Mismatch Cycles

Isaac Baley[*]   Ana Figueiredo[†]   Robert Ulbricht[‡]

January 31, 2020

Abstract

This paper studies the dynamics of skill mismatch over the business cycle. We build a tractable directed search model, in which workers differ in skills along multiple dimensions and sort into jobs with heterogeneous skill requirements along those dimensions. Skill mismatch arises due to information and labor market frictions. Estimated to the U.S., the model replicates salient business cyclic properties of mismatch. We show that job transitions in and out of bottom job rungs, combined with career mobility of workers, are important to account for the empirical behavior of mismatch. The predicted career dynamics provide a novel narrative for the scarring effect of unemployment. The model suggests significant welfare costs associated with mismatch due to learning frictions.

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

JEL Classification: E24, E32, J24, J64.

---

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

"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] 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, which reduces mismatch among ongoing work-relations. On the other hand, we find that mismatch among new hires goes up in recessions, which is caused by a simultaneous increase in over- and underqualification among workers hired for low-complexity jobs. These patterns are consistent with direct evidence on the cyclicality of mismatch, which we document among workers in the NLSY79.

The logic behind our theoretical findings is caused 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 among the respective job

---

[1]See Yamaguchi (2012), Lindenlaub (2017), and Lise and Postel-Vinay (2018) for related calibration strategies using the same combination of NLSY79 and O*NET.

ladder in response to belief revisions, transitions into new career paths (to jobs that employ previously untried skills) tend to increase mismatch as a consequence of higher uncertainty. Accordingly, the cyclicality of mismatch is closely entangled with the business cycle dynamics of career mobility. Specifically, our model predicts that career mobility—using a task-based definition—is countercyclical.[2] This is because workers that are fired from the bottom rungs of a given career path will optimally seek to find jobs utilizing a different skill set rather than re-applying to jobs that they know to be underqualified for. In that sense, the two opposing forces shaping the cyclicality of mismatch are in fact both manifestations of a *cleansing* of underqualified workers, which increases career mobility in recessions and in turn causes the increase in mismatch among new hires.

Our model provides a novel narrative for the "scarring effect of unemployment", complementing recent explanations by Jarosch (2015), Jung and Kuhn (2019), and Huckfeldt (2019). In line with empirical evidence, workers that are displaced from their career suffer large and persistent earnings losses, even after they have been re-employed. This is because rebuilding a career in a new sector implies high levels of uncertainty, which slows down a worker's climb through the job ladder. In the calibrated model, this translates into persistent wage losses that significantly outlast unemployment: While 10 years after displacement, unemployment among displaced workers is less than 5 percent above the control group, wages among the displaced continue to be depressed by about 17 percent.

We conclude our paper with two counterfactual exercises that evaluate the welfare cost induced by information frictions in our model. First, we compute the cost associated with "within-career" mismatch, due to a suboptimal allocation of workers to job-rungs. On average, we find that labor productivity under the optimal rung-allocation is 7.0 percent higher than in equilibrium. Decomposing this gains by business cycle state, the output gap amounts to 7.4 percent in expansions compared to 6.4 percent in recessions, indicating that mismatch is *dampening* the business cycle. In the second counterfactual exercise, we quantify the implicit friction on career mobility imposed by imperfect information regarding the gains of pursuing a new career. We find that if workers were able to churn careers and instantaneously learn the relevant skills, then in order to induce the same career-mobility pattern as in equilibrium we would need to subject them to an explicit switching cost equivalent to 10 months of the average output per worker.

---

[2]We verify this prediction using a model-consistent measure of career-mobility that is task-based. Kambourov and Manovskii (2008) and Huckfeldt (2019) document similar patterns using occupation- and industry-based measures of career-mobility.

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 recent theoretical works by Lise and Postel-Vinay (2018) and Lindenlaub and Postel-Vinay (2017) that also emphasize the irreducibility of worker heterogeneity into a single unidimensional index. 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.[3] Second, both papers focus on steady states, whereas our framework allows for aggregate shocks and is tractable enough to explore out of steady state dynamics, which is at the core of our exploration.

Finally, our model incorporates learning à la Jovanovic (1979, 1984). Our paper particularly relates to more recent works, in which learning is about worker skills, rather than a match-specific productivity term (e.g., Groes, Kircher and Manovskii, 2013, Papageorgiou, 2014, and Wee, 2016). In our model, this implies that the assessment of future match qualities varies with the prior work experience of workers and, in particular, leads to countercyclical fluctuations in uncertainty. Relatedly, Acharya and Wee (2020) explore a complementary mechanism that similarly gives rise to countercyclical uncertainty that reduces matching efficiency in recessions.[4]

In support of the empirical relevance of skill uncertainty, we provide direct evidence for imperfect information about worker skills exploiting worker forecasts about their own future occupation. We document that the forecast errors entailed in these forecasts can be systematically predicted by measures of worker ability that have 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.[5]

---

[3]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) consider a similar approach as motivation for an empirical exploration of multidimensional skill mismatch.

[4]See also Straub and Ulbricht (2012, 2014), Senga (2016), and Baley and Blanco (2019) for further mechanisms, outside a labor market context, that give rise to countercyclical fluctuations in uncertainty at the firm-level.

[5]Fredriksson, Hensvik and Skans (2018) also provide indirect evidence pointing to information frictions

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

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 describe the predicted business cycle dynamics of mismatch and contrast them with the data. Section 5 explores the consequences of job displacement. Section 6 studies the welfare consequences of information frictions. 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 *skill types* and sort into jobs that are characterized by the employed skill type and are further differentiated by the intensity they make use of a given skill ("skill requirement"). 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

using Swedish administrative data.

and across workers i. Abilities are not observed (directly), but their distribution is public information.

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

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

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

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

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

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

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

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

---

[6]In fact, this interpretation could be made exact with two slight changes to the environment: (i) time is discrete, (ii) the penalty on underqualification is given by $g(r - a_{i,k} - \sigma\epsilon_{i,t})$ where $\epsilon_{i,t} \sim \mathcal{N}(0, 1)$ is i.i.d. across i and t. Here g can be any smooth 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 monotonically increasing in $a_{i,k}$, 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.

follow a diffusion given by the usual Kalman-Bucy filter,

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

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

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.[7] Employed workers have the opportunity to search the labor market at rate $\kappa \in [0, 1]$ and can search for jobs within their current career path (i.e., the skill type k of the aspired job must match their current job). Vacancies are created by an infinite supply of potential firms, which can open a vacancy in any submarket $\omega \equiv (k, r, x, \hat{a}_k, \Sigma_k)$ at flow costs c.

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

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

---

[7]The exception are workers that are exogenously forced to switch careers (introduced below), which would otherwise prefer to re-apply to their old career. The reason why the no recall assumption does not pose much of a restriction otherwise is that k lays 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 that career choices are dependent on the aggregate productivity state in a way where workers prefer to explore a new career over a certain skill k for a given productivity z, but would prefer k over a new career after a change in z.

In addition to an endogenous separation choice (further detailed below), worker–firm pairs separate at an exogenous rate $\delta > 0$. Moreover, every time a worker enters the unemployment pool (endogenously or exogenously), she switches careers with an exogenous probability $\epsilon \in (0,1)$. If hit by such a career-shock, the worker is forever prevented from applying to any submarket involving the skill type k of their previous career.

Remark on notion of careers In our terminology, the label *career* refers to a set of jobs that utilize similar skills. Our definition differs from previous approaches that have defined careers based on occupation- or industry-codes. While related, such definitions would be misleading in our case as distinct occupations may share very similar skill mixes, whereas others may bundle together jobs with distinct skills.[8] For a consistent interpretation of the model, one should therefore think of careers in terms of skill-mixes when mapping the model to the data. Our calibration of the model in Section 3 aims to do so by employing a task-based definition of careers.

2.2 Equilibrium Characterization

Notation To converse notation, we suppress i subscripts from all variables going forward. All functions are indexed with a time subscript t to express their potential dependence on the aggregate state (with the exception of 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

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

Unemployed worker problem As 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 she

---

[8]For instance, using the methodology described in Section 3, we find that the skill mix of an economist is very similar to the ones of financial managers, actuaries, and loan counselors, which all constitute different occupations at the 3-digit level. By contrast, career definitions based on 2-digit occupation codes, bundle together many occupations with vastly different skill mixes.

entered unemployment. The value of being unemployed *conditional* on searching for jobs of skill type k, denoted by $U_t(\hat{a}_k, \Sigma_k, z)$, is therefore given by:

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

The flow value of being unemployed is comprised of three 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 she is searching (maximized subject to the θ–x frontier defined by (2)), and (iii) 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 lays 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

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

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

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

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

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

8

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 the uncertainty-induced risk 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 forth term captures the change in value induced by exogenous separation, in which case the worker–firm pair obtains the separation value J^{sep} (defined below). Finally, the last term captures the change in value induced by aggregate productivity shocks.

In the event of separation, the joint separation value is given by

$$J_t^{\text{sep}}(\hat{a}_k, \Sigma_k, z) = \epsilon \mathcal{U}_t(a_0, S_0, z) + (1 - \epsilon)\mathcal{U}_t(\hat{a}_k, \Sigma_k, z),$$

reflecting the possibility that the worker switches careers for exogenous reasons with probability ϵ (the continuation value for the firm is zero given free entry). Combining, the joint value of the worker–firm pair, as determined by the optimal continuation choice, is given by

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

Job ladder   We now explore which jobs workers search for as a function of the belief $(\hat{a}_k, \Sigma_k)$. Substituting the θ–x frontier defined by the entry decision of firms (2) into (3) and (5), it is immediate that the choice of task-complexity always maximizes the joint surplus,

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

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

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 6-step job ladder corresponding to $\mathcal{R} = \{0.5, 1, 1.5 \ldots, 3\} \times S_0^{1/2}$. Workers pursuing a new career, search for jobs with the lowest
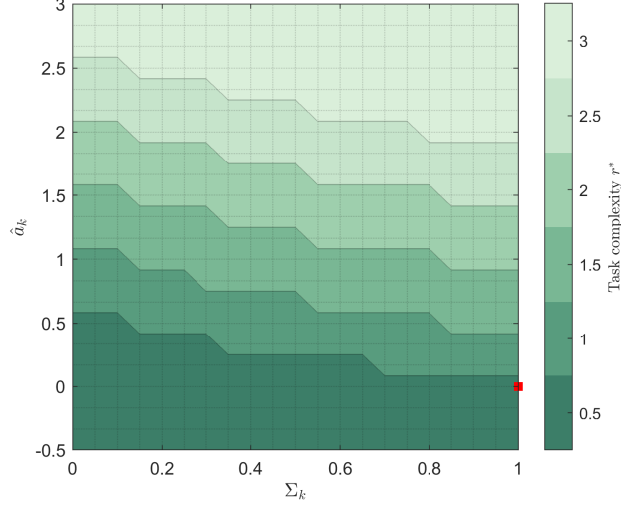
9

Figure 1: Job ladder. *Notes.*—The graph shows the task complexity r^* chosen as a function of expected ability $\hat{a}_k$ and uncertainty Σ_k. The red square marks the unconditional prior (a_0, S_0) for untried skill types. $\hat{a}_k$, $\Sigma_k^{1/2}$ 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.

complexity, $r^*(a_0, S_0, z_H) = 0.5 S_0^{1/2}$ (indicated by the red square in the plot). As workers become more optimistic regarding their skills in a given career k, they apply to more complex jobs (indicated by lighter shades of green). Reflected in the downward-sloping thresholds, there is generally an option value associated with higher uncertainty that induces workers to apply to jobs for which they expect to be underqualified. This is because $J(\hat{a}_k, \cdot, \cdot, \cdot)$ is truncated below by the option to separate, whereas higher levels of $\hat{a}_k$ generally increase J as it reduces the likelihood of being underqualified.[9] Relatedly, 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.

It remains to characterize the value of 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

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

---

[9]This prediction is consistent with our data on mismatch (introduced below), in which workers are on average 33 percent more underqualified than overqualifed. Our uncertainty-based account of this evidence complements an existing body of work that explains overshooting of workers for difficult jobs with overconfidence (Dubra, 2004) and the anticipation of future skill acquisition (Lise and Postel-Vinay, 2018).
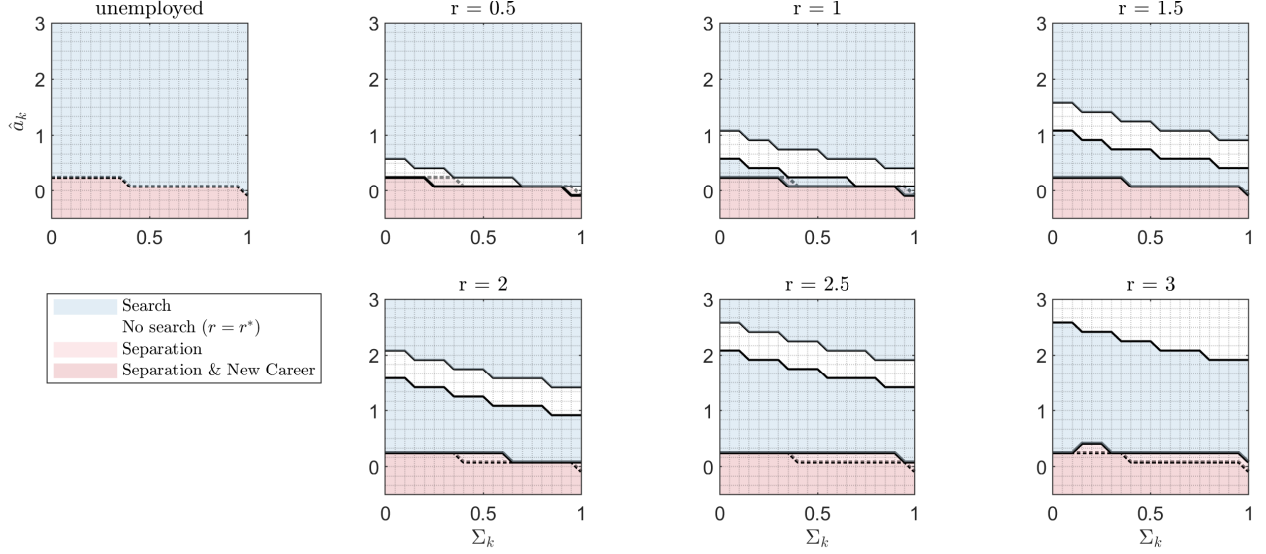
10

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 are denominated in units of $S_0^{1/2}$ (S_0 for Σ_k). The figure is plotted for $z = z_H$; the case where $z = z_L$ looks similar. See Section 3 for a detailed description of the parametrization.

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 $r = r^*$. In words: employed workers only search for jobs that are better matches (in expectations), which again is an immediate implication of bilateral efficiency.

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 job with complexity $r \in \mathcal{R}$). Unemployed workers change careers whenever $\hat{a}_k$ is small (indicated by the red area below the dotted threshold). Otherwise they search for jobs in their current career (with a job finding rate that is increasing in $\hat{a}_k$; not indicated in the plot). Employed workers are characterized by a separation threshold (black solid lines above the light red area), below which they separate (with or without career switch). 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 blue area bordered by the separation region below and the no-search region (in white) above. If $\hat{a}_k$ falls into the upper blue area, they aspire to *climb up* the job ladder instead.

11

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 $(\hat{a}, \Sigma)$.[10] Based on the search and separation policies above, we can characterize two Kolmogorov forward equations, one for Γ and one for Υ, which together with the process for z fully describe the dynamics in this economy. While the construction of these equations is standard, their precise expression is slightly protracted. We therefore confine their presentation to Appendix A.

Equilibrium and block-recursivity An equilibrium is a joint value function satisfying equation (6), an unemployed value function satisfying equation (4), lifetime utilities x satisfying the free entry condition (2), and a distribution of worker–firm pairs and unemployed workers evolving according to equations (13) and (14) (stated in Appendix A).

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

3 Calibration

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

---

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

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

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

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

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

The key observation here 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 here, it suffices to classify occupations into careers so that job requirements $\{\boldsymbol{q}_{k,r}\}$ are (approximately) orthogonal across k.[11] 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 angle between two skill vectors $\boldsymbol{q}_1$ and $\boldsymbol{q}_2$,

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

Then any job transition from a job with $\boldsymbol{q}_1$ to a job with $\boldsymbol{q}_2$ is treated as a career-switch if and only if $\varphi(\boldsymbol{q}_1, \boldsymbol{q}_2) \geq \bar{\varphi}$ for some $\bar{\varphi}$ (below, $\bar{\varphi}$ is chosen so that the average correlation in requirements for career-switches is zero).[12] To account for variations in economic relevance

---

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

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

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.

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.[13]

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

Residual correlation in skills across careers A brief comment is in order before we proceed to the empirical implementation of our approach. We have argued that an orthogonal classification of careers allows for an exact mapping of our model to the data. One may nevertheless ask: how does the model map to the data when there is some residual correlation in skills across careers?

In theory, residual correlation would allow workers to learn about their ability to execute previously untried careers (although their ability to do so is likely limited in practice[14]).

---

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

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

14

To the extend that workers can predict their performance in untried careers, we speculate that workers would guide their search towards occupations for which they are best-qualified based on prior work-experience. Using the notation of our model, we could capture this by re-interpreting a_0 as the conditional mean of the best-perceived career based on prior information, and S_0 as the residual uncertainty surrounding the required skill bundle. 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 of a_0 (and decrease of S_0) throughout the lifetime of a worker (as workers become better at predicting at which skills they excel). To the extend that we do not focus on life-cycle dynamics, but on business cycle dynamics, we do not consider this a big abstraction.

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, 2018), which we interpret as mathematics, verbal, social, and technical skills.[15] To make them comparable, we normalize each skill dimension in terms of percentile ranks.[16] See Appendix C.1 for details on the construction of our skill measure.

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.[17] As detailed above, we associate a job transition from $\boldsymbol{q}_{i,t}$ to $\boldsymbol{q}_{i,t+1}$ as a career-switch if the angle between the two skill vectors, $\varphi(\boldsymbol{q}_{i,t}, \boldsymbol{q}_{i,t+1})$, is larger than $\bar{\varphi}$. The 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.[18] Using this strategy, we set $\bar{\varphi} = 12.58°$ which implies that 44.1 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 and

---

[15]Guvenen et al. (2020) and Lise and Postel-Vinay (2018) 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).

[16]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.

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

[18]The small correlation in skills for career-switchers contrasts strongly with an average correlation of .94 among job-switchers that are classified as within-career transitions.

Visschers (2014), Carrillo-Tudela et al. (2016), and Huckfeldt (2019).

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. See Appendix C.3 for details.

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

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

Here v_j are "market weights", obtained from the regression coefficients on each of the four mismatch dimensions in a Mincer regression (normalized so $\sum_{j=1}^{4} v_j = 1$).[19] Intuitively, the weights ensure that our mismatch measure is not driven by skills that are economically irrelevant. Similarly, we define positive mismatch, measuring overqualification, and negative mismatch, measuring underqualification, as

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

3.2   Parametrization of the Model

Assigned parameters We parametrize the model at a monthly frequency. The discount rate ρ is set to $\log(1.05)/12$ corresponding to an annual discount rate of 5%. We choose isoelastic contact rate functions, $p(\theta) = \theta^\gamma$ and $q(\theta) = \theta^{\gamma-1}$, where in line with the evidence surveyed in Petrongolo and Pissarides (2001) the elasticity of matches to vacancies, γ, is set to 0.4.

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. (2017).[20] We choose to set the relative search intensity κ based on direct evidence as opposed to targeting the

---

[19]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.

[20]Holzer (1987) and Faberman et al. (2017) document a relative time spend on search activities of 0.48 and 0.51, respectively.

job-to-job rate, because job-to-job transitions are clearly caused by many factors not present in the model, including relocation shocks, rent seeking purposes, and random fluctuations in match-quality. If we would force the model to match the empirical job-to-job rate, we would effectively require learning about skills to account for these other forces, overstating the importance of learning for job-to-job mobility.[21]

We specify the set of potential task-complexities, $\mathcal{R}$, using a six-point grid given by $\{0.5, 1, \ldots, 3\} \cdot S_0^{1/2}$, 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.[22] Worker abilities are normalized around $a_0 = 0$. We approximate beliefs about worker skills using a 64-point grid for $\hat{a}_k$ on $[-3, 7] \cdot S_0^{1/2}$ and a 21-point grid for Σ_k on $[0, 1] \cdot S_0$. Finally, we normalize log productivity in recessions to 0, and choose transition rates for z in order to match the monthly switching intensities between recessions and expansions in the U.S., where recessions are periods with an unemployment rate above its unconditional average of about 6.5%.

Target moments We calibrate the remaining parameters jointly 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 $\mathbb{E}[b/y]$ equal to .71 as found by Hall and Milgrom (2008). Finally, we identify z_H (relative to z_L) from an average recession–expansion difference in unemployment amounting to 2.8 p.p. in the US.

To identify the speed of learning, parametrized by σ, we target an average slope of the empirical separation hazard over the first 18 months of employment, $\log(haz_1/haz_{18})$, of .338 as found in the NLSY79 sample. 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 exogenous career-shock ϵ to ensure consistency of the model with an average propensity to switch careers of 44.1 percent, as documented above in the NLSY79.

---

[21]In our calibration, the monthly job-to-job worker flows are .011 compared to an empirical rate of .03, suggesting that learning can account for roughly one third of the observed job-to-job mobility.

[22]Adding an extra grid point at 0 has no effect as no search is directed to such submarkets in our calibration. Similarly, adding an extra grid point at $3.5 \cdot S_0^{1/2}$ does not change the results as it attracts only a relative mass of 0.004 workers at the ergodic distribution.

Table 1: Targeted moments

Fitted Moments	Data	Model	Origin
$\mathbb{E}_L[U] - \mathbb{E}_H[U]$	.028	.027	BLS
$\mathbb{E}[\text{UE rate}]$	.425	.434	Shimer (2012)
$\mathbb{E}[\text{EU rate}]$	.035	.035	Shimer (2012)
$\mathbb{E}[b/y]$	.710	.712	Hall and Milgrom (2008)
$\mathbb{E}[\log(haz_1/haz_{18})]$	.338	.331	NLSY79
$\mathbb{E}[\chi = 1]$	.441	.441	NLSY79, O*NET
$\mathbb{E}_L[\chi = 1] - \mathbb{E}_H[\chi = 1]$	.069	.064	NLSY79, O*NET
$\mathbb{E}[|a_k - r|]$	.280	.283	NLSY79, O*NET

Notes.—The notation $\mathbb{E}[\cdot]$ denotes unconditional expectations, computed at the ergodic distribution of the model. $\mathbb{E}_L[\cdot]$ and $\mathbb{E}_H[\cdot]$ denote expectations conditional on the aggregate state being in a recession or expansion, respectively. U denotes the aggregate unemployment rate, EU and UE are monthly transition rates, y is output per worker–firm pair, haz_x is the separation hazard after x months of employment, and χ is an indicator evaluating to unity if workers switch careers when entering the unemployment pool.

Similarly, 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 variance of skills, S_0, we exploit a close relation between S_0 and the scale of mismatch. To see this link, suppose for a moment that output would be linear (using the same specification as in (1), but with y on the left-hand side instead of $\log y$). In this case, we can express (z_L, z_H, b, c, σ) in units of $S_0^{1/2}$ so that all policy functions become scale-neutral: any change in S_0 and an accordant rescaling of (z_L, z_H, b, c, σ) will simply re-scale r and a_k (and hence mismatch) but will not change any policy functions or any other target moment. While this exact independence between S_0 and the rest of our calibration strategy breaks down in the log-linear model, the intuition that S_0 is most closely related to the scale of mismatch continues to hold. We therefore target the average mismatch in our NLSY79 sample, given by .280, to identify S_0.

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. The values of z_H, b and c have immediate interpretations given their calibration targets. The return on skills, η, is estimated to be relatively high in the sense that it induces workers to err on the the side of being underqualified when facing uncertainty about their skills. This is illustrated in panel (a) of Figure 4 where we plot the ergodic distribution over $\hat{a}_k - r$ (see also Figure 1).

The estimated standard deviation of skills is 0.476, which implies a significant amount of

Table 2: Summary of parameters

Parameter	Description	Value
Assigned		
ρ	Monthly discount rate	$\log(1.05)/12$
γ	Matching elasticity	0.4
κ	Relative search intensity of employed	0.5
a_0	Unconditional skill mean	0
z_L	Aggregate log-productivity in recessions	0
$\lambda_{z_L}, \lambda_{z_H}$	Poisson rates of productivity shock	0.0128, 0.0172
Estimated		
z_H	Aggregate log-productivity in expansions	0.138
b	Home production utility	0.959
c	Flow cost of vacancies	0.529
η	Return on task complexity	0.720
$S_0^{1/2}$	Standard deviation of skills	0.476
σ	Standard deviation of signal noise	5.776
δ	Exogenous separation rate	0.021
ϵ	Exogenous career-switch propensity	0.063

equilibrium dispersion in (log) value added per worker, seen in panel (c) of Figure 4.

The standard deviation of the signal is $\sigma = 5.776$, implying an median uncertainty amounting to $0.35 \cdot S_0$, or roughly one third of the prior uncertainty. The average uncertainty $\mathbb{E}[\Sigma_k]$ amounts to $0.42 \cdot S_0$, so the distribution over Σ_k is right-skewed (see panel e). Not surprisingly, however, despite the overall right-skew, the distribution of Σ_k has also a concentration of mass at $\Sigma_k = S_0$, reflecting the reset in learning after workers switch careers.

Interestingly, the distribution over beliefs is censored slightly below 0 (see panel d), reflecting the option to switch careers whenever workers become pessimistic about their skills (see also Figure 2). By comparison, the true distribution of skills in workers pursued careers is much more dispersed as can be seen from the same plot. The dispersion is particularly high around the truncation point at the left of the distribution over $\hat{a}$, because $\hat{a} \approx 0$ is positively correlated with high levels of uncertainty due to workers just having started their careers. By contrast, the dispersion to the right of the distribution of $\hat{a}_k$ is much lower, since high levels of $\hat{a}_k$ are correlated with low levels of uncertainty as extreme belief revisions are more likely the more signals one obtains.[23]

Finally, the estimated value for δ is 0.021, so that exogenous separations account for 59%

---

[23]More formally, $\mathrm{Var}[\hat{a}_k - a_0]$ is directly proportional to the measure of signals received, which is negatively related to uncertainty. The negative correlation is further reinforced, because workers optimally stay with careers that they think they are good at.
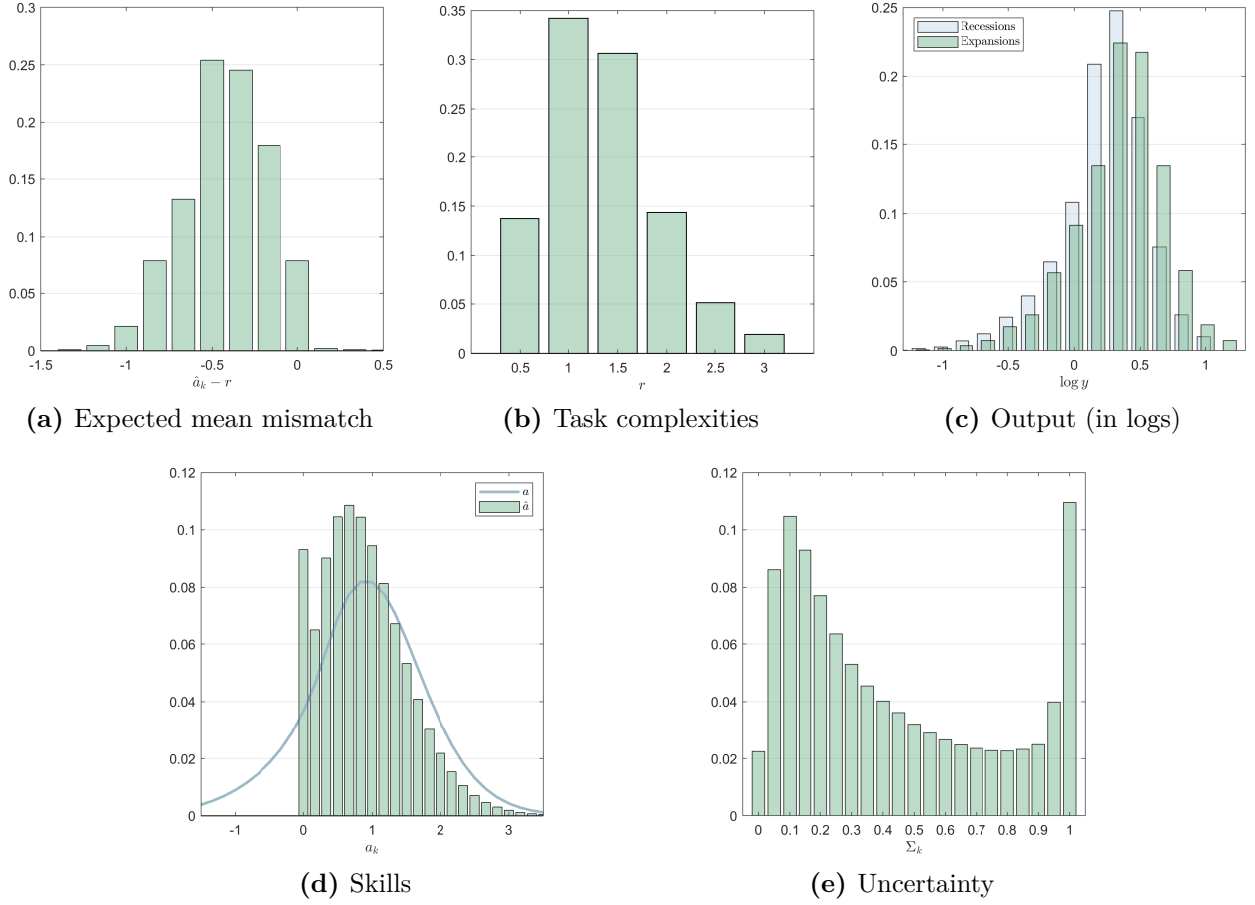
(a) Expected mean mismatch **(b)** Task complexities **(c)** Output (in logs)

(d) Skills **(e)** Uncertainty

Figure 4: Ergodic distributions. *Notes.*—Apart from output, all variables are denominated in units of $S_0^{1/2}$.

of all separations. The estimated value for ϵ is 0.063, implying that exogenous career shocks account for 14% of all career mobility.

3.3 Direct Evidence for Learning About Skills

We conclude this section 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 predictive of future occupations, $\mathbb{E}[\boldsymbol{q}_{i,t+60}|\boldsymbol{a}_i] = \boldsymbol{a}_i$, one would then predict the forecast error to be given by

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

The main premise of our test is that under the null hypothesis that workers know their skills, the forecast error should be orthogonal to the predicted error $\text{pe}_{i,t,j}$. Note that the null is a direct implication of workers knowing their skills aside from being aware of their own forecast, 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.[24]

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

$$\sum_{j=1}^{4} \text{fe}_{i,t,j} = \beta_0 + \beta_1 \sum_{j=1}^{4} \text{pe}_{i,t,j} + \epsilon_{i,t}. \tag{11}$$

Our estimate for β_1 is given by .550 with a standard error of .006. Table 3 further reports variations of our test where we separately estimate (11) for each skill dimension,

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

In all cases, we reject the null hypothesis that $\beta_1 = \cdots = \beta_4 = 0$, which we interpret as evidence in support of workers having to learn their own skills. In particular, we find that both $\text{pe}_{i,t,j}$ and $\text{pe}_{i,t,\text{math}}$ are significantly correlated with $\text{fe}_{i,t,j}$ for all skill dimensions j. The findings are consistent with anecdotal evidence given in Guvenen et al. (2020), which suggests that workers are unaware of the 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.

4   Mismatch Cycles

In this section, we study the macro-dynamics of mismatch in the model and in the data. We begin by presenting reduced-form evidence about the cyclicality of mismatch. We then show that the model quantitatively captures the empirical impact of a recession on mismatch. Finally, we use the model to explore the mechanism at work and demonstrate that job

---

[24]This is because any variable that is realized at date t should be orthogonal to workers' expectation error under full information.

Table 3: Direct evidence for learning

Dependent variable:	$\sum_j \mathrm{fe}_j$ (1)	$\mathrm{fe}_{\mathrm{math}}$ (2)	$\mathrm{fe}_{\mathrm{verb}}$ (3)	$\mathrm{fe}_{\mathrm{tech}}$ (4)	$\mathrm{fe}_{\mathrm{soc}}$ (5)
$\sum_k \mathrm{pe}_k$	.550***				
	(.006)				
$\mathrm{pe}_{\mathrm{math}}$		.461***	.317***	.239***	.154***
		(.034)	(.035)	(.035)	(.037)
$\mathrm{pe}_{\mathrm{verb}}$		.046	.192***	−.058*	.132***
		(.033)	(.033)	(.034)	(.033)
$\mathrm{pe}_{\mathrm{tech}}$		.036	−.012	.309***	−.173***
		(.032)	(.032)	(.032)	(.034)
$\mathrm{pe}_{\mathrm{soc}}$		.032*	.059***	−.039**	.290***
		(.019)	(.019)	(.019)	(.017)
R-squared	.311	.317	.288	.258	.183
Obs.	19203	19203	19203	19203	19203

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.

and career mobility at the bottom rungs of the job ladder are key for generating mismatch dynamics in the model.

4.1 Mismatch Cycles in the Data

We first explore the relation between mismatch and the U.S. business cycle in the data. We do so by estimating the following empirical specification:

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

Here $m_{i,t}$ is the mismatch of worker i at time t as defined in Section 3.1; $JS_{i,t}$ and $UE_{i,t}$ are dummies indicating job stayers and new hires from unemployment[25]; $\mathrm{recession}_t$ is an indicator that evaluates to unity if the aggregate unemployment rate is above its long-run 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.[26] Note that job-to-job

---

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

[26]Notice that our specification of fixed effects implies some redundancy among our control variables; but as usual in linear regression frameworks this does not bias our main coefficients of interest.

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 ($\beta_1 + \beta_2$)	−.315***	.010	−.325***
	(.130)	(.095)	(.091)
New hires ($\beta_1 + \beta_3$)	.589*	.414*	.175
	(.315)	(.227)	(.207)
Total cyclicality	−.270**	.035	−.305***
	(.130)	(.095)	(.090)

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).

transitions are the omitted category here and are absorbed by β_1.[27]

Table 4 reports the estimated business cycle effects. Looking at job stayers, mismatch declines in recessions by an average of .315 percentage points, which corresponds to 1.13% 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 (.589 percentage points, or 1.93% of the average mismatch among new hires). Specifically, we find that unemployed workers finding a job in a recession are on average more overqualified *and* more underqualified compared to workers findings jobs 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 labor force. Aggregate mismatch is, therefore, primarily determined by the cleansing effect of recessions, comprising roughly acyclical dynamics of overqualification and procyclical dynamics of underqualification.

4.2 Model vs Data

We now compare the estimated effects of a recession to their model analogue. Specifically, we compute the impact of a recession on a variable as the difference in conditional means, $\mathbb{E}_L[\cdot] - \mathbb{E}_H[\cdot]$, computed at the ergodic distribution of the estimated model. The results are

---

[27]As our model does *not* imply any systematic cyclicality in mismatch among job-to-job movers, we do not focus on job-to-job transitions here. In our regression, we do not find a significant cyclicality in mismatch among job-to-job movers (see Table 8 in the Appendix for details).

Table 5: Cyclicality of mismatch: model versus data

	$m_{i,t}^+$		$m_{i,t}^-$	
	Data	Model	Data	Model
Job stayers	.010	−.049	−.325	−.361
	(.095)		(.091)	
New hires	.414	.430	.175	.525
	(.227)		(.207)	
Total cyclicality	.035	−.019	−.305	−.307
	(.095)		(.090)	

Notes.—Data columns replicate the coefficients and standard errors reported in Table 4. The model columns report the difference in conditional means, $\mathbb{E}_L[\cdot] - \mathbb{E}_H[\cdot]$, computed at the ergodic distribution of the estimated model. All coefficients are multiplied by 100 (so that mismatch ranges from 0 to 100).

reported in Table 5 alongside their empirical counterparts.

Overall, the model does a fairly good job at replicating the estimated coefficients, the exception being the cyclicality of m^- among new hires where the model overpredicts the increase in underqualification. Otherwise, the model closely matches the procyclicality in m^- and the acyclicality in m^+ among job stayers. Similarly, the model closely matches the total cyclicality in both m^- and m^+, and correctly predicts that the total cyclicality is by-and-large driven by the cleansing effect. Finally, the model also matches the increase in overqualification among new hires. With the aforementioned exception, all model moments are within a one standard deviation of their empirical counterparts.

4.3 Understanding the Key Forces in the Model

We now use the model to simulate the response to a negative shock in aggregate productivity, and use it to explore the forces driving mismatch in the model. To implement the simulation, we consider an economy that has been in the high productivity state z_H indefinitely and then shifts to the low productivity state z_L for the duration of our simulation (36 months). Panel (a) of Figure 5 shows the response of average mismatch across workers ($\times 100$). The responses are normalized relative to the steady state with $z = z_H$, at which we initialize the simulation.

Cleansing In line with the ergodic moments, the overall response of mismatch is procyclical (see dashed lines). Upon impact, negative mismatch declines by .22 percentage points, and subsequently remains permanently reduced over the course of the recession. Intuitively, the decline in negative mismatch elicits from a *cleansing* of underqualified workers whose employment becomes unprofitable after the adverse shock to match surplus. In Figure 6
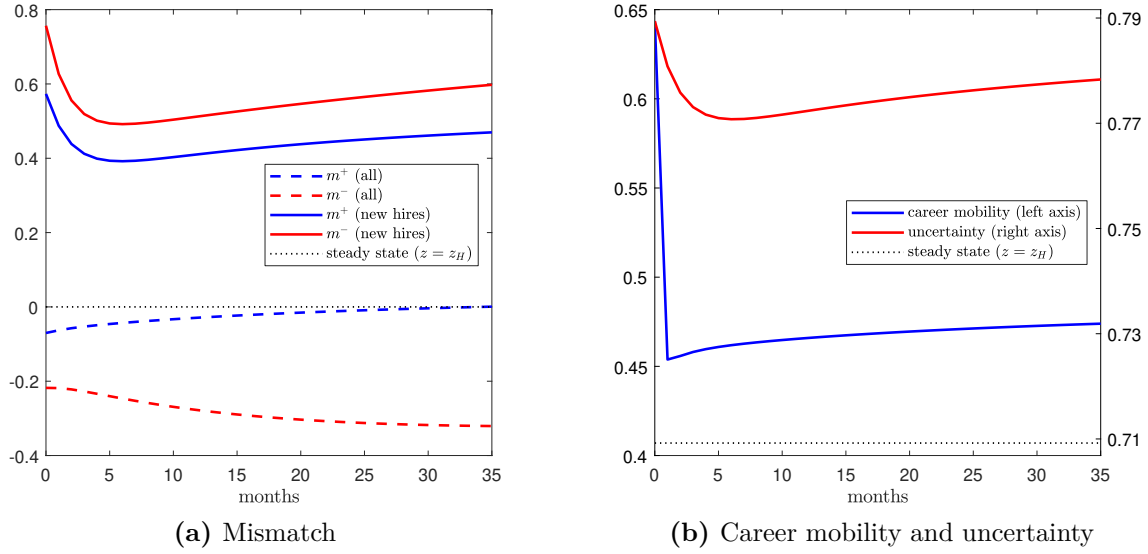
(a) Mismatch

(b) Career mobility and uncertainty

Figure 5: Dynamic response to a negative productivity shock. *Notes.*—Panel (a) shows responses in overall mismatch and mismatch among new hires ($\times 100$), normalized relative to the steady state where $z = z_H$. Panel (b) shows the response in the propensity to switch careers (left axis) and average uncertainty among unemployed workers (right axis).

we decompose the resulting layoffs by job rung. Out of all layoffs in the first month of the recession, 79.1% are from the bottom rung of the job ladder, and 95.6% are from the bottom two rungs. Compared to the ergodic distribution (Figure 4b), cleansing is hence disproportionately present among the bottom two rungs. This is because expected mismatch is disproportionately high in those rungs as workers only move to higher rungs once there is sufficient evidence for their qualification, at which point extreme revisions in beliefs are unlikely.[28]

For a more explicit exposition, Table 6 further decomposes layoffs by beliefs and mismatch. Workers that loose their job after the negative productivity shock have an average expected skill of $\hat{a}_k = .11 \cdot S_0^{1/2}$ compared to an average skill requirement of $r = .63 \cdot S_0^{1/2}$. At the same time, they are characterized by an high average uncertainty of $\Sigma_k = .81$. In terms of mismatch, this corresponds to a negative mismatch that is 8.60 percentage points larger than the average mismatch at the $z = z_H$ steady state at which we initialized the simulation. Interestingly, these layoffs are also characterized by a high positive mismatch (3.13 percentage points above the $z = z_H$ steady state), which explains the initial decline in positive mismatch seen in the left panel of Figure 5. The reason for the decline in positive mismatch is precisely the

---

[28]The logic goes back again to the negative correlation between uncertainty and expected skills/job rungs discussed in Footnote 23. A low probability of large belief revisions in turn implies that $r^*(\hat{a}_{k,t+\Delta}, \Sigma_{k,t+\Delta}, z)$ is likely to be the same as $r^*(\hat{a}_{k,t}, \Sigma_{k,t}, z)$, giving little cause for cleansing.
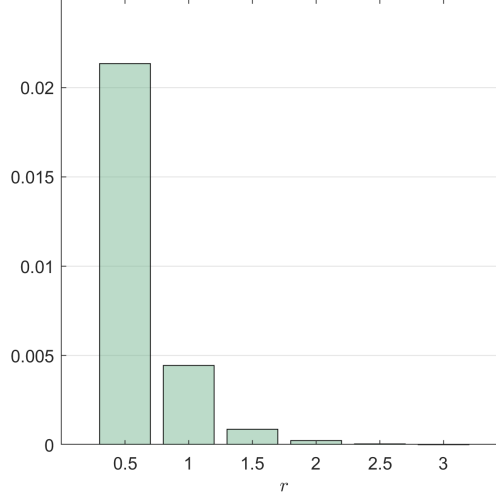
Figure 6: Layoffs by job rung. *Notes.*—The figure shows the job-rung distribution of workers transitioning from employment into unemployment within the first month of the recession.

Table 6: Beliefs and mismatch statistics for layoffs

	Mean	25 percent	50 percent	75 percent
$\hat{a}_{i,t}$	.11	.00	.00	.17
$\Sigma_{i,t}$	.81	.95	.95	.95
$\Delta m_{i,t}^{+}$	3.13	.74	4.70	4.70
$\Delta m_{i,t}^{-}$	8.60	8.86	8.90	8.91

Notes.—The table reports the mean and 25, 50, 75 percentiles for layoffs within the first month of the recession. $\hat{a}$ and Σ are denoted in units of $\sqrt{S_0}$ and S_0. Δm^{+} and Δm^{-} are multiplied by 100 and normalized relative to their respective steady state values where $z = z_H$.

high uncertainty: Even though cleansed workers are *expected* to be underqualified, their skill estimates are surrounded by enough uncertainty so that their *true* skills exceed r often enough to result in above average overqualification. Unlike the persistent decline in underqualification, the positive mismatch, however, rises again in the sequel as workers eventually end up in new careers with initially positive signals regarding their skills, leading to the overall acyclical response in m^{+} seen in the ergodic moments.

Sullying   The procyclical response of overall mismatch stands in contrast to the counter-cyclical response of mismatch among new hires from unemployment: Unemployed workers that find a job in a recession are on average more overqualified *and* more underqualified compared to workers finding jobs in expansions (see solid lines in Figure 5a). The driving force that generates this "sullying" among new hires is precisely the cleansing of underqualified workers. This is because workers that are cleansed out of the bottom rung of their job ladder optimally direct their job search to a new career rather than re-applying to the same job

26

ladder for which they are underqualified (see blue line in panel b). Cleansing therefore directs unemployed workers to search for jobs that utilize unfamiliar skills, resulting in an increased uncertainty among unemployed (red line in panel b), which in turn increases mismatch among new hires.

4.4 Labor Productivity Puzzle

Before exploring further applications of our model, we point out an interesting implication of the procyclical mismatch for labor productivity. To draw inference about labor productivity, we need to take a stand on the nature of the aggregate "productivity" shock z. One possibility is the literal interpretation as shock to productive efficiency, in which case overall labor productivity is procyclical in our simulation. However, owing to the partial equilibrium nature of the model, we can alternatively interpret z as demand shock to the real price of labor output. In this case, aggregate labor productivity, $A_t^{\text{eff}} \equiv e^{-z_t} y_t/(1 - U_t)$, is entirely determined by skill mismatch,

$$A_t^{\text{eff}} = \frac{1}{1 - U_t} \int_{\text{employed}} \exp\Big(\eta r_{i,t} - \max\{r_{i,t} - a_{i,k_t}, 0\}\Big) \, \mathrm{d}i,$$

which is *countercyclical*: Evaluated at the ergodic distribution, A_t^{eff} is .63% higher in recessions than in expansions.

The finding 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, 2019). Through the lens of the model, we would precisely expect such development when productivity shocks are diminishing and business cycles become increasingly demand-driven, consistent with the household balance sheet narrative of the Great Recession.

5 Scarring Effect of Unemployment

Previous literature has documented a large and persistent impact of a job loss on future wages and earnings (e.g., Davis and von Wachter, 2011; Jarosch, 2015), specifically when the job loss is accompanied by occupational displacement (Huckfeldt, 2019). In this section, we offer a natural narrative for the "scarring effect of unemployment" based on the time it takes displaced workers to rebuild their careers. In line with the evidence in Huckfeldt (2019), earnings losses are in large realized through wage losses and are concentrated among workers that are separated from their job *and* are displaced from their career.
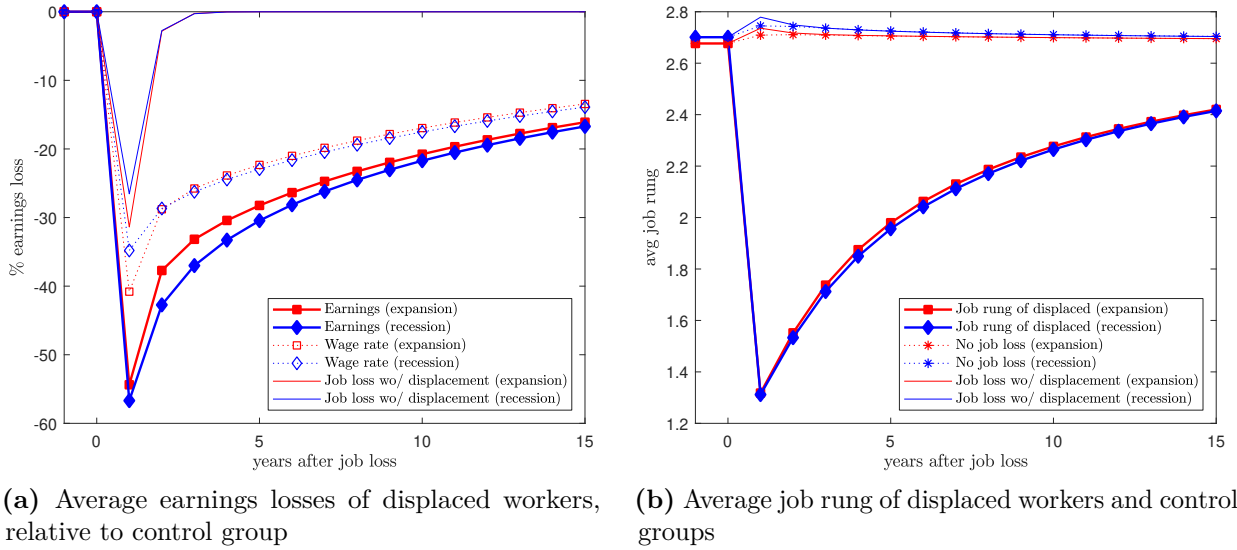
(a) Average earnings losses of displaced workers, relative to control group

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

Figure 7: Effect of career displacement vs. job loss. *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 (on a scale of 1 to 6) for workers displaced from their career, workers separated from their job without career displacement, and workers without job loss.

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 life time utility x. We do so by following Schaal (2017) and choosing the unique wage scheme under which employed workers find it sequentially optimal to pursue the contracted continuation and search policies, even in the absence of any contractual commitment. The unique wage arrangement with this properties effectively pays workers their expected marginal product, adjusted for the cost of recruitment which is loaded onto workers at the instant of hiring (see Appendix D for details).

Career displacement vs. job loss Figure 7(a) shows the earnings and wage path of a worker that is displaced from her current career at $t = 0$, conditional on the business cycle state at $t = 0$. Relative to the counterfactual of no job loss, earnings are reduced by roughly 55 percent one year after the displacement, and continue to be depressed by about 20 percent ten years later. While initially a significant share of the earnings loss is explained by a slow rate of reemployment (after one year, 37 percent of the workers displaced during recessions and 26 percent of the ones displaced during expansions are unemployed), most of the long-run "scar" is due to a persistent decline in wages.

The reason behind this long-run "scar" on wages is that displaced workers—who previously

occupied jobs at all rungs of the job ladder— must rebuild their careers in new sectors, which in the model means exploring new abilities for which uncertainty is very high. Slowed down by the time it takes to learn about these abilities, it therefore takes time for workers to identify alternative career paths and learn where to locate on the new job ladder. Panel (b) of Figure 7 illustrates this by plotting the average job rungs of displaced workers in the sequel to their job loss. While workers that are separated from their job without career displacement are able to immediately re-enter the labor market at their previous job rungs (with little consequences for earnings[29]), workers that are displaced from their career enter the labor market at the bottom job rung and take years to advance (on average) to their previous rungs. The prolonged impact of this long climb through the job rungs on earnings is able to fully 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, 2019) to 15–20 percent (Jarosch, 2015).

Aggregate response to a "China shock"   Recent literature has also provided evidence on the aggregate consequences of job displacement, using evidence from a trade shock that led to mass layoffs in manufacturing due to increased competition from Chinese imports. Similar to the evidence regarding the impact of job displacement at the micro level, the literature has documented large and persistent effects of this aggregated displacement on wages and productivity (e.g., Autor, Dorn and Hanson, 2013, 2016), whereas the impact on unemployment has been transient (Bloom et al., 2019). To explore the aggregate consequences of job displacement, we implement a "China shock" as an aggregate career displacement shock for a random 1 percent of the labor force. Figure 8 shows that this shock generates a persistent decline in aggregate productivity, which outlasts the direct impact on unemployment. Interestingly, 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. This is in line with the evidence in Autor, Dorn and Hanson (2013), which documents that the wage reductions following an aggregate displacement shock to manufacturing were not realized in manufacturing, but indeed are concentrated outside that sector.

---

[29]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.
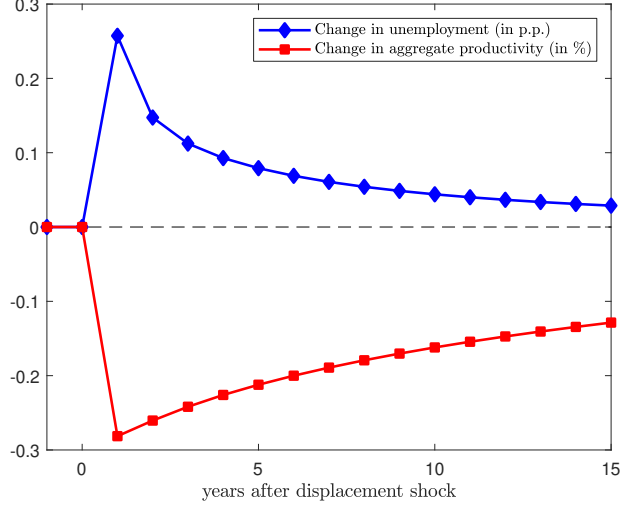
Figure 8: Impact of an aggregate displacement shock. *Notes.*—The figure shows the impact of a career displacement of 1 percent of the labor force on aggregate productivity (in percent) and unemployment (in percentage points). Both responses are averages computed at the ergodic process for the aggregate productivity state.

6   Welfare Consequences of Learning

This section quantifies the consequences of information frictions in terms of output losses, cross-sectional distributions, aggregate volatility, and implicit frictions in career mobility. We do so in two exercises. First, we study the gains from reallocating workers across job rungs according to their true skills, keeping their employment status and career choice fixed. Second, we study the implicit frictions imposed by learning on career mobility in the counterfactual environment where workers can churn careers and learn skills instantaneously.

6.1   Reallocating Job Rungs

Our first benchmark eliminates mismatch along the job ladder by reallocating workers to the job rung that maximizes output given their true skill, fixing employment statuses and career choices at the baseline equilibrium. The optimal allocation of job rungs to workers is given by

$$
r^{\mathrm{ce}}(a_k) = \begin{cases} r_{a_k}^- & \text{if } a_k \leq \eta r_{a_k}^- + (1-\eta) r_{a_k}^+ \\ r_{a_k}^+ & \text{else,} \end{cases}
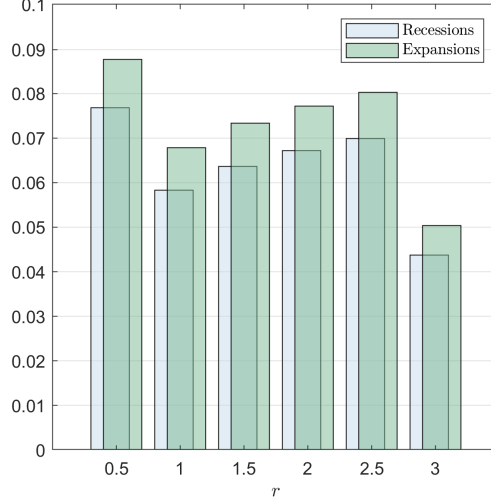$$

30

Figure 9: Gains from eliminating "within-career" mismatch, decomposed by job rung and business cycle state. *Notes.*—The figure shows the average output gain per worker from reallocating to the efficient job rung $r^{\text{ce}}(a_k)$, while keeping fixed their employment status and career choice. The gains are denominated in terms of unconditional average output per worker, $\mathbb{E}[y_{i,t}]/\mathbb{E}[1 - U_t]$.

where $r_{a_k}^-$ and $r_{a_k}^+$ define the lower and upper bound of a_k in $\mathcal{R}$.[30] Figure 9 presents the output gains per worker induced by such reallocation. The gains range from 4.4% for workers at the top rung in recessions to 8.8% for workers at the bottom rung in expansions. Aggregating across rungs and business cycle states, output under the optimal rung allocation is 7.0% higher than in equilibrium.

Across the job ladder, the gains show a non-monotonic pattern, which is the consequence of two opposing forces. On the one hand, uncertainty is negatively correlated with rungs, implying that a larger fraction of workers is optimally reallocated at bottom rungs. On the other hand, skills are complementary with task complexity, so that the output gains are larger at higher rungs. (The gains at the top rung are limited due to the imposed upper bound on $\mathcal{R}$; see also Footnote 30.)

Interestingly, the gains are higher in expansions than in recessions across all rungs. Aggregated across rungs, output under the optimal rung allocation is 7.4% higher in expansions compared to 6.4% in recessions. In other words, efficiency losses due to learning have a *dampening* effect on the business cycle. Intuitively, this is because cleansing reduces mismatch

---

[30]Here, we extend $\mathcal{R}$ beyond its lower boundary using the same grid as in the interior; i.e., we reallocate workers to the task space $\{\ldots, -.5, 0, .5, 1, \ldots, 3\} \cdot S_0^{1/2}$. We do so to avoid constraining the gains by an arbitrary lower bound on $\mathcal{R}$ (recall that the lower bound on $\mathcal{R}$ did not constrain equilibrium search). We do, however, impose the same upper bound on $\mathcal{R}$ as in equilibrium, because—albeit being inconsequential until now—this bound was indeed binding in equilibrium for a measure of .004 of workers. Efficiency gains from reallocating workers beyond this bound would therefore be a mechanical consequence of broadening the task space.

in recessions so that there is less potential for reallocating workers across rungs.

6.2 Implicit Friction on Career Mobility

In the model, workers only slowly learn the return on pursuing a career as they first have to find a job and then only learn from noisy signals about the relevant skill. Information frictions therefore impose an implicit cost on exploring new careers that reduces career mobility.

We now assess the magnitude of this implicit cost on exploring new careers. We do so by considering a fictitious career-switching problem in which workers can instantaneously churn careers and learn the relevant skill at *infinite speed* subject to an explicit switching cost $\xi_{i,t}$. For any given worker, we then calculate the magnitude of the explicit switching cost $\xi_{i,t}$ that keeps them indifferent between accessing the fictitious churning technology and sticking to their equilibrium career choice. Intuitively, our approach replaces the implicit information friction on career mobility by an explicit switching cost $\xi_{i,t}$, which we design so as to impose the same career mobility patterns for all workers.

Specifically, let $X_{i,t}$ denote the current value of a worker; $\mathcal{U}(\hat{a}_k, \Sigma_k, z)$ if unemployed, and $J(\hat{a}_k, \Sigma_k, r, z)$ if employed.[31] Then the marginal benefit of exploring a new career and learning the relevant skill instantaneously, $(\hat{a}, \Sigma) = (a, 0)$, is given by

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

To preempt workers from assessing the churning technology it hence suffices to set $\xi_{i,t} = \tilde{\xi}_{i,t}$. Table 7 reports the result (denominated in the *economy-wide* average monthly output per worker). The implicit friction is largest for low-skilled workers as they benefit the most from exploring new careers. It ranges from the equivalent of 18 months of output for workers at the bottom rung of the job ladder to less than one fifth of monthly output at the top rung.[32] Averaged across workers and business cycle states, the implicit friction evaluates to the equivalent of 10.02 months of average output per worker.

---

[31]We use the joint worker–firm value J for employed workers to be consistent with joint surplus maximization as imposed throughout the paper.

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

Table 7: Mobility Costs Implicit in Learning

	Recession	Expansion
Unemployed	16.03	15.37
$r = 0.5 \cdot S_0^{1/2}$	17.60	18.12
$r = 1.0 \cdot S_0^{1/2}$	12.85	13.27
$r = 1.5 \cdot S_0^{1/2}$	7.23	7.42
$r = 2.0 \cdot S_0^{1/2}$	2.91	2.96
$r = 2.5 \cdot S_0^{1/2}$	0.93	0.95
$r = 3.0 \cdot S_0^{1/2}$	0.18	0.19

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

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.

We have further used the model to quantify the welfare cost associated with information-driven mismatch, concluding that the potential gains from eliminating such mismatch are substantial. Moreover, our empirical results in Section 3.3 show that we—as researchers—can systematically outpredict workers when it comes to their ability, suggesting that, at least in part, these welfare gains can be feasibly implemented by some policy. We leave it to future research to explore policies that could materialize the potential efficiency gains that we have documented.

Our framework is among the first that incorporates multidimensional sorting into an equilibrium model with labor market frictions (see also, Lise and Postel-Vinay, 2018; Lindenlaub and Postel-Vinay, 2017). It is distinguished from the existing literature by its analytical tractability, which opens the door to an analysis of aggregate shocks. Our framework delivers rich predictions regarding job and career mobility, which may be interesting for further investigation.

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.

Baley, Isaac, and Andrés Blanco. 2019. "Firm uncertainty cycles and the propagation of nominal shocks." *American Economic Journal: Macroeconomics*, 11(1): 276–337.

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. 2014. "Unemployment and Endogenous Reallocation over the Business Cycle." *Working Paper.*

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.

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

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

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

Dubra, Juan. 2004. "Optimism and overconfidence in search." *Review of Economic Dynamics*, 7(1): 198–218.

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

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

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

Gali, Jordi, and Thijs van Rens. 2019. "The Vanishing Procyclicality of Labor Productivity." *Warwick Economic Research Paper.*

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.

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

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

Huckfeldt, Christopher. 2019. "Understanding the Scarring Effect of Recessions." *working paper.*

Jarosch, Gregor. 2015. "Searching for Job Security and the Consequences of Job Loss." *working paper.*

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.

Kambourov, Gueorgui, and Iourii Manovskii. 2008. "Rising Occupational and Industry Mobility in the United States: 1968-97." *International Economic Review*, 49(1): 41–79.

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

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

Lise, Jeremy, and Fabien Postel-Vinay. 2018. "Multidimensional Skills, Sorting, and Human Capital Accumulation." *Working Paper.*

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.

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.*

Papageorgiou, Theodore. 2014. "Learning your comparative advantages." *Review of Economic Studies*, 81(3): 1263–1295.

Petrongolo, Barbara, and Christopher A Pissarides. 2001. "Looking into the Black Box: A Survey of the Matching Function." *Journal of Economic Literature*, 39(2): 390–431.

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.

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

Senga, Tatsuro. 2016. "A New Look at Uncertainty Shocks: Imperfect Information and Misallocation." *working paper.*

Shimer, Robert. 2012. "Reassessing the ins and outs of unemployment." *Review of Economic Dynamics*, 15(2): 127–148.

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

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

Straub, Ludwig, and Robert Ulbricht. 2012. "Credit Crunches, Information Failures, and the Persistence of Pessimism." *mimeo, Toulouse School of Economics.*

Straub, Ludwig, and Robert Ulbricht. 2014. "Endogenous Uncertainty and Credit Crunches." *working paper.*

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.

A Kolmogorov Forward Equations

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

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

$$\dot{\Gamma}_t(\hat{a}, \Sigma, r) = \dot{\Gamma}_t^{\text{learn}}(\hat{a}, \Sigma, r) + \dot{\Gamma}_t^{\text{ee}}(\hat{a}, \Sigma, r) + \dot{\Gamma}_t^{\text{ue}}(\hat{a}, \Sigma, r) - \dot{\Gamma}_t^{\text{eu}}(\hat{a}, \Sigma, r). \tag{13}$$

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

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

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

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

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

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

Finally, the fourth term defines separations into unemployment,[33]

$$\dot{\Gamma}_t^{\text{eu}}(\hat{a}, \Sigma, r) = \begin{cases} \delta \Gamma_t(\hat{a}, \Sigma, r) & \text{if } J_t^{\text{act}}(\hat{a}, \Sigma, r, z) > J^{\text{sep}}(\hat{a}, \Sigma, r, z) \\ \lim_{\pi \to \infty} \pi \Gamma_t(\hat{a}, \Sigma, r) & \text{else.} \end{cases}$$

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

$$\dot{\Upsilon}_t(\hat{a}, \Sigma) = \dot{\Upsilon}_t^{\text{cs}}(\hat{a}, \Sigma) + \dot{\Upsilon}_t^{\text{eu}}(\hat{a}, \Sigma) - \dot{\Upsilon}_t^{\text{ue}}(\hat{a}, \Sigma). \tag{14}$$

---

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

Here, the first term defines net changes in (current-career) beliefs due to agents switching careers,[34]

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

where $\chi(\hat{a}, \Sigma, z) \in \{0, 1\}$ is an indicator evaluating to unity if workers switch careers $(U_t(a_0, S_0, z) > U_t(\hat{a}, \Sigma, z))$. The second term defines gross inflows into unemployment, taking into account that workers switch careers at an exogenous probability ϵ,

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

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

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

B   Examples of General Production Function

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

Complementary-skill case   Let

$$F(z(t), \boldsymbol{q}_{k,r}, \boldsymbol{a}_i) \equiv \exp \left[z(t) + \sum_{j=1}^J \left(\eta q_{k,r,j} - \max \left\{ q_{k,r,j} - \frac{q_{k,r,j} a_{i,j}}{\sum_{j=1}^J q_{k,r,j}}, 0 \right\} \right) \right]. \quad (15)$$

Substituting $r = \sum_{j=1}^J q_{k,r,j}$ and $w_{k,j} = q_{k,r,j} / (\sum_{j=1}^J q_{k,r,j})$, we can rewrite (15) in more accessible form

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

---

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

which clearly collapses into (1) for an orthogonal weighting scheme; e.g.,[35]

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

Substitutible-skill case   Let

$$F(z(t), \boldsymbol{q}_{k,r}, \boldsymbol{a}_i) \equiv \exp\left[z(t) + \eta \sum_{j=1}^{J} q_{k,r,j} - \max\left\{ \sum_{j=1}^{J} q_{k,r,j} - \frac{\sum_{j=1}^{J} q_{k,r,j} a_{i,j}}{\sum_{j=1}^{J} q_{k,r,j}}, 0 \right\} \right], \qquad (16)$$

which can be rewritten more compactly as

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

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

C   Measuring Job Requirements, Employment Transitions, and Worker Skills

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

C.1   Job Requirements

Following Guvenen et al. (2020), we measure skill requirements using 26 O*NET descriptors from the Knowledge, Skills and Abilities categories that were identified by the Defense Manpower Data Center (DMDC) to be related to each ASVAB category, augmented by six descriptors linked to social skills.[36] As in Guvenen et al. (2020), we link those O*NET descriptors to ASVAB test category based on the relatedness score provided by DMDC. The verbal skill requirement is then defined as the first principal component of Word Knowledge and

---

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

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

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

C.2   Employment Transitions

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

Sample selection   As the NLSY79 is well-known and requires little description, we focus in the following on describing the sample selection used in this paper. We focus on the subsample of males and females from the so-called cross-sectional sample, which is designed to represent the non-institutionalized civilian segment of the U.S. in 1979.[37] As is standard in the literature, we drop individuals who were more than two years in the military force, individuals with a weak labor market attachment (spending more than 10 years out of the labor force), individuals that were already working in 1979, and those that do not have information on the Armed Services Vocational Aptitude Battery (ASVAB) test scores.

C.3   Worker Skills and Mismatch

Worker skills   We measure workers skills using ASVAB test scores available in the NLSY79 (see Appendix C.2 for a description of our subsample). The ASVAB is a general test that measures knowledge and skills in 10 different components that was taken by survey participants when first entering the survey.[38] As in Guvenen et al. (2020), we focus on a subset of seven components (arithmetic reasoning, mathematics knowledge, paragraph comprehension, word knowledge, mechanical comprehension, general science and electronics information) which are

---

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

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

linked to math, verbal and technical skills, and are combined using Principal Components Analysis. For the social dimension, we proceed in the same fashion using the individual scores in two different tests provided by the NLSY79: the Rotter Locus of Control Scale and the Rosenberg Self-Esteem Scale.[39] To adjust for differences in test-taking age, before proceeding with PCA, we normalize the mean and the variance of each test score according to their age-specific values. Then, once we have the raw scores in each skill dimension, we convert these into percentile ranks.

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

D   Computing Wages Without Commitment by Workers

This appendix details the computation of wages used for the exploration in Section 5. Following Schaal (2017) we adopt the unique wage scheme that induces equilibrium search and job continuation policies to be self-enforcing for workers (without requiring a contractual commitment).

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

---

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

given by

$$
\rho W_t^{\mathrm{act}}(\hat{a}_k, \Sigma_k, r, z) = w_{i,t} + \tilde{\Lambda}_t(\hat{a}_k, \Sigma_k, r, z) +
$$
$$
+ \max_{x,r} \left\{ \kappa p(\theta_t(\omega, z)) \left(x - W_t(\hat{a}_k, \Sigma_k, r, z) \right) \right\} +
$$
$$
+ \delta \left(J_t^{\mathrm{sep}}(\hat{a}_k, \Sigma_k, z) - W_t(\hat{a}_k, \Sigma_k, r, z) \right) +
$$
$$
+ \lambda_z \left(W_t(\hat{a}_k, \Sigma_k, r, -z) - W_t(\hat{a}_k, \Sigma_k, r, z) \right), \quad (17)
$$

where

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

and

$$
W_t(\hat{a}_k, \Sigma_k, r, z) = \max \left\{ W_t^{\mathrm{act}}(\hat{a}_k, \Sigma_k, r, z), J_t^{\mathrm{sep}}(\hat{a}_k, \Sigma_k, z) \right\}.
$$

Here, the separation value for the worker equals the joint value of separation, J_t^{sep}, as firms' separation values are fixed at zero due to free entry. Absent contractual commitments, workers' on-the-job search maximizes (17) subject to (2). Rearranging the associated first-order condition, we have

$$
\theta = p'^{-1} \left(\frac{c}{J_t(\hat{a}_k, \Sigma_k, r^*, z) - W_t(\hat{a}_k, \Sigma_k, r, z)} \right). \quad (18)
$$

Comparing (18) with (9), we conclude that for search to be self-enforcing, the worker value of the relationship must match the joint value whenever she is actively searching. Accordingly, the unique self-enforcing wage scheme is given by

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

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

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

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

E Additional Tables

Table 8: Cyclicality in mismatch among job-to-job transitions in the data

Dependent variable ($\times 100$):	$m_{i,t}$	$m_{i,t}^{+}$	$m_{i,t}^{-}$
	(1)	(2)	(3)
Job-to-job transitions (β_1)	.376	.517**	$-.141$
	(.313)	(.236)	(.218)

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).