, and a lifetime utility x implicit in the employment contracts offered by firms to workers. Workers direct their search towards these submarkets. Specifically, unemployed workers have the opportunity to search the labor market at rate 1 and can search any submarket. For simplicity, we rule out recall of previously abandoned skill types but notice that the assumption imposes little restrictions on workers' search policies in practice.[7] Employed workers have the opportunity to search the labor market at rate $\kappa \in [0, 1]$ and can search for jobs within their current career path (i.e., the skill type k of the aspired job must match their current job). Vacancies are created by an infinite supply of potential firms, which can open a vacancy in any submarket $\omega \equiv (k, r, x, \hat{a}_k, \Sigma_k)$ at flow costs c.

Workers searching in a given submarket and vacancies posted in that submarket come together through a frictional matching process. In particular, a worker searching in submarket ω meets a vacancy at rate $p(\theta_t(\omega, z))$ where $\theta_t(\omega, z)$ denotes the vacancy-to-worker ratio of submarket ω. Similarly, a vacancy posted in submarket ω meets a worker at rate $q(\theta_t(\omega, z)) = p(\theta_t(\omega, z))/\theta_t(\omega, z)$. As usual, we assume that p is twice differentiable, strictly increasing and concave; q is strictly decreasing; and $p(0) = q(\infty) = 0$, $p(\infty) = q(0) = \infty$.

When a firm and a worker meet in a submarket, the firm offers the worker a wage contract worth x in lifetime utility and hires the worker. Following Menzio and Shi (2010, 2011), we assume that the underlying contract space is complete, so that separations are bilaterally efficient. In particular, endogenous job separations as well as the search policies of employed workers are taken so as to maximize the joint value of the relationship.

In addition to an endogenous separation choice (further detailed below), worker–firm pairs separate at an exogenous rate $\delta > 0$. Moreover, independent of their current employment status, workers switch careers at an exogenous rate $\epsilon > 0$. If hit by such a career-shock, workers are forever prevented from applying to any submarket involving the skill type k of their previous career.

Remark on notion of careers In our terminology, the label *career* refers to a set of jobs that utilize similar skills. Our definition differs from previous approaches that have defined careers based on occupation- or industry-codes. While related, such definitions would be misleading in our case as distinct occupations may share very similar skill mixes, whereas

[7]The exception are workers that are exogenously forced to switch careers (introduced below), which would otherwise prefer to re-apply to their old career. The reason why the no recall assumption does not pose much of a restriction otherwise is that k lies in a continuum. In particular, absent aggregate shocks, workers would never find it optimal to return to skill types that they have previously abandoned. The restriction therefore merely rules out recall after aggregate productivity shocks. For the calibration introduced in the next section, workers indeed never find it optimal to do so if given the chance.

others may bundle together jobs with distinct skills.[8] For a consistent interpretation of the model, one should therefore think of careers in terms of skill-mixes when mapping the model to the data. Our calibration of the model in Section 3 aims to do so by employing a skill-based definition of careers.

2.2 Equilibrium Characterization

Notation To converse on notation, we suppress i subscripts from all variables going forward. All value functions are indexed with a time subscript t to express their potential dependence on the aggregate state (except for their dependence on aggregate productivity z, which is kept as explicit argument).

Vacancy creation By free entry, the value of creating a vacancy must be zero in every submarket. Let $J_t(\hat{a}_k, \Sigma_k, r, z)$ denote the joint value of a worker–firm pair. The zero profit condition reads $c = q(\theta_t(\omega, z))(J_t(\hat{a}_k, \Sigma_k, r, z) - x)$. Rearranging, this pins down the market tightness as a function of the firm's share of the surplus, $\theta_t(\omega, z) = f_\theta(J_t(\hat{a}_k, \Sigma_k, r, z) - x)$, where

$$f_\theta(V) \equiv \begin{cases} q^{-1}\left(c/V\right) & V \geq 0 \\ 0 & \text{else.} \end{cases} \tag{2}$$

Unemployed worker problem Because there is no learning during unemployment, the belief about an unemployed worker's skills, $\{\hat{a}_k, \Sigma_k\}_{k\in[0,1]}$, remains at the same value at which they entered unemployment. The value of being unemployed *conditional* on searching for jobs of skill type k, denoted by $U_t(\hat{a}_k, \Sigma_k, z)$, is therefore given by:

$$\begin{aligned} \rho U_t(\hat{a}_k, \Sigma_k, z) = b + \max_{x,r} \left\{p(\theta_t(\omega, z))\left(x - U_t(\hat{a}_k, \Sigma_k, z)\right)\right\} + \\ + \epsilon\left(U_t(a_0, S_0, z) - U_t(\hat{a}_k, \Sigma_k, z)\right) + \\ + \lambda_z\left(U_t(\hat{a}_k, \Sigma_k, -z) - U_t(\hat{a}_k, \Sigma_k, z)\right). \end{aligned} \tag{3}$$

The flow value of being unemployed is comprised of four terms: (i) the utility flow of home production, (ii) the product between the job finding rate and the excess utility, $x - U$, promised to the worker in the submarket they are searching (maximized subject to the θ–x frontier

[8] For instance, using the methodology described in Section 3, we find that the skill mix of an economist is very similar to the ones of actuaries, financial managers, and mathematicians and statisticians, which all constitute different occupations at the 3-digit level (see Appendix F.1). Defining careers based on 2-digit occupation codes instead bundles together many occupations with vastly different skill mixes.

defined by (2)), (iii) the product between the exogenous career switching rate and the induced value change when starting a new career with $(\hat{a}_{k'}, \Sigma_{k'}) = (a_0, S_0)$, and (iv) the product between the arrival rate of aggregate productivity shocks and the corresponding change in value (here, "$-z$" denotes the complementary state of z).

Intuitively, $U_t(\hat{a}_k, \Sigma_k, z)$ measures an unemployed worker's value of searching in career k. It remains to solve for the optimal career choice of unemployed workers. Fortunately, the problem is simplified by our assumption that k lies in a continuum, which implies that the choice of skill types is stationary as workers never run out of new careers to explore. Accordingly, unemployed workers effectively face the choice between searching within their current career path, summarized by the belief $(\hat{a}_k, \Sigma_k)$, or starting a new career k' with $(\hat{a}_{k'}, \Sigma_{k'}) = (a_0, S_0)$. The *unconditional* value of being unemployed is then given by

$$\mathcal{U}_t(\hat{a}_k, \Sigma_k, z) = \max\left\{U_t(\hat{a}_k, \Sigma_k, z), U_t(a_0, S_0, z)\right\}. \tag{4}$$

Joint surplus maximization Next, consider the worker–firm pair's joint continuation choice and the search policy of employed workers. As long as the relationship remains active, its flow value is given by

$$\begin{aligned}\rho J_t^{\text{act}}(\hat{a}_k, \Sigma_k, r, z) = e^{z+\eta r}\,\mathbb{E}_t[e^{-\max\{r-a_k,0\}}] + \Lambda_t(\hat{a}_k, \Sigma_k, r, z) +\\ + \max_{x,r}\left\{\kappa p(\theta_t(\omega, z))\,(x - J_t(\hat{a}_k, \Sigma_k, r, z))\right\} +\\ + \delta\,(\mathcal{U}_t(\hat{a}_k, \Sigma_k, z) - J_t(\hat{a}_k, \Sigma_k, r, z)) +\\ + \epsilon\,(\mathcal{U}_t(a_0, S_0, z) - J_t(\hat{a}_k, \Sigma_k, r, z)) +\\ + \lambda_z\,(J_t(\hat{a}_k, \Sigma_k, r, -z) - J_t(\hat{a}_k, \Sigma_k, r, z))\,.\end{aligned} \tag{5}$$

Here the first term corresponds to the expected output flow of the worker–firm pair. Using $a_k \sim \mathcal{N}(\hat{a}_k, \Sigma_k)$, we can explicitly compute the expected loss from underqualification as $\mathbb{E}_t[e^{-\max\{r-a_k,0\}}] = \psi(\hat{a}_k - r, \sqrt{\Sigma_k})$ with

$$\psi(x, s) \equiv e^{x+s^2/2}\Phi\left(-x/s - s\right) + \Phi\left(x/s\right),$$

where $\Phi(\cdot)$ is the standard Normal cdf. The second term in (5) captures how J changes as uncertainty declines over the course of the relationship (first term of Λ) as well as how uncertainty affects the value itself (second term of Λ),

$$\Lambda_t(\hat{a}_k, \Sigma_k, r, z) \equiv \left(\frac{\Sigma_k}{\sigma}\right)^2\left(-\frac{\partial J_t(\hat{a}_k, \Sigma_k, r, z)}{\partial \Sigma_k} + \frac{1}{2}\frac{\partial^2 J_t(\hat{a}_k, \Sigma_k, r, z)}{\partial \hat{a}_k^2}\right).$$

The third term in (5) captures changes in the joint value due to the worker moving to a better-matched job (where the maximization is again subject to the θ–x frontier defined in (2)). The fourth and fifth terms capture the change in value induced by exogenous separation and exogenous career switching, in which cases the worker–firm pair obtains, respectively, $\mathcal{U}_t(\hat{a}_k, \Sigma_k, z)$ and $\mathcal{U}_t(a_0, S_0, z)$. Here we used that the post-separation value for the firm is zero given free entry. The last term captures the change in value induced by aggregate productivity shocks.

Finally, accounting for endogenous separations, the joint value of the worker–firm pair is given by

$$J_t(\hat{a}_k, \Sigma_k, r, z) = \max\left\{J_t^{\text{act}}(\hat{a}_k, \Sigma_k, r, z), \mathcal{U}_t(\hat{a}_k, \Sigma_k, z)\right\}. \tag{6}$$

Job ladder We next explore workers' submarket choice as a function of the belief $(\hat{a}_k, \Sigma_k)$. Substituting the θ–x frontier defined by (2) into (3) and (5), it is immediate that the choice of task-complexity always maximizes the joint value,

$$r^*(\hat{a}_k, \Sigma_k, z) = \arg\max_{r \in \mathcal{R}} J_t(\hat{a}_k, \Sigma_k, r, z). \tag{7}$$

For employed workers, this is a direct consequence of bilateral efficiency. For unemployed workers, it is similarly in their best interest to maximize the joint value because the firms' share is fixed by the free entry condition, making the worker effectively residual claimant on the value.

Figure 1 illustrates the resulting job ladder using the parametrization described in Section 3. The figure displays the choice of r as a function of $\hat{a}_k$ and Σ_k. As the search policies are very similar for both realizations of aggregate productivity, we only plot them for the case where $z = z_H$. In the adopted parametrization, there is a 7-step job ladder corresponding to $\mathcal{R} = \{0, 0.5, 1, \ldots, 3\} \times S_0^{1/2}$. Workers pursuing a new career, search for jobs with the lowest complexity, $r^*(a_0, S_0, z_H) = 0$ (indicated by the red square in the plot). As workers become more optimistic regarding their skills in a given career k, they apply to more complex jobs (indicated by lighter shades of green). There is no search towards job rungs below the one chosen by career-switchers, as such jobs would be dominated by the option to pursue a new career.

The effect of uncertainty is more ambiguous: While high uncertainty leads workers at the bottom of the expected skill distribution to apply for jobs for which they expect to be overqualified, it leads high expected-skill workers to apply for jobs for which they are on average underqualified.[9] This is because for high expected-skill workers the expected value

[9]This prediction is consistent with our data on mismatch (introduced below), in which workers at the

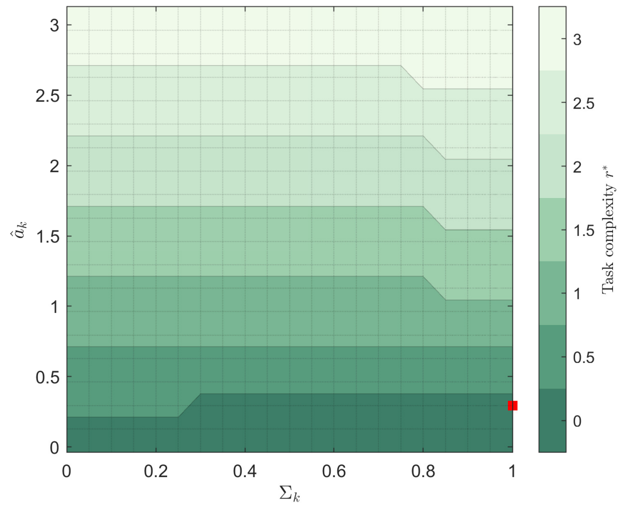


Figure 1: Job ladder. *Notes.*—The graph shows the task complexity r^* chosen as a function of expected ability $\hat{a}_k$ and uncertainty Σ_k. The red square marks the unconditional prior (a_0, S_0) for untried skill types. Values for $\hat{a}_k$, $\Sigma_k^{1/2}$ and r are denominated in units of $S_0^{1/2}$. The graph is plotted for $z = z_H$; the case where $z = z_L$ looks similar. See Section 3 for a description of the parametrization.

of learning is nearly symmetric in good and bad news, making *expected contemporaneous output* the primary determinant of r^*, which for the calibrated value of η is maximized when workers are expected to be slightly underqualified (whenever $\Sigma_k > 0$).[10] By contrast, for low expected-skill workers, being overqualified entails a positive option value due to the relative ease to adjust job rungs upwards via on-the-job search, whereas being underqualified at the bottom job rung entails job loss and career switching.

It remains to characterize the lifetime utility x chosen by workers that are actively searching for new jobs. From (2), x is decreasing in market tightness θ, creating a trade-off for the worker to search in submarkets with higher job finding rates p versus searching in submarkets with higher utility x. Maximizing (3) subject to the θ–x frontier defined by (2), the market tightness chosen by unemployed workers is given by

$$\theta = p'^{-1}\left(\frac{c}{J_t(\hat{a}_k, \Sigma_k, r^*, z) - U_t(\hat{a}_k, \Sigma_k, z)}\right) \tag{8}$$

with r^* as in (7). Similarly, maximizing (5) subject to (2), the market tightness chosen by

bottom job rungs are systematically overqualified, whereas workers in upper job rungs are systematically underqualified (see Appendix H.3).

[10] In general, $\arg\max_r \mathbb{E}[y_t] > \hat{a}$ if and only if $\eta > \bar{\eta}(\Sigma_k)$, with $\bar{\eta}(\Sigma_k) = 1 - \left(\sqrt{2/\pi} \cdot \Phi(-\Sigma_k^{1/2})/\phi(\Sigma_k^{1/2}) + 1\right)^{-1} < 0.5$ for all $\Sigma_k > 0$.

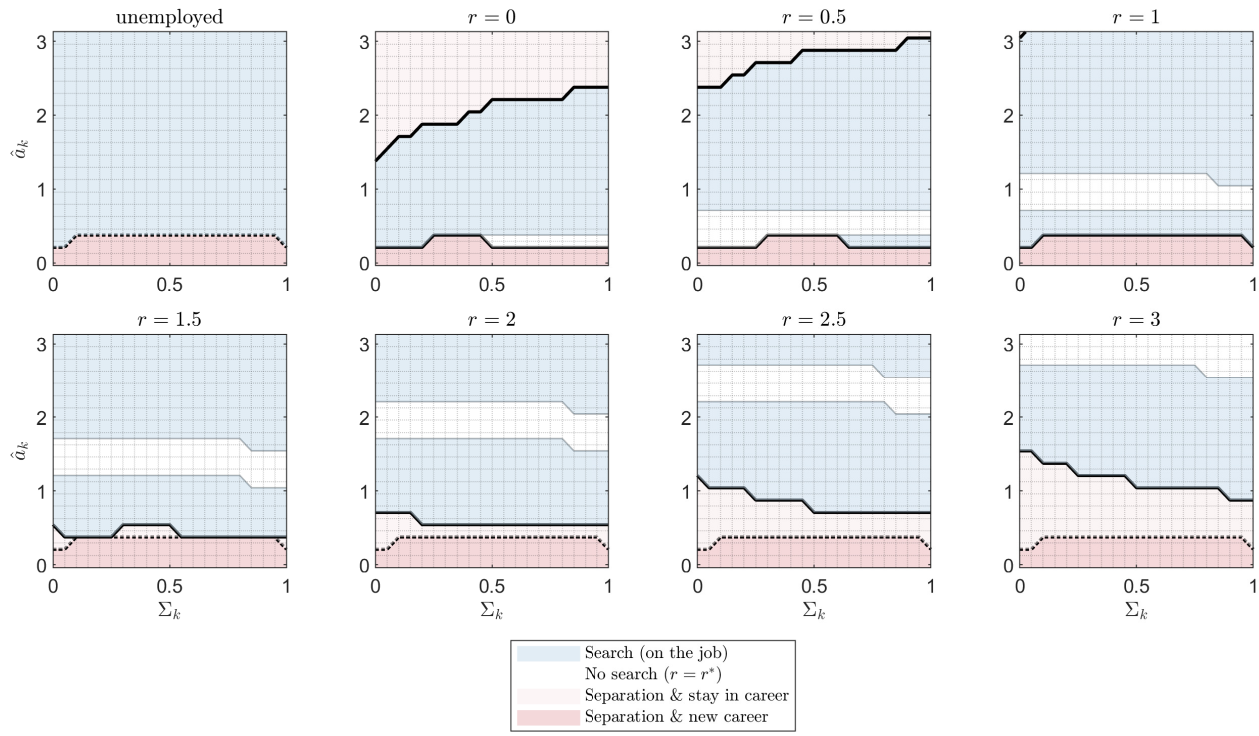


Figure 2: Search and separation policies. *Notes.*—The figure shows search policies as a function of expected ability $\hat{a}_k$, uncertainty Σ_k, and the employment state (unemployed/employed in job with complexity r). Values for $\hat{a}_k$, $\Sigma_k^{1/2}$ and r are denominated in units of $S_0^{1/2}$. The figure is plotted for $z = z_H$. See Section 3 for a detailed description of the parametrization.

employed workers is given by

$$\theta = p'^{-1}\left(\frac{c}{J_t(\hat{a}_k, \Sigma_k, r^*, z) - J_t(\hat{a}_k, \Sigma_k, r, z)}\right). \tag{9}$$

Note that by properties of p the last expression evaluates to zero whenever $r = r^*$. That is, due to bilateral efficiency, employed workers only search for jobs that are better matches (in expectation).

Figure 2 illustrates the search and separation policies of workers as a function of beliefs $(\hat{a}_k, \Sigma_k)$ and current employment status (unemployed or employed in a job with complexity $r \in \mathcal{R}$). Unemployed workers change careers whenever $\hat{a}_k$ is small (indicated by the red area below the dotted threshold). Otherwise they search for jobs in their current career (with a job finding rate that is increasing in $\hat{a}_k$; not indicated in the plot). Employed workers are characterized by a separation threshold (black solid lines), below and above which they separate (with or without career switch[11]). Workers in continuing relationships actively search

[11] Workers separate from their jobs without career switch when the gains from increasing the job finding rate outweigh the cost of unemployment, as observed for workers whose current job rung is far from their

for better matched jobs whenever $r \neq r^*(\hat{a}_k, \Sigma_k, z)$. Specifically, they aspire to *climb down* the job ladder if $\hat{a}_k$ falls into the blue area bordered by the separation region below and the no-search region (in white) above. If $\hat{a}_k$ falls into the upper blue area, they aspire to *climb up* the job ladder instead.

Distributional dynamics The aggregate state in this economy consists of the triplet (z, Γ, Υ), where Γ is the distribution over active worker–firm pairs $(\hat{a}, \Sigma, r)$ and Υ is the distribution over unemployed workers $(\hat{a}, \Sigma)$.[12] Based on the search and separation policies above, we can characterize two Kolmogorov forward equations, one for Γ and one for Υ, which together with the process for z fully describe the dynamics in this economy. While the construction of these equations is standard, their precise expression is slightly protracted. We therefore confine their presentation to Appendix A.

Equilibrium and block-recursivity An equilibrium is a joint value function satisfying equation (6), an unemployed value function satisfying equation (4), lifetime utilities x satisfying the free entry condition (2), and a distribution of worker–firm pairs and unemployed workers evolving according to equations (15) and (16) (stated in Appendix A).

As usual, directed search together with bilateral efficiency and free entry imply that the unique equilibrium is block-recursive (e.g., Menzio and Shi, 2010, 2011; Schaal, 2017). This is because free entry of firms implies that the market tightness in each submarket is only a function of the joint surplus rather than depending on the distribution of workers across submarkets (see equations (8) and (9)). Hence, given that job finding rates are independent of cross-sectional distributions, so are the search problems of workers and the corresponding value functions (3) and (5). Absent any other cross-sectional dependence, we conclude that the only aggregate dependence of $\mathcal{U}$ and J is through z. On this account, we drop the time-subscript t from all value functions going forward.

3 Calibration

This section describes the parametrization of the model. Following the literature, we use a set of standard moments to identify parameters common to labor search models. To inform ourselves about parameters unique to our model, we use a combination of moments

desired one.

[12] Due to the symmetry in k discussed above, there is no need to keep track of the distribution of workers across k separately.

constructed using data from the U.S. Department of Labor's O*NET project together with a worker-level panel from the 1979 National Longitudinal Survey of Youth (NLSY79).

3.1 Measuring Careers and Mismatch in the Data

In the model, careers are each associated with a unique skill type. In the sequel, we argue that when matched with an adequate empirical definition of careers, this simple notion of careers is *isomorphic* to a more general version of our model, in which each job utilizes a mix of different skill types. Specifically, provided that skill-mixes are orthogonal to one another for a given career classification, such a more general model of skill utilization can always be reduced to the simple model introduced in Section 2. Motivated by this observation, we measure career-mobility in the data as job transitions between occupations that are characterized by sufficiently orthogonal skill-mixes based on its O*NET descriptors.

Model-consistent measure of careers To guide our interpretation of the data, consider the following generalization of our model, in which each job utilizes a mix of different skill types. Output per worker–firm pair is given by

$$y_{i,k,r}(t) = F(z(t), \boldsymbol{q}_{k,r}, \boldsymbol{a}_i),$$

where $\boldsymbol{a}_i \equiv (a_{i,1}, \ldots, a_{i,J})$ defines a vector of skills for each worker i over J basic aptitudes. Similarly, $\boldsymbol{q}_{k,r} \equiv r \cdot (w_{k,1}, \ldots, w_{k,J})$ defines a requirement vector over the same aptitudes for a given job. As before, jobs are classified in terms of their task complexity r and a particular skill mix, indexed by $k \in \{1, \ldots, K\}$. The difference is that each k now maps into a vector of weights $(w_{k,1}, \ldots, w_{k,J})$ over the J basic aptitudes, normalized to sum to unity, as opposed to a unique skill type.

The key observation is that—with an appropriate classification of careers—the more general model outlined here can be (approximately) collapsed into the one developed in Section 2. Specifically, to make our simple model consistent with the more general production technology outlined, it suffices to classify occupations into careers so that job requirements $\{\boldsymbol{q}_{k,r}\}$ are (approximately) orthogonal across k.[13] With this in mind, we interpret two occupations observed in the data as different careers if their requirement vectors are "sufficiently orthogonal".

[13] Here we tacitly assume that K is sufficiently large so that workers do not "run out of careers" during their lifetime. We also assume that F collapses to (1) when $\{\boldsymbol{q}_{k,r}\}$ are orthogonal across k. See Appendix B for two examples where skills are perfect complements and perfect substitutes.

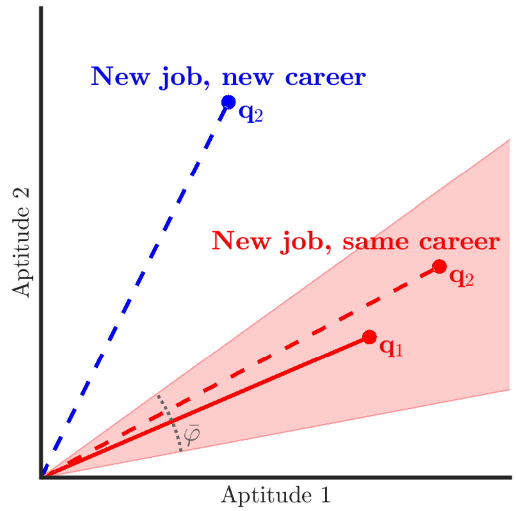


Figure 3: Schematic illustration of empirical measure of careers for $J = 2$. Job transitions from $\boldsymbol{q}_1$ to jobs within the $\bar{\varphi}$-cone are interpreted as transitions up and down the *same* job ladder; transitions to jobs outside the $\bar{\varphi}$-cone are interpreted as career switches.

Specifically, let $\varphi : \mathbb{R}^J \times \mathbb{R}^J \to [0, \pi/2]$, define the angle between two skill vectors $\boldsymbol{q}_1$ and $\boldsymbol{q}_2$,

$$\varphi(\boldsymbol{q}_1, \boldsymbol{q}_2) = \cos^{-1}\left(\frac{\boldsymbol{q}_1 \cdot \boldsymbol{q}_2'}{\|\boldsymbol{q}_1\| \, \|\boldsymbol{q}_2\|}\right).$$

Then any job transition from a job with $\boldsymbol{q}_1$ to a job with $\boldsymbol{q}_2$ is treated as a career switch if and only if $\varphi(\boldsymbol{q}_1, \boldsymbol{q}_2) \geq \bar{\varphi}$ for some $\bar{\varphi}$ (below, $\bar{\varphi}$ is chosen so that the average correlation in requirements for career switches is zero).[14] To account for variations in economic relevance across the J skill dimensions, we weigh them using a set of market weights when computing $\varphi(\boldsymbol{q}_1, \boldsymbol{q}_2)$ in our empirical implementation.[15]

Figure 3 illustrates our empirical approach to measuring career switches for the case where $J = 2$. Starting from job $\boldsymbol{q}_1$, transitions into jobs within the cone defined by $\bar{\varphi}$ (depicted by the red shaded area) are interpreted as transitions up and down the *same* job ladder (i.e., changes in r with a negligible variation in the skill mix k). Transitions to jobs outside the $\bar{\varphi}$-cone are interpreted as career switches (i.e., transitions with a significant change in the skill-mix k). Appendix F.1 provides examples for occupations inside and outside the $\bar{\varphi}$-cones of "economists" and "dental assistants".

[14]See also Gathmann and Schönberg (2011) for a similar approach used to measure occupational distance.

[15]Specifically, let $v_1, \ldots, v_J$ denote a set of weights (further described below). Then $\varphi(\boldsymbol{q}_1, \boldsymbol{q}_2)$ is computed using the weighted dot product $\boldsymbol{q}_1 \cdot \boldsymbol{q}_2' \equiv \sum_j v_j q_{1,j} q_{2,j}$.

Residual correlation in skills across careers We have argued that an orthogonal classification of careers allows for an exact mapping of our model to the data. In Appendix H, we provide evidence that given our classification learning is indeed uncorrelated across careers. Nevertheless, one may ask about the implications if this weren't the case.

In theory, if skills were correlated across careers, workers could partially predict their performance in previously untried careers (although their ability to do so is likely limited in practice[16]). This would allow them to direct their search towards occupations for which they believe to be most qualified. Using the notation of our model, we could capture this by re-interpreting a_0 as the conditional mean of the best-perceived career, and S_0 as the residual uncertainty. As long as skills are not perfectly correlated, the model would still give rise to an increase in uncertainty and mismatch after career switches, not changing its fundamental dynamics. The main addition compared to the uncorrelated skill case would be a likely increase in a_0 (and decrease in S_0) over the life cycle of a worker, reflecting that workers become better at predicting their strengths with additional experience.

Measuring skill requirements and careers Our empirical measure of skill requirements is based on the O*NET project, which describes occupations using a list of 277 descriptors relating to required worker attributes and skills. We follow the literature and reduce the large set of descriptors to $J = 4$ dimensions using Principal Components (Guvenen et al., 2020; Lise and Postel-Vinay, 2020), which we interpret as mathematics, verbal, social, and technical skills.[17] To make them comparable, we normalize each skill dimension in terms of percentile ranks.[18] Appendix E describes the data in more detail.

To identify career moves, we merge our skill measures with the NLSY79. Let $\boldsymbol{q}_{i,t} = (q_{i,t,1}, \ldots, q_{i,t,4})$ denote the four-dimensional skill measure associated with the job held by worker i at date t.[19] As detailed above, we associate a job transition from $\boldsymbol{q}_{i,t}$ to $\boldsymbol{q}_{i,t+1}$ with a career switch if the angle between the two skill vectors, $\varphi(\boldsymbol{q}_{i,t}, \boldsymbol{q}_{i,t+1})$, is larger than $\bar{\varphi}$. The

[16] In practice, the ability of a worker to predict performance across careers is likely to be impaired by a lack of information regarding the precise importance of skills in each career. For instance, suppose skills are perfect substitutes as in (18) in Appendix B; i.e., skills enter production through the linear index $\boldsymbol{w}_k \boldsymbol{a}_i'$. In this case, learning about the linear index $\boldsymbol{w}_k \boldsymbol{a}_i'$ is a sufficient statistic for the current career, but cannot easily projected across careers without knowing both $\boldsymbol{w}_k$ and $\boldsymbol{w}_{k'}$.

[17] Guvenen et al. (2020) and Lise and Postel-Vinay (2020) reduce worker requirements to only three dimensions. We add the technical component as it has been shown to be an important determinant for labor market outcomes (Prada and Urzúa, 2017).

[18] To make our measure of skill requirements comparable with our measure of worker skills (described below), we compute the percentile ranks based on the distribution of requirements among jobs observed in the NLSY79 sample.

[19] We map 2010 SOC codes used by O*NET to classify occupations into Census codes used by NLSY79 using standard crosswalk files.

threshold $\bar{\varphi}$ is chosen so that the average correlation in requirements (across skill dimensions) is zero for career moves: $\sum_{j=1}^{4} v_j \text{Corr}(q_{i,t,j}, q_{i,t+1,j}) = 0$, where $\{v_j\}$ is a set of market weights described below.[20] Using this strategy, we set $\bar{\varphi} = 14.8°$ which implies that 42.1 percent of all job transitions in the NLSY79 sample are career switches. The propensity to switch careers is comparable to the numbers obtained by Fujita and Moscarini (2017), Carrillo-Tudela et al. (2016), Carrillo-Tudela and Visschers (2020), and Huckfeldt (2021).

Measuring worker skills and mismatch Following Guvenen et al. (2020) we define mismatch based on the absolute difference in skill requirements and worker skills. For this purpose, we measure worker skills based on six ASVAB scores available from the NLSY79 sample, individual scores on the Rotter locus-of-control scale, and the Rosenberg self-esteem scale. We follow a similar procedure as for skill requirements to reduce those scores into a four-dimensional measure of worker abilities in math, verbal, social and technical skills.

Let $\boldsymbol{a}_i = (a_{i,1}, \ldots, a_{1,4})$ denote the skill vector of worker i. The mismatch between worker i and their current occupation is then given by:

$$m_{i,t} \equiv \sum_{j=1}^{4} v_j \left| a_{i,j} - q_{i,t,j} \right| . \tag{10}$$

Here v_j are "market weights", obtained from the regression coefficients on each of the four mismatch dimensions in a Mincer regression (normalized so $\sum_{j=1}^{4} v_j = 1$).[21] The weights ensure that our mismatch measure is not driven by skills that are economically irrelevant. Similarly, we define positive mismatch, measuring overqualification, and negative mismatch, measuring underqualification, as

$$m_{i,t}^{+} \equiv \sum_{j=1}^{4} v_j \max\{a_{i,j} - q_{i,t,j}, 0\} \qquad m_{i,k}^{-} \equiv \sum_{j=1}^{4} v_j \max\{q_{i,t,j} - a_{i,j}, 0\}.$$

3.2 Parametrization of the Model

Assigned parameters We parametrize the model at a monthly frequency. The discount rate ρ is set to $\log(1.05)/12$ corresponding to an annual discount rate of 5%.

The relative search intensity of employed workers, κ, is set to 0.5, consistent with the

[20]The zero correlation in skills target for career-switchers contrasts strongly with an average correlation of .89 among job-switchers that are classified as within-career transitions.

[21]Specifically, we regress $\log \text{wage}_{i,t}$ on math, verbal, technical, and social mismatch, controlling for a quadratic polynomial in age and worker fixed effects. The resulting weights are .58, .14, .09, .19 for math, verbal, technical, and social, respectively.

relative search effort documented in Holzer (1987) and Faberman et al. (2017).[22] We choose to set the relative search intensity κ based on direct evidence as opposed to targeting the job-to-job rate, because job-to-job transitions are clearly caused by many factors not present in the model, including relocation shocks, rent seeking purposes, and random fluctuations in match-quality. If we would force the model to match the empirical job-to-job rate, we would effectively require learning about skills to account for these other forces, overstating the importance of learning for job-to-job mobility.[23]

We specify the set of potential task-complexities, $\mathcal{R}$, using a seven-point grid given by $\{0, 0.5, \dots, 3\} \cdot S_0^{1/2}$, denoted in standard deviations of a_k. The boundaries of the grid are chosen so that adding additional grid points has no impact on the results.[24] We approximate beliefs about worker skills using a 61-point grid for $\hat{a}_k$ on $[-3, 7] \cdot S_0^{1/2} + a_0$ and a 21-point grid for Σ_k on $[0, 1] \cdot S_0$. Finally, we normalize log productivity in recessions to 0, and choose transition rates for z in order to match the monthly switching intensities between recessions and expansions in the U.S., where recessions are periods with an unemployment rate above its unconditional average of about 6.5%.

Target moments We calibrate the remaining parameters using the method of moments with weights chosen to minimize the relative distance between model and empirical moments. All model moments are computed at the ergodic distribution. As usual, all parameters are identified jointly. In the following we provide a heuristic mapping from moments to parameters to guide intuition.

Following the literature, we target worker flows in and out of unemployment as documented by Shimer (2012) to identify the exogenous separation rate δ and the flow cost of vacancy creation c. We identify b by targeting a replacement ratio of $b/\mathbb{E}[y]$ equal to .71 as found by Hall and Milgrom (2008). Following Menzio and Shi (2010) and Schaal (2017), we choose CES contact rate functions, $p(\theta) = (1 + \theta^{-\gamma})^{-1/\gamma}$ and $q(\theta) = (1 + \theta^{\gamma})^{-1/\gamma}$. The matching function parameter γ is set to match an elasticity of UE flows with respect to the aggregate vacancy–unemployment ratio of .28 as estimated by Shimer (2005). Finally, we identify z_H (relative to z_L) from an average recession–expansion difference in unemployment amounting to 2.8 p.p. in the US.

To identify the speed of learning, parametrized by σ, we target an average slope of the

[22]Conditional on searching for jobs, Holzer (1987) and Faberman et al. (2017) document a relative time spent on search activities among employed workers of 0.48 and 0.51, respectively.

[23]In our calibration, the monthly job-to-job worker flows are .021.

[24]Adding an extra grid point at $-0.5 \cdot S_0^{1/2}$ has no effect as no search is directed to such submarkets in our calibration. Adding an extra grid point at $3.5 \cdot S_0^{1/2}$ does not change the results as it attracts only a negligible mass of 0.005 workers at the ergodic distribution.

Table 1: Targeted moments

Fitted Moments	**Data**	**Model**	**Origin**
$\mathbb{E}_L[U] - \mathbb{E}_H[U]$	.028	.027	BLS
$\mathbb{E}[\text{UE rate}]$	.425	.424	Shimer (2012)
$\mathbb{E}[\text{EU rate}]$	.035	.035	Shimer (2012)
$b/\mathbb{E}[y]$	.710	.708	Hall and Milgrom (2008)
$\epsilon_{UE/\theta}$	.280	.281	Shimer (2005)
$\mathbb{E}[\log(haz_3/haz_{18})]$	1.37	1.37	NLSY79
$\mathbb{E}[\chi = 1]$	.422	.421	NLSY79, O*NET
$\mathbb{E}_L[\chi = 1] - \mathbb{E}_H[\chi = 1]$	.069	.065	NLSY79, O*NET
$\mathbb{E}_{NC}[m^-]$	.096	.096	NLSY79, O*NET
$\mathbb{E}_{NC}[m^+]$	.201	.201	NLSY79, O*NET

Notes.—The notation $\mathbb{E}[\cdot]$ denotes unconditional expectations, computed at the ergodic distribution of the model. $\mathbb{E}_L[\cdot]$, $\mathbb{E}_H[\cdot]$ and $\mathbb{E}_{NC}[\cdot]$ denote expectations conditional on the aggregate state being in a recession and expansion, and conditional on the first job in a new career. U denotes the aggregate unemployment rate, EU and UE are monthly transition rates, y is output per worker–firm pair, $\epsilon_{UE/\theta}$ is the elasticity of the UE rate with respect to the aggregate vacancy–unemployment ratio, haz_x is the separation hazard after x months of employment, χ is an indicator evaluating to unity if workers switch careers during a job transition (this includes both EE' and EUE' transitions), and m^- and m^+ denote negative and positive mismatch.

empirical separation hazard between the 3rd and 18th month of employment, $\log(haz_3/haz_{18})$, of 1.37 as found in the NLSY79 sample.[25] Intuitively, a high speed of learning (low values of σ) allows worker–firm pairs to quickly identify whether a match is profitable, implying a steep decline in the separation hazard over time. By contrast, if learning is slow, worker–firm pairs will keep revising their beliefs for a prolonged time, reflected in a flattening of the hazard curve.

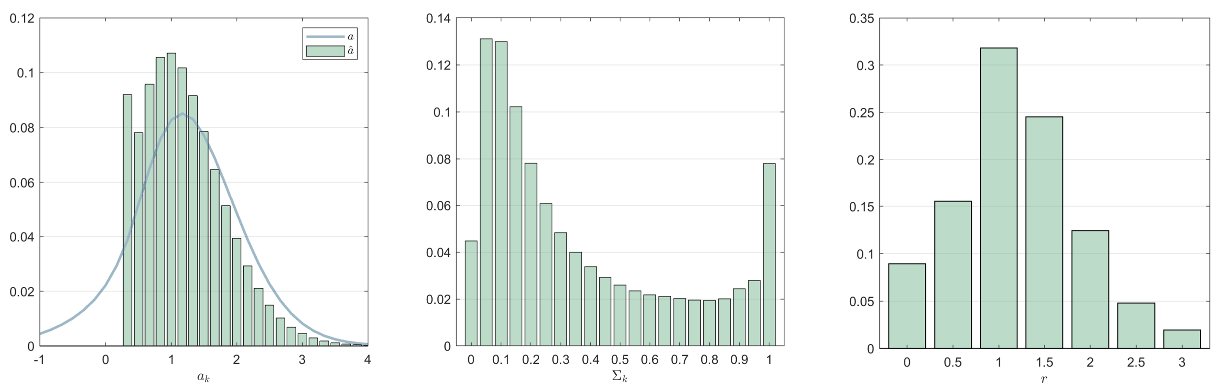
Next, we use the arrival rate of exogenous career-shocks, ϵ, to ensure consistency of the model with an average propensity to switch careers of 42.1 percent, as documented above in the NLSY79. Relatedly, we use the technology parameter η to match the empirical cyclicality in career mobility, which we find to be 6.9 percentage points higher in recessions compared to expansions.

Finally, to identify the prior mean and variance of skills, a_0 and S_0, we match the positive and negative mismatch of workers in the first job of a new career. This captures that total mismatch in the first job after a career-switch is closely linked to the prior uncertainty S_0, whereas the ratio between over- and underqualification pins down a_0 relative to the "entry"

[25]We measure the slope starting after the 3rd month of employment as the first three months are often subject to explicit or implicit probationary agreements, which the model abstracts from.

Table 2: Summary of parameters

Parameter	Description	Value
Assigned		
ρ	Monthly discount rate	$\log(1.05)/12$
κ	Relative search intensity of employed	0.5
z_L	Aggregate log-productivity in recessions	0
$\lambda_{z_L}, \lambda_{z_H}$	Poisson rates of productivity shock	0.0128, 0.0172
Estimated		
z_H	Aggregate log-productivity in expansions	0.290
b	Home production utility	0.981
c	Flow cost of vacancies	0.007
γ	Matching function parameter	0.510
η	Return on task complexity	0.497
a_0	Unconditional mean of skills	0.105
$S_0^{1/2}$	Standard deviation of skills	0.357
σ	Standard deviation of signal noise	2.505
δ	Exogenous separation rate	0.012
ϵ	Exogenous career switching rate	0.003



(a) Skills **(b)** Uncertainty **(c)** Job rungs

Figure 4: Ergodic distribution of individual state variables. *Notes.*—Values for a_k, $\hat{a}_k$, $\Sigma_k^{1/2}$ and r are denominated in units of $S_0^{1/2}$.

job rung $r^*(a_0, S_0, z)$. We note that according to the data, workers starting a new career are on average overqualified.

Estimation results Table 1 reports the data targets alongside the corresponding moments in the calibrated model. The model fits the data almost perfectly.

The calibrated parameters are listed in Table 2. Figure 4 shows the implied ergodic distribution of individual state variables: mean beliefs about the currently employed skill (along with their true realization a), uncertainty, and task complexities. The distribution of mean beliefs is censored slightly below a_0, reflecting the option to switch careers whenever workers become pessimistic about their skills. Moreover, comparing the distribution of $\hat{a}_k$ with the true distribution of currently pursued skills a_k, the latter is more dispersed, especially around $\hat{a}_k = a_0$. This is because uncertainty is highest at the beginning of a career and is negatively correlated with $|\hat{a}_k - a_0|$ as extreme belief revisions are more likely the more information is observed.

The distribution of uncertainty is visibly right-skewed with a median uncertainty of $0.25 \cdot S_0$ and a mean of $0.36 \cdot S_0$. Not surprisingly, however, despite the overall right-skew, the distribution of Σ_k also has a concentration of mass at $\Sigma_k = S_0$, reflecting the reset in learning after workers switch careers.

Finally, the distribution of job rungs is hump shaped, with a median job rung of $1 \cdot S_0^{1/2}$ and a mean of $1.2 \cdot S_0^{1/2}$.

4 Dynamics at the Worker-level

We are now ready to study the equilibrium allocation of workers to jobs and how it evolves over time. In this section, we do so, focusing on the micro-level dynamics of workers. We begin with a random simulation illustrating the labor market dynamics of a single worker. Next, we highlight how workers' career choice and their progression through the job ladder are both shaped by considerable inertia. Finally, we show how this inertia carries over to earnings and generates a significant unemployment scar after job displacement.

4.1 Sample Path for a Single Worker

In the model, the allocation of workers to jobs is governed by an interaction of learning, career choice, and workers' progression through job rungs. Figure 5 illustrates this interaction by simulating a 10-year sample path for a single worker, while keeping the aggregate state fixed at $z = z_L$. There are no exogenous separation nor displacement shocks realized throughout the path. At $t = 0$, the worker is unemployed and initial beliefs are $(\hat{a}_k, \Sigma_k) = (a_0, S_0)$.

Given the initial belief, the worker directs their search at $t = 0$ towards the bottom job rung ($r = 0$). Once matched, they start revising their belief, resulting in declining uncertainty (fourth panel), revisions to their mean estimate (solid black line in the first panel), and revisions to expected mismatch (second panel). Over time, these revisions lead to a reallocation in jobs via on-the-job search, job separations and career changes.

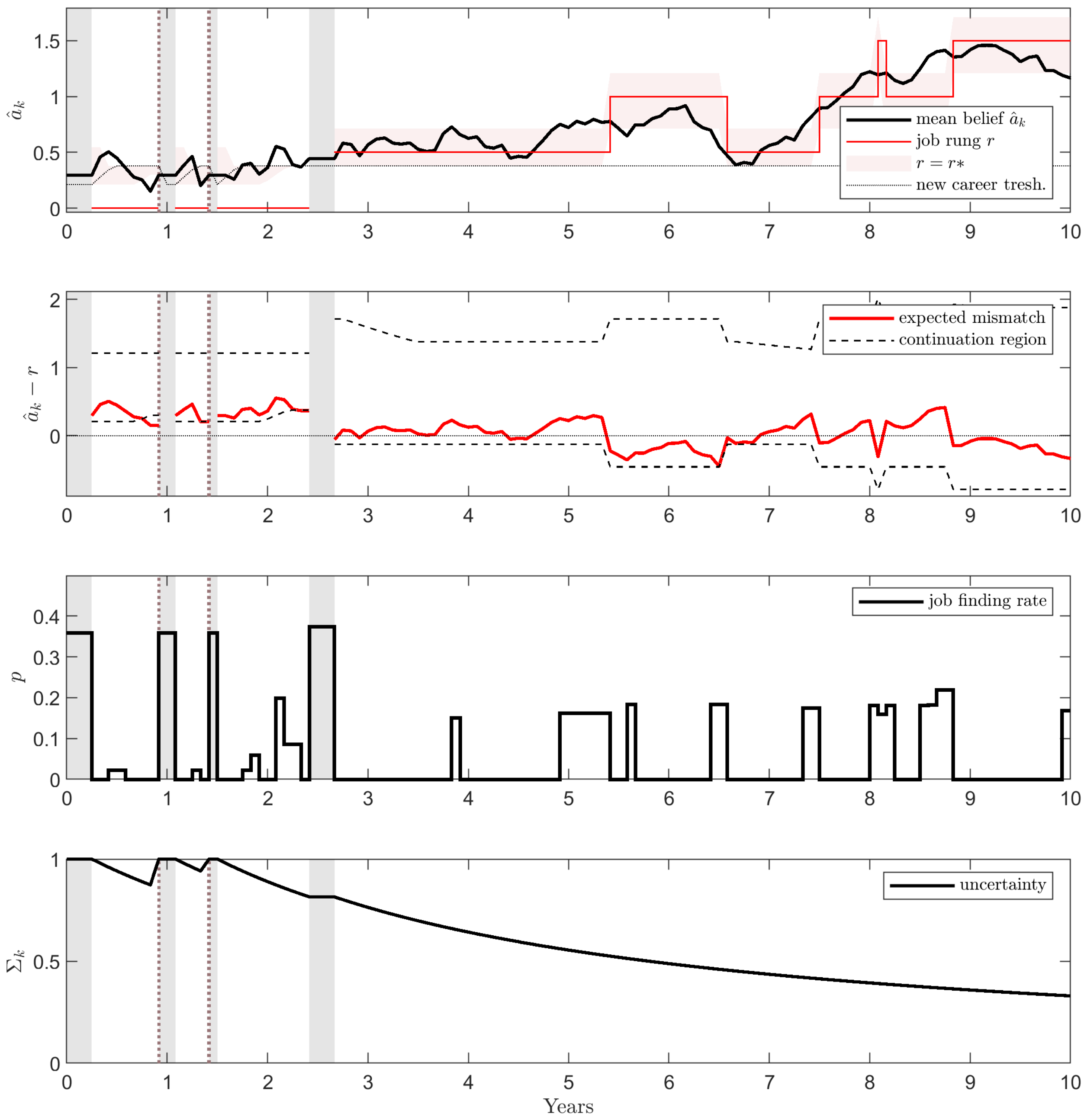


Figure 5: Sample path for a single worker. *Notes.*—The figure shows a random career path for a single worker, initialized without a job and with $(\hat{a}_k, \Sigma_k) = (a_0, S_0)$. Throughout, the aggregate state is fixed at $z = z_L$. Vertical gray bands depict unemployment spells. Vertical purple dotted lines depict career switches.

Specifically, the worker engages in on-the-job search whenever their desired job rung r^* differs from the current job rung r (red line in panel 1). Graphically, on-the-job search episodes occur whenever the mean belief falls outside the lightly red shaded bands in panel 1 (which indicate that $r = r^*(\hat{a}_k, \Sigma_k, z)$). For instance, starting at about 5 years, the upward-revision in $\hat{a}_k$ leads the worker to attempt to climb up the job ladder, which they succeed at about 5.5 years. Further successful job-to-job transition can be seen in years 6–9, during which the worker goes through an additional five job-to-job transitions (inducing changes in the job rung as seen in the first panel).

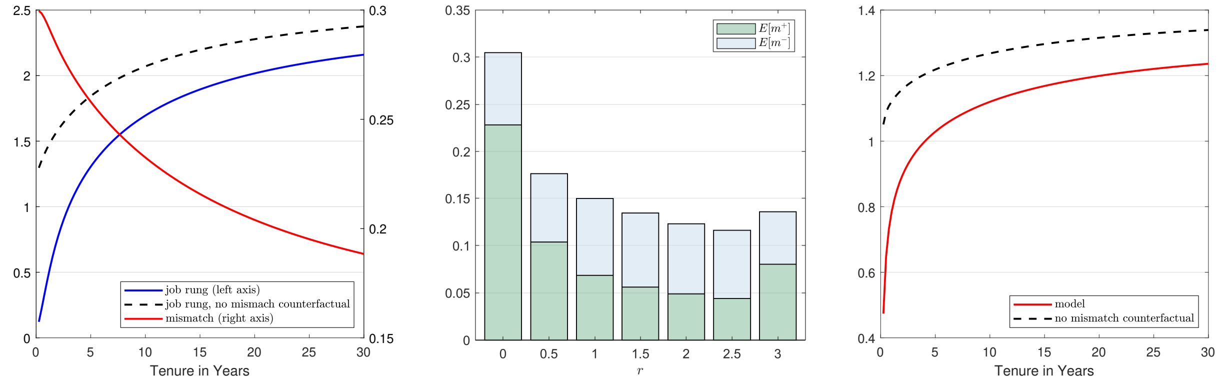
Endogenous job separations occur whenever mismatch falls outside the black dashed lines in the second panel, as observed after about 0.9, 1.4 and 2.4 years.[26] Once the worker is unemployed, they direct their search towards a new career whenever $\hat{a}_k$ falls below the thin dotted threshold in panel 1, as observed for the first two of the three unemployment spells (indicated by the vertical purple dotted lines at the beginning of the corresponding unemployment spell). In these cases, the belief resets to $(\hat{a}_k, \Sigma_k) = (a_0, S_0)$, and the worker directs their search to the bottom job rung of the new career. By contrast, the third separation after 2.4 years occurs because the gains from climbing the job ladder are sufficiently large so that increasing the job finding rate (by moving to unemployment) outweighs the cost of being temporarily unemployed. During this final unemployment spell the worker hence directs their search to a higher job rung within the same career.

4.2 Inertia in Job Rungs, Mismatch and Earnings

The sample path in Figure 5 demonstrates that the allocation of workers to jobs is subject to inertia, both within and across careers. The inertia reflects, on the one hand, the time needed to learn about any subsisting mismatch and, on the other hand, its slow dissolution due to search frictions. We next explore the consequences of this inertia for workers' progression through job rungs, mismatch and earnings.

Inertia in job rungs and mismatch We begin by highlighting inertia in workers' progression through job rungs. Figure 6a plots the average job rung as a function of workers' tenure in a given career. The average job rung increases in tenure for two reasons: (i) the climbing of the job ladder of high ability workers; (ii) the selection out of a carrier by low ability workers. Both forces are subject to inertia. Absent frictions, workers would always pursue a career with $a_k \geq r^{\max}$ and would always be employed at the top job rung $r^{\max} = 3 \cdot S_0^{1/2}$, yielding a

[26]The separation thresholds can be equivalently expressed in terms of $\hat{a}$ as we have done in Figure 2.



(a) Job rung and mismatch **(b)** Mismatch by job rung **(c)** Earnings

Figure 6: Relationship between tenure, job rungs, mismatch, and earnings. *Notes.*—Tenure is defined as the time since the last career change. The no mismatch counterfactuals show fictitious profiles for average job rungs and earnings where the distribution of abilities evolves as in equilibrium, but where $r = r^*(a, 0, z)$ at all times. All conditional expectations are computed at the ergodic distribution.

flat relationship between job rungs and tenure.[27] This starkly contrasts with the *slow* climb through the job rungs seen in Figure 6a.

To assess the relative importance of the two sources of inertia, we contrast the model's evolution of job rungs with a counterfactual where there is no mismatch conditional on skills; that is, $r = r^*(a, 0, z)$. To make the counterfactual comparable, we evaluate it using exactly the same distribution of skills (conditional on tenure) as emerges in equilibrium. By construction, the counterfactual only reflects the selection effect, which in our calibration explains about 50 percent of the overall increase in job rungs with tenure.

The slow reallocation of job rungs immediately translates to mismatch being persistent as well. Moreover, as seen in Figure 6b, this naturally translates into a negative correlation between mismatch and job rungs.[28] Interestingly, despite the overall decline in mismatch across job rungs, there is a relative increase in the contribution of underqualification among higher job rungs, which is driven by the diminishing option value of being overqualified as discussed in the context of Figure 1.

Inertia in earnings Having documented inertia in job rungs and mismatch, we next look at its impact on earnings. Because wages are not uniquely determined by the bilaterally

[27] Career mobility is subject to inertia because evaluating the prospects of a career takes time due to the information friction and reduces the returns to trying out new careers given the anticipation of mismatch. In Appendix C, we assess the cost of this implicit friction, finding that on average it amounts to 4.75 months of average output per worker.

[28] There is a slight increase in mismatch at the highest job due to an increase in "overqualification" among workers whose skills exceed the top job rung $r^{\max}$.

efficient labor contracts explored so far, we first have to take a stand on the wage arrangement that firms use to deliver a worker's promised lifetime utility x. We do so by following Schaal (2017) and choosing the unique wage scheme under which employed workers find it sequentially optimal to pursue the contracted continuation and search policies, even in the absence of any contractual commitment. The unique wage arrangement with this properties effectively pays workers their expected marginal product, adjusted for the cost of recruitment which is loaded onto workers at the instant of hiring (see Appendix D for details).

Earnings are inversely related to mismatch through its adverse impact on labor productivity. For underqualified workers this is due to the direct penalty on production. For overqualified workers this is due to the opportunity cost of operating a task complexity that is too low. In either case, earnings are again subject to strong inertia in both the reallocation of job rungs within career and an inefficiently low propensity to switch careers. Figure 6c plots the resulting earnings profile in tenure along with the no mismatch counterfactual.[29] The slow climb through the job ladder gives rise to a steep wage ladder that spans many years. Over the course of the first 10 years, earnings increase by a factor of 2.4, of which about 70 percent are explained by a reduction in mismatch and about 30 percent by a shift in skills.

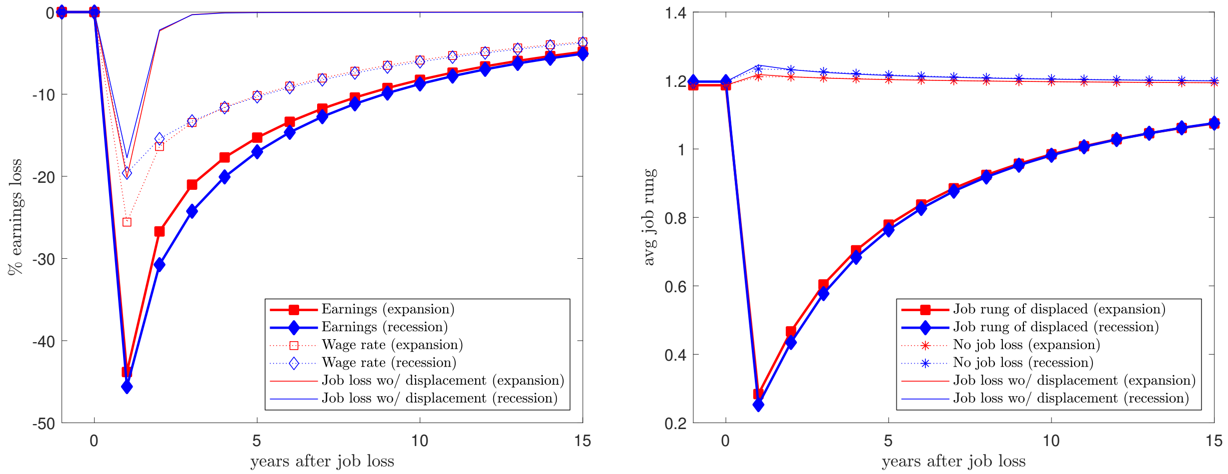
4.3 Scarring Effect of Unemployment

Previous literature has documented a large and persistent impact of involuntary job loss on future wages and earnings (e.g., Davis and von Wachter, 2011; Jarosch, 2021), especially when the job loss is accompanied by occupational displacement (Huckfeldt, 2021). In this section, we offer a narrative for the "scarring effect of unemployment" based on the inertia in mismatch and earnings. In line with the evidence in Huckfeldt (2021), earnings losses are in large parts realized through wage losses and are concentrated among workers that are separated from their job *and* are displaced from their career.

Career displacement vs. job loss Figure 7a shows the earnings and wage path of a worker with at least 3 years prior job tenure that is displaced from their current career at $t = 0$, conditional on the business cycle state at $t = 0$.[30] Relative to the counterfactual of no job loss, earnings are reduced by roughly 47 percent one year after the displacement, and

[29] To increase comparability, we keep both the distribution of abilities and recruiting costs fixed at their equilibrium level. That is, counterfactual earnings are reduced by the same recruiting cost as in the model, so that the difference in earnings solely reflects the increase in labor productivity due to a lack of mismatch.

[30] The restriction to workers with 3 years of prior job tenure parallels the selection made by Davis and von Wachter (2011) and Jarosch (2021) in their empirical studies. Absent the tenure requirement, the earnings loss from displacement amounts to 44, 16 and 8 percent after one, five and ten years, respectively.



(a) Average earnings losses of displaced workers, relative to control group

(b) Average job rung of displaced workers and control groups

Figure 7: Response to a displacement shock. *Notes.*—Panel (a) shows the earnings and wage losses by workers displaced from their career at $t = 0$, conditional on the business cycle state at $t = 0$. All responses are as a percentage relative to the counterfactual of no job loss. Panel (b) shows the corresponding average job rungs (in units of $S_0^{1/2}$) for workers displaced from their career, workers separated from their job without career displacement, and workers without job loss.

continue to be depressed by about 19 percent five years later, and by about 10 percent ten years later. While initially a significant share of the earnings loss is explained by a slow rate of reemployment (after one year, 36 percent of the workers displaced during recessions and 28 percent of the ones displaced during expansions are unemployed), most of the long-run "scar" is due to a persistent decline in wages.

The reason behind this long-run "scar" on wages is that displaced workers—who previously occupied jobs at all rungs of the job ladder— must rebuild their careers in new sectors, which is subject to inertia as described above. Figure 7b illustrates this by plotting the average job rungs of displaced workers in the sequel to their job loss. While workers that are separated from their job without career displacement are able to immediately re-enter the labor market at their previous job rungs (with little consequences for earnings[31]), workers that are displaced from their career enter the labor market at the bottom job rung and take years to advance (on average) to their previous rungs. The prolonged impact of this long climb through the job rungs on earnings is able to account for the evidence in the literature, which 5–10 years after displacement documents earnings losses relative to the control group ranging from 5–10 percent (Davis and von Wachter, 2011; Huckfeldt, 2021) to 15–20 percent (Jarosch, 2021).

[31]There is a small and temporary earnings loss for workers that are separated without career displacement due to the job loss itself and the recruiting cost that is loaded onto starting wages.

Table 3: Cyclicality of mismatch in the model

Mismatch measure	$m_{i,t}$	$m^-_{i,t}$	$m^+_{i,t}$	$e_{i,t}$
Cyclical difference (in %)	-0.46	-2.58	1.53	0.36

Notes.—The table reports the difference in conditional means between recessions and expansions, computed at the ergodic distribution and denominated in percent of the ergodic mean.

5 Aggregate Fluctuations in Mismatch

In this section, we study the macro-dynamics of mismatch and its implications for aggregate productivity. We also present reduced-form evidence on the cyclicality of mismatch in the data.

5.1 Mismatch Cycles in the Model

5.1.1 Aggregate Productivity Shocks

We begin by computing the cyclical difference in mismatch, defined by the difference in conditional means between recessions and expansions, $\mathbb{E}_L[\cdot] - \mathbb{E}_H[\cdot]$. Table 3 reports the results. The model predicts procyclical fluctuations in underqualification (negative mismatch being 2.58 percent lower in recessions than in expansions) and countercyclical fluctuations in overqualification (positive mismatch being 1.53 percent higher in recessions than in expansions). Combined, total mismatch is mildly procyclical, being on average 0.43 percent lower in recessions than in expansions.

To assess the impact of these mismatch fluctuations on output, we compute the component of labor productivity affected by the endogenous selection into job rungs,

$$e_{i,k,r}(t) \equiv \exp\left(\eta r - \max\left\{r - a_{i,k}, 0\right\}\right),$$

which we call "labor efficiency". Conditional on skills, labor efficiency is decreasing in both over- and underqualification, capturing the impact of both types of mismatch in *output units*.[32] Using this measure, we find that the decline in mismatch translates to an increase in average labor efficiency by 0.36 percent in recessions compared to expansions.

Labor efficiency vs labor productivity We note that the countercyclicality in labor efficiency does not immediately translate into predictions regarding aggregate labor productivity.

[32]Labor efficiency is decreasing in both types of mismatch, reflecting the direct impact on labor productivity (if workers are underqualified) and the opportunity cost of choosing a task complexity that is too low (if workers are overqualified).

To draw inference about aggregate labor productivity, we first need to take a stand on the nature of the aggregate "productivity" shock z. One possibility is the literal interpretation as shock to productive efficiency. In this case overall labor productivity is given by $\exp(z)\mathbb{E}[e_{i,t}]$, which is procyclical in our calibration. However, owing to the partial equilibrium nature of the model, we can alternatively interpret z as a demand shock to the real price of labor output.[33] In this case, aggregate labor productivity is entirely determined by the endogenous labor efficiency $\mathbb{E}[e_{i,t}]$ and is hence countercyclical.

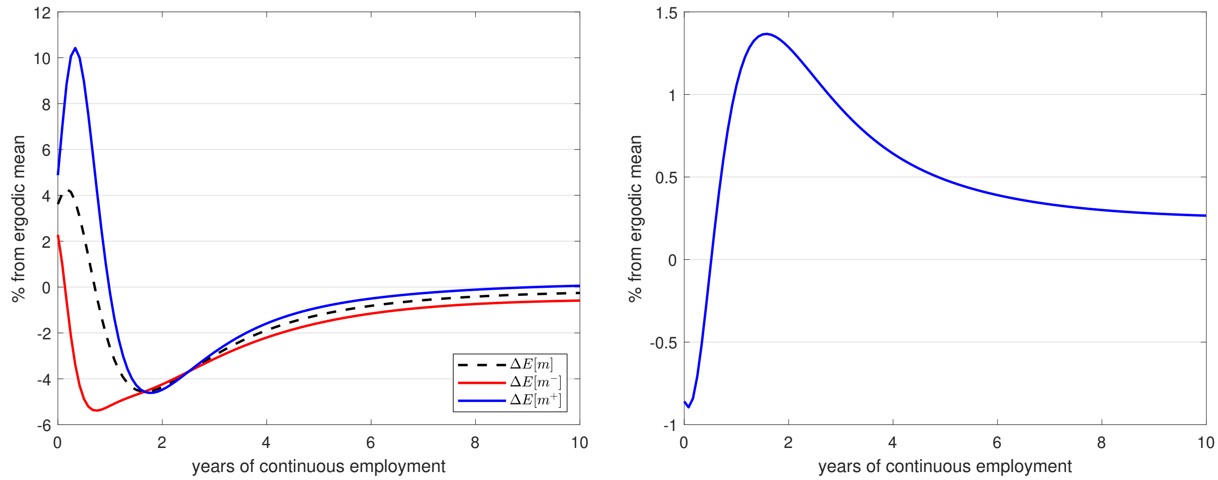
This flexibility in interpreting z suggests a new narrative for the "labor productivity puzzle"; namely the fact that labor productivity has become less procyclical in the U.S., and actually rose in 2008-09 during the Great Recession (e.g., Mulligan, 2011; McGrattan and Prescott, 2012; Gali and van Rens, 2021). Through the lens of the model, we would precisely expect such development when productivity shocks are diminishing and business cycles have become increasingly demand-driven, consistent with findings in Hazell et al. (2020) as well as with the household balance sheet narrative of the Great Recession (Mian, Rao and Sufi, 2013).

Cyclicality by employment tenure The overall cyclicality in mismatch is the result of opposing effects operating at different tenure levels. To isolate these effects, we next break down the cyclicality by the time a worker has been continuously employed since their last unemployment spell ("employment tenure"). Figure 8 shows the decomposition.[34] Among new hires, both over- and underqualification is significantly heightened in recessions. Combined, total mismatch among new hires increases by about 4 percent in recessions, which translates to a 0.9 percent decline in labor efficiency.

The increase in mismatch among new hires contrasts starkly with the cyclicality in mismatch among workers with an employment tenure of more than 6 months, for which mismatch is reduced in recessions, raising their labor efficiency. For instance, workers that have been continuously employed for 2 years have an average labor efficiency that is 1.3 percent higher in recessions compared to expansions.

[33] Here we tacitly assume that the real price of labor output fluctuates relative to b and c, either because b and c are defined in real terms as in Walsh (2005) and Christiano, Eichenbaum and Trabandt (2015, 2016), or because of sectoral heterogeneity.

[34] Formally, we again compute the cyclical difference in mismatch as the difference in conditional means, but now do so conditional on employment tenure τ, resulting in $\mathbb{E}[\cdot|z_L, \tau] - \mathbb{E}[\cdot|z_H, \tau]$. The conditioning on employment tenure is imposed at the same instant in which we condition on the business cycle state and does *not* fix the business cycle state at the instant of hiring.



(a) Mismatch

(b) Labor efficiency

Figure 8: Mismatch cyclicality by employment tenure. *Notes.*—The figure plots the cyclicality of m, m^-, m^+ and e conditional on being continuously employed for τ years; i.e., $\Delta\mathbb{E}[\cdot|\tau] = \mathbb{E}[\cdot|z_L, \tau] - \mathbb{E}[\cdot|z_H, \tau]$. All cyclical differences are denominated in percentage deviations from their ergodic mean $\mathbb{E}[\cdot]$.

Understanding the mechanism: cleansing and sullying We next explore the forces driving mismatch in the model. Figure 9 shows the impact of the business cycle on separation policies and career mobility. When moving from an expansion to a recession, the continuation region shrinks, resulting in a "cleansing" of workers. Quantitatively, most of this cleansing is concentrated towards workers close to the lower separation threshold, especially at the bottom job rungs. This is because workers generally attempt to resolve expected mismatch via on-the-job search so that in equilibrium little mass is actually distributed across the cleansing region. The one exception to this is workers with skill estimates below the career-switching threshold. These workers anticipate to change careers once they lose their jobs, but hold on to their job as long as they can to avoid the utility loss of becoming unemployed. Because these workers do not engage in on-the-job search, they make up most of the mass inside the cleansing region (84 percent at the ergodic distribution). Figure 10a breaks down the distribution of workers inside the cleansing region by job rungs, confirming that virtually all of the cleansing is concentrated at the bottom job rungs.

Figure 10b further decomposes workers cleansed out in recessions by the relative prevalence of positive and negative mismatch, $\mathbb{E}[m^+]/\mathbb{E}[m]$ and $\mathbb{E}[m^-]/\mathbb{E}[m]$. Workers cleansed from the bottom job rung are more likely to be overqualified, even at the lower separation threshold, reflecting the general tendency to be overqualified at the bottom job rung (c.f. Figure 1). However, even though these workers are *expected* to be overqualified, their skill estimate is surrounded by enough uncertainty so that there is also significant underqualification among

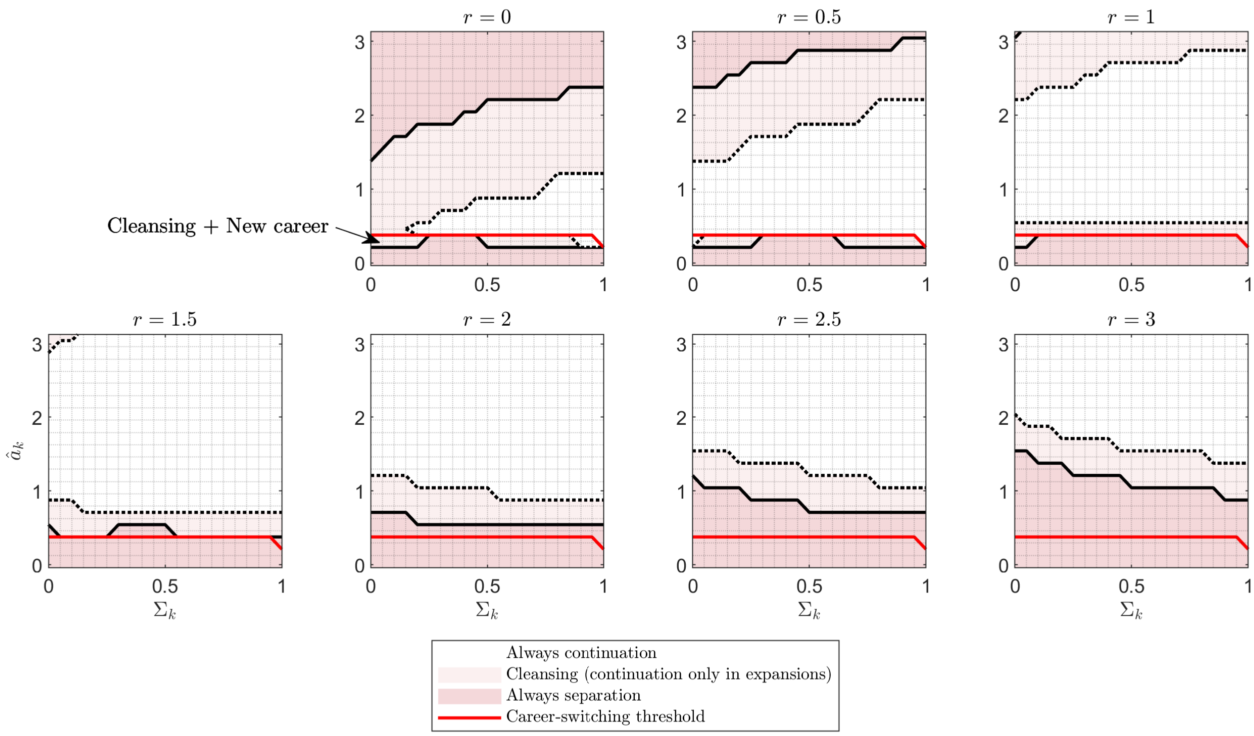
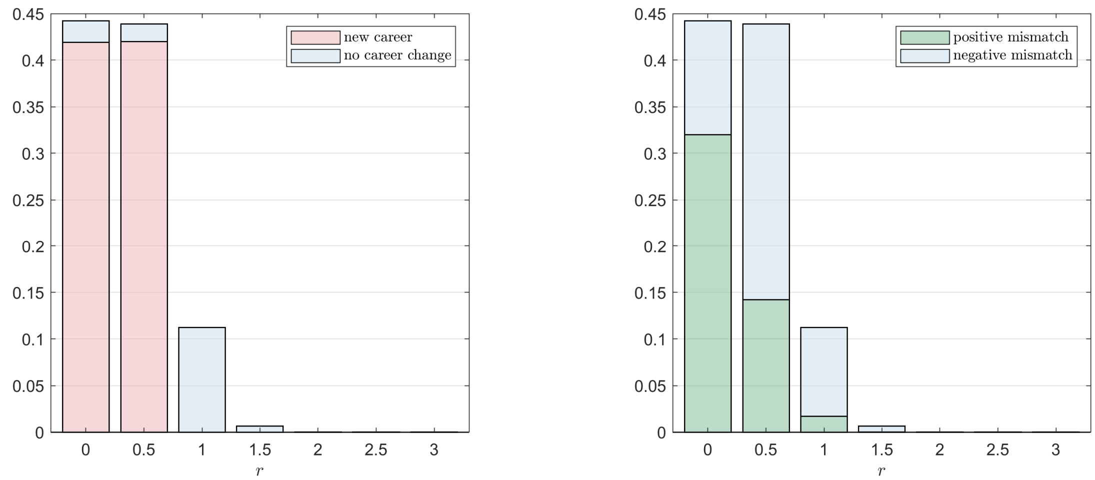


Figure 9: Separation policies: expansions vs recessions. *Notes.*—The figure shows how the continuation region contracts during recessions. The career-switching threshold is plotted for $z = z_L$. Values for $\hat{a}_k$, $\Sigma_k^{1/2}$ and r are denominated in units of $S_0^{1/2}$.

workers cleansed from the bottom job rung. Similarly, workers cleansed from higher job rungs are more likely to be underqualified but again with significant uncertainty. Importantly, regardless of the type of mismatch, mismatch is overall more pronounced among cleansed workers than on average: evaluated at the ergodic distribution, positive mismatch is 50 percent higher among cleansed workers compared to the average worker in an expansion, and negative mismatch is 59 percent higher. This explains the decline in mismatch among workers with high employment tenure seen in Figure 8. (The cyclicality of labor efficiency is hump-shaped in tenure, because at low-tenure levels, the match is less likely to pre-date a given aggregate shock, whereas at very high tenure levels, few workers are mismatched to begin with.)

As highlighted above, the procyclical mismatch among workers with high employment tenure stands in contrast with the countercyclical mismatch among new hires. The logic behind this "sullying" among new hires is precisely the cleansing of workers with loose attachment to their career. Not only does it explain why the model is able to generate the countercyclical career mobility seen in the data but, as a consequence, it also implies increased mismatch among the endogenously displaced workers, similar to the one documented for exogenously displaced workers in Section 4.3.



(a) By career attachment **(b)** By mismatch type

Figure 10: Distribution of cleansing by job rung, career attachment and mismatch. *Notes.*—The figure shows the job-rung distribution of workers inside the cleansing region, evaluated at the ergodic distribution. Panel (a) further decomposes the distribution by workers' attachment to their current career. Panel (b) further decomposes by the proportion of positive to negative mismatch in the dissolved jobs.

5.1.2 Sectoral displacement shocks

Before proceeding to the data, we use our model to briefly explore the implication of sectoral shocks, which displace a nonzero mass of workers from their career. One can view such shocks as a reduced form approximation to structural change or to recessions that disproportionately affect certain sectors such as leisure and hospitality during the 2020/21 pandemic. We implement such a sectoral shock as an aggregate displacement of a random mass of 1 percent of the labor force from their career. For simplicity, we assume that the shock affects all workers in a sector proportionately, regardless of their employment status. In this pure form, the shock acts as a prototypical "sullying shock", without the countervailing impact of cleansing on mismatch. Accordingly, it induces a *countercyclical* mismatch response.

In light of recent empirical literature, it is interesting to highlight two features of the simulated response (shown in Figure 11). First, aggregate productivity (or, equivalently, labor efficiency given that z is unshocked) is persistently reduced, outlasting the immediate impact on unemployment. Second, these productivity losses are realized in sectors not originally affected by the shock. This is because displaced workers must rebuild their careers in *new* sectors, which persistently reduces labor productivity below its long-run potential, even after re-employment. Both features are in line with evidence on the aggregate consequences of job displacement following a trade shock that led to mass layoffs in manufacturing due to increased competition from Chinese imports. In particular, the literature has documented

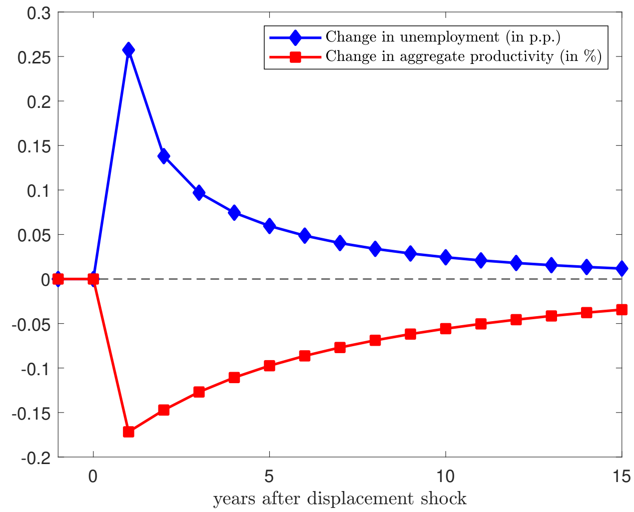


Figure 11: Aggregate impact of a sectoral displacement shock. *Notes.*—The figure shows the responses in aggregate productivity (in percent) and unemployment (in percentage points) to a displacement shock that affects 1 percent of the labor force. The responses are averages over the ergodic process for aggregate productivity z.

large and persistent effects of this displacement on wages and productivity (e.g., Autor, Dorn and Hanson, 2013, 2016), whereas its impact on unemployment has been transient (Bloom et al., 2019). As predicted by the model, Autor, Dorn and Hanson (2013) document that the wage reductions following an aggregate displacement shock to manufacturing were not realized in manufacturing, but indeed are concentrated outside that sector.

5.2 Mismatch Cycles in the Data

We next explore the relation between mismatch and the U.S. business cycle in the data, using the empirical mismatch measure introduced in Section 3.1. We do so by estimating the following empirical specification:

$$m_{i,t} = \beta_0 + (\beta_1 + \beta_2 JS_{i,t} + \beta_3 UE_{i,t}) \times \text{recession}_t + \\ + \gamma \times (JS_{i,t}, UE_{i,t}, x_{i,t}) + \delta_i + \delta_{m_t} + \delta_{y_t} + \epsilon_{i,t}. \quad (11)$$

Here $m_{i,t}$ is the mismatch of worker i at time t; $JS_{i,t}$ and $UE_{i,t}$ are dummies indicating job stayers and new hires from unemployment[35]; recession$_t$ is an indicator that evaluates to unity if the aggregate unemployment rate is above its unconditional average of about 6.5%; $x_{i,t}$ is a

[35] Job stayers are defined as all workers that have the same employer at date t as in the previous month. New hires are defined as all newly hired workers that reported to be not working, unemployed or out of the labor force in the previous month.

Table 4: Cyclicality of mismatch in the data

Dependent variable (×100):	$m_{i,t}$ (1)	$m^{+}_{i,t}$ (2)	$m^{-}_{i,t}$ (3)
Job stayers ($\beta_1 + \beta_2$)	−.293**	.017	−.310***
	(.127)	(.091)	(.085)
New hires ($\beta_1 + \beta_3$)	.646**	.501**	.145
	(.291)	(.209)	(.185)
Total cyclicality	−.250**	.045	−.292***
	(.126)	(.091)	(.084)

Notes.—Standard errors clustered at the worker level are in parenthesis. Asterisks, *, **, ***, indicate coefficients that are significantly different from 0 at the 10%, 5%, 1% level, respectively. Dependent variables are multiplied by 100 (so mismatch ranges from 0 to 100).

set of individual controls, including a quadratic polynomial in age, the region of residence, and a full set of 1-digit occupation and industry dummies; and δ_i, δ_{m_t}, δ_{y_t} are individual, month, and 5-yearly fixed effects, respectively. Here, job-to-job transitions are the omitted category and are absorbed by β_1.[36] We note that the inclusion of individual fixed effects controls for compositional changes in the workforce over the business cycles (e.g., Solon, Barsky and Parker, 1994).

Table 4 reports the estimated business cyclicality. Looking at job stayers, mismatch declines in recessions by an average of .293 percentage points, which corresponds to 1.01% of the unconditional average in mismatch. Decomposing the decline into positive and negative mismatch (columns 2 and 3), we find that the decline is entirely driven by layoffs of underqualified workers, whereas mismatch due to overqualification is acyclical.

The procyclicality of mismatch among job stayers stands in contrast to the cyclicality among newly employed workers, which is countercyclical (.646 percentage points, or 2.31% of the average mismatch among new hires). Decomposing the mismatch, we find that the overall cyclicality is largely driven by unemployed workers finding a job in recessions being on average more overqualified compared to workers finding a job in expansions.

Looking at the total cyclicality (third row), we find that overall mismatch is procyclical. Intuitively, even though new hires are significantly more mismatched during recessions, they only constitute a small fraction of the workforce. Aggregate mismatch is, therefore, primarily determined by the cleansing effect of recessions, comprising roughly acyclical dynamics of overqualification and procyclical dynamics of underqualification.

[36]As our model does not imply any robust prediction for the cyclicality in mismatch among job-to-job movers, we do not focus on job-to-job transitions here. See Table 12 in Appendix G.3 for details on the implied mismatch cyclicality among job-to-job transitions.

Comparison to the model The strong presence of a cleansing effect in the data lends support to the baseline version of our model in which business cycles are driven by aggregate productivity shocks. Using the baseline model to compute the analog to the empirical moments in Table 4, we obtain

$$\Delta\mathbb{E}_{JS}[m^+] = .086 \qquad\qquad \Delta\mathbb{E}_{JS}[m^-] = -.208$$
$$\Delta\mathbb{E}_{UE}[m^+] = .399 \qquad\qquad \Delta\mathbb{E}_{UE}[m^-] = .175,$$

where Δ denotes the difference in conditional means, $\mathbb{E}_L[\cdot] - \mathbb{E}_H[\cdot]$, computed at the ergodic distribution. Overall, the model does a fairly good job at replicating the estimated coefficients, the exception being the cyclicality of m^+ among job stayers, for which the model predicts a small countercyclical response as opposed to the acyclical one in the data. Otherwise, the model captures well the strong cleansing effect on underqualified workers, as well as the sullying effect among new hires, which has a more pronounced effect on overqualification.

6 Suggestive Evidence

We conclude the paper by providing direct evidence towards the learning friction at the core of this paper and towards its implications for career mobility and mismatch. Appendix H contains additional supportive evidence towards the assumptions and mechanism of the model.

6.1 Learning About Skills

We begin by providing direct evidence for workers having imperfect information about their skills as modeled here. We do so using a NLSY79 survey question that asks workers about their expected occupation in 60 months. Based on the reported forecasts, we construct forecast errors between a worker's realized occupation in 60 months and their prediction:

$$\mathrm{fe}_{i,t,j} \equiv q_{i,t+60,j} - \hat{q}_{i,t+60,j},$$

where $\hat{q}_{i,t+60,j}$ is the requirement in skill j associated with the predicted occupation. Suppose an econometrician observes a noisy measure of a worker's skills $\boldsymbol{a}_i$. Hypothesizing that skills are indeed predictive of future occupations, $\mathbb{E}[\boldsymbol{q}_{i,t+60}|\boldsymbol{a}_i] = \boldsymbol{a}_i$, one would then predict the forecast error regarding the utilization of skill j to be given by

$$\mathrm{pe}_{i,t,j} \equiv a_{i,j} - \hat{q}_{i,t+60,j}.$$

Table 5: Direct evidence for learning

Dependent variable:	$\sum_j \text{fe}_j$	fe_j			
		math	verbal	technical	social
	(1)	(2)	(3)	(4)	(5)
$\sum_j \text{pe}_j$	.562***				
	(.020)				
pe_j		.556***	.471***	.331***	.482***
		(.019)	(.020)	(.020)	(.020)
R-squared	.321	.331	.257	.155	.251
Obs.	1575	1575	1575	1575	1575

Notes.—Robust standard errors are in parenthesis. Asterisks, *, **, ***, indicate coefficients that are significantly different from 0 at the 10%, 5%, 1% level, respectively.

Importantly, $\text{pe}_{i,t,j}$ is fully realized at the time the forecasts are surveyed. The main premise of our test is that under the null hypothesis that workers know their skills, the forecast error should therefore be orthogonal to the predicted error $\text{pe}_{i,t,j}$. Note that the orthogonality test follows immediately from the null of workers knowing their own skills, and holds regardless whether or not the econometric conjecture $\mathbb{E}[\boldsymbol{q}_{i,t+60}|\boldsymbol{a}_i] = \boldsymbol{a}_i$ is correct. Moreover, while the goodness of our measure for worker skills affects the power of the test, it is inconsequential for its validity.[37]

We assess the hypothesis of full information by estimating the following specification:

$$\sum_{j=1}^{4} \text{fe}_{i,t,j} = \beta_0 + \beta_1 \sum_{j=1}^{4} \text{pe}_{i,t,j} + \epsilon_{i,t}. \tag{12}$$

Our estimate for β_1 is given by .56 with a standard error of .02. Table 5 further reports variations of our test where we separately estimate (12) for each skill dimension,

$$\text{fe}_{i,t,j} = \beta_0 + \beta_1 \text{pe}_{i,t,j} + \epsilon_{i,t,j}.$$

In all cases, we reject the null hypothesis that workers have full information about their skills. The findings are consistent with anecdotal evidence given in Guvenen et al. (2020), which suggests that workers are unaware of their own ASVAB test scores, and with recent work by Conlon et al. (2018) who document substantial forecast errors regarding labor market outcomes using the Survey of Consumer Expectations of the NY Fed.

[37]This is because any variable that is realized at date t should be orthogonal to workers' expectation error under full information. This holds true independently of the remainder of workers' information structure and regardless of whether $\boldsymbol{a}_i$ is a noisy measure itself. See Chahrour and Ulbricht (2021) for a formal proof.

We also note that $\beta_1 > 0$ in all specifications, indicating that learning has the expected effect: Suppose, for instance, that a worker underestimates their future use of math skills. Then our estimate indicates that over time, as the worker learns about their skills, they indeed end up in a career that is more math-intense than initially predicted.

To sum up, our estimates (i) reject the null that workers perfectly know their skills, and (ii) support the prediction that, as workers learn about their skills, their occupation-choices are skill-driven.

6.2 Career Mobility and Mismatch

Skills predict career mobility Our model predicts that workers seek to switch careers when their belief estimate about current skills, $\hat{a}_k$, falls below a certain threshold. Lacking data on $\hat{a}_k$, we can not directly explore this prediction in the data. Still, because $\hat{a}_k$ is centered around the true skill a_k, we can use our skill measure to proxy for $\hat{a}_k$. To do so, define $a_i(k) \equiv (w_{k,1}, \ldots w_{k,J}) \times \boldsymbol{a}_i'$ as the suitability of worker i's skills for their current career k, defined by their skills weighted by the normalized skill requirements, $\{w_{k,j}\}$, introduced in Section 3.1. We then estimate the following specification in the sample of all job transitions in the NLSY79:

$$\text{career switch}_{i,t} = \beta_0 + \beta_1 a_i(k_{i,t-1}) + \gamma x_{i,t} + \delta_{m_t} + \delta_{y_t} + \epsilon_{i,t}, \tag{13}$$

where $\text{career switch}_{i,t}$ is a dummy that equals 1 if the transition entails a career switch; $x_{i,t}$ is a set of worker controls, including a quadratic polynomial in age, the region of residence, and race, gender and education dummies; and δ_{m_t} and δ_{y_t} are month and 5-yearly fixed effects. We estimate $\hat{\beta}_1 = -0.071$, implying that a lower skill index for the job prior to the transition indeed raises the propensity of career switching, consistent with the predictions of the model.

Career mobility predicts mismatch Our model further predicts that workers that switch careers are on average more mismatched in their new job compared to non-switchers. Moreover, because mismatch is caused by uncertainty, we not only expect it to be higher on average among switchers but further expect it to have a higher variance.

We explore these predictions by comparing switchers with non-switchers, using again the same sample of all job transition in the NLSY79. Specifically, we estimate the impact on average mismatch using the following specification:

$$m_{i,t} = \beta_0 + \beta_1 \text{career switch}_{i,t} + \gamma x_{i,t} + \delta_{m_t} + \delta_{y_t} + \epsilon_{i,t}, \tag{14}$$

Table 6: Career mobility and mismatch

Dependent variable:	career switch$_{i,t}$	$m_{i,t}$	s.d.$(m_{i,t})$
	(1)	(2)	(3)
$a_i(k_{i,t-1})$	$-.071^{***}$		
	(.023)		
career switch$_{i,t}$		$.957^{***}$	$.605^{***}$
		(.308)	(.178)

Notes.—Standard errors clustered at the worker level are in parenthesis. Asterisks, *, **, ***, indicate coefficients that are significantly different from 0 at the 10%, 5%, 1% level, respectively. Mismatch in columns 2 and 3 is multiplied by 100 (so it ranges from 0 to 100). Column 3 reports the second stage of a conditional heteroskedasticity model, using the residuals $\epsilon_{i,t}$ from column 2 to compute s.d.$(m_{i,t}) = |\epsilon_{i,t}|$. See the main text for a description of the controls.

using the same set of controls as in (13). We find $\hat{\beta}_1 = .957$, implying that average mismatch among career-switchers is indeed significantly higher than among non-switchers. To explore the impact on the variance of mismatch, we use a conditional heteroskedasticity model using the residuals from (14) to compute s.d.$(m_{i,t}) = |\epsilon_{i,t}|$. The second stage is specified as follows,

$$\text{s.d.}(m_{i,t}) = \beta_0 + \beta_1 \text{career switch}_{i,t} + \gamma x_{i,t} + \delta_{m_t} + \delta_{y_t} + \zeta_{i,t},$$

using again the same set of controls as in (13) and (14). As predicted by the model, we find a positive and statistically significant effect that increases the standard deviation of mismatch by .605 for career-switchers compared to non-switchers.

7 Conclusion

This paper studies the business cyclicality of worker–occupation mismatch in a quantitative business cycle model with labor market and information frictions. We estimate the model using U.S. data. We find that aggregate mismatch is procyclical among job stayers and countercyclical among new hires, with the former force being overall dominating. These patterns are consistent with direct evidence on the cyclicality of mismatch. We have also shown that the model predicts a scarring effect of job displacement that is sufficiently large to account for empirical evidence on the unemployment scar.

Our framework is among the first that incorporates multidimensional sorting into an equilibrium model with labor market frictions (see also, Lise and Postel-Vinay, 2020; Lindenlaub and Postel-Vinay, 2017). It is distinguished from the existing literature by its analytical tractability, which opens the door to an analysis of aggregate shocks. Our framework delivers rich predictions regarding job and career mobility.

References

Acharya, Sushant, and Shu Lin Wee. 2020. "Rational Inattention in Hiring Decisions." *American Economic Journal: Macroeconomics*, 12(1): 1–40.

Autor, D., and D. Dorn. 2013. "The Growth of Low Skill Service Jobs and the Polarization of the U.S. Labor Market." *American Economic Review*, 103(5): 1553–1597.

Autor, David H., David Dorn, and Gordon H. Hanson. 2013. "The China Syndrome: Local Labor Market Effects of Import Competition in the United States." *American Economic Review*, 103(6): 2121–2168.

Autor, David H., David Dorn, and Gordon H. Hanson. 2016. "The China Shock: Learning from Labor-Market Adjustment to Large Changes in Trade." *Annual Review of Economics*, 8(1): 205–240.

Barlevy, Gadi. 2002. "The Sullying Effect of Recessions." *Review of Economic Studies*, 69: 65–96.

Barnichon, Regis, and Yanos Zylberberg. 2019. "Under-Employment and the Trickle-Down of Unemployment." *American Economic Journal: Macroeconomics*, 11(2): 40–78.

Bloom, Nicholas, Kyle Handley, André Kurmann, and Philip Luck. 2019. "The Impact of Chinese Trade on US Employment: The Good, The Bad, and The Apocryphal." *Working Paper.*

Bowlus, Audra J. 1995. "Matching Workers and Jobs: Cyclical Fluctuations in Match Quality." *Journal of Labor Economics*, 13(2): 335–350.

Carrillo-Tudela, Carlos, and Ludo Visschers. 2020. "Unemployment and Endogenous Reallocation over the Business Cycle." *CEPR Discussion Paper No. DP14697.*

Carrillo-Tudela, Carlos, Bart Hobijn, Powen She, and Ludo Visschers. 2016. "The Extent and Cyclicality of Career Changes: Evidence for the U.K." *European Economic Review*, 84: 18–41.

Chahrour, Ryan, and Robert Ulbricht. 2021. "Robust Predictions for DSGE Models with Incomplete Information." *Boston College Working Paper No 925.*

Christiano, Lawrence J., Martin S. Eichenbaum, and Mathias Trabandt. 2015. "Understanding the Great Recession." *American Economic Journal: Macroeconomics*, 7(1): 110–167.

Christiano, Lawrence J., Martin S. Eichenbaum, and Mathias Trabandt. 2016. "Unemployment and Business Cycles." *Econometrica*, 84(4): 1523–1569.

Conlon, John, Laura Pilossoph, Matthew Wiswall, and Basit Zafar. 2018. "Labor Market Search With Imperfect Information and Learning." *NBER Working Paper*.

Crane, Leland, Henry Hyatt, and Seth Murray. 2018. "Cyclical Labor Market Sorting." *working paper*.

Davis, Steven J., and Till von Wachter. 2011. "Recessions and the Costs of Job Loss." *Brookings Papers on Economic Activity*, , (2): 1–72.

Faberman, R. Jason, Andreas Mueller, Ayşegül Şahin, and Giorgio Topa. 2017. "Job Search Behavior among the Employed and Non-Employed." *IZA Working Paper No. 10960*.

Fredriksson, Peter, Lena Hensvik, and Oskar Nordström Skans. 2018. "Mismatch of Talent: Evidence on Match Quality, Entry Wages, and Job Mobility." *American Economic Review*, 108(11): 3303–3338.

Fujita, Shigeru, and Giuseppe Moscarini. 2017. "Recall and Unemployment." *American Economic Review*, 107(12): 3875–3916.

Gali, Jordi, and Thijs van Rens. 2021. "The Vanishing Procyclicality of Labor Productivity." *Economic Journal*, 131(633): 302–326.

Gathmann, Christina, and Uta Schönberg. 2011. "How General Is Human Capital?" *Journal of Labor Economics*, 28(1): 1–49.

Gertler, Mark, Christopher Huckfeldt, and Antonella Trigari. 2020. "Unemployment Fluctuations, Match Quality, and the Wage Cyclicality of New Hires." *The Review of Economic Studies*, 87(4): 1876–1914.

Groes, Fane, Philipp Kircher, and Iourii Manovskii. 2013. "The U-Shapes of Occupational Mobility." *Review of Economic Studies*, 82(2): 659–692.

Guvenen, Fatih, Burhan Kuruscu, Satoshi Tanaka, and David Wiczer. 2020. "Multidimensional Skill Mismatch." *American Economic Journal: Macroeconomics*, 12(1): 210–44.

Hall, Robert E, and Paul R Milgrom. 2008. "The Limited Influence of Unemployment on the Wage Bargain." *American Economic Review*, 98(4): 1653–1674.

Haltiwanger, John, Henry Hyatt, Erika McEntarfer, and Matthew Staiger. 2021. "Cyclical Worker Flows: Cleansing vs. Sullying." *NBER Working Paper No. 28802.*

Hazell, Jonathon, Juan Herreño, Emi Nakamura, and Jón Steinsson. 2020. "The Slope of the Phillips Curve: Evidence from U.S. States." *NBER Working Paper No. 28005.*

Heckman, James J., Jora Stixrud, and Sergio Urzua. 2006. "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior." *Journal of Labor Economics*, 24(3): 411–482.

Holzer, Harry J. 1987. "Job Search by Employed and Unemployed Youth." *Industrial and Labor Relations Review*, 40(4): 601–611.

Huckfeldt, Christopher. 2021. "Understanding the Scarring Effect of Recessions." *working paper.*

Jarosch, Gregor. 2021. "Searching for Job Security and the Consequences of Job Loss." *NBER Working Paper No. 28481.*

Jovanovic, Boyan. 1979. "Job Matching and the Theory of Turnover." *Journal of Political Economy*, 87(5): 972–990.

Jovanovic, Boyan. 1984. "Matching, Turnover, and Unemployment." *Journal of Political Economy*, 92(1): 108–122.

Jung, Philip, and Moritz Kuhn. 2019. "Earnings Losses and Labor Mobility over the Life Cycle." *Journal of the European Economic Association*, 17(3): 678–724.

Lindenlaub, Ilse. 2017. "Sorting Multidimensional Types: Theory and Application." *The Review of Economic Studies*, 84: 718–789.

Lindenlaub, Ilse, and Fabien Postel-Vinay. 2017. "Multidimensional Sorting Under Random Search." *working paper.*

Lise, Jeremy, and Fabien Postel-Vinay. 2020. "Multidimensional Skills, Sorting, and Human Capital Accumulation." *American Economic Review*, 110(8): 2328–76.

Lise, Jeremy, and Jean-Marc Robin. 2017. "The Macrodynamics of Sorting between Workers and Firms." *American Economic Review*, 107(4): 1104–1135.

McGrattan, Ellen R., and Edward C. Prescott. 2012. "The Great Recession and Delayed Economic Recovery: A Labor Productivity Puzzle?" *Government policies and the delayed economic recovery*, 115–154.

Menzio, Guido, and Shouyong Shi. 2010. "Block Recursive Equilibria for Stochastic Models of Search on the Job." *Journal of Economic Theory*, 145(4): 1453–1494.

Menzio, Guido, and Shouyong Shi. 2011. "Efficient Search on the Job and the Business Cycle." *Journal of Political Economy*, 119(3): 468–510.

Menzio, Guido, Irina A. Telyukova, and Ludo Visschers. 2016. "Directed Search over the Life Cycle." *Review of Economic Dynamics*, 19: 38–62.

Mian, Atif, Kamalesh Rao, and Amir Sufi. 2013. "Household Balance Sheets, Consumption, and the Economic Slump." *The Quarterly Journal of Economics*, 128(4): 1687–1726.

Mortensen, D. T., and Christopher A Pissarides. 1994. "Job Creation and Job Destruction in the Theory of Unemployment." *The Review of Economic Studies*, 61(3): 397–415.

Moscarini, Giuseppe. 2001. "Excess Worker Reallocation." *Review of Economic Studies*, 68(3): 593–612.

Mulligan, Casey. 2011. "Rising Labor Productivity During the 2008-9 Recession." *NBER Working Paper 17584*.

Neal, Derek. 1999. "The Complexity of Job Mobility among Young Men." *Journal of Labor Economics*, 17(2): 237–261.

Papageorgiou, Theodore. 2014. "Learning Your Comparative Advantages." *Review of Economic Studies*, 81(3): 1263–1295.

Prada, María F., and Sergio Urzúa. 2017. "One Size Does Not Fit All: Multiple Dimensions of Ability, College Attendance, and Earnings." *Journal of Labor Economics*, 35(4): 953–991.

Şahin, Ayşegül, Joseph Song, Giorgio Topa, and Giovanni L. Violante. 2014. "Mismatch Unemployment." *American Economic Review*, 104(11): 3529–3564.

Schaal, Edouard. 2017. "Uncertainty and Unemployment." *Econometrica*, 85(6): 1675–1721.

Shimer, Robert. 2005. "The Cyclical Behavior of Equilibrium Unemployment and Vacancies." *American Economic Review*, 95(1): 25–49.

Shimer, Robert. 2012. “Reassessing the Ins and Outs of Unemployment.” *Review of Economic Dynamics*, 15(2): 127–148.

Solon, Gary, Robert B Barsky, and Jonathan A Parker. 1994. “Measuring the Cyclicality of Real Wages: How Important Is Composition Bias.” *Quarterly Journal of Economics*, 109(1): 1–25.

Speer, Jamin D. 2017. “Pre-Market Skills, Occupational Choice, and Career Progression.” *Journal of Human Resources*, 52(1): 187–246.

Stigler, George J. 1962. “Information in the Labor Market.” *Journal of Political Economy*, 70(5, Part 2): 94–105.

Walsh, Carl E. 2005. “Labor Market Search, Sticky Prices, and Interest Rate Policies.” *Review of Economic Dynamics*, 8(4): 829–849.

Wee, Shu Lin. 2016. “Delayed Learning and Human Capital Accumulation: The Cost of Entering the Job Market During a Recession.” *working paper*.

Yamaguchi, Shintaro. 2012. “Tasks and Heterogeneous Human Capital.” *Journal of Labor Economics*, 30(1): 1–53.

Online Appendix to "Mismatch Cycles"

A Kolmogorov Forward Equations

Let $p_U(\hat{a}, \Sigma, z)$ and $p_E(\hat{a}, \Sigma, r, z)$ define the job finding rates of unemployed and employed workers as given by (8) and (9).

Active relationships The distribution over active relationships, $\Gamma_t(\hat{a}, \Sigma, r)$, is characterized by the following PDE:

$$\dot{\Gamma}_t(\hat{a}, \Sigma, r) = \dot{\Gamma}_t^{\text{learn}}(\hat{a}, \Sigma, r) + \dot{\Gamma}_t^{\text{ee}}(\hat{a}, \Sigma, r) + \dot{\Gamma}_t^{\text{ue}}(\hat{a}, \Sigma, r) - \dot{\Gamma}_t^{\text{eu}}(\hat{a}, \Sigma, r) - \epsilon \Gamma_t(\hat{a}, \Sigma, r). \tag{15}$$

Here, the first term defines distributional dynamics driven by changes in beliefs, given by

$$\dot{\Gamma}_t^{\text{learn}}(\hat{a}, \Sigma, r) = \left(\frac{\partial}{\partial \Sigma} + \frac{1}{2} \frac{\partial^2}{\partial \hat{a}^2} \right) \left[\left(\frac{\Sigma}{\sigma} \right)^2 \Gamma_t(\hat{a}, \Sigma, r) \right].$$

The second term, defines reallocation dynamics due to job-to-job transitions,

$$\dot{\Gamma}_t^{\text{ee}}(\hat{a}, \Sigma, r) = -p_E(\hat{a}, \Sigma, r, z)\Gamma_t(\hat{a}, \Sigma, r) + \sum_{r' \in \mathcal{R}} p_E(\hat{a}, \Sigma, r', z)\Gamma_t(\hat{a}, \Sigma, r') \cdot \mathbf{1}_{r = r^*(\hat{a}, \Sigma, z)},$$

where $\mathbf{1}_C$ denotes the indicator function for a given condition C. The third term, defines the incoming flow of new hires out of unemployment,

$$\dot{\Gamma}_t^{\text{ue}}(\hat{a}, \Sigma, r) = p_U(\hat{a}, \Sigma, z)\Lambda_t(\hat{a}, \Sigma) \cdot \mathbf{1}_{r = r^*(\hat{a}, \Sigma, z)}.$$

The fourth term defines separations into unemployment,[38]

$$\dot{\Gamma}_t^{\text{eu}}(\hat{a}, \Sigma, r) = \left(\delta + \lim_{\pi \to \infty} \pi \chi^{\text{sep}}(\hat{a}, \Sigma, r, z) \right) \Gamma_t(\hat{a}, \Sigma, r)$$

where $\chi^{\text{sep}}(\hat{a}, \Sigma, r, z) \in \{0, 1\}$ is an indicator evaluating to unity when the value of the match becomes negative ($J_t^{\text{act}}(\hat{a}, \Sigma, r, z) \leq \mathcal{U}(\hat{a}, \Sigma, z)$). Finally, the fifth term defines exogenous career switches.

[38]Note that for the endogenous separations case, the rate of outflows equals ∞ as long as $\Gamma_t(\hat{a}, \Sigma, r) \neq 0$ for the corresponding states, implying that the only possible limit is $\Gamma_t(\hat{a}, \Sigma, r) = 0$ for any states $(\hat{a}, \Sigma, r)$ outside the continuation region.

Unemployed Similarly, the distribution over unemployed workers, $\Upsilon_t(\hat{a},\Sigma)$, is characterized by the following PDE:

$$\dot{\Upsilon}_t(\hat{a},\Sigma) = \dot{\Upsilon}^{\text{cs}}_t(\hat{a},\Sigma) + \dot{\Upsilon}^{\text{eu}}_t(\hat{a},\Sigma) - \dot{\Upsilon}^{\text{ue}}_t(\hat{a},\Sigma). \tag{16}$$

Here, the first term defines net changes in (current-career) beliefs due to agents switching careers,[39]

$$\begin{aligned}\dot{\Upsilon}^{\text{cs}}_t(\hat{a},\Sigma) = -\left(\epsilon + \lim_{\pi\to\infty} \pi\chi^{\text{cs}}(\hat{a},\Sigma,z)\right)\Upsilon_t(\hat{a},\Sigma) + \\ + \iint \left(\epsilon + \lim_{\pi\to\infty} \pi\chi^{\text{cs}}(\hat{a},\Sigma,z)\right)\Upsilon_t(\hat{a}',\Sigma')\,\mathrm{d}(\hat{a}',\Sigma')\cdot \mathbf{1}_{(\hat{a},\Sigma)=(a_0,S_0)},\end{aligned}$$

where $\chi^{\text{cs}}(\hat{a},\Sigma,z) \in \{0,1\}$ is an indicator evaluating to unity when switching careers is optimal ($U_t(a_0,S_0,z) > U_t(\hat{a},\Sigma,z)$). The second term defines gross inflows into unemployment, including those from exogenous career switches,

$$\dot{\Upsilon}^{\text{eu}}_t(\hat{a},\Sigma) = \int \dot{\Gamma}^{\text{eu}}_t(\hat{a},\Sigma,r)\,\mathrm{d}r + \iiint \epsilon\Gamma_t(\hat{a},\Sigma,r)\,\mathrm{d}(\hat{a}',\Sigma',r)\cdot \mathbf{1}_{(\hat{a},\Sigma)=(a_0,S_0)}.$$

Finally, the third term defines the outflows from unemployment due to workers finding jobs,

$$\dot{\Upsilon}^{\text{ue}}_t(\hat{a},\Sigma) = p_U(\hat{a},\Sigma,z)\Lambda_t(\hat{a},\Sigma).$$

Transmission of aggregate shocks The aggregate productivity state z_t affects the cross-sectional distribution through three channels: (1) its direct impact on job finding rates $p_U(\hat{a},\Sigma,z)$ and $p_E(\hat{a},\Sigma,r,z)$, (2) its direct impact on the separation and career switching thresholds $\chi^{\text{sep}}(\hat{a},\Sigma,r,z)$ and $\chi^{\text{cs}}(\hat{a},\Sigma,z)$, and (3) its direct impact on the desired job rung $r^*(\hat{a},\Sigma,z)$. These direct effects translate into shifts in $\dot{\Gamma}^{\text{ee}}_t$, $\dot{\Gamma}^{\text{eu}}_t$, $\dot{\Gamma}^{\text{ue}}_t$, $\dot{\Upsilon}^{\text{eu}}_t$, $\dot{\Upsilon}^{\text{ue}}_t$ and $\dot{\Upsilon}^{\text{cs}}_t$, which in turn propagate to Γ_t and Υ_t according to (15) and (16). In particular, an aggregate shock to z_t manifests itself both through a discrete shift in the cross-sectional distributions Γ_t and Υ_t upon impact and by alternating their subsequent evolution $\dot{\Gamma}_t$ and $\dot{\Upsilon}_t$.

For the calibration from Section 3, the direct effects are sizable for p_U, p_E and χ^{sep}, whereas the direct effects on χ^{cs} and r^* are negligible.[40] Specifically, the direct effects on p_U, p_E and

[39]Note that the rate of workers switching careers equals ∞ as long as $\Upsilon_t(\hat{a},\Sigma) \neq 0$ for the corresponding states. The only possible limit is therefore given by $\Upsilon_t(\hat{a},\Sigma) = 0$ for any states $(\hat{a},\Sigma)$ in which workers switch careers. Accordingly, the corresponding switching rates, defining the inflow into (a_0,S_0), equal the inflow into the switching states from employment.

[40]The cyclicality of career-mobility is entirely driven through the distributional shift in Υ_t caused by the shift in the separation threshold χ^{sep}.

χ^{sep} imply strong procyclical fluctuations in the job finding rate from unemployment and the job-to-job mobility rate, and countercyclical fluctuations in the separation rate (reflecting the contraction in the continuation region depicted in Figure 9).[41]

B Examples of General Production Function

This appendix provides two examples of a general production technology $F(z, \boldsymbol{q}, \boldsymbol{a})$ that collapses into (1) when $\boldsymbol{q}_{k,r}$ are orthogonal.

Complementary-skill case Let

$$F(z(t), \boldsymbol{q}_{k,r}, \boldsymbol{a}_i) \equiv \exp\left[z(t) + \sum_{j=1}^{J}\left(\eta q_{k,r,j} - \max\left\{q_{k,r,j} - \frac{q_{k,r,j}a_{i,j}}{\sum_{j=1}^{J} q_{k,r,j}}, 0\right\}\right)\right]. \tag{17}$$

Substituting $r = \sum_{j=1}^{J} q_{k,r,j}$ and $w_{k,j} = q_{k,r,j}/(\sum_{j=1}^{J} q_{k,r,j})$, we can rewrite (17) in more accessible form

$$\log y_{i,k,r} = z(t) + \sum_{j=1}^{J} w_{k,j}\left(\eta r - \max\{r - a_{i,j}, 0\}\right),$$

which clearly collapses into (1) for an orthogonal weighting scheme; e.g.,[42]

$$\begin{bmatrix}\boldsymbol{w}_1' & \boldsymbol{w}_2' & \cdots & \boldsymbol{w}_K'\end{bmatrix} = \boldsymbol{I}_K.$$

Substitutible-skill case Let

$$F(z(t), \boldsymbol{q}_{k,r}, \boldsymbol{a}_i) \equiv \exp\left[z(t) + \eta\sum_{j=1}^{J} q_{k,r,j} - \max\left\{\sum_{j=1}^{J} q_{k,r,j} - \frac{\sum_{j=1}^{J} q_{k,r,j}a_{i,j}}{\sum_{j=1}^{J} q_{k,r,j}}, 0\right\}\right], \tag{18}$$

which can be rewritten more compactly as

$$\log y_{i,k,r} = z(t) + \eta r - \max\left\{r - \sum_{j=1}^{J} w_{k,j}a_{i,j}, 0\right\}.$$

Again, it is easy to verify that $y_{i,k,r}$ collapses into (1) for an orthogonal weighting scheme.

[41] Evaluated at the ergodic distribution, the cyclical differences between expansions and recessions are: 9.5 percentage points (pp.) for the monthly job finding rate from unemployment, 0.4 pp. for the monthly job-to-job mobility rate, and −0.7 pp. for the separation rate.

[42] Here, we tacitly set $K = J$, for ease of exposition. Weighting schemes other than the identity scheme may require a redefinition of skill types, but can equally be reduced to (1) for an appropriate definition of skills as long as $\{\boldsymbol{a}_k\}$ are orthogonal across the adopted career classification $\{k\}$.

Table 7: Inertia in career mobility

	Recession	Expansion
Unemployed	8.85	8.98
$r = 0$	9.33	9.95
$r = 0.5 \cdot S_0^{1/2}$	7.89	8.46
$r = 1.0 \cdot S_0^{1/2}$	4.97	5.21
$r = 1.5 \cdot S_0^{1/2}$	2.34	2.42
$r = 2.0 \cdot S_0^{1/2}$	0.88	0.90
$r = 2.5 \cdot S_0^{1/2}$	0.26	0.27
$r = 3.0 \cdot S_0^{1/2}$	0.04	0.04

Notes.—The table reports the implicit cost on career mobility induced by mismatch, denominated in monthly average output per worker, $\mathbb{E}[y_{i,t}]/\mathbb{E}[1 - U_t]$.

C Inertia in Career Mobility

As alluded to in the main text, inertia not only marks workers' reallocation across job rungs within careers, but also their career choice. This is because evaluating the prospects of a career takes time due to the information friction and reduces the returns to trying out new careers given the anticipation of mismatch. In what follows, we assess the magnitude of this implicit cost on exploring new careers. We do so by considering a fictitious career-switching problem in which workers can instantaneously churn careers and learn the relevant skill at *infinite speed* subject to an explicit switching cost $\xi_{i,t}$. For any given worker, we then calculate the magnitude of the explicit switching cost $\xi_{i,t}$ that keeps them indifferent between accessing the fictitious churning technology and sticking to their equilibrium career choice. Intuitively, our approach replaces the implicit information friction on career mobility (and the cost of entailing mismatch) by an explicit switching cost $\xi_{i,t}$, which we design so as to impose the same career mobility patterns for all workers.

Specifically, let $X_{i,t}$ denote the current unemployment value $\mathcal{U}(\hat{a}_k, \Sigma_k, z)$ if a worker is currently unemployed, and the joint worker–firm value $J(\hat{a}_k, \Sigma_k, r, z)$ if they are employed. Then the marginal benefit of exploring a new career and learning the relevant skill instantaneously, $(\hat{a}, \Sigma) = (a, 0)$, is given by

$$\tilde{\xi}_{i,t} = \int_{-\infty}^{\infty} \max\left\{\mathcal{U}_t(a, 0) - X_{i,t}, 0\right\} \,\mathrm{d}\Phi\left(\frac{a - a_0}{\sqrt{S_0}}\right).$$

To preempt workers from assessing the churning technology it hence suffices to set $\xi_{i,t} = \tilde{\xi}_{i,t}$. Table 7 reports the result (denominated in the *economy-wide* average monthly output per worker). The implicit friction is largest for low-skilled workers as they benefit the most from exploring new careers. It ranges from the equivalent of 10 months of output for workers at the bottom rung of the job ladder to about one work day of at the top rung.[43] Averaged across workers and business cycle states, the implicit friction evaluates to the equivalent of 4.75 months of average output per worker.

D Wages Without Commitment by Workers

This appendix details the computation of wages used for the exploration in Section 4. Following Schaal (2017) we adopt the unique wage scheme that induces equilibrium search and job continuation policies to be self-enforcing for workers (without requiring a contractual commitment).

Let w_t denote the wage of worker i at date t, and let W_t define the expected lifetime utility of an employed worker that is delivered by the contracted process for $\{w_t\}$. Notice that the characterization so far only pins down $W_t = x_t$ during hiring but does not determine how the promised hiring utility, x_t, is delivered across states and throughout the duration of the work-relationship. In analogue to (5), the expected utility flow of an active relationship is given by

$$\begin{aligned}\rho W_t^{\text{act}}(\hat{a}_k, \Sigma_k, r, z) = w_{i,t} + \tilde{\Lambda}_t(\hat{a}_k, \Sigma_k, r, z) + \\ + \max_{x,r} \{\kappa p(\theta_t(\omega, z)) (x - W_t(\hat{a}_k, \Sigma_k, r, z))\} + \\ + \delta (\mathcal{U}_t(\hat{a}_k, \Sigma_k, z) - W_t(\hat{a}_k, \Sigma_k, r, z)) + \\ + \epsilon (U_t(a_0, S_0, z) - W_t(\hat{a}_k, \Sigma_k, r, z)) + \\ + \lambda_z (W_t(\hat{a}_k, \Sigma_k, r, -z) - W_t(\hat{a}_k, \Sigma_k, r, z)), \end{aligned} \tag{19}$$

where

$$\tilde{\Lambda}_t(\hat{a}_k, \Sigma_k, r, z) \equiv \left(\frac{\Sigma_k}{\sigma}\right)^2 \left(-\frac{\partial W_t(\hat{a}_k, \Sigma_k, r, z)}{\partial \Sigma_k} + \frac{1}{2}\frac{\partial^2 W_t(\hat{a}_k, \Sigma_k, r, z)}{\partial \hat{a}_k^2}\right)$$

and

$$W_t(\hat{a}_k, \Sigma_k, r, z) = \max\left\{W_t^{\text{act}}(\hat{a}_k, \Sigma_k, r, z), \mathcal{U}_t(\hat{a}_k, \Sigma_k, z)\right\}.$$

Absent contractual commitments, workers' on-the-job search maximizes (19) subject to (2).

[43] The implicit friction is slightly larger for workers at the bottom job rung than for unemployed workers due to the presence of exogenously laid off workers among the unemployed who have strong incentives to retain their current career.

Rearranging the associated first-order condition, we have

$$\theta = p'^{-1}\left(\frac{c}{J_t(\hat{a}_k, \Sigma_k, r^*, z) - W_t(\hat{a}_k, \Sigma_k, r, z)}\right). \tag{20}$$

Comparing (20) with (9), we conclude that for search to be self-enforcing, the worker value of the relationship must match the joint value whenever they are actively searching. Accordingly, the unique self-enforcing wage scheme is given by

$$w_{i,t} = e^{z+\eta r}\mathbb{E}_t[e^{-\max\{r-a_k,0\}}] = e^{z+\eta r}\psi(\hat{a}_k - r, \sqrt{\Sigma_k});$$

i.e., workers are compensated their marginal product at each instant of an ongoing work-relationship. Moreover, because $W_t^{\text{act}} = x_t$ must hold at hiring, workers must reimburse firms for their recruitment cost at the instant of hiring, implying a one-time reduction in wages equal to

$$J(\hat{a}_k, \Sigma_k, r, z) - x = c/q\left(\theta(\hat{a}_k, \Sigma_k, r, z)\right).$$

Finally, noticing that the described wage arrangement implies $W_t^{\text{act}} = J_t^{\text{act}}$ at any instant of an ongoing relationship, we conclude that workers' job continuation/separation choices are also aligned with the bilaterally efficient ones observed under commitment.

E Measuring Job Requirements, Employment Transitions, and Worker Skills

This appendix details the measurement of job requirements, employment transitions, and worker skills.

E.1 Job Requirements

Following Guvenen et al. (2020), we measure skill requirements using 26 O*NET descriptors from the Knowledge, Skills and Abilities categories that were identified by the Defense Manpower Data Center (DMDC) to be related to each ASVAB category, augmented by six descriptors linked to social skills.[44] As in Guvenen et al. (2020), we link those O*NET

[44]The descriptors used are the following: oral comprehension, written comprehension, deductive reasoning, inductive reasoning, information ordering, mathematical reasoning, number facility, reading comprehension, mathematics skill, science, technology design, equipment selection, installation, operation and control, equipment maintenance, troubleshooting, repairing, computers and electronics, engineering and technology, building and construction, mechanical, mathematics knowledge, physics, chemistry, biology, english language,

descriptors to ASVAB test category based on the relatedness score provided by DMDC. The verbal skill requirement is then defined as the first principal component of Word Knowledge and Paragraph Comprehension, the math requirement is that of Math Knowledge and Arithmetic Reasoning, and the technical requirement is the first principal component of Electronics Info, General Science, and Mechanical comprehension. For the social dimension, we also collapse the six O*NET descriptors into a single dimension defined by the first principal component. Finally, we normalize all requirements by converting them into percentile ranks based on the distribution of occupations in our NLSY79 sample (see below).

E.2 Employment Transitions

Employment histories We infer employment histories from the NLSY79 Work History Data File, which is a nationally representative panel of workers who are followed from first entry into the labor market. We aggregate the available employment data, which is recorded at a weekly frequency, to a monthly frequency by focusing on the job for which an individual worked the most hours in a given month.

Sample selection As the NLSY79 is well-known and requires little description, we focus in the following on describing the sample selection used in this paper. We focus on the subsample of males and females from the so-called cross-sectional sample, which is designed to represent the non-institutionalized civilian segment of the U.S. in 1979.[45] As is standard in the literature, we drop individuals who were more than two years in the military force, individuals with a weak labor market attachment (spending more than 10 years out of the labor force), individuals that were already working in 1979, and those that do not have information on the Armed Services Vocational Aptitude Battery (ASVAB) test scores.

E.3 Worker Skills and Mismatch

Worker skills We measure workers skills using ASVAB test scores available in the NLSY79 (see Appendix E.2 for a description of our subsample). The ASVAB is a general test that measures knowledge and skills in 10 different components that was taken by survey participants when first entering the survey.[46] As in Guvenen et al. (2020), we focus on a subset of seven

social perceptiveness, coordination, persuasion, negotiation, instructing, service orientation.

[45]The NLSY79 also contains supplemental samples that oversample ethnic minorities, economically disadvantaged people, and the military, none of which we include in our analysis.

[46]The components are arithmetic reasoning, mathematics knowledge, paragraph comprehension, word knowledge, general science, numerical operations, coding speed, automotive and shop information, mechanical comprehension, and electronics information.

components (arithmetic reasoning, mathematics knowledge, paragraph comprehension, word knowledge, mechanical comprehension, general science and electronics information) which are linked to math, verbal and technical skills, and are combined using Principal Components Analysis. For the social dimension, we proceed in the same fashion using the individual scores in two different tests provided by the NLSY79: the Rotter Locus of Control Scale and the Rosenberg Self-Esteem Scale.[47] To adjust for differences in test-taking age, before proceeding with PCA, we normalize the mean and the variance of each test score according to their age-specific values. Then, once we have the raw scores in each skill dimension, we convert these into percentile ranks.

Mismatch We merge the panel of worker-level data with the occupation data using using three-digit Census occupational codes. Note that O*NET uses SOC codes from 2010, which are more detailed than the occupational codes in the NLYS79, based on the three-digit Census occupation codes. Hence several occupations in NLSY79 have more than one score. Using a crosswalk to identify each SOC code with a Census code, we take an unweighted average over all the SOC codes that map to the same code in the census three-digit level occupation classification. We then proceed to construct mismatch as defined in the main body of the paper.

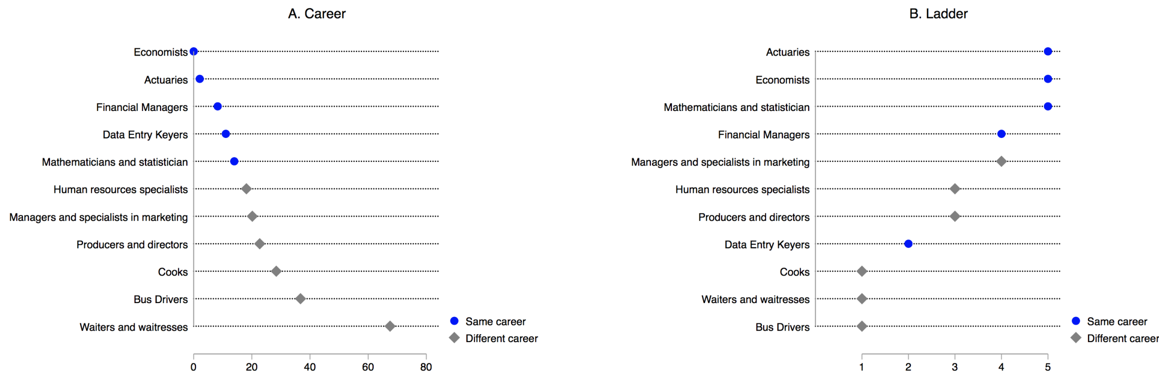
F On Skill-based Definition of Careers

This appendix examines in further detail our skill-based definition of careers. We present examples that illustrate how our definition classifies occupations across different careers; we assess the prevalence of radical vs. gradual career switches; we compare our skill-based measure of career mobility with alternative measures; and finally, we show that the cyclicality of career mobility is primarily driven by job transitions that go through unemployment.

F.1 Illustrative Examples of Career Mobility

We begin presenting two examples of career mobility identified through our angular measure. The first example fixes the requirement vector of the 3-digit occupation $\boldsymbol{q}_1$ = "Economist" and considers a selection of 3-digit occupational titles with requirement vectors $\boldsymbol{q}_2$. According to our

[47] The Rotter Locus of Control Scale measures the degree of control individuals feel they possess over their life, and the Rosenberg Self-Esteem Scale aims at reflecting the degree of approval or disapproval towards oneself. These measures have been commonly used in previous works as measures of non-cognitive skills (Speer, 2017; Lise and Robin, 2017; Guvenen et al., 2020). For more details, see Heckman, Stixrud and Urzua (2006).



(a) By angular distance to Economist

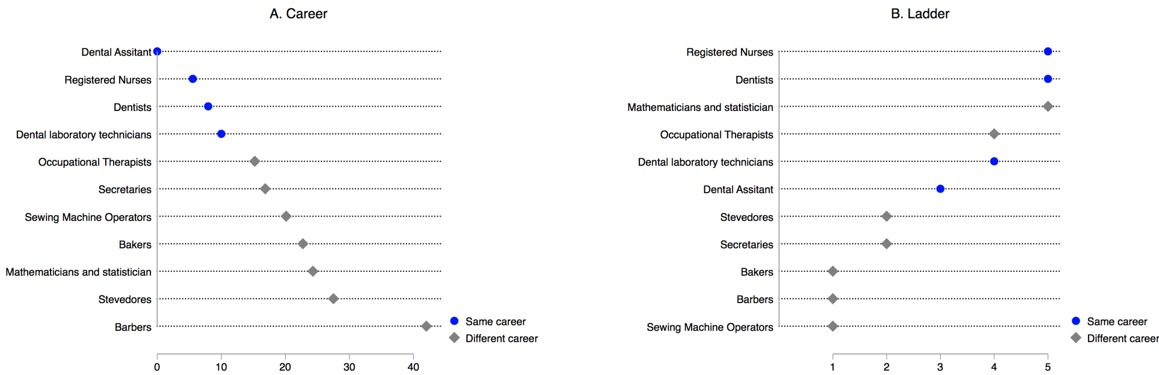
(b) By job rung quintiles

Figure 12: Examples of occupations inside and outside "Economist" cone. *Notes.*—Blue dots correspond to occupations classified within the same career.

skill-based criterion, these occupations are classified within the same career as an "Economist" if their angular distance $\varphi(\boldsymbol{q}_1, \boldsymbol{q}_2)$ is smaller than the calibrated threshold $\bar{\varphi} = 14.8°$. Figure 12a plots the angular distances for these occupations relative to an "Economist". By definition, the angular distance from "Economist" to "Economist" is zero. In this example, "Actuaries", "Financial Managers", "Data Entry Keyers" and "Mathematicians and Statisticians" fall inside the "Economist" cone (blue dots) and, hence, transitions from "Economist" to any of these occupations are classified within the same career. In contrast, transitions to occupations that fall outside this cone (gray diamonds), such as "Cooks", are classified as a career switch relative to a "Economist".

To assess movement up and down the job ladder, Figure 12b plots the same set of occupational titles according to their position in the job ladder, measured by the corresponding quintile in the job rung distribution. Within the "Economist" cone, "Actuaries"and "Mathematicians and Statisticians" are top-tier occupations (5th quintile) while "Financial Managers" (4th quintile) and "Data Entry Keyers" (2nd quintile) are lower-tier occupations. Changing jobs to any of the latter occupations would entail a movement down the job ladder within the same career.

Figures 13a and 13b present a second illustrative example for the occupation $\boldsymbol{q}_1$ = "Dental Assistant". According to our skill-based definition, if a "Dental Assistant" becomes a "Dental laboratory technician", a "Dentist" or a "registered nurse", this is interpreted as a movement up the job ladder: the skill-mix required by any of those occupations is fairly similar to a "Dental Assistant", but the task complexity is increased. In contrast, if a "Dental Assistant" becomes, say, a "Baker", this is interpreted as a career switch.



(a) By angular distance to Dental Assistant

(b) By job rung quintiles

Figure 13: Examples of occupations inside and outside "Dental Assistant" cone. *Notes.*—Blue dots correspond to occupations classified within the same career.

Table 8: Gradual job transitions

N	4	5	6	7	8
$\Pr[\varphi(\boldsymbol{q}_0, \boldsymbol{q}_N) < \bar{\varphi}]$	0.82	0.82	0.82	0.82	0.81

Notes.—The table shows the fraction of transition paths for which the final job falls within the cone of the original job; i.e., $\varphi(\boldsymbol{q}_0, \boldsymbol{q}_N) < \bar{\varphi}$.

F.2 Gradual Career Transitions

Our approach to measuring career mobility identifies large changes in the occupation requirements that occur at distinct points of time. One implication of this approach is that career switches if broken down to a sequence of small steps may not constitute a distinct career switch at any point of time.

In the following, we explore the empirical prevalence of such "gradual" career transitions. To do so, fixing an integer N, we first construct the sample of all N consecutive job transitions in the NLSY79 which do not constitute a career-transition according to our measure. That is, letting $\boldsymbol{Q}_N \equiv \{\boldsymbol{q}_s\}_{s\in\{0,1,\dots,N\}}$ denote the job requirements of $N+1$ consecutive jobs of a given worker, our sample contains the universe of all $\boldsymbol{Q}_N$ such that each individual transition satisfies $\varphi(\boldsymbol{q}_{s-1}, \boldsymbol{q}_s) < \bar{\varphi}$ for all $s \in \{1, 2, \dots, N\}$. Equipped with this sample, we then re-apply our criterion to the initial and final job, and compute the fraction of samples for which $\varphi(\boldsymbol{q}_0, \boldsymbol{q}_N) < \bar{\varphi}$. Table 8 reports the results for different values of N. In all cases, we find a moderate prevalence of gradual career transitions of 18–19%. By contrast, for the majority of within-career job sequences the final job falls within the cone of the initial job.

Table 9: Cyclicality of career mobility under alternative definitions of a career

	skill-based	1-digit	2-digit	3-digit	1-digit (SOC)	k-means
career mobility	.42	.50	.50	.61	.50	.45
excess cyclicality	.07	.04	.04	.03	.04	.04
corr. with skill-based	1.00	.64	.64	.69	.64	.75

Notes.—First row shows the unconditional career switching propensity (in percent). Second row shows the cyclicality of career mobility, computed as the difference in career mobility in recessions to expansions (in p.p.). Third row shows the correlation of different career mobility measures with our skill-based definition.

F.3 Comparison With Alternative Definitions of Careers

Here we compare our skill-based measure of career mobility with alternative measures. In particular, we compare it with the following alternative criteria to define careers: 1-digit, 2-digit, or 3-digit occupational codes from Autor and Dorn (2013); 1-digit occupations from the Standard Occupational Classification (SOC); and a classification derived from a k-means algorithm that groups occupations into different careers such that the angular distance to the average skill-requirement in a career is minimized (specifically, we choose the number of clusters to be $k = 6$ that delivers an unconditional career mobility rate of 45%, closely matching the career mobility rate of 42.2% obtained under our skill-based definition). The exercise considers the universe of job transitions and for each transition determines whether or not it is registered as a career transition according to these alternative criteria.

Table 9 summarizes the comparison. Overall, we see significant differences across these classifications. While all measures are moderately correlated with our baseline measure, with correlations ranging from .64 to .75, there are significant differences in the average propensities to switch careers, ranging from .42 to .61. Interestingly, however, despite these differences, all measures imply countercyclical career mobility.

F.4 Cyclicality in Career Mobility By Transition Type

According to our model, the cyclicality of career mobility is intrinsically tied to job transitions through unemployment. We note that this prediction is not driven by our restriction on career transitions.[48] This is because it is precisely the workers that are cleansed from their jobs that

[48] Our model assumes that workers can switch careers exclusively through a spell of unemployment. While in reality, of course, some career switches occur through job-to-job transitions, this assumption is meant to capture that switching careers is more costly and time intensive than other job-to-job transitions. For instance, professional networks are naturally centered around current and past careers, facilitating within-career switches or even giving rise to entirely unsolicited offers. By contrast, career-switching arguably requires a more active

Table 10: Cyclicality of career mobility by type of job transition

Transition type	Fraction of all job transitions	Fraction of all career switches	Excess cyclicality
EUE'	.56	.48	.08
EE'	.44	.52	−.01
Total (EUE' + EE')	1.00	1.00	.07

Notes.—Job transitions and career switches by type of job transition. EUE' refers to job transitions that undergo an unemployment spell. EE' refers to direct job-to-job transitions. Excess cyclicality is computed as career switching rate in recessions minus expansions.

cause the increase in career switching during recessions.

To assess this implication of the model, we decompose the empirical cyclicality of career mobility into its cyclicality among transitions through unemployment (EUE') and job-to-job transitions (EE'). Table 10 shows the decomposition. Consistent with the predictions of the model, the overall cyclicality (+0.07 percentage points in recessions) is exclusively driven by countercyclicality in EUE' transitions (+0.08 percentage points), whereas the propensity to switch careers among EE' transitions is roughly acyclical (−0.01 percentage points).

G On Mismatch Cyclicality

This appendix presents additional empirical results and robustness checks on the cyclicality of mismatch. We show a time series for aggregate mismatch; we examine the cyclical properties of mismatch using alternative business cycle indicators; we assess mismatch cyclicality for job-to-job movers; we present a robustness check using the two cohorts of the NLSY data; and finally, we show the cyclical properties of mismatch for each of the four underlying skill dimensions (math, verbal, technical and social skills).

G.1 Aggregate mismatch cyclicality

Figure 14 plots time series of aggregate mismatch and the unemployment rate. To construct the aggregate mismatch series (left scale), we residualize mismatch with respect to the controls from the baseline regression (11), and then compute a symmetric 2-quarter moving average to smooth the series from seasonal fluctuations.

By construction, the mismatch series is centered around zero. For most of the sample period, we observe a negative correlation between mismatch and the unemployment rate, with

search. Our restriction on career-switching captures this, in reduced form, by forcing employed workers to quit their job and search "full time" when seeking a career change.

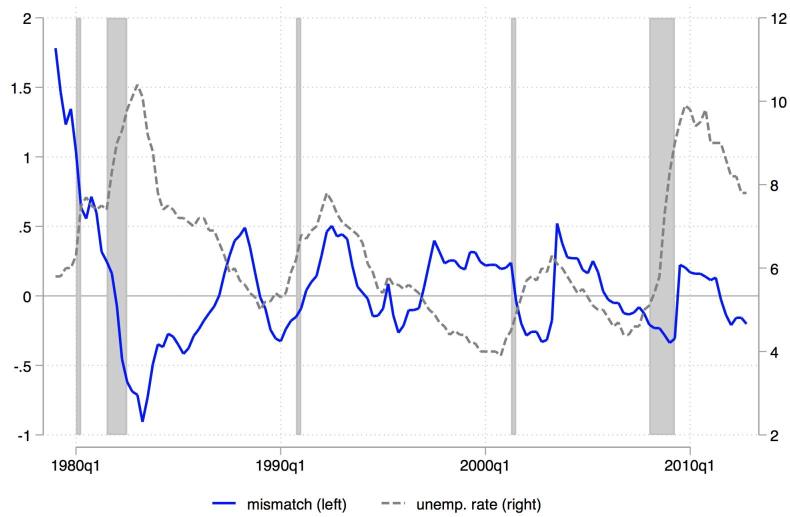


Figure 14: Time series of aggregate mismatch and the unemployment rate. *Notes.*—The figure shows time series of residualized mismatch (left scale, controlled by the baseline regression controls) and the unemployment rate (right scale, in percentage points). Shaded regions correspond to NBER-defined recessions.

two notable exceptions: the period around the 1990-1991 recession, and the years prior to the Great Recession where both series are declining.

For comparison, the shaded regions indicate NBER-defined recessions. Both the unemployment rate and mismatch are lagging the NBER-defined recessions. We further explore this in Appendix G.2.

G.2 Alternative cyclical indicators

Our baseline recession indicator defines recessions as times when the unemployment rate exceeds its long-term average of about 6.5%. Using an unemployment-defined cyclical indicator is natural for our purpose because, by definition, fluctuations in mismatch are tied to job flows, especially in and out of unemployment.

Here, we examine the cyclical properties of mismatch using alternative business cycle indicators: (A) unemployment rate, (B) HP-filtered unemployment rate, (C) NBER recession indicator (applied to all months within a quarter), and (D) 4-quarter lag of the NBER recession indicator. We repeat the main regression for mismatch cyclicality in (11) using these four alternative measure of the business cycle.

Table 11 presents the results. In Panels A and B we use unemployment-related indicators and confirm our baseline results: total mismatch is procyclical; the procyclicality of total mismatch is primarily driven by underqualified workers being laid-off in recessions; and new

Table 11: Mismatch cyclicality: Alternative cyclical indicators

Dependent variable ($\times$100):	$m_{i,t}$ (1)	$m^+_{i,t}$ (2)	$m^-_{i,t}$ (3)
Panel A: Unemp. rate			
Job stayers ($\beta_1 + \beta_2$)	−.167***	−.047	−.120***
	(.056)	(.041)	(.038)
New hires ($\beta_1 + \beta_3$)	.247**	.140*	.107*
	(.100)	(.072)	(.065)
Total cyclicality	−.147***	−.037	−.110***
	(.055)	(.041)	(.037)
Panel B: Unemp. rate deviations			
Job stayers ($\beta_1 + \beta_2$)	−.123	−.013	−.110**
	(.084)	(.064)	(.054)
New hires ($\beta_1 + \beta_3$)	.371	.172	.198
	(.239)	(.178)	(.149)
Total cyclicality	−.091	−.006	−.084
	(.082)	(.063)	(.053)
Panel C: NBER			
Job stayers ($\beta_1 + \beta_2$)	−.075	−.060	−.015
	(.112)	(.083)	(.075)
New hires ($\beta_1 + \beta_3$)	−.124	−.287	.162
	(.332)	(.244)	(.198)
Total cyclicality	−.075	−.064	−.011
	(.109)	(.081)	(.072)
Panel D: NBER lagged			
Job stayers ($\beta_1 + \beta_2$)	−.198**	−.065	−.133**
	(.090)	(.067)	(.059)
New hires ($\beta_1 + \beta_3$)	1.061***	.695***	.366*
	(.329)	(.249)	(.204)
Total cyclicality	−.150*	−.046	−.104*
	(.090)	(.066)	(.059)

Notes.—Standard errors clustered at the worker level are in parenthesis. Asterisks, *, **, ***, indicate coefficients that are significantly different from 0 at the 10%, 5%, 1% level, respectively. Dependent variables are multiplied by 100 (so mismatch ranges from 0 to 100). Panel A and B include yearly fixed effects.

Table 12: Cyclicality in mismatch among job-to-job transitions in the data

Dependent variable (×100):	$m_{i,t}$ (1)	$m^{+}_{i,t}$ (2)	$m^{-}_{i,t}$ (3)
Job-to-job transitions (β_1)	.265	.442**	−.177
	(.293)	(.215)	(.220)

Notes.—Standard errors clustered at the worker level are in parenthesis. Asterisks, *, **, ***, indicate coefficients that are significantly different from 0 at the 10%, 5%, 1% level, respectively. Dependent variables are multiplied by 100 (so mismatch ranges from 0 to 100).

hires from unemployment have countercyclical fluctuations in mismatch.

Next, we examine mismatch cyclicality using the NBER-defined indicators. When using the contemporaneous NBER indicator (Panel C) we obtain insignificant coefficients. In contrast, when using the lagged NBER indicator (Panel D) the coefficients are highly significant and comparable in size to our baseline results. These results are explained by the lag in unemployment compared to the NBER recession indicator as visible in Figure 14.[49] As argued above, this matters, because fluctuations in mismatch are intrinsically tied to job flows, explaining why it is the lagged NBER indicator that is significantly correlated with mismatch.

In summary, we conclude that mismatch contemporaneously correlates with unemployment measures, while it correlates with the lagged NBER indicator.

G.3 Job-to-Job Transitions

Our model has sharp predictions for mismatch cyclicality among job stayers and new hires from unemployment, which are corroborated in the data and reported in the main body in Table 4. Here we supplement the analysis with empirical observations of mismatch cyclicality for job-to-job movers. Table 12 shows the impact of a recession on mismatch among job-to-job movers, as captured by β_1 in specification (11). We obtain significant countercyclical fluctuations in positive mismatch. This observation is consistent with procyclical upgrading of match quality driven by job-to-job transitions, as examined by Gertler, Huckfeldt and Trigari (2020). While we see a significant increase in overqualification during recessions, we do not see a significant impact on underqualification or total mismatch among job-to-job movers.

[49] The contemporaneous correlation between the unemployment rate and the NBER indicator is 0.12, while the cross-autocorrelation between the unemployment rate and the lagged NBER indicator is 0.5.

Table 13: Mismatch cyclicality in the data: NLSY79 and NLSY97

Dependent variable (×100):	$m_{i,t}$ (1)	$m^{+}_{i,t}$ (2)	$m^{-}_{i,t}$ (3)
Job stayers ($\beta_1 + \beta_2$)	$-.267$**	$-.009$	$-.258$***
	(.114)	(.080)	(.076)
New hires ($\beta_1 + \beta_3$)	.552**	.487***	.065
	(.236)	(.168)	(.185)
Total cyclicality	$-.238$**	.013	$-.251$***
	(.113)	(.080)	(.075)

Notes.—Standard errors clustered at the worker level are in parenthesis. Asterisks, *, **, ***, indicate coefficients that are significantly different from 0 at the 10%, 5%, 1% level, respectively. Dependent variables are multiplied by 100 (so mismatch ranges from 0 to 100).

G.4 Two cohorts: NLSY79 and NLSY97

For our baseline estimates, we only use data from the NLSY 1979 cohort. The reason is that the 1997 cohort does not contain data on the Rotter Locus of Control Scale and the Rosenberg Self-Esteem Scale, which we use to construct a measure of social ability and which is found to be a key predictor of labor market outcomes by Guvenen et al. (2020) and Lise and Postel-Vinay (2020). Here, we replicate the mismatch cyclicality results extending the data to include the 1997 cohort. Due to the lack of data needed to measure social ability in the 1997 survey, mismatch now only comprises math, verbal, and technical skills. Table 13 presents the results which are analogue to those in Table 4 in the main text. Overall, the coefficients are very similar across both samples. We conclude that our results are robust to including one or two cohorts.

G.5 By Skill Dimension

In the main text, we assess the cyclicality of a mismatch index, defined in (10), which aggregates mismatch across four skill dimensions using market weights. Here, we examine the cyclical properties of mismatch for each skill dimension: math, verbal, technical and social. We do so by running the same empirical specification as in (11) but separately by skill. Table 14 is the analog to Table 4 in the main text. As before, we report the cyclicality of total, positive, and negative mismatch, and report coefficients separately for job stayers, new hires from unemployment, and the totality of workers.

Overall, mismatch cyclicality by skill dimension has the same cyclical properties as total mismatch. For each skill dimensions, we consistently obtain procyclical mismatch among job stayers (first row of each panel), that is, mismatch decreases for job stayers in recessions

Table 14: Mismatch cyclicality in the data: By skill dimension

Dependent variable (×100):	$m_{i,t}$ (1)	$m^{+}_{i,t}$ (2)	$m^{-}_{i,t}$ (3)
Panel A: Math			
Job stayers ($\beta_1 + \beta_2$)	−.379*	−.030	−.349*
	(.021)	(.791)	(.001)
New hires ($\beta_1 + \beta_3$)	.693*	.535*	.159
	(.066)	(.040)	(.492)
Total cyclicality	−.334*	−.002	−.332*
	(.041)	(.983)	(.002)
Panel B: Verbal			
Job stayers ($\beta_1 + \beta_2$)	−.298*	−.013	−.311*
	(.062)	(.900)	(.002)
New hires ($\beta_1 + \beta_3$)	.986*	.741*	.245
	(.007)	(.002)	(.274)
Total cyclicality	−.235	.052	−.288*
	(.139)	(.625)	(.005)
Panel C: Technical			
Job stayers ($\beta_1 + \beta_2$)	−.273	.028	−.301*
	(.102)	(.823)	(.005)
New hires ($\beta_1 + \beta_3$)	.006	.089	−.083
	(.987)	(.740)	(.726)
Total cyclicality	−.247	.044	−.290*
	(.137)	(.719)	(.007)
Panel D: Social			
Job stayers ($\beta_1 + \beta_2$)	−.044	.152	−.196*
	(.784)	(.156)	(.086)
New hires ($\beta_1 + \beta_3$)	.551	.413*	.138
	(.139)	(.098)	(.575)
Total cyclicality	−.012	.167	−.179
	(.938)	(.119)	(.114)

Notes.—Standard errors clustered at the worker level are in parenthesis. Asterisks, *, **, ***, indicate coefficients that are significantly different from 0 at the 10%, 5%, 1% level, respectively. Dependent variables are multiplied by 100 (so mismatch ranges from 0 to 100).

and the decline is entirely driven by layoffs of underqualified workers (those with negative mismatch, column 3). Additionally, we obtain countercyclical mismatch among new hires from unemployment (second row of each panel), in this case driven by more overqualified workers finding jobs in recessions than in expansions (those with positive mismatch, column 2). Finally, total mismatch is procyclical (third row of each panel), as before. While all skill dimensions show similar cyclical properties, math and verbal skills are the ones with the highest statistical significance.

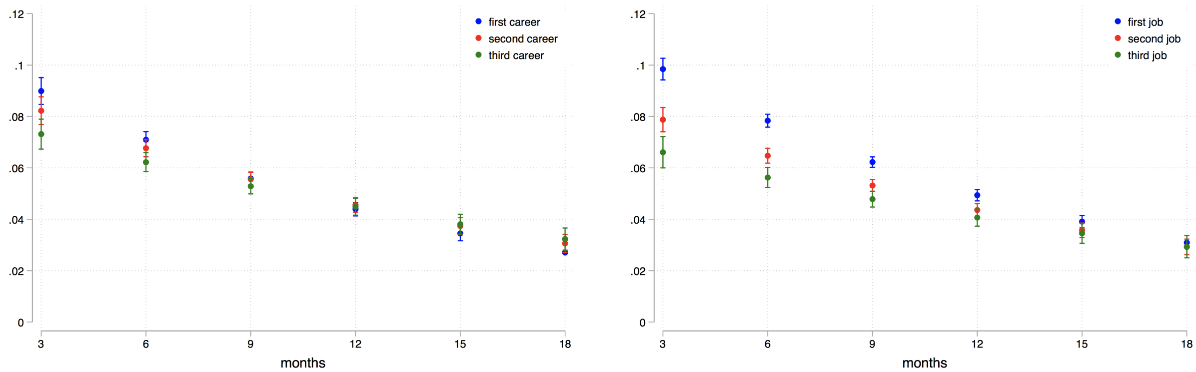
H Across-career vs. Within-career Experiences

This appendix provides further suggestive evidence on the assumptions, mechanisms, and implications of our learning model. Our model assumes that learning is geared towards workers' ability in their *current* career. Moreover, for simplicity, we further assume that ability is uncorrelated across careers (but have noted that this assumption is not essential). Using the NLSY data, we present various pieces of evidence that validate these assumptions and, furthermore, corroborate key implications of the model on the difference between career switches and job transitions within a career.

H.1 Evidence from Job Separation Hazards

In the model, we make the simplifying assumption that skills are independent *across* careers. This assumption implies that the separation hazard should be independent of the number of careers previously held by a worker. Figure 15a shows that this is indeed the case in the data. It plots the job separation hazard conditional on the number of careers held. Corroborating the independence assumption, there are no significant differences between the separation hazards for the first, second, and third career.

In contrast, the model implies that learning *within* careers is a relevant factor and thus one would expect that job separation hazards would depend on prior work experience *within* that same career. Figure 15b confirms this prediction by plotting the job separation hazard conditional on the number of jobs held by a worker within the same career. We observe that the separation hazard for the first job in a career is significantly larger than for subsequent jobs in the same career; moreover, the separation hazard declines at a steeper rate for the first job in a career, consistent with uncertainty being highest at the beginning of a career.



(a) By number of careers previously held

(b) By number of jobs within same career

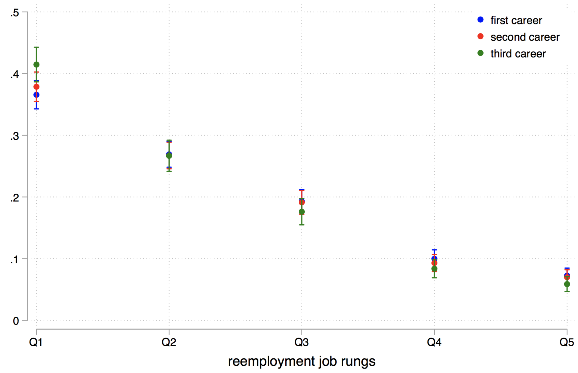
Figure 15: Job separation hazards. Panel Afor subsequent jobs within the same career. *Notes.*—Separation hazards include EU and EE' transitions and assume a linear baseline hazard. Error bars indicate 95% confidence intervals. All moments are residualized with respect to race, gender, education, region of residence, and initial age at the start of the job spell.

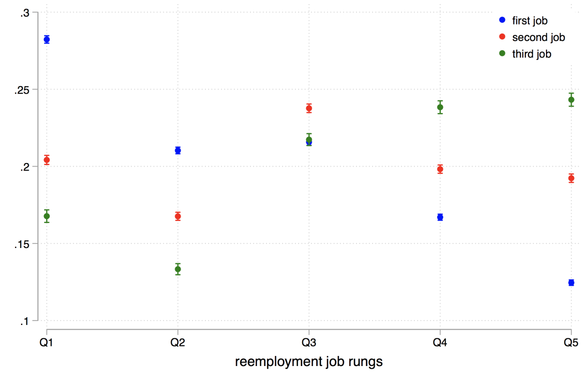
H.2 Evidence from Distributions of Reemployment Job Rungs

Next, we use the distribution of reemployment job rungs to provide additional indirect evidence. As with the separation hazards, uncorrelated learning across careers implies that the distribution of reemployment job rungs should be independent of the number of careers. This is indeed confirmed in Figure 16a, where we observe that the likelihood to start at any job rung is independent of the number of careers previously held.

In particular, regardless of the number of careers held before, upon a career switch a worker is always more likely to start at the bottom of the job ladder, consistent with the predictions of the model. Figure 17 further substantiates this finding, showing that the tendency of workers to start at the bottom job rung after career-switches holds *independently of their job rung in the previous career*. This fact supports our prediction that career switches entail restarting learning about untried skills and thus workers optimally aim for jobs at the bottom of the new job ladder.

Regarding *within* career transitions, Figure 16b shows that the distribution of reemployment job rungs within a career is affected by the number of previously held jobs, consistent with learning within careers. This distribution is initially skewed towards the lowest job rung and becomes increasingly skewed towards the highest job rung as career tenure increases. These observations are consistent with the predictions of the model: Short-tenure workers are more likely to start at the bottom of the job ladder while long-tenure workers are more likely to get reemployed at higher rungs as explored in Section 4.





(a) By number of careers previously held

(b) By number of jobs within same career

Figure 16: Distribution of reemployment job rungs. *Notes.*—Error bars indicate 95% confidence intervals. Include EU and EE' transitions. All moments are residualized with respect to race, gender, education, region of residence, a quadratic polynomial in age, and month and 5-year fixed effects.

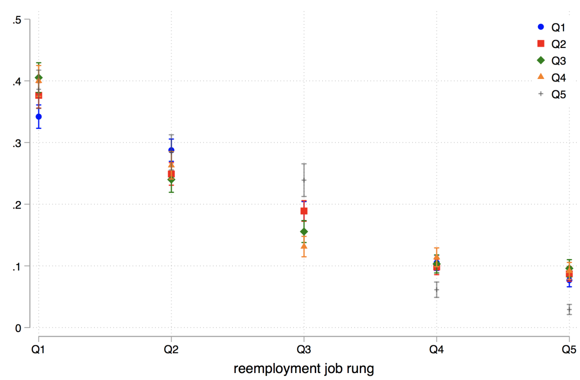


Figure 17: Reemployment job rungs for career switchers, conditional on previous position in the job ladder (by rung quintiles). *Notes.*—Error bars indicate 95% confidence intervals. Transitions include EUE' and EE' transitions. All moments are residualized with respect to age, gender, race, education, region, industry, and 5-year fixed effects

Table 15: Empirical distribution of mismatch across job rungs

Quintile of $r_{i,t}$	$m_{i,t}$	$m_{i,t}^{+}$	$m_{i,t}^{-}$
Q1	2.010	7.780	−5.770
Q2	−0.720	2.090	−2.800
Q3	−0.960	−1.470	0.520
Q4	−1.550	−4.280	2.730
Q5	0.730	−6.070	6.810

Notes.—Mismatch is residualized with respect to region, a quadratic polynomial in age, and individual, month and 5-yearly fixed effects.

H.3 Evidence from the Distribution of Mismatch Across Job Rungs

Finally, our model predicts that, with the exception of the highest job rung, mismatch is declining in job rungs. Moreover, the decline is driven by a decline in overqualification, whereas underqualification becomes relatively more important at higher job rungs (c.f. Figure 6b). To explore this prediction, we use the generalized model introduced in Section 3.1 to assign a task complexity $r_{i,t}$ to each job. We then compute the average mismatch (residualized with respect to region, a quadratic polynomial in age, and individual, month and 5-yearly fixed effects) for each quintile of the task complexity distribution. Table 15 reports the results. Consistent with the model, total mismatch is declining across job rungs with the exception of the highest job rung, and the decline is driven by overqualification.