Mismatch Cycles[*]

Isaac Baley

Universitat Pompeu Fabra, CREi, Barcelona School of Economics, and CEPR

Ana Figueiredo

Erasmus School of Economics and Tinbergen Institute

Robert Ulbricht

Boston College

March 16, 2022

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.

"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 1.4 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 due to increased 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 Krolikowski (2017), Jung and Kuhn (2019), Jarosch (2021), and Huckfeldt (2022).[2] 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.[3] In addition, we provide indirect evidence towards the model's mechanism. First, career mobility is predicted by the suitability of workers' skills for their

---

[2]In the literature, the unemployment scar is explained by multiple channels. In all studies, part of the scar is attributed to a selection effect arising from the progression of employed workers to "better" jobs through on-the-job search. In Jung and Kuhn (2019), Jarosch (2021), and Huckfeldt (2022) an additional part of the scar is explained by human capital depreciating during unemployment. Finally, most related to our mechanism, Jarosch (2021) complements these channels with a high separation rate from bottom job rungs, which also adds inertia to workers' progression through the job ladder.

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

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 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 theoretical works by Lise and Postel-Vinay (2020) and Lindenlaub and Postel-Vinay (2020) that also emphasize the irreducibility of worker heterogeneity into a single unidimensional index.[4] 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.[5] 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 due to career mobility. 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.[6] 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

---

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

[5]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.

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

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 (2021) who provide direct evidence that overall sorting is countercyclical, Bowlus (1995) who provides indirect evidence 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[7]).

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,

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

$$\mathrm{d}\Sigma_{i,k}(t) = -\left(\frac{\Sigma_{i,k}}{\sigma}\right)^2 \mathrm{d}t.$$

---

[7]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.

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 $(\hat{a}_{i,k}, \Sigma_{i,k})$, 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.[8] 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, with $p(0) = p'(\infty) = 0$ and $p'(0) = \infty$.

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.

---

[8]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.

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 others may bundle together jobs with distinct skills.[9] 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 conserve 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 given by:

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

---

[9]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 Online Appendix G.1). Using 2-digit occupation codes, on the other hand, bundles together many occupations with vastly different skill mixes.

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 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 a 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

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

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 follows from 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 black square in the plot). Workers that are more optimistic regarding their skills in their current career, apply to more complex jobs (indicated by lighter shades of gray). 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

10

overqualified, it leads high expected-skill workers to apply for jobs for which they are on average underqualified.[10] This is because for high expected-skill workers the expected value 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 $\Sigma_k > 0$).[11] 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 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, given 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 switch careers whenever $\hat{a}_k$ is small (indicated by the dark gray 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

---

[10]This prediction is consistent with our data on mismatch (introduced below), in which workers at the bottom job rungs are systematically overqualified, whereas workers in upper job rungs are systematically underqualified (see Online Appendix I.3).

[11]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 $\Sigma_k > 0$.

separate (with or without career-switch[12]). Workers in continuing relationships actively search 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 light-shaded area bordered by the separation region below and the no-search region above. If $\hat{a}_k$ falls into the upper light-shaded 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 $(\hat{a}, \Sigma)$.[13] 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 Online Appendix A.

Equilibrium and block-recursivity   An equilibrium is a joint worker–firm 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 (A.1) and (A.2) in Online 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

---

[12]Workers may separate from their jobs, yet continue with their current career, because the gains from increasing the job finding rate may outweigh the cost of unemployment, as observed for workers whose current job rung is far from their desired one.

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

inform ourselves about parameters unique to our model, we use a combination of moments 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 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 $\boldsymbol{a}_i \equiv (a_{i,1}, \dots, a_{i,J})$ defines a vector of skills for each worker i over J basic aptitudes. Similarly, $\boldsymbol{q}_{k,r} \equiv r \cdot (w_{k,1}, \dots, w_{k,J})$ 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 $k \in \{1, \dots, K\}$. The difference is that each k now maps into a vector of weights $(w_{k,1}, \dots, w_{k,J})$ 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.[14] With this in mind, we interpret two occupations observed in the data as different careers if their requirement vectors are "sufficiently orthogonal". Specifically, let $\varphi : \mathbb{R}^J \times \mathbb{R}^J \to [0, \pi/2]$, define the angular distance between two skill vectors

---

[14]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 Online Appendix B for two examples where skills are perfect complements and perfect substitutes.

$\boldsymbol{q}_1$ 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).[15] 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.[16]

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 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). Online Appendix G.1 provides examples for occupations inside and outside the $\bar{\varphi}$-cones of "economists" and "dental assistants".

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 Online Appendix I, 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[17]). 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.

---

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

[16]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 $\boldsymbol{q}_1 \cdot \boldsymbol{q}_2' \equiv \sum_j v_j q_{1,j} q_{2,j}$.

[17]In practice, predicting workers' performance across careers is likely impaired by uncertainty about the *importance* of skills for different careers. For instance, suppose skills enter production through a 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, yet it cannot easily projected across careers without additional knowledge about both $\boldsymbol{w}_k$ and $\boldsymbol{w}_{k'}$.

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.[18] To make them comparable, we normalize each skill dimension in terms of percentile ranks.[19] Online Appendix F 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.[20] 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 angular distance between the two skill vectors, $\varphi(\boldsymbol{q}_{i,t}, \boldsymbol{q}_{i,t+1})$, exceeds $\bar{\varphi}$. The threshold $\bar{\varphi}$ is chosen so that the average correlation in requirements (across skill dimensions) is zero for career moves: $\sum_{j=1}^{4} v_j \mathrm{Corr}(q_{i,t,j}, q_{i,t+1,j}) = 0$, where $\{v_j\}$ is a set of market weights described below.[21] 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 (2021), and Huckfeldt (2022).

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:

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

---

[18]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).

[19]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.

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

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

Here v_j are "market weights", obtained from the regression coefficients on each of the four mismatch dimensions in a Mincer regression (normalized so $\sum_{j=1}^{4} v_j = 1$).[22] 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

$$m_{i,t}^{+} \equiv \sum_{j=1}^{4} v_j \max\{a_{i,j} - q_{i,t,j}, 0\} \qquad\qquad m_{i,k}^{-} \equiv \sum_{j=1}^{4} v_j \max\{q_{i,t,j} - a_{i,j}, 0\}. \qquad (11)$$

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 relative search effort documented in Holzer (1987) and Faberman et al. (2022).[23] 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 motives, 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.[24]

We specify the set of potential task-complexities, $\mathcal{R}$, using a seven-point grid given by $\{0, 0.5, \ldots, 3\} \cdot S_0^{1/2}$, 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.[25] 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%.

---

[22]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.

[23]Conditional on searching for jobs, Holzer (1987) and Faberman et al. (2022) document a relative time spent on search activities among employed workers of 0.48 and 0.51, respectively.

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

[25]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.

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 percentage points in the US.

To identify the speed of learning, parametrized by σ, we target an average slope of the 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.[26] 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" job rung $r^*(a_0, S_0, z)$. We note that according to the data, workers starting a new career are on average overqualified.

---

[26]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.

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 large 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 $0.35 \cdot S_0$. 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 $1.0 \cdot S_0^{1/2}$ 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 that illustrates 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 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.

Specifically, the worker engages in on-the-job search whenever their desired job rung r^* differs from the current job rung r (gray dashed line in panel 1). Graphically, on-the-job search episodes occur whenever the mean belief falls outside the lightly 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 experiences five additional 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 gray dashed lines in the second panel, as observed after about 0.9, 1.4 and 2.4 years.[27] 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 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 career by low ability workers.

---

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

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 flat relationship between job rungs and tenure.[28] 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 increase in job rungs with tenure.

The slow reallocation of job rungs causes mismatch to be persistent as well. Moreover, as shown in Figure 6b, this naturally translates into a negative correlation between mismatch and job rungs.[29] 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 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 Online Appendix C 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

---

[28]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 Online Appendix C, we assess the cost of this implicit friction, finding that on average it amounts to 4.7 months of average output per worker.

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

resulting earnings profile in tenure along with the no mismatch counterfactual.[30] 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, most of which is explained by the eventual outflow of low-ability workers.

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, 2022). 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 (2022), earnings losses are in large part 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 $t = 0$.[31] Relative to the counterfactual of no job loss, earnings are reduced by roughly 47 percent one year after the displacement, and 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 logic 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 rung of displaced workers in the sequel of 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[32]), workers that are displaced

---

[30]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.

[31]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. Without the tenure requirement, the earnings loss from displacement amounts to 45, 16 and 8 percent after one, five and ten years, respectively.

[32]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.

from their career enter the labor market at the bottom rung and take years to advance 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, 2022) to 15–20 percent (Jarosch, 2021).

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   Cleansing and sullying

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 2.7 percent smaller in recessions than in expansions—, and countercyclical fluctuations in overqualification—positive mismatch being 1.6 percent larger in recessions than in expansions. Combined, total mismatch is mildly procyclical, being on average 0.5 percent smaller in recessions than in expansions.

The overall cyclicality in mismatch is the result of opposing effects operating at different tenure levels. This is illustrated in Figure 8, which breaks down the cyclicality by the time a worker has been continuously employed since their last unemployment spell ("employment tenure"). Among new hires (workers with zero employment tenure), both over- and under-qualification are significantly heightened in recessions. Combined, total mismatch among new hires increases by about 4 percent in recessions, exposing a recessionary "sullying" effect among newly employed.

The increase in mismatch among newly employed contrasts starkly with the cyclicality in mismatch among workers with an employment tenure of more than 9 months, for whom mismatch is reduced in recessions, evocative of a recessionary "cleansing" effect. For instance, among workers that have been continuously employed for 2 years, both over- and underqualification are about 4 percent smaller in recessions compared to expansions.

5.1.2 Understanding the mechanism

Through the lens of the model, sullying and cleansing are two sides of a single mechanism. To illustrate, Figure 9 compares the separation/continuation policy of worker–firms across business cycle states. In the one extreme, if a worker is well-matched to their current job, the match is continued in both business cycle states (white regions in the graph). In the other extreme, if a worker is excessively mismatched, the match is always terminated (dark gray regions). Finally, most relevant for us, when (expected) mismatch is between these two extremes, the match is continued in expansions but terminated in recessions (light gray regions). This is the natural consequence of matches being more productive in expansions, prompting worker–firms to tolerate a higher level of mismatch in expansions, and triggering the cleansing of workers with above-average levels of mismatch during recessions.

Interestingly, it is precisely the cleansing of mismatched workers that induces the sullying among new hires. This is because workers who lose their jobs during recessions tend to switch careers, which in turn comes with higher skill uncertainty and higher levels of mismatch.

To appreciate this link from cleansing to sullying, consider the career-switching threshold depicted by the solid black line in Figure 9.[33] *Unemployed* workers with a skill estimate $\hat{a}_k$ below this threshold switch careers. Importantly, however, even if a worker anticipates to switch careers once they lose their job, they may still prefer to remain employed in the former career pro tempore to avoid the cost of finding a new job. Graphically, such workers with a "loose career attachment" exist whenever the career-switching threshold (solid black line) lies above the separation threshold for $z = z_L$ (dotted gray line), as it does for the bottom three job rungs. Workers in this region of the graph are precisely the ones switching careers after being cleansed out in recessions and are responsible for the sullying effect.

For sullying to be quantitatively important, workers with a loose career attachment must make up a significant fraction of the cleansing region. Figure 10a shows the composition of workers inside the cleansing region, confirming that indeed the majority of cleansed workers switch careers upon separation (86 percent at the ergodic distribution).[34] The figure further reveals that virtually all of the cleansing is concentrated at the bottom job rungs, consistent with that mismatch is more prevalent at the bottom job rungs.

Figure 10b sheds additional light on the mechanism. First, it compares cleansed jobs

---

[33]The threshold is depicted for $z = z_L$, which is the relevant one here. The threshold for $z = z_H$ is virtually identical (differing in a single grid point).

[34]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 are precisely the workers with a loose career attachment for whom applying to other jobs within the same career has no value.

with surviving ones in terms of their mismatch. Both over- and underqualification are more pronounced in cleansed jobs. Hence cleansing indeed has a procyclical impact on mismatch. Second, when cleansed workers re-apply to jobs within their current career, they are naturally less mismatched in their new job. When switching careers, however, mismatch in the new career may initially exceed the one in the cleansed job. This is the case for overqualification, echoing that our calibration matches the empirically strong prevalence of overqualification among career-switchers. In particular, the initial increase in overqualification is sufficiently large for positive mismatch to be overall countercyclical (Table 3). By contrast, underqualification among career-switchers is less than in cleansed jobs so that underqualification is overall procyclical.

In sum, both the cleansing and sullying during recessions are driven by transitions in and out of bottom job rungs. The combination of these two effects explains the tenure profile of the mismatch cyclicality documented in Figure 8: At low tenure levels, a match is likely to be created within the current aggregate state so that sullying becomes the dominant factor. At high tenure levels, a match is likely to pre-date the current aggregate state so that cleansing becomes the dominant factor. Finally, at very high levels of tenure, few workers are mismatched to begin with, resulting in a negligible cyclical impact on mismatch.

5.1.3 Consequences for aggregate productivity

We next assess how the cyclical sorting patterns discussed so far affect output. To do so, we examine the endogenous component of labor productivity determined by the selection of workers into job rungs and careers,

$$e_{i,k,r}(t) \equiv \frac{y_{i,k,r}(t)}{\exp\left(z(t)\right)},$$

which we call "labor efficiency". Using this measure, we find that the sorting of workers into jobs and careers translates to an increase in labor efficiency of about 0.4 percent in recessions compared to expansions. The black crossed line in Figure 11 exscinds the overall cyclicality of labor efficiency into its cyclicality at different employment tenures. Echoing the sullying among new hires, average labor efficiency among new hires is about 1 percent smaller in recessions than in expansions. By contrast, average labor efficiency among workers who have been continuously employed for two years is about 1.5 percent larger in recessions.

It is instructive to decompose (log) labor efficiency using (1) and (11) as follows:

$$\log e_{i,k,r}(t) = \eta a_{i,k} - \eta m^+_{i,k,r} - (1-\eta)m^-_{i,k,r}.$$

24

The decomposition identifies three endogenous components that determine labor efficiency. Conditional on skills $a_{i,k}$, labor efficiency decreases in both over- and underqualification, reflecting the direct impact of mismatch on output (if underqualified) and the opportunity cost of switching to a higher task complexity (if overqualified). Additionally, labor efficiency further varies with the career choice, which determines the skills employed for production (circled gray line in Figure 11). Interestingly, the skill cyclicality is determined by the same two forces determining the mismatch cyclicality: Among new hires, the rise in career mobility naturally translates into lower average skills. In our calibration, this explains about two thirds of the overall drop in labor efficiency among new hires. By contrast, the cleansing of highly mismatched workers during recessions shifts the composition of the workforce towards more skilled workers. The reason is again that mismatch is more pronounced among workers with low career tenures, which is correlated with lower than average skills. In our calibration, this explains about three quarters of the overall rise in labor efficiency among workers who have been continuously employed for two years.

We note that the countercyclicality in labor efficiency does not immediately translate into predictions regarding aggregate labor productivity. 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 a shock to productive efficiency. In this case, overall labor productivity is given by $\exp(z)\mathbb{E}[e_{i,k,r}]$, 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.[35] In this case, aggregate labor productivity is entirely determined by the endogenous labor efficiency $\mathbb{E}[e_{i,k,r}]$ 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. (2022) as well as with the household balance sheet narrative of the Great Recession (Mian, Rao and Sufi, 2013).

---

[35]Here we tacitly assume that the real price of labor output fluctuates relative to home production b and the vacancy cost 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.

5.1.4 Sectoral displacement shocks

So far, shocks to the aggregate labor product affected all workers equally. We now use our model to explore the case where shocks are directed to certain careers. While stylized, one can view this exercise as an approximation to structural change or to recessions that disproportionately affect certain sectors such as leisure and hospitality during the 2020/21 pandemic.

Specifically, we consider a "sectoral shock" that displaces 1 percent of the labor force from their career. For simplicity, we assume that all workers in the affected careers are displaced, regardless of their employment status. In this pure form, the shock acts as a prototypical "sullying shock", forcing all affected workers to switch careers while shutting down any compositional cleansing effects. 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 12). First, aggregate productivity (or, equivalently, labor efficiency given that the displacement shock is uncorrelated with z) 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 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.

5.1.5 On the relative importance of cleansing vs. sullying

Given its calibration, the model predicts that cleansing dominates sullying for the *average* recession. Before examining mismatch in the data, we briefly comment on the relative contribution of cleansing and sullying for other recessions. The model suggests that the relative importance of the two forces hinges on both the *scale* and the *scope* of recessions.

Regarding the scale, Online Appendix E compares impulse responses of mismatch and labor efficiency across aggregate shocks of different sizes. The analysis suggests that the

relative importance of sullying is declining in the scale of a recession. The reason is that once the shock is of a certain scale, further scaling it up will only amplify cleansing but will not induce any additional sullying. For instance, for small shocks with an initial unemployment response of 1 percentage points, sullying roughly cancels the cleansing effect on mismatch after 12 months into the recession. By contrast, for large shocks with an initial unemployment response of 5 percentage points, cleansing strongly dominates, reducing mismatch by about 8.6 percent after 12 months.

At the same time, as evident from our extension on sectoral shocks, the relative importance of the two forces hinges critically on the scope of a recession. In the case of sectoral shocks, recessions are of small scope (they only affect a few careers), but of large magnitude (in our stylized implementation, they fully shut down all affected careers). In this prototypical form, sectoral shocks do not induce any cleansing and, instead, lead to recessions marked by especially pronounced sullying.

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:

$$
m_{i,t} = \beta_0 + (\beta_1 + \beta_2 JS_{i,t} + \beta_3 UE_{i,t}) \times \text{recession}_t +
$$
$$
+ \gamma \times (JS_{i,t}, UE_{i,t}, x_{i,t}) + \delta_i + \delta_{m_t} + \delta_{y_t} + \epsilon_{i,t}. \quad (12)
$$

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[36]; 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 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.[37] 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

---

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

[37]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 H.3 in Online Appendix H.4 for details on the implied mismatch cyclicality among job-to-job transitions.

Parker, 1994).

Table 4 reports the estimated business cyclicality. Looking at job stayers, mismatch declines in recessions by an average of 0.29 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 (0.65 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.

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

$$\Delta\mathbb{E}_{JS}[m^+] = .094 \qquad\qquad \Delta\mathbb{E}_{JS}[m^-] = -.210$$
$$\Delta\mathbb{E}_{UE}[m^+] = .416 \qquad\qquad \Delta\mathbb{E}_{UE}[m^-] = .148,$$

where Δ denotes the difference in conditional means, $\mathbb{E}_L[\cdot] - \mathbb{E}_H[\cdot]$, computed at the ergodic distribution.[38] 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.

---

[38]In analog to Table 4, the cyclical differences are multiplied by 100.

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. Online Appendix I 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:

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

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

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

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.[39]

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

$$\sum_{j=1}^{4} \text{fe}_{i,t,j} = \beta_0 + \beta_1 \sum_{j=1}^{4} \text{pe}_{i,t,j} + \epsilon_{i,t}. \tag{13}$$

[39]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.

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 (13) 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.

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 $a_i(k) \equiv (w_{k,1}, \dots w_{k,J}) \times \boldsymbol{a}_i'$ as the suitability of worker i's skills for their current career k, determined 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:

$$\text{career switch}_{i,t} = \beta_0 + \beta_1 a_i(k_{i,t-1}) + \gamma x_{i,t} + \delta_{m_t} + \delta_{y_t} + \epsilon_{i,t}, \tag{14}$$

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:

$$m_{i,t} = \beta_0 + \beta_1 \text{career switch}_{i,t} + \gamma x_{i,t} + \delta_{m_t} + \delta_{y_t} + \epsilon_{i,t}, \tag{15}$$

using the same set of controls as in (14). 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 (15) 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 (14) and (15). 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, 2020). It is distinguished from the existing literature by its analytical tractability, which opens the door to the 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, 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. 2021. "Unemployment and Endogenous Reallocation over the Business Cycle." *Edinburgh School of Economics Discussion Paper Series No. 302.*

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." *American Economic Journal: Macroeconomics*, forthcoming.

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. 2021. "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. 2022. "Job Search Behavior among the Employed and Non-Employed." *Econometrica*, forthcoming.

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.

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. 2022. "The Slope of the Phillips Curve: Evidence from U.S. States." *Quarterly Journal of Economics*, forthcoming.

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

Huckfeldt, Christopher. 2022. "Understanding the Scarring Effect of Recessions." *American Economic Review*, forthcoming.

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.

Krolikowski, Pawel. 2017. "Job Ladders and Earnings of Displaced Workers." *American Economic Journal: Macroeconomics*, 9(2): 1–31.

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

Lindenlaub, Ilse, and Fabien Postel-Vinay. 2020. "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.

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.

Table 1: Targeted moments

Fitted Moments	Model	Data	Origin
$\mathbb{E}_L[U] - \mathbb{E}_H[U]$	.028	.028	BLS
$\mathbb{E}[\text{UE rate}]$	.425	.425	Shimer (2012)
$\mathbb{E}[\text{EU rate}]$	.035	.035	Shimer (2012)
$b/\mathbb{E}[y]$	.706	.710	Hall and Milgrom (2008)
$\epsilon_{UE/\theta}$	.280	.280	Shimer (2005)
$\mathbb{E}[\log(haz_3/haz_{18})]$	1.37	1.37	NLSY79
$\mathbb{E}[\chi = 1]$	.421	.422	NLSY79, O*NET
$\mathbb{E}_L[\chi = 1] - \mathbb{E}_H[\chi = 1]$	.069	.069	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.

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	0.0172, 0.0128
Estimated		
z_H	Aggregate log-productivity in expansions	0.301
b	Home production utility	0.985
c	Flow cost of vacancies	0.007
γ	Matching function parameter	0.514
η	Return on task complexity	0.494
a_0	Unconditional mean of skills	0.105
$S_0^{1/2}$	Standard deviation of skills	0.357
σ	Standard deviation of signal noise	2.463
δ	Exogenous separation rate	0.012
ϵ	Exogenous career switching rate	0.003

Table 3: Cyclicality of mismatch in the model

Mismatch measure	$m_{i,t}$	$m_{i,t}^{-}$	$m_{i,t}^{+}$
Cyclical difference (in %)	-0.48	-2.70	1.61

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.

Table 4: Cyclicality of mismatch in the data

Dependent variable (×100):	$m_{i,t}$	$m_{i,t}^{+}$	$m_{i,t}^{-}$
	(1)	(2)	(3)
Job stayers $(\beta_1 + \beta_2)$	−0.292**	0.017	−0.309***
	(0.127)	(0.091)	(0.085)
New hires $(\beta_1 + \beta_3)$	0.648**	0.502**	0.146
	(0.291)	(0.209)	(0.185)
Total cyclicality	−0.249**	0.043	−0.292***
	(0.126)	(0.091)	(0.084)

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).

Table 5: Direct evidence for learning

Dependent variable:	$\sum_j \text{fe}_j$	fe_j			
		math	verbal	technical	social
	(1)	(2)	(3)	(4)	(5)
$\sum_j \text{pe}_j$	0.562***				
	(0.020)				
pe_j		0.556***	0.471***	0.331***	0.482***
		(0.019)	(0.020)	(0.020)	(0.020)
R-squared	0.321	0.331	0.257	0.155	0.251
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.

Table 6: Career mobility and mismatch

Dependent variable:	career switch$_{i,t}$	$m_{i,t}$	s.d.$(m_{i,t})$
	(1)	(2)	(3)
$a_i(k_{i,t-1})$	-0.071^{***}		
	(0.023)		
career switch$_{i,t}$		0.957^{***}	0.605^{***}
		(0.308)	(0.178)

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.
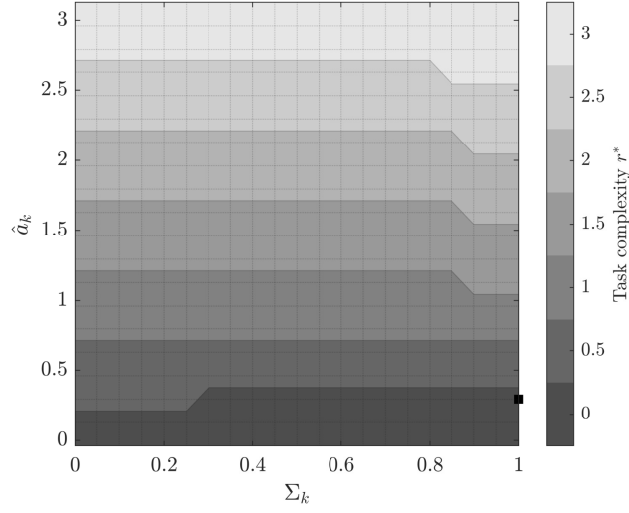
Figure 1: Job ladder. *Notes.*—The graph shows the task complexity r^* chosen as a function of expected ability $\hat{a}_k$ and uncertainty Σ_k. The black square marks the unconditional prior (a_0, S_0) for untried skill types. Values for $\hat{a}_k$, $\Sigma_k^{1/2}$ and r are denominated in units of $S_0^{1/2}$. 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.
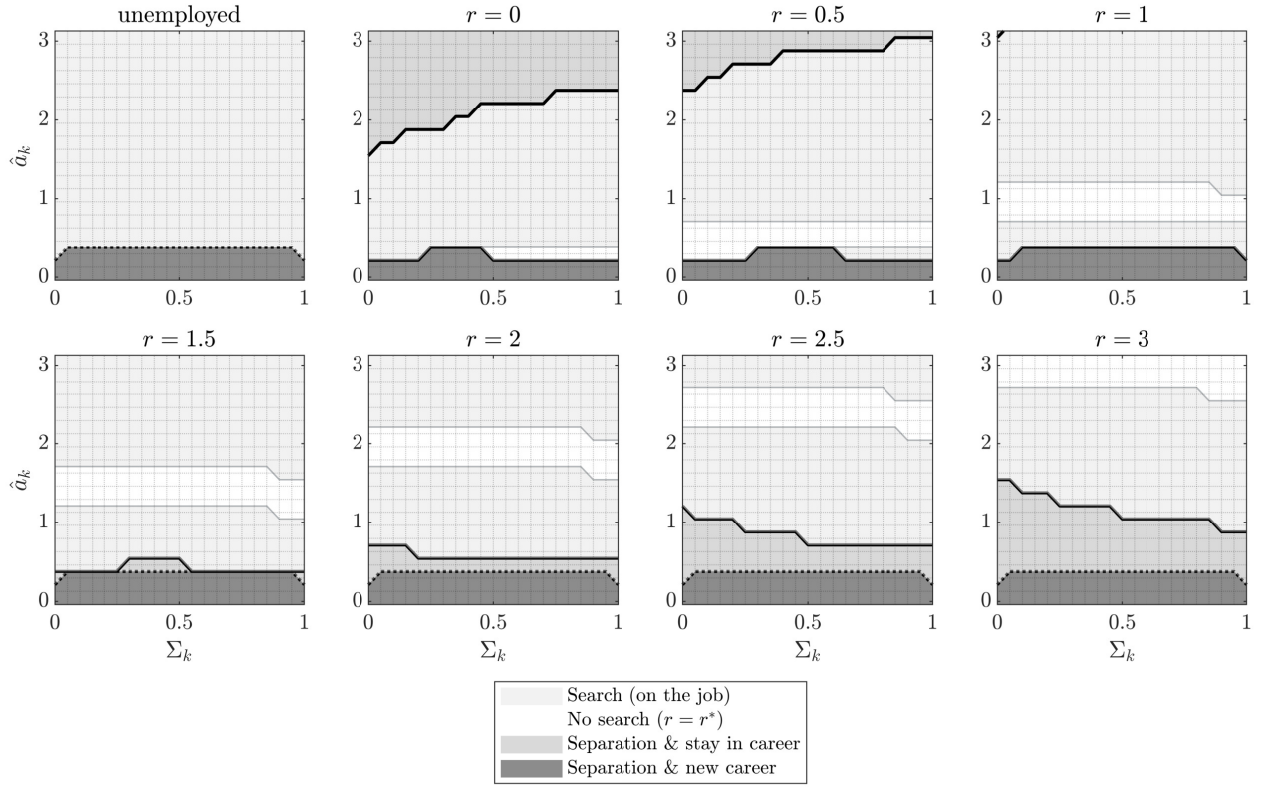
Figure 2: Search and separation policies. *Notes.*—The figure shows search policies as a function of expected ability $\hat{a}_k$, uncertainty Σ_k, and the employment state (unemployed/employed in job with complexity r). Values for $\hat{a}_k$, $\Sigma_k^{1/2}$ and r are denominated in units of $S_0^{1/2}$. The figure is plotted for $z = z_H$. See Section 3 for a detailed description of the parametrization.
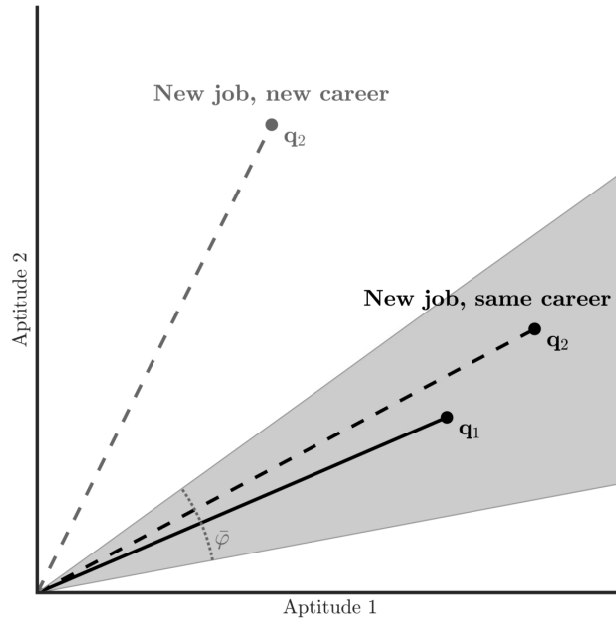
Figure 3: Schematic illustration of empirical measure of careers for $J = 2$. Job transitions from $\boldsymbol{q}_1$ to jobs within the $\bar{\varphi}$-cone are interpreted as transitions up and down the *same* job ladder; transitions to jobs outside the $\bar{\varphi}$-cone are interpreted as career switches.
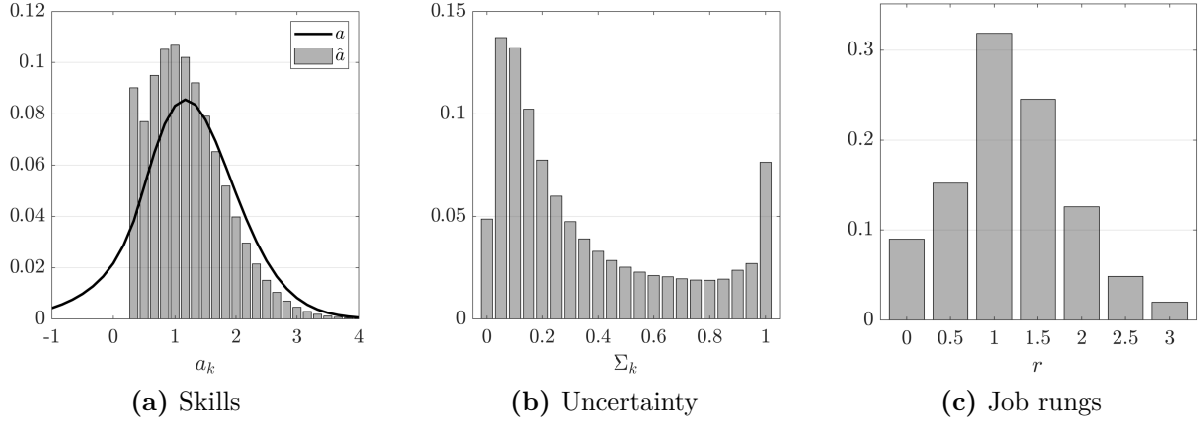
Figure 4: Ergodic distribution of individual state variables. *Notes.*—Values for a_k, $\hat{a}_k$, $\Sigma_k^{1/2}$ and r are denominated in units of $S_0^{1/2}$.
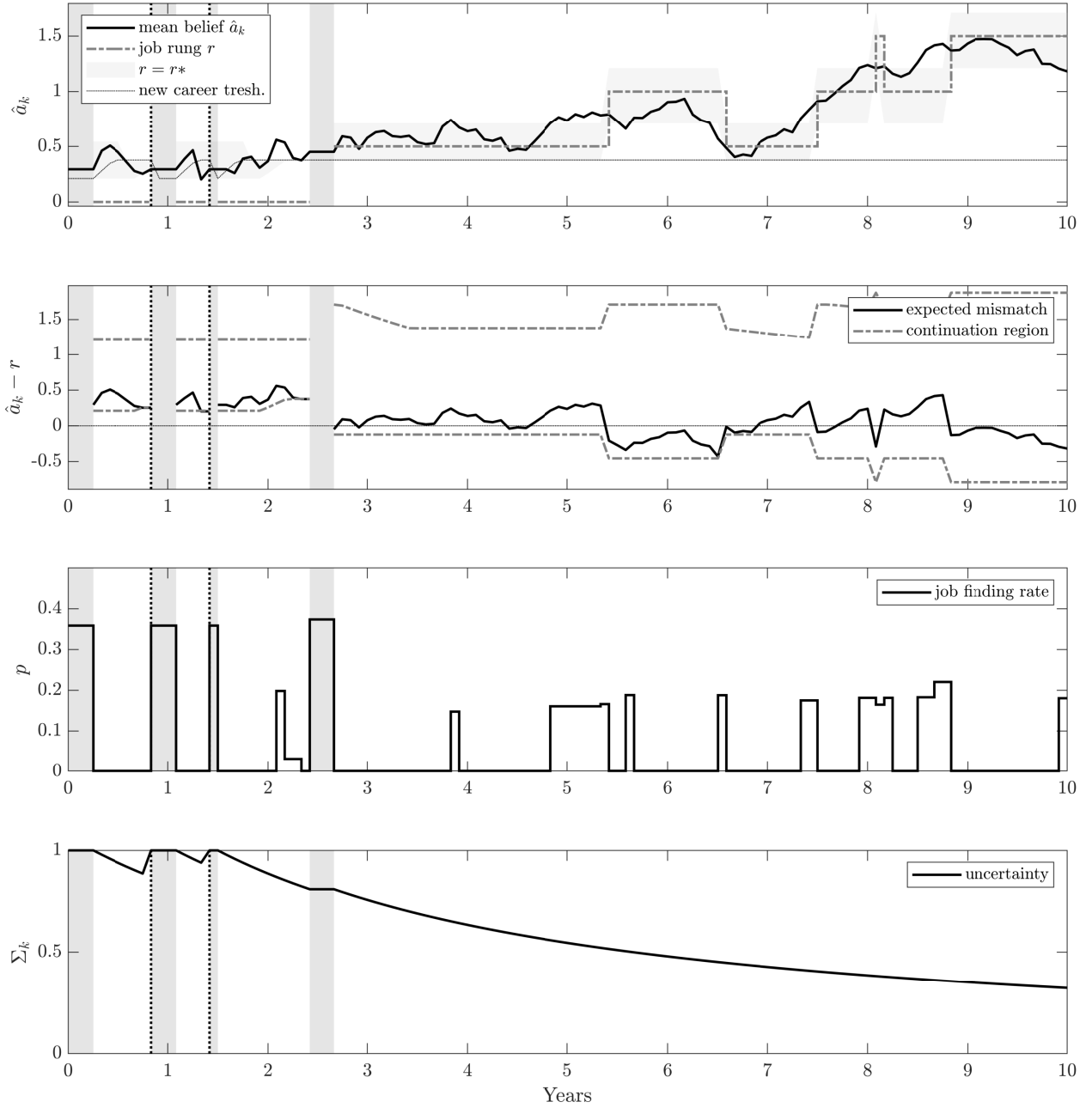
Figure 5: Sample path for a single worker. *Notes.*—The figure shows a random career path for a single worker, initialized without a job and with $(\hat{a}_k, \Sigma_k) = (a_0, S_0)$. Throughout, the aggregate state is fixed at $z = z_L$. Vertical gray bands depict unemployment spells. Vertical dotted lines depict career switches.

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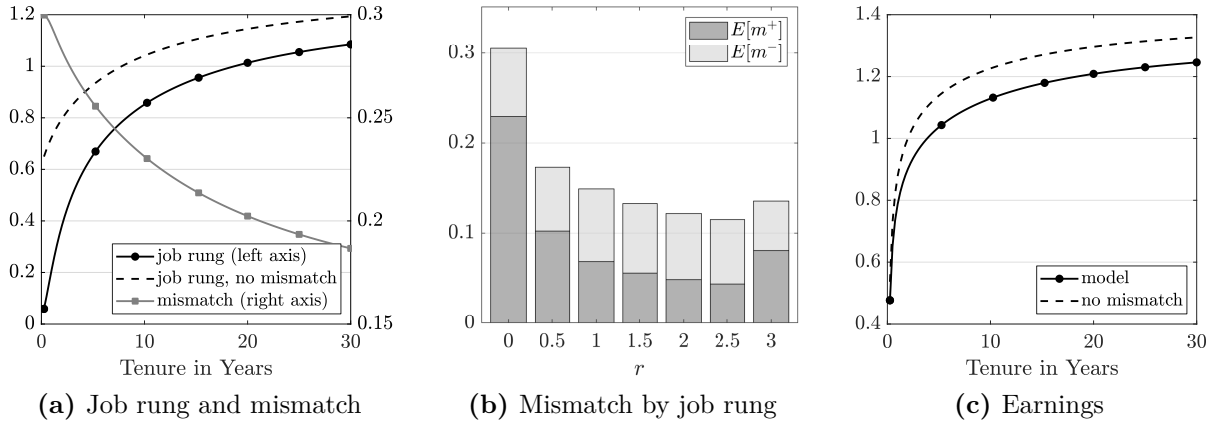
Figure 6: Relationship between tenure, job rungs, mismatch, and earnings. *Notes.*—Tenure is defined as the time since the last career change. The no mismatch counterfactuals show fictitious profiles for average job rungs and earnings where the distribution of abilities evolves as in equilibrium, but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

48
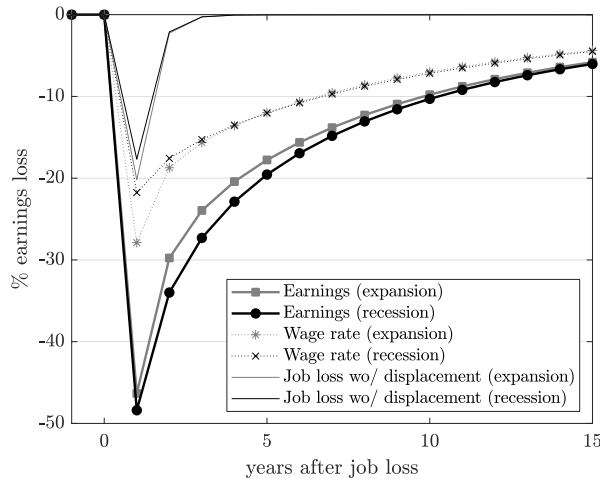
(a) Average earnings losses of displaced workers, relative to control group
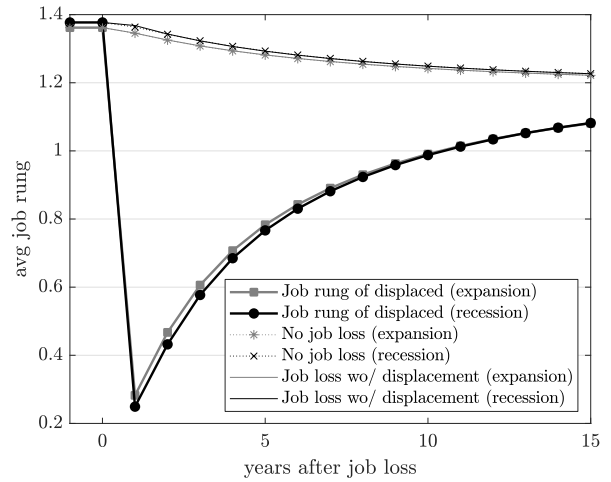
(b) Average job rung of displaced workers and control groups

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.
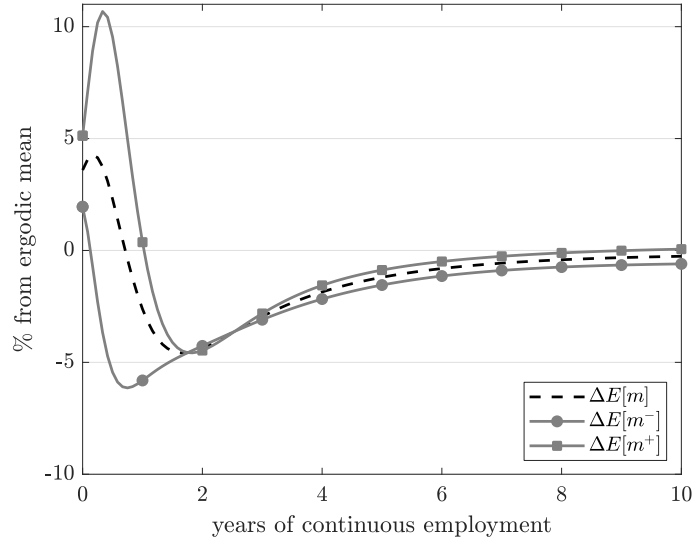
Figure 8: Mismatch cyclicality by employment tenure. *Notes.*—The figure plots the cyclicality of m, m^- and m^+ conditional on being continuously employed for τ years; i.e., $\Delta\mathbb{E}[\cdot|\tau] = \mathbb{E}[\cdot|z_L,\tau] - \mathbb{E}[\cdot|z_H,\tau]$. All cyclical differences are denominated in percentage deviations from their ergodic mean $\mathbb{E}[\cdot]$.
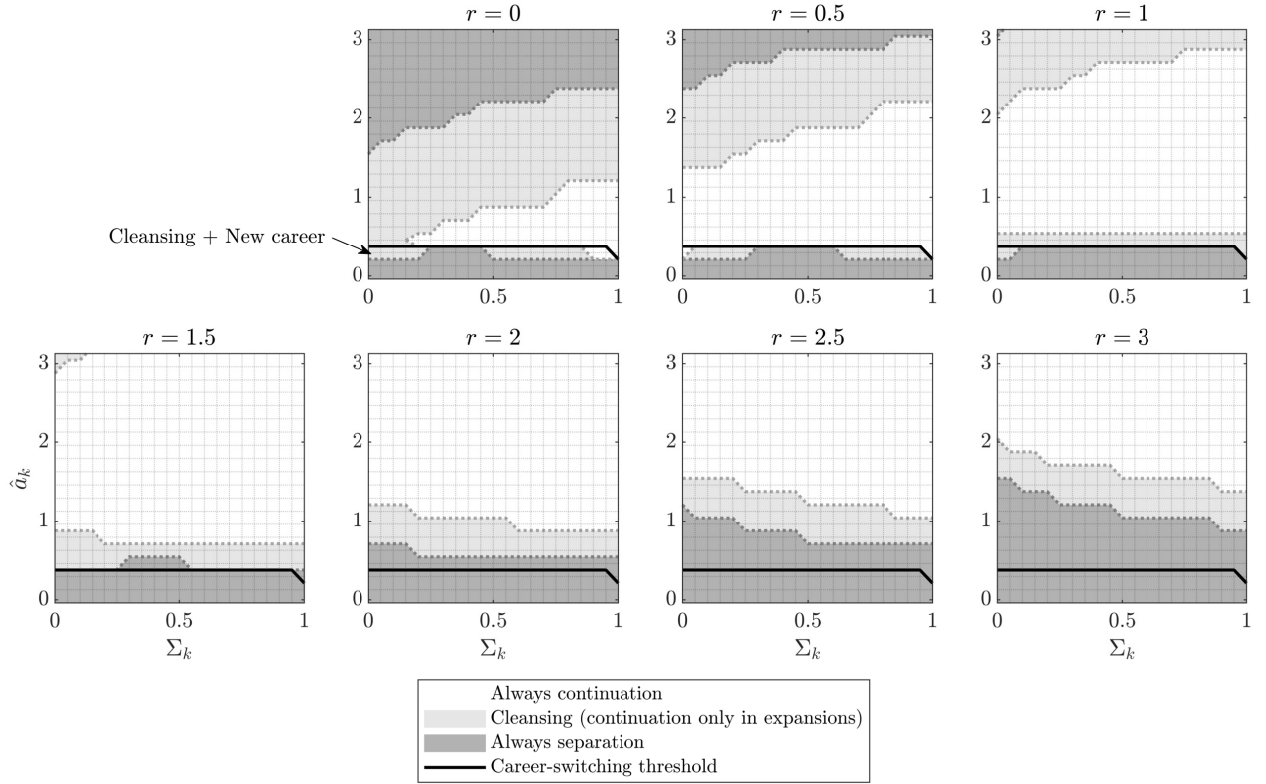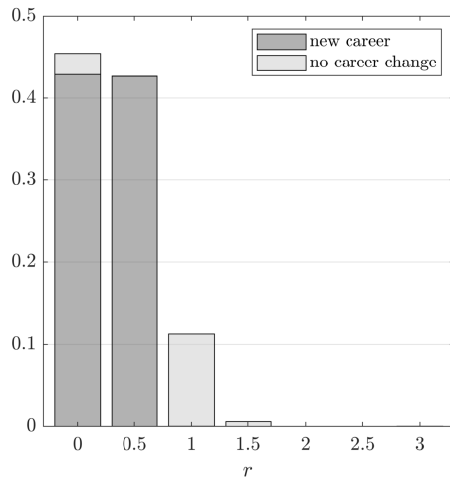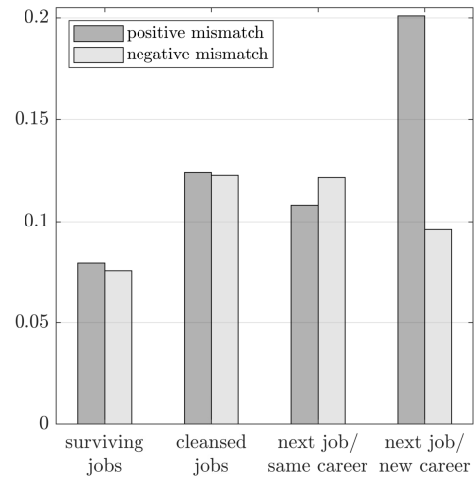
Figure 9: Separation policies: expansions vs recessions. *Notes.*—The figure shows how the continuation region contracts during recessions. The career-switching threshold is plotted for $z = z_L$. Values for $\hat{a}_k$, $\Sigma_k^{1/2}$ and r are denominated in units of $S_0^{1/2}$.

(a) Composition of cleansed workers

(b) Mismatch of cleansed vs next jobs

Figure 10: Composition and mismatch of cleansed workers. *Notes.*—Panel (a) shows the distribution of cleansed workers over job rungs and career mobility. Panel (b) shows mean mismatch for cleansed matches compared to surviving matches, as well as mismatch in the next job conditional on career mobility.
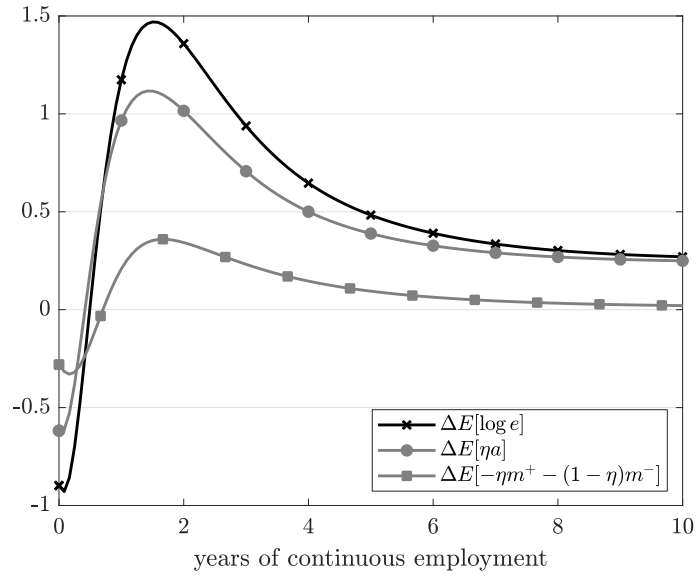
Figure 11: Cyclicality of labor efficiency by employment tenure. *Notes.*—The figure plots the cyclicality of $\log e$ and its constituents conditional on being continuously employed for τ years; i.e., $\Delta\mathbb{E}[\cdot|\tau] = \mathbb{E}[\cdot|z_L, \tau] - \mathbb{E}[\cdot|z_H, \tau]$.
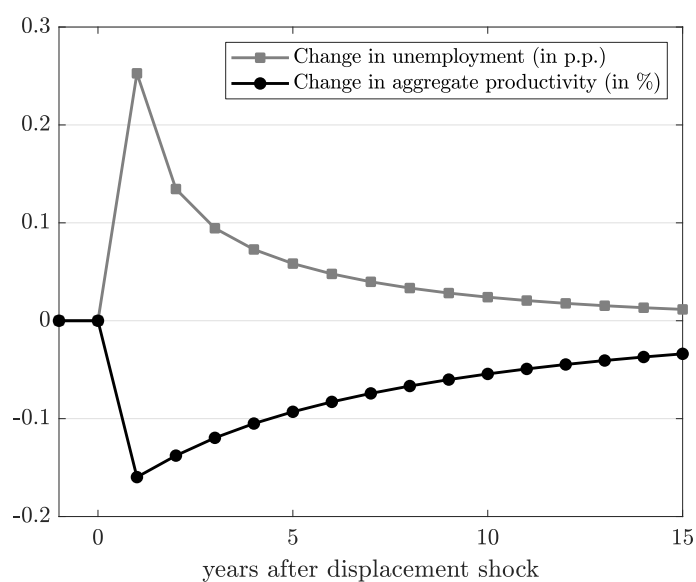
Figure 12: Aggregate impact of a sectoral displacement shock. *Notes.*—The figure shows the responses in aggregate productivity (in percent) and unemployment (in percentage points) to a shock that displaces 1 percent of the labor force from their career. The responses are averages over the ergodic process for aggregate productivity z.