American Economic Journal: Macroeconomics 2020, 12(1): 210–244
https://doi.org/10.1257/mac.20160241

Multidimensional Skill Mismatch†

By Fatih Guvenen, Burhan Kuruscu, Satoshi Tanaka, and David Wiczer*

What determines the earnings of a worker relative to his peers in the same occupation? What makes a worker fail in one occupation but succeed in another? More broadly, what are the factors that determine the productivity of a worker-occupation match? To help answer questions like these, we propose an empirical measure of multidimensional skill mismatch that *is based on the discrepancy between the portfolio of skills required by an occupation and the portfolio of abilities possessed by a worker for learning those skills. This measure arises naturally in a dynamic model of occupational choice and human capital accumulation with multidimensional skills and Bayesian learning about one's ability to learn skills. Not only does mismatch depress wage growth in the current occupation, it also leaves a scarring effect—by stunting skill acquisition—that reduces wages in future occupations. Mismatch also predicts different aspects of occupational switching behavior. We construct the empirical analog of our skill mismatch measure from readily available US panel data on individuals and occupations and find empirical support for these implications. The magnitudes of these effects are large: moving from the worst- to best-matched decile can improve wages by 11 percent per year for the rest of one's career.* (*JEL* E24, J24, J31, J41)

What determines the earnings of a worker relative to his peers in the same occupation? What makes a worker fail in one occupation but succeed in another? More broadly, what are the factors that determine the productivity of a worker-occupation match? Each of these questions highlights a different aspect of the career search process that all workers go through in the labor market. Given their clear importance, economists have long sought answers to these questions. Going

* Guvenen: University of Minnesota, 4-101 Hanson Hall, 1925 Fourth Street South, Minneapolis, MN 55455, Fedearl Reserve Bank of Minneapolis, and NBER (email: guvenen@umn.edu); Kuruscu: University of Toronto, 150 St. George Street, Toronto, ON M5S 3G7, Canada (email: burhan.kuruscu@utoronto.ca); Tanaka: University of Queensland, Level 6 Colin Clark Building, Blair Drive, QLD 4072, Australia (email: s.tanaka@uq.edu.au); Wiczer: Stony Brook University, 100 Nicolls Road, Social & Behavioral Sciences Building, 6th Floor, Stony Brook, NY 11794 (email: david.wiczer@stonybrook.edu). Richard Rogerson was coeditor for this article. For comments and discussions, we thank Joe Altonji, Alessandra Fogli, Tim Kautz, Philipp Kircher, Jeremy Lise, Fabien Postel-Vinay, Rob Shimer, Kjetil Storesletten, José-Víctor Ríos-Rull, Carl Sanders, and Uta Schönberg as well as the participants at the 2011 and 2015 SED conferences, 2014 NBER Summer Institute, 2014 Barcelona GSE Summer Forum, 2015 CIREQ Conference on Information Frictions, 2018 Mannheim Macro Workshop, Universities of Copenhagen, Hitotsubashi, Tokyo, UCSB, Western Ontario, Erlangen-Nuremberg, Australian National, New South Wales, and the German Institute for Employment Research (IAB). The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Banks of Minneapolis or St. Louis. Guvenen acknowledges financial support from the National Science Foundation. Kuruscu acknowledges financial support from Social Sciences Humanities Research Council of Canada.

† Go to https://doi.org/10.1257/mac.20160241 to visit the article page for additional materials and author disclosure statement(s) or to comment in the online discussion forum.

back (at least) to the classic Roy (1951) model, a key idea has been that workers differ in their skills (or productivity) in different kinds of occupations and self-select into those where they have a comparative advantage.[1]

In this paper, we propose an empirical measure of multidimensional skill mismatch for each worker-occupation pair and show that it is very informative about the questions raised above as well as in predicting some other key occupation-related outcomes. That skills are multidimensional (e.g., cognitive, physical, verbal, social, and so on) is neither a new idea nor a controversial one. It is also supported by a wide range of empirical research in psychometry, education, and economics.[2] The contributions of this paper are twofold. The first contribution is to show that our mismatch measure emerges naturally in a dynamic model of occupational choice, multidimensional skill accumulation, and Bayesian learning about abilities to acquire those skills. The framework we propose essentially extends the Roy model by imposing specific structure on key components. We derive structural equations from this model that show how wages and occupation changes depend on the skill mismatch measure. Our second contribution is to construct the empirical counterpart of the skill mismatch measure using data from both workers *and* occupations, and examine the extent to which it is informative about current and future wages and occupational switching.

The specifics of the model are as follows. Each occupation produces output according to a technology described by the vector of *skill requirements* that is different for each occupation. These requirements determine, simultaneously, how much output is produced by a worker with a given ability vector and how much human capital the worker accumulates during the period.[3] The technology is specified such that, for each worker, there is a unique (and interior) optimal level of investment in each skill type depending on his abilities, and thus an optimal occupation choice. Workers enter the economy with imperfect information about their true abilities to acquire different types of skills and learn over time in a Bayesian fashion. Hence, in each period, the occupation they choose is different from the optimal one under full information. Skill mismatch measures the distance between the portfolio of skills required by an occupation and the portfolio of abilities possessed by a worker.

Our first contribution is to derive a Mincer-style wage equation that shows that the current wage is negatively related to mismatch at a worker's occupation *and* to the *cumulative* mismatch the worker experienced at his past occupations. Therefore, skill mismatch not only depresses current wages but also leaves a long-lasting impact on future wages—even after the worker leaves the mismatched occupations. In the empirical analysis, we find strong empirical support for this prediction. It is worth noting that the derivation of this equation does not depend on the precise friction that creates mismatch, so the conclusion just described does not rely on mismatch arising from Bayesian learning (as opposed to, say, search frictions).

[1] See Sanders and Taber (2012) for a recent survey on this literature.

[2] Developmental psychologist Howard Gardner formally proposed his theory of "multiple intelligences" in a 1983 book, and the subsequent literature provided supporting evidence. It is worth mentioning that Gardner found particular motivation for multiple intelligences in the proliferation of occupations (see, e.g., Gardner 2011, xxii).

[3] The distinction between abilities and skills is that the former is fixed for a worker whereas skills are accumulated, and at a rate that depends on the worker's ability to learn each skill and the occupation they work at.

A second set of predictions we are interested in concerns the relationship between skill mismatch and occupational switching, for which the particular friction that creates mismatch becomes relevant. The one we assume—imperfect information about learning abilities—is not an implausible or uncommon one.[4] With Bayesian learning, workers update their beliefs about their abilities every period, which causes them to optimally switch occupations and reduce their mismatch, in turn reducing occupational mobility.

We find empirical support for four key predictions of this model. First, mismatch reduces both the level of wages *and* the slope of wage growth with occupational tenure. Second, as already noted, current wages depend negatively on *cumulative* mismatch in previous occupations. Third, the probability of switching occupations increases with our mismatch measure. Fourth, occupational switches are directional: workers who are overqualified in a given skill dimension tend to switch to occupations that are more skill intensive in that dimension and vice versa. Furthermore, conditional on switching, the reduction in mismatch is proportional to the level of mismatch in the previous occupation.

To construct the empirical measure of skill mismatch, we use the 1979 National Longitudinal Survey of Youth (NLSY79), which contains information on the occupation and wage histories of thousands of individuals since their youth. It also contains test scores from an occupational placement test—the Armed Services Vocational Aptitude Battery (ASVAB)—and several measures of noncognitive skills.[5] On the occupation side, we use the US Department of Labor's O*NET project for data on skill requirements. Combining these data from both sides, we can measure mismatch along three skill dimensions: (cognitive) math skills, (cognitive) verbal skills, and (noncognitive) social skills.

We incorporate these (contemporaneous and cumulative) mismatch measures into a Mincerian wage regression and also interact them with occupational tenure, along with a long list of controls. Consistent with our model, we find robustly negative coefficients on mismatch and its interaction with occupational tenure. The estimates imply that, after 10 years of occupational tenure, the average wage of workers in the top decile of the current mismatch distribution is 6.8 percent lower than those in the bottom decile (i.e., best-matched workers). Cumulative past mismatch has a similarly large effect: wages are 9.4 percent lower for workers in the top decile of the cumulative mismatch distribution compared with those in the bottom decile. These long-lasting effects of mismatch are consistent with the human capital accumulation channel captured by our model but would be missed by theories where match quality

[4] Among many examples, see for example Farber and Gibbons (1996), Gibbons et al. (2005), Antonovics and Golan (2012), and Papageorgiou (2014). One difference however is that these studies assumed learning was about worker's comparative advantage (or what we would call skills here), whereas in our model it is about the *abilities* to learn those skills.

[5] We interpret workers' test scores as corresponding to (noisy measures of) abilities in our model. Although these scores probably reflect a combination of abilities and accumulated skills, this distinction is probably not critical for our purposes because the two are highly correlated at young ages. Huggett, Ventura, and Yaron (2011) estimate that this correlation exceeds 0.85 at age 20. Since these tests are taken at the beginning of workers' careers, we interpret them as abilities. Another question is why workers do not choose their ideal occupation if they know their ASVAB scores for each ability type. This is because, first, the NLSY respondents were not told their exact test score—they were given a fairly wide range—and second, these test scores are themselves noisy measures of individuals' true underlying abilities as discussed further later.

only affects the current match. Overall, eliminating all mismatch would raise the average wage of workers by 11 percent every year.

Turning to occupational switching behavior, the data reveal some interesting patterns. A hazard model for occupational switching shows that the probability of switching is increasing in current mismatch. The magnitudes are fairly large: the switching probability is about 3.4 percentage points higher for a worker at the ninetieth percentile of the mismatch distribution relative to another worker at the tenth percentile. This gap is about one-fifth of the average switching probability in our sample. In addition, occupational transitions tend to "correct" previous mismatches: if a worker is overqualified along a certain skill dimension, the next occupation, on average, has higher skill requirements *in that dimension* (as well as in other skill dimensions but to a lesser extent). Furthermore, consistent with our theory, the step size of these switches are proportional to the level of current mismatch.

Finally, we extend our wage regressions to distinguish mismatch for overqualified and under-qualified workers. We find that both those who were overqualified and those who were under-qualified in previous occupations have lower wages today. This implication is consistent with our model but inconsistent with a standard Ben-Porath model with multidimensional skills, as we discuss in Section IB.

Our paper contributes to an active literature on the skill content of occupations. One strand of this literature uses data on occupation characteristics (e.g., from O*NET) and explores how they are related to the wages in those occupations (Ingram and Neumann 2006, Poletaev and Robinson 2008, Gathmann and Schönberg 2010, Bacolod and Blum 2010, and Yamaguchi 2012). A second strand focuses on worker-side information to define mismatch. Among these, Fredriksson, Hensvik, and Skans (2018) defines mismatch as the gap between the skills of a new hire and his experienced peers in the same occupation and establishment. Perry, Wiederhold, and Ackermann-Piek (2014) reviews this strand of the literature. The main difference of the mismatch measure we propose it that it combines information from both sides, worker skills and occupational skill requirements. We then show that this measure is a significant predictor of current and future wages, even after controlling for worker skills and occupational skill requirements separately.

Our paper also has useful points of contact with an emerging literature that aims to understand how skill accumulation is affected by a worker's interaction with her peers. Among these Herkenhoff et al. (2018) and Jarosch, Oberfield, and Rossi-Hansberg (2018) study learning from coworkers at the same firm and how this affects a worker's future wages, and Akcigit et al. (2018) studies how the quality of an inventor (measured by patents) is influenced by the quality of inventors he or she interacted in the past (either as team member, coworker, resident in same area, etc.). All three papers find that past interactions with better peers raise a worker's wages and productivity. Our analysis is related but differs in context: we find that working at an occupation with higher skill requirements than a worker possesses (negatively mismatched) *reduces* his future wage growth; perhaps surprisingly, the effect is symmetric: being positively mismatched (overqualified in some skills) also reduces future wages. While our results cannot be directly compared to theirs given the different contexts, these results collectively point to the setting where learning takes place (workplace, occupation, etc.) as a key factor in skill accumulation.

Lise and Postel-Vinay (2016) and Sanders (2014) are two contemporaneous papers that are most closely related. They both estimate models with multidimensional skills, sorting, and human capital accumulation, using data similar to ours (NLSY79 and O*NET).[6] Sanders (2014) focuses is on the effects of a worker's uncertainty about her skill level, so he classifies job tasks based on how much workers may know about their skills to perform that task. Instead, we focus on different *classes* of skills (such as math, verbal, and social) and emphasize uncertainty about the *ability* to learn each skill. Sanders does not directly examine the effects of mismatch on labor market outcomes. Lise and Postel-Vinay (2016) study a model where skill mismatch results from search frictions and estimate mismatch by matching model and data moments, whereas we first construct the empirical skill mismatch measure directly from data and *then* test the relationships predicted by the model. These are complementary approaches, each with distinct advantages. Our findings in Section IV lend support to learning but do not provide evidence against search frictions as an additional source of mismatch. Finally, Groes, Kircher, and Manouskii (2015) studied how the probability of occupational switching depends on mismatch, defined as the gap between a worker's wage from the average of his coworkers. Our findings on switching are consistent with theirs. They did not analyze the effects on wages.

Finally, our paper also has useful points of contact with the literature that studies career outcomes in the presence of comparative advantage or employer/worker learning or both. Because different sectors can reward skills differently, as workers learn about their skills they switch toward sectors that maximize their comparative advantage (Gibbons et al. 2005, Antonovics and Golan 2012, Gervais et al. 2016, and Papageorgiou 2014).[7]

The paper proceeds as follows. In Section I we present our model. In Section II we describe our data, and Section III describes our methodology. Section IV presents the results, and finally, we conclude in Section V.

I. Model

In this section, we present a life cycle model of occupational choice and human capital accumulation. It provides a natural framework to analyze how skill mismatch arises and, more importantly, how it affects current and future wages of a worker in different occupations.[8] The model delivers some testable implications that we will empirically study in Section IV.

Briefly, the structure of the labor market builds upon Rosen (1972), wherein workers supply labor services to firms (where we think of each firm as a distinct occupation), and firms in turn offer different training/learning opportunities.

[6] Speer (2017) is also methodologically relevant to our work by combining worker- and occupation-side skill measures to study how gender discrimination and layoffs shape occupational choice.

[7] In a slightly different context, Farber and Gibbons (1996) and Altonji and Pierret (2001) investigate the extent of employer learning about worker abilities. They show that an interaction term between ability and job tenure included in a Mincer wage regression is positive, so wages reflect abilities more closely over time, which they interpret as evidence of employers learning about workers' abilities during their relationship.

[8] Throughout the paper we use the terms "human capital" and "skill" interchangeably.

Consequently, a worker's compensation from a job/occupation is not only a wage but also a rise in his skill level. Our model introduces two key features into this framework. First, human capital is multidimensional, and the rate at which a worker acquires each new skill depends both on his learning ability of that skill and the occupation's requirements for investment in that skill, meaning an allocation choice in the spirit of Roy (1951). Second, workers enter the labor market without full knowledge about their learning abilities; instead, they have a prior belief about their abilities that they update over time in a Bayesian fashion. This information friction is the source of skill mismatch between workers and their occupations. We now turn to the details of the model.

A. *Environment*

Each worker lives for T periods and supplies one unit of labor inelastically in the labor market. The objective of a worker is to maximize the expected present value of earnings/wages:

$$\mathbb{E}_0\left[\sum_{t=1}^{T}\beta^{t-1}w_t\right],$$

where β is the subjective time discount factor.

Occupations and Technology.—There is a continuum of occupations, each using n types of skills, indexed with $j \in \{1,2,\ldots,n\}$. Occupations differ in their *skill intensity* (alternatively called *skill requirement*) of each skill type, denoted with the vector $\mathbf{r} = (r_1, r_2, \ldots, r_n) \geq 0$, which is fixed over time. With a slight abuse of notation, we will use $\mathbf{r}_t = (r_{1,t}, r_{2,t}, \ldots, r_{n,t})$ to denote the particular occupation chosen by the worker in period t. Just as skills are multidimensional, so are the abilities to accumulate them, which we stack in a vector, $\mathbf{A} \equiv (A_1, A_2, \ldots, A_n)$.

A worker starts period t with the skill portfolio $\mathbf{h}_t = (h_{1,t}, h_{2,t}, \ldots, h_{n,t})$, chooses an occupation $\mathbf{r}_t$, accumulates new skills at that occupation, and then produces output with these upgraded skills. The technology for skill accumulation is a key ingredient in our model and is given by

$$k_{j,t} \equiv h_{j,t} + (A_j + \varepsilon_{j,t})r_{j,t} - r_{j,t}^2/2, \tag{1}$$

where $k_{j,t}$ is the upgraded skill of type j, which is used in production in period t and determines next period's starting human capital level (as explained in a moment), and $\varepsilon_{j,t}$ is a random disturbance term (whose role will become clear later). One way to think about this structure is that, in a given period, workers first go through training and then produce output with their new/upgraded skills.

This specification has two important features. First, the occupation's skill requirement, $r_{j,t}$, affects human capital accumulation non-monotonically; the linear positive term captures the benefit of an occupation whereas the negative quadratic captures the costs of skill acquisition, such as additional training required at high-skill jobs. As will become clear later, this non-monotonicity ensures that optimal occupational choice is an interior one (i.e., workers do not all flock to the occupations with the highest $r_{j,t}$). Second, with the formulation in (1), the linear benefit term is proportional to the worker's ability $(A_j + \varepsilon_{j,t})$, whereas the cost term is independent of ability, which gives rise to comparative advantage and sorting by ability level—workers choose occupations with higher skill requirements only in dimensions where their ability is relatively high. These two features will come to play important roles in what follows.

The output of a worker in a given occupation is the sum of his end-of-period skills, and perfect competition among occupations for workers ensures that a worker's wage equals the output he produces:

$$w_t = \sum_j k_{j,t}\left(h_{j,t}, A_j; r_{j,t}, \varepsilon_{j,t}\right). \tag{2}$$

At first blush, this wage equation (2) may appear a bit peculiar as there is no explicit price multiplying each skill (as would be the case in a standard human capital model, e.g., Ben-Porath 1967 or Rosen 1972), which makes it seem as if there is no variation in the value of skills across occupations or even a relative price of each skill at the aggregate level. This is not the case however: the $k_{j,t}$s already embed how productive each (newly produced) skill is in each occupation (through equation (1)). So, two workers with the same human capital portfolio at the beginning of the period, $\mathbf{h}_t$, who work at different occupations during the period will end up with different amounts of $k_{j,t}$ and therefore receive different levels of compensation. Thus, a worker's wage depends on his human capital vector $\mathbf{h}_t$, learning ability, $\mathbf{A}$, occupation $\mathbf{r}_t$, and stochastic disturbance, $\varepsilon_{j,t}$. The occupation-specific skill intensity, $\mathbf{r}_t$, plays a dual role, determining the relative price of each skill in each occupation in addition to also affecting the rate of skill accumulation in each occupation.[9]

Finally, the beginning-of-period human capital in period $t+1$ is equal to $k_{j,t}$ adjusted for depreciation:

$$h_{j,t+1} = (1-\delta)k_{j,t} = (1-\delta)\left(h_{j,t} + (A_j + \varepsilon_{j,t})r_{j,t} - r_{j,t}^2/2\right), \tag{3}$$

where the depreciation rate δ is the same for all skill types and occupations.

B. *Empirical Wage Equation and Skill Mismatch*

We can now derive our first key equation, which shows how a worker's wage depends on his abilities and the history of occupations he has been employed at. To this end, first, combining equations (1)–(2) and rearranging yields

$$w_t = \sum_{j=1}^{n}\left(h_{j,t} + \frac{A_j^2}{2} - \frac{(A_j - r_{j,t})^2}{2}\right) + \sum_{j=1}^{n} r_{j,t}\varepsilon_{j,t}. \tag{4}$$

[9] In this sense, our model fully captures the spirit of comparative advantage as in the Roy model but does so in a way that both retains tractability, which will allow us to derive a number of key predictions and deliver implications of mismatch as will become clear in a moment. The generality afforded by this formulation is also in the same spirit as the "skill weights" approach in Lazear (2009).

This expression makes clear that a worker's wage depends positively on his (beginning-of-period) human capital and ability and negatively on $\left(A_j - r_{j,t}\right)^2$, which is simply the deviation between his ability level and his occupation's skill requirement—or what we will call skill *mismatch*. Next, using equation (3) and repeatedly substituting for human capital (and setting $\delta \equiv 0$ for clarity) yields[10]

$$h_{j,t} = h_{j,1} + \frac{A_j^2}{2}(t-1) - \sum_{s=1}^{t-1} \frac{\left(A_j - r_{j,s}\right)^2}{2} + \sum_{s=1}^{t-1} r_{j,s}\,\epsilon_{j,s}, \tag{5}$$

which shows that human capital grows with experience at a rate that is proportional to ability (second term) and, more importantly for our purposes, is depressed by the degree of mismatches at all his past occupations. Finally, substituting this expression into (4) delivers the key empirical wage equation that we will use in the empirical analysis:

$$w_t = \sum_j h_{j,1} + \underbrace{\frac{1}{2}\sum_{j=1}^{n} A_j^2 \times t}_{\text{ability}\times\text{experience}} - \underbrace{\frac{1}{2}\sum_{j=1}^{n}\sum_{s=1}^{t}\left(A_j - r_{j,s}\right)^2}_{\text{mismatch}} + \sum_{j=1}^{n}\sum_{s=1}^{t} r_{j,s}\,\varepsilon_{j,s}. \tag{6}$$

Two remarks are in order. First, the wage inherits the two properties of human capital noted above; in particular, its growth is proportional to a weighted average of the worker's ability portfolio, and it is depressed by the history of past mismatches.[11] This long-lasting negative effect of past mismatches on future wages is the one of the key implications of the model that we will test in the data. Second, notice that we did not yet specify how workers choose occupations or what causes mismatch. So, the empirical implications of this wage equation do not depend on those specifics. Instead, they come from the specification of technology that determines how wages and human capital accumulation depend on abilities and occupations.[12]

Turning to the wage equation, we rearrange it slightly to distinguish between the effect of current mismatch and past cumulative mismatch on wages. To this end, let t^c denote the period in which the worker switched to his current occupation (i.e., $r_{j,s} = r_{j,t^c}$ for $s \geq t^c$), so his tenure in the current occupation is $t - t^c + 1$. We can rewrite equation (6) as

$$\begin{aligned} w_t = \sum_j h_{j,1} &+ \underbrace{\frac{1}{2}\sum_j A_j^2 \times t}_{\text{ability}\times\text{experience}} - \underbrace{\frac{1}{2}\sum_j \left(A_j - r_{j,t^c}\right)^2}_{\text{current mismatch}} \times \underbrace{\left(t - t^c + 1\right)}_{\text{current tenure}} \\ &- \frac{1}{2}\underbrace{\sum_{s=1}^{t^c-1}\sum_j \left(A_j - r_{j,s}\right)^2}_{\text{cumulative past mismatch}} + \sum_{s=1}^{t}\sum_j r_{j,s}\,\varepsilon_{j,s}, \end{aligned} \tag{7}$$

[10] The analogous expression with $\delta > 0$ is given in online Appendix A.

[11] With positive depreciation, mismatches that are farther in the past will be discounted in calculating cumulative mismatch. See equation (11) in online Appendix A.

[12] It might seem that if the model is correctly specified and we were to estimate the wage equation (6) from data, we should get coefficients of $1/2$ on the ability and mismatch terms. But this would only be true if A_j and $r_{j,s}$ were cardinal variables that had exact analogs in the data, which is not the case. Therefore, the model's prediction is more about the significance of these variables for determining wages rather than about particular values for the estimated coefficients.

which shows the separate effects of current mismatch and past cumulative mismatches, which we will map into the data in Section IV.

C. *Information Structure and Bayesian Learning*

Each worker draws ability A_j from a normal distribution at the beginning of his life: $A_j \sim \mathcal{N}\left(\mu_{A_j}, \sigma_{A_j}^2\right)$. The worker does not know his true A_j but observes a signal, $\hat{A}_{j,1} = A_j + \eta_j$, where $\eta_j \sim \mathcal{N}\left(0, \sigma_{\eta_j}^2\right)$. So, his prior beliefs are normally distributed with mean $\hat{A}_{j,1}$ (unbiased) and precision $\lambda_{j,1} \equiv 1/\sigma_{\eta_j}^2$.

Each period, the worker observes $A_j + \varepsilon_{j,t}$, where $\varepsilon_{j,t} \sim \mathcal{N}\left(0, \sigma_{\varepsilon_j}^2\right)$. This is equivalent to saying the worker observes each of his skills each period and knows his initial skill vector, $\mathbf{h}_1$. Then, given his current beliefs, the worker updates his belief about A_j. The worker's belief at the beginning of each period is normally distributed. Let $\hat{A}_{j,t}$ be the mean, $\lambda_{j,t}$ be the precision of this distribution at the beginning of period t, and λ_{ε_j} be the precision of $\varepsilon_{j,t}$. After observing $A_j + \varepsilon_{j,t}$, the worker updates his belief according to the following recursive Bayesian formula:

$$\hat{A}_{j,t+1} = \frac{\lambda_{j,t}}{\lambda_{j,t+1}} \hat{A}_{j,t} + \frac{\lambda_{\varepsilon_j}}{\lambda_{j,t+1}} \left(A_j + \varepsilon_{j,t}\right), \tag{8}$$

where $\lambda_{j,t+1} = \lambda_{j,t} + \lambda_{\varepsilon_j}$.[13]

D. *Worker's Problem*

Given the current beliefs about his abilities, the problem of the worker in period t is given as follows:

$$V_t\left(\mathbf{h}_t, \hat{\mathbf{A}}_t\right) = \max_{\{r_{j,t}\}} E_t\left[\sum_j k_{j,t} + \beta V_{t+1}\left(\mathbf{h}_{t+1}, \hat{\mathbf{A}}_{t+1}\right)\right],$$

subject to (1), (3), and (8). Since occupations are represented by a vector of skill intensities, this problem yields a choice of occupation in the current period that then determines not only current wages but also future human capital levels. The expectation in the worker's problem is taken with respect to the distribution of his beliefs about A_j (for $j = 1, \ldots, n$), given by $\mathcal{N}\left(\hat{A}_{j,t}, 1/\lambda_{j,t}\right)$, and the distribution of $\varepsilon_{j,t}$, given by $\mathcal{N}\left(0, \sigma_{\varepsilon_j}^2\right)$.

[13] Notice that to the extent that abilities are correlated, each new observation about one ability also provides information about other abilities, which we implicitly assumed away here. Taking this additional information into account is possible but will turn the learning problem into a multidimensional Kalman filtering problem, which makes the solution of the model analytically less tractable, so we do not consider it here. It is a potential extension in future work. That said, the extent to which this cross updating of beliefs empirically matters depends on the magnitude of the correlation between abilities (c.f., Guvenen 2007), which is low for two out of the three ability pairs we study later and high between the other pair (see Table 1). Finally, we also assume that a worker's occupation and its skill requirements do not influence the quality of signal a worker gets about his ability. These assumptions allow us to derive the analytical results.

PROPOSITION 1: *The optimal solution to the worker's problem is characterized by the following two functions*:

(*i*) *occupational choice*: $r_{j,t} = \hat{A}_{j,t}$;

(*ii*) *value function*:

$$V_t(\mathbf{h}_t, \hat{\mathbf{A}}_t) = \left(\sum_{s=t}^{T} \beta^{s-t}\right)\left(\sum_{j=1}^{n}\left(h_{j,t} + \hat{A}_{j,t}^2/2\right)\right) + B_t(\hat{\mathbf{A}}_t),$$

where B_t is a known time-varying function that does not affect the worker's choices.[14]

Three remarks about this solution are in order. First, given that $r_{j,t} = \hat{A}_{j,t}$, skill *mismatch* in dimension j can be written either as $(A_j - \hat{A}_{j,t})^2$ or $(A_j - r_{j,t})^2$. In the empirical section, we use the worker's test scores that proxy A_js and his occupation's skill intensities that correspond to $r_{j,t}$s in order to construct our mismatch measure. Second, since A_js enter into the worker's objective function linearly, the solution only depends on the worker's expectation of A_j, which is $\hat{A}_{j,t}$. Third, the worker's human capital and wage depend both on his belief $\hat{A}_{j,t}$ and also his true ability A_j and the shock $\varepsilon_{j,t}$. Thus, his realized wage and human capital will be different from his own expectations of these two variables.

Relationship to the Standard Human Capital Framework.—Before moving further, it is instructive to point out three key ways our model differs from the standard Ben-Porath formulation. The first two should be obvious by now. First, we introduce multidimensional skills and abilities; and second, here, skill accumulation depends not only on a worker's learning abilities (A_j) but also on his occupation. This latter feature is in the spirit of Rosen (1972).

Third, and less obvious, the incentives that drive skill acquisition are quite different here compared to the Ben-Porath framework. To see this, notice that in the Ben-Porath model (assuming perfect information and one type of skill, and interpreting r_t as human capital investment), the current wage would be $w_t = h_t - r_t^2/2$, and the next period's human capital would be $h_{t+1} = h_t + Ar_t$. Thus, the choice of r_t that maximizes the current wage is zero, whereas the one that maximizes future human capital is infinite, so there is a strong intertemporal trade-off that determines the choice of r_t. In contrast, in our model, the analogs of the two equations are $w_t = h_t + (Ar_t - r_t^2/2)$ and $h_{t+1} = h_t + \left(Ar_t - r_t^2/2\right)$, where both equations have the common term $Ar_t - r_t^2/2$. Because of this symmetry, the same interior choice of r_t maximizes both current wage and future human capital; the intertemporal trade-off disappears.[15] This makes our setup more similar to a learning-by-doing model where more work today raises both the current wage and future human capital.

Fourth, in our framework, both overqualified and under-qualified workers (i.e., those with $r_j \neq A$) experience lower rates of skill accumulation, which is clear from the law of motion for h_{t+1}. In contrast, in the Ben-Porath version, if a worker

[14] The proof of this proposition for the general case with positive depreciation is in online Appendix A.
[15] We show this result more formally in online Appendix A.2.

is under-qualified today (i.e., $A < r_j$), he would earn lower wages today but higher wages in the future because he over-accumulates skills (compared to his optimal occupation) in the higher r occupation (recall $h_{t+1} = h_t + Ar_t$). Consequently, in a wage regression, negative *past* mismatch (reflecting the extent to which a worker was under-qualified in the past) will have a positive effect on the current wage. In contrast, in our setup both positive and negative past mismatches reduce current wages. Our empirical findings in Table 6 show that both positive and negative past mismatches reduce current wages, providing support for our specification.

Two Issues with Implementation.—Before concluding this section, we discuss two practical issues that arise when the wage equation (7) is estimated. The first one is that the error term in (7) is correlated with mismatch measures because the mean of a worker's posterior beliefs about his abilities is correlated with past shocks.[16] As a result, our estimates of the coefficient on mismatch will be biased. Fortunately, as we state formally in Lemma 1, we can determine the sign of the bias, which will become useful in interpreting our results.

Specifically, first, current mismatch and past shocks are positively correlated. If we observe a high (low) wage in a period due to positive (negative) shocks, we will also observe a high (low) mismatch. Because we expect wages and mismatch to be negatively correlated and wages and errors to be positively correlated, the direction of the bias will be toward zero and the true effect of mismatch on wages should be stronger than the effect we estimate in our empirical analysis.[17] The following lemma summarizes this result.

LEMMA 1: *Let* $M_{j,t} \equiv \sum_{s=1}^{t}(A_j - r_{j,s})^2$ *and* $\Omega_{j,t} \equiv \sum_{s=1}^{t} r_{j,s}\varepsilon_{j,s}$. *Then,* $\operatorname{cov}(M_{j,t}, \Omega_{j,t}) > 0$. *Therefore, the estimated coefficient of mismatch provides a lower bound (in magnitude) for the true effect.*

Another issue that concerns the empirical estimation of the wage equation is that we do not directly observe A_js. Instead, we will use workers' ASVAB test scores, which are noisy signals about their true abilities. To illustrate how this might affect our estimates, let $\tilde{A}_j \equiv A_j + \nu_j$, where $\nu_j \sim \mathcal{N}(0, \sigma_{\nu_j}^2)$, denote the test scores. To see how using $\tilde{A}_j$ instead of A_j in the estimation affects our results, insert $A_j = \tilde{A}_j - \nu_j$ into (6), which gives

$$w_t = \sum_j h_{j,1} + \frac{1}{2}\sum_j \tilde{A}_j^2 \times t - \frac{1}{2}\sum_{s=1}^{t}\sum_j \left(\tilde{A}_j - r_{j,s}\right)^2 + \sum_{s=1}^{t}\sum_j r_{j,s}\left(\varepsilon_{j,s} - \nu_j\right).$$

In the following lemma, we show that estimating this equation delivers estimates of the coefficients on both the ability term and mismatch that are biased toward zero.

[16] This is a direct implication of Bayesian learning and can be seen by repeatedly substituting equation (8) backward.

[17] "True" mismatch effect is the one when workers are randomly assigned to occupations and therefore mismatch happens exogenously.

LEMMA 2 (Measurement Error and Attenuation Bias): *Let* $\tilde{\Delta}_{j,t} = \tilde{A}_j^2 \times t$ *and* $\tilde{M}_{j,t} \equiv \sum_{s=1}^{t} (\tilde{A}_j - r_{j,s})^2$, $\tilde{\Omega}_{j,t} \equiv \sum_{s=1}^{t} r_{j,s}(\varepsilon_{j,s} - \nu_j)$. *Then*:

(*i*) $\text{cov}\left(\tilde{\Delta}_{j,t}, \tilde{\Omega}_{j,t}\right) < 0$: *therefore*, *the estimated coefficient of ability-experience interaction provides a lower bound for the true effect.*

(*ii*) $\text{cov}\left(\tilde{M}_{j,t}, \tilde{\Omega}_{j,t}\right) > 0$: *therefore*, *the estimated coefficient of mismatch provides a lower bound for the true effect.*

Lemmas 1 and 2 establish that the coefficients we obtain in the empirical analysis provide *lower bounds* on the effects of mismatch on wages.

E. *Occupational Switching*

We now turn to workers' occupational switching decisions and how they relate to past and current mismatch. Note that workers' beliefs are unbiased at any point in time, so mean beliefs over the population are equal to mean abilities. However, each worker will typically over- or underestimate his abilities in a given period. Over time, beliefs will become more precise and converge to his true abilities. Thus, workers choose occupations with which they are better matched and mismatch declines. The following lemma formalizes this simple result.

LEMMA 3 (Mismatch by Labor Market Experience): *Average mismatch is given by* $E\left[(A_j - r_{j,t})^2\right] = 1/\lambda_{j,t}$. *Since the precision* $\lambda_{j,t}$ *increases with labor market experience*, *average mismatch declines with experience.*

The occupational switching decision is closely linked to mismatch. To illustrate this point, assume that an occupational switch occurs if a worker chooses an occupation whose skill intensities fall outside a certain neighborhood of the skill intensities of his previous occupation in at least one skill dimension. More formally, letting $\kappa_j > 0$ be a positive number, an occupational switch occurs in period t if $r_{j,t} > r_{j,t-1} + \kappa_j$ or $r_{j,t} < r_{j,t-1} - \kappa_j$ for some j. The following two propositions characterize the patterns of occupational switches.

PROPOSITION 2 (Probability of Occupational Switching): *The probability of occupation switching increases with current mismatch and declines with age.*

Mismatch is higher when the mean of a worker's belief is further away from his true ability. In that case, conditional on labor market experience, each observation causes a bigger update of the mean of a worker's belief. Because occupational switching is related to the change in the mean belief, the probability of switching increases with mismatch.[18] Moreover, conditional on mismatch, if the precision of beliefs is higher, the probability of switching occupations will be lower since each

[18] Since $r_{j,t} = \hat{A}_{j,t}$, notice that occupational switch would occur if $\hat{A}_{j,t} - \hat{A}_{j,t-1} > \kappa_j$ or $\hat{A}_{j,t} - \hat{A}_{j,t-1} < -\kappa_j$ for some j.

observation will update the belief by a smaller amount. Since the precision of beliefs increases and the worker's occupational choice converges to his ideal occupation with experience (i.e., mismatch declines), the probability of switching occupation declines. In Table 7, we show empirical evidence in support of this prediction.

We now turn to some key predictions of our model for the "direction" of occupational switching that we test in the data in Section IVC. In particular, we can show that occupational switches tend to be in the direction of reducing existing mismatch. That is, workers who are overqualified (positive mismatch) in a certain skill j will, on average, switch to an occupation with a higher requirement of skill j, thereby reducing the amount by which they are overqualified. And the opposite applies for skill dimensions along which they are negatively mismatched (under-qualified). The following proposition formalizes this result.

Introducing some notation, let $\pi_{j,t}^{up} \equiv \Pr\left(r_{j,t+1} - r_{j,t} > \kappa_j\right)$ denote the probability that a worker's occupation next period will have skill requirement j that is higher than his current occupation. We refer to this as "moving up." Similarly, define the probability of moving down: $\pi_{j,t}^{down} \equiv \Pr\left(r_{j,t+1} - r_{j,t} < -\kappa_j\right)$.

LEMMA 4 (Direction of Occupational Switches): *If the worker is overqualified in skill j, that is, $A_j - r_{j,t} > 0$, then:*

(*i*) *the probability of moving up in skill j is larger than the probability of moving down: $\pi_{j,t}^{up} > \pi_{j,t}^{down}$, and*

(*ii*) *the probability of moving up in skill j increases with the extent of overqualification: $\partial \pi_{j,t}^{up} / \partial\left(A_j - r_{j,t-1}\right) > 0$.*

A worker has positive mismatch (is overqualified) for his occupation in skill dimension j if he chose an occupation with a lower skill-j intensity than his ability. This would happen if he underestimates his ability in dimension j. For such a worker, a new observation, on average, increases his expectations of his ability, and as a result, he becomes more likely to switch to an occupation with a higher skill-j intensity. While the proposition is stated in terms of upward mobility of overqualified workers, the opposite is also true: under-qualified workers are more likely to move to occupations with lower skill intensities.

In addition to predicting the likelihood of switching, the model also delivers a prediction for the *size* of the change in the skill space *conditional* on switching, and how it varies with the level of mismatch in the current occupation. The following proposition states the result.

PROPOSITION 3: *Conditional on switching occupations, the reduction in mismatch in skill j is proportional to the level of mismatch in the previous occupation: $r_{j,t} - r_{j,t-1} \sim N\left(\left(\lambda_{\epsilon_j}/\lambda_{j,t}\right)\left(A_j - r_{j,t-1}\right), \left(\lambda_{\epsilon_j}/\lambda_{j,t}\right)^2 \sigma_{\epsilon_j}^2\right)$.*

In addition to providing a clear testable prediction (that we will test in IVC), this proposition also provides a way to distinguish information frictions from search frictions as the source of mismatch. More concretely, with undirected search (e.g., as in Lise and Postel-Vinay 2016), an overqualified worker will switch to occupations that require higher skills as he samples more occupations, *but* conditional on switching the reduction in mismatch will be independent of his initial level of mismatch. Thus, in our model the average change in skills upon switching is an increasing function of initial mismatch, whereas with undirected search it is a flat line. We will revisit this discussion in the empirical analysis.

II. Data

The main source of data for our analysis is the NLSY79, which tracks a nationally representative sample of individuals who were 14–22 years of age on January 1, 1979. It contains detailed information on earnings, employment, and occupational titles for each job, of each worker. In addition, all respondents took the ASVAB test at the start of the survey. The respondents were also given a behavioral test to elicit their social attitudes (e.g., self-esteem, willingness to engage with others, etc.).

Using these test scores, we construct three ability measures for each worker: *verbal*, *math*, and *social*. We then link these ability measures to the skill requirements of his main occupation (expressed in a way that is comparable to these abilities). The latter are constructed from O*NET data on occupations (as described later). Combining these two pieces of information allows us to create a measure of skill mismatch for each worker-occupation pair. We now provide a brief description of the data sources and methods, and relegate further details of sample selection, variable construction, and sample statistics to online Appendix C.

A. *NLSY*79

We use the Work History Data File of the NLSY79 to construct yearly panels from 1978–2010, providing up to 33 years of labor market information for each individual. We restrict our analysis to males and focus on the nationally representative sample of 3,003 individuals. We exclude individuals who were already working when the sample began so as to avoid the left truncation in their employment history. Such truncation would pose problems for our empirical measures, which require the complete work history to be recorded for each individual. We further drop individuals that are weakly attached to the labor force. The complete description of our sample selection is in online Appendix C. Our final sample runs from 1978–2010 and includes 1,992 individuals and 44,591 individual-year observations.

Measurement error in occupational switching has received particular attention, and we address it by dropping transitions that immediately revert (which often indicates incorrect coding in the middle year) and conditioning occupation switches on simultaneous employer switches. The latter condition corresponds to one of the filters used by Kambourov and Manovskii (2009), which Carrillo-Tudela and Visschers (2014) shows is quite important because the majority of miscoded occupational switches are within employers. Annual occupational mobility in our sample is 15.94 percent compared with 18.48 percent reported in Kambourov and

Manovskii (2008) who use the Panel Study of Income Dynamics (PSID) for the period 1968–1997. Online Appendix C contains further details.

Data on Workers' Abilities.—The version of the ASVAB taken by NLSY79 respondents had 10 component tests.[19] We focus on the following four components on verbal and math abilities that can be linked to skill counterparts: word knowledge, paragraph comprehension, arithmetic reasoning, and mathematics knowledge. We followed Altonji, Bharadwaj, and Lange (2012) to process the ASVAB scores. In particular, because age differences can have a systematic effect on the score and respondents ranged from age 14–21, we normalize the mean and variance of each test score by their age-specific values following these authors.

The NLSY79 included three attitudinal scales that measure a respondent's noncognitive abilities. We focus on two of these measures: the Rotter locus of lontrol scale and Rosenberg self-esteem scale. Both were administered early in the sample, 1979 and 1980, respectively. The Rosenberg scale measures a respondent's feelings about oneself, his self-worth and satisfaction. The Rotter scale elicits a respondent's feelings about his autonomy in the world, the primacy of his self-determination rather than chance. These scores were previously used by Heckman, Stixrud, and Urzua (2006), and Bowles, Gintis, and Osborne (2001) reviews evidence on the effects of noncognitive abilities on earnings. Just as with the ASVAB scores, we equalized the mean and variance across ages. We call this dimension of a noncognitive ability *social* ability hereafter.

Occupational Skill Requirements.—The Department of Labor's O*NET project aims to characterize the mix of knowledge, skills, and abilities that are used to perform the tasks that make up an occupation. It includes information on 974 occupations that can be mapped into the 292 occupation categories included in the NLSY79. For each of these occupations, analysts at O*NET give a score for the importance of each of 277 descriptors.[20] We use 26 of these descriptors that are most related to the ASVAB component tests, a choice dictated by our measures that relate ASVAB to O*NET and described below, and another 6 descriptors related to the social skills. The complete list is in online Appendix C.2.

B. *Creating Verbal*, *Math*, *and Social Components*

Information about workers' abilities and occupational skill requirements in verbal and math fields are aggregated in two steps. First, we convert the O*NET skills into four ASVAB test categories using the crosswalk created by the Defense Manpower Data Center (DMDC).[21] The DMDC selected 26 O*NET descriptors that were

[19] These 10 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.

[20] We use the analysts database, version 4.0, which does not include data from survey respondents and should yield a more consistent picture across occupations.

[21] To increase the ASVAB's general appeal, the ASVAB Career Exploration Program was established by the US Department of Defense to provide career guidance to high school students. As part of the program, they created a mapping between ASVAB test scores and O*NET occupation requirements (OCCU-Find).

particularly relevant and assigned each a *relatedness score* to each ASVAB category test. For each ASVAB category test, we create an O*NET analog by summing the 26 descriptors and weighting them by this relatedness score. The result is that each occupation gets a set of scores that are comparable to the ASVAB categories, each a weighted average of the 26 original O*NET descriptors.

Second, after normalizing each dimension's standard deviation to be 1, we reduce these four ASVAB categories into two composite dimensions, verbal and math, by applying principal component analysis (PCA). The verbal score is the first principle component of word knowledge and paragraph comprehension, and the math score is that of math knowledge and arithmetic reasoning. Because the scale of these principal components is somewhat arbitrary, we convert all four scores (verbal worker ability, math worker ability, verbal occupation requirement, math requirement) into percentile ranks among individuals or occupations.[22]

Likewise, we create a single index of social ability both on workers' and occupations' sides. From the O*NET, we reduce the six O*NET descriptors to a single dimension by taking the first principal component after scaling each dimension's standard deviation to be 1. For the worker's side, we first take the negative of the Rotter scale, because a lower score implies more feeling of self-determination. After scaling both NLSY79 measures to have a standard deviation of 1, we take the first principal component. Both occupation- and worker-side data are then converted into percentile rank scores.

Table 1 reports (a) the correlation between workers' verbal, math, and social ability scores for 1,992 individuals in our sample and (b) the correlation between workers' abilities and the corresponding skill requirements in their current occupation for 44,591 observations in our sample. In panel A, the highest correlation between ability scores is 0.78 (math and verbal), and the others are 0.27 and 0.30, indicating that each component carries useful independent information. Panel B provides a rough measure of the sorting of workers into occupations by skills. While the correlations are fairly low (reflecting both mismatch but also measurement error) they are all positive. Workers with strong math skills tend to sort into occupations with generally high skill requirements. A worker's social skills have a relatively low correlation with occupation requirements along every dimension.

III. Empirical Methodology

In this section, we introduce our empirical measure of (contemporaneous) skill mismatch as well as *cumulative mismatch*, to analyze the persistent effects of past mismatch on current wages. We also present two additional statistics, called *positive* and *negative mismatch*, to analyze the effects of over- and under-qualification at a given occupation. These measures are then incorporated into the wage equation developed in Section I and estimated in the next section.

[22] The rank scores of skills among occupations are calculated by weighting each occupation by the number of observations of individuals in that occupation in NLSY79.

Table 1—Correlations among Ability and Skill Requirement Scores

	Panel A. Worker ability			*Panel B. Occupational skill requirement*		
Workers' ability	Verbal	Math	Social	Verbal	Math	Social
Verbal	1.00			0.37	0.34	0.35
Math	0.78	1.00		0.44	0.40	0.35
Social	0.30	0.27	1.00	0.13	0.11	0.16

Note: The correlations reported in panel A are computed using data on 1,992 individuals in our sample, and those in panel B are computed using 44,591 worker-occupation pairs observed during the sample period.

A. *An Empirical Measure of Skill Mismatch*

(*Contemporaneous*) *Mismatch*.—Specifically, as in Section IB, $\tilde{A}_{i,j}$ is the *measured* ability of individual i in skill dimension j, and $\tilde{r}_{c,j}$ is the *measured* skill requirement of occupation (or *c*areer) c in the same dimension. Let $q\left(\tilde{A}_{i,j}\right)$ and $q\left(\tilde{r}_{c,j}\right)$ denote the corresponding percentile ranks of the worker ability and the occupation skill requirements. To define our measure, we take the difference in each skill dimension j between worker abilities and occupational requirements. We sum the absolute value of each of these differences using weights $\{\omega_j\}$ to obtain[23]

$$m_{i,c} \equiv \sum_{j=1}^{n} \left\{ \omega_j \times \left| q\left(\tilde{A}_{i,j}\right) - q\left(\tilde{r}_{c,j}\right) \right| \right\}.$$

The weights are chosen to be the factor loadings from the first principal component, normalized to sum to 1.[24] To help understand magnitudes in our analysis, we rescale our mismatch measure so that its standard deviation is equal to 1. Table B.1 in online Appendix B shows descriptive statistics for the mismatch measure that reveal that the prevalence of mismatch is not specific to a particular educational group, race, or industry.

Cumulative Past Mismatch.—A key idea that we will explore in this paper is whether a poor match between a worker and his current occupation can have persistent effects that last beyond the current job. To this end, we construct a measure of *cumulative* mismatch as follows. Consider a worker who has worked at p different occupations as of period t, whose indices are given

[23] We use absolute deviations instead of another metric like quadratic deviations, as would be suggested by the quadratic mismatch terms that appear in equation (7). This is because our measures $q\left(\tilde{A}_{i,j}\right)$ and $q\left(\tilde{r}_{c,j}\right)$ are ordinal rather than cardinal. In Section I, we derived the quadratic form knowing the cardinal values. Given that we can only measure ranks, absolute deviations are the more robust measure of distance. Having said this, we also tried with quadratic distance, and the results were substantively unchanged.

[24] That is, we apply principal component analysis (PCA) to the set of absolute values of differences, $\left\{ \left| q(\tilde{A}_{i,j}) - q(\tilde{r}_{c,j}) \right| \right\}_{j=1}^{n}$, and obtain the first principle component. The weights for the first principle component through PCA turned out to be (verbal, math, social) $=$ (0.43, 0.43, 0.12). We do not know a priori the relative importance of each skills dimension to wages, which could have been a preferable basis for weighting. However, our results were little changed when we used other reasonable weights, like the one that sets an equal weight for all dimensions.

by the vector $\{c(1), c(2), \ldots, c(p)\}$. The tenure in each of these matches is given by the vector $\{\hat{T}_{c(1)}, \hat{T}_{c(2)}, \ldots, \hat{T}_{c(p-1)}, T_{c(p),t}\}$, where $\hat{T}_{c(s)}$ denotes total tenure in the past occupation $c(s)$, and $T_{c(p),t}$ is the tenure in the current occupation at period t. These must add up to the total experience of the worker at period t: $\hat{T}_{c(1)} + \hat{T}_{c(2)} + \cdots + \hat{T}_{c(p-1)} + T_{c(p),t} = E_t$. Cumulative mismatch is defined as the *average* mismatch in the $p-1$ previous occupations:

$$(9)\ \bar{m}_{i,t} \equiv \frac{m_{i,c(1)}\hat{T}_{c(1)} + m_{i,c(2)}\hat{T}_{c(2)} + \cdots + m_{i,c(p-1)}\hat{T}_{c(p-1)}}{\hat{T}_{c(1)} + \hat{T}_{c(2)} + \cdots + \hat{T}_{c(p-1)}} = \frac{\sum_{s=1}^{p-1} m_{i,c(s)}\hat{T}_{c(s)}}{\sum_{s=1}^{p-1}\hat{T}_{c(s)}}.$$

Each past mismatch value is weighted by its corresponding $\hat{T}_{c(s)}$, so the duration the worker was exposed to an occupation determines its influence on average. This variable is the empirical analog of the cumulative mismatch term in equation (7) and represents the lingering effect of previous mismatches on the current wage. If occupational match quality only had an effect within a given match (as in, e.g., Jovanovic 1979 or Mortensen and Pissarides 1994), this variable would have no effect on later wages. On the other hand, if dynamic decisions, such as human capital accumulation, are important and mismatch depresses them, then poor matches in past occupations can reduce current wages.

Positive versus Negative Mismatch.—Equation (7) in Section I tells us mismatch may reduce a worker's wages for two reasons: a worker's ability may exceed the occupational requirement, and/or his ability does not meet the occupational requirement. To analyze these positive and negative effects of mismatch separately, we introduce two additional measures. We call them *positive mismatch* and *negative mismatch*, which are defined as

$$m_{i,c}^{+} \equiv \sum_{j=1}^{n}\omega_j \max\left[q\left(\tilde{A}_{i,j}\right) - q\left(\tilde{r}_{c,j}\right), 0\right] \text{ and } m_{i,c}^{-} \equiv \sum_{j=1}^{n}\omega_j \min\left[q\left(\tilde{A}_{i,j}\right) - q\left(\tilde{r}_{c,j}\right), 0\right],$$

respectively. These definitions mean that $m_{i,c} = m_{i,c}^{+} + \left(-m_{i,c}^{-}\right)$. That is, we decompose our mismatch measure into a part where some of the worker's abilities are overqualified (positive mismatch) and a part where some of them are under-qualified (negative mismatch). We can also define *positive cumulative mismatch* and *negative cumulative mismatch* based on these two measures by applying the definition of cumulative mismatch in Section IIIA.

B. *Empirical Specification of the Wage Equation*

Based on our theory in Section I, we augment the standard Mincer wage regression with measures of mismatch to investigate whether current or cumulative mismatch (or both) matters for current wages. To the extent that current mismatch matters for the level of wages, it can be viewed as a useful, measurable proxy for occupational match quality, which has been treated as an unobservable component of the regression residual by much of the extant literature.[25] Furthermore, if cumulative mismatch or the interaction between match quality and tenure turns out to matter for current wages, then this would provide evidence that match quality affects human capital accumulation and life-cycle wage dynamics.

We will estimate the empirical analog of the wage equation (7) with a few modifications to consider the wage equation for individual i who is working with employer l in occupation c at time t:

$$\ln w_{i,l,c,t} = \underbrace{\gamma_1 m_{i,c}}_{\text{current mismatch}} + \underbrace{\gamma_2 \left(m_{i,c} \times T_{i,c,t}\right)}_{\text{current mismatch} \times \text{tenure}} + \underbrace{\gamma_3 \overline{m}_{i,t}}_{\text{cumulative past mismatch}} \\ + \gamma_4 \bar{A}_i + \gamma_5 \left(\bar{A}_i \times T_{i,c,t}\right) + \gamma_6 \bar{r}_c + \gamma_7 \left(\bar{r}_c \times T_{i,c,t}\right) \\ + \Phi_1\left(J_{i,l,t}\right) + \Phi_2\left(T_{i,c,t}\right) + \Phi_3\left(E_{i,t}\right) + \alpha_4 OJ_{i,t} + X'_{i,t}\beta + \theta_{i,l,c,t}. \tag{10}$$

In the above equation, we have our mismatch measure, $m_{i,c}$, and its interaction term with occupational tenure, $m_{i,c} \times T_{i,c,t}$. We also include our cumulative mismatch measure, $\overline{m}_{i,t}$, to capture the potential scarring effects of previous mismatches on current wages. This form comes from equation (7) in our model, which suggests that there should be dynamic effects from mismatch.

In the second line of equation (10), $\bar{A}_i$ is the ability of worker i averaged across skill dimensions, and $\bar{r}_c$ is the skill requirement of occupation c averaged over skill dimensions.[26] We also include their interactions with occupational tenure. These variables are important to include because we might worry that our match quality measures are just proxies for an individual effect from worker or occupation. Equation (7) would suggest that we include $\bar{A}_i \times E_{i,t}$ instead of $\bar{A}_i \times T_{i,c,t}$. We have also estimated the equation including $\bar{A}_i \times E_{i,t}$ term instead and obtained very similar results (see online Appendix J.5). We have chosen to present the results with $\bar{A}_i \times T_{i,c,t}$ in the regression as our baseline because we are principally concerned with the coefficient on $m_{i,c} \times T_{i,c,t}$ and, by including $\bar{A}_i \times T_{i,c,t}$, we want to convince readers that the mismatch terms are not simply capturing ability, as $m_{i,c}$ is correlated with $\bar{A}_i$ by construction.

In the last line of equation (10), we have employer tenure, $J_{i,l,t}$, occupational tenure, $T_{i,c,t}$, labor market experience, $E_{i,t}$, and a dummy variable that indicates a continuing job, $OJ_{i,t}$, where Φ_1, Φ_2, and Φ_3 denote polynomials.[27] Finally, when estimating equation (10), we include one-digit level occupation and industry dummies and a vector of education and demographic characteristics, $X_{i,t}$.

The last term, $\theta_{i,l,c,t}$, corresponds to the accumulated informational noise in the model, and as shown above in Lemma 1, the correlation of this term with mismatch biases the estimated coefficient on mismatch terms toward zero. To the extent that

[25] See, for example, Altonji and Shakotko (1987), Topel (1991), Altonji and Williams (2005), Kambourov and Manovskii (2009).

[26] More precisely, $\bar{A}_i$ is the average of the percentile rank scores of the measured worker's abilities, $\left\{q\left(\tilde{A}_{i,j}\right)\right\}_{j=1}^{n}$, and $\bar{r}_c$ is that of the measured occupational requirements, $\left\{q\left(\tilde{r}_{c,j}\right)\right\}_{j=1}^{n}$. Both $\bar{A}_i$ and $\bar{r}_c$ are again converted into percentile rank scores among individuals or among occupations.

[27] We use a second-order polynomial for $\Phi_1(\cdot)$ and third-order polynomials for $\Phi_2(\cdot)$ and $\Phi_3(\cdot)$.

this correlation is large, the true effect of mismatch on wages will be larger than our estimates in the next section indicate.

In addition, $\theta_{i,l,c,t}$ may also include unobserved individual- and match-specific factors and a serial correlation structure. To deal with serial correlation, we allow for an AR(1) structure for the residual, $\theta_{i,l,c,t} = \rho_\theta \theta_{i,l,c,t-1} + \varepsilon_{i,l,c,t}$.[28] Although this structure is not as general as an unrestricted correlation structure, it provides large gains in efficiency (see Cameron and Miller 2015 for a detailed discussion of this point). We will conduct sensitivity analyses allowing for more general error structures later.

Instrumenting Tenure Variables.—As Proposition 2 established, the probability of occupational switching is a function of match quality. Consequently, occupational tenure is endogenous in a wage regression because both tenure and wages are functions of match quality.[29] This particular endogeneity problem is well recognized in the literature, going back to Altonji and Shakotko (1987), who proposed and implemented an instrumental variables estimator to avoid the resulting bias. To be clear, if our empirical mismatch measure captured all relevant aspects of mismatch as well as unobservable individual heterogeneity that mattered for occupational tenure, including this mismatch measure, which we do, would take care of the endogeneity problem. But clearly, our measure is a proxy and very likely misses some aspects of mismatch that matter for both wages and tenure, so to preempt an endogeneity bias, we follow Altonji and Shakotko's (1987) IV approach to instrument for experience and tenure variables. This method has been used extensively in the previous literature (see, among others, Topel 1991 and Altonji and Williams 2005). Because our regressors also include variables that interact with tenure, we create a corresponding instrument replacing tenure with its instrument. Employer tenure, labor market experience, and the dummy variable for a continuing job are also instrumented in the same manner.[30]

C. *Workers' Information Set*

Before concluding this section, it is important to discuss why workers in our NLSY sample might be uncertain about their abilities, as assumed in our model, even after they have taken the ASVAB, Rotter, and Rosenberg tests. There are at least two reasons for this uncertainty. First, and most importantly, the NLSY respondents were not told their rank in the test but were rather given a relatively broad range where their score landed. For example, a respondent knew he scored 10 out of 25 on mathematics knowledge, but was only told that his score corresponded to a rank between the twentieth and fortieth percentiles. Just as in our theoretical model, this is a noisy signal

[28] This assumption is common in the income dynamics literature and was also used in Kambourov and Manovskii (2009).

[29] This endogeneity bias is separate from the bias discussed in Lemma 1 that showed how the mismatch term was correlated with errors.

[30] Similar to an occupational match component, an employer match component is potentially correlated with employer tenure and the dummy variable of a continuation of a job. An individual-specific component is potentially correlated with labor market experience.

centered around the true mean. As the econometrician, we see the entire NLSY79 sample, so we can compute the worker's precise rank.

Second, as the econometrician, we can process these test scores to extract more information than what the respondents could. For example, we removed age effects from the test scores, which affects the scores the respondents see but is probably not economically relevant. Similarly, by taking the first principal component from several related tests, we are, statistically speaking, uncovering the underlying ability from several tests that are individually noisy measures. Not knowing the population-level correlations, the respondents could not possibly do the same analysis.

IV. Empirical Results

In this section, we discuss the empirical evidence using our mismatch measures. We will first relate mismatch to wages by incorporating it into a Mincer regression and then study its relationship to switching probability and the direction of switching. We find that the mismatch-tenure interaction and cumulative past mismatch are quite important in the determination of wages. Mismatch also increases the probability of a switch, and once one does switch, it predicts whether a worker will move up or down in the skills required by his occupation. Furthermore, the reduction in mismatch upon switching is proportional to the level of mismatch in the previous occupation.

A. *Mismatch and Wages*

Equation (7) of our model suggested a direct link between mismatch, the history of mismatch, and wages. With this motivation in mind, we operationalized it in the regression in equation (10). Table 2 presents the main coefficients from these wage regressions and the rest are in online Appendix B. The first column includes our measure of mismatch into a standard wage regression. The next adds its interaction with occupational tenure. The full baseline specification is in the third column, where we introduce our measure of cumulative mismatch. As we discussed in the previous section, we instrument all the tenure variables in the columns labeled "IV-GLS" and show the results without instrumenting in the "GLS" columns. In columns 3 and 6, there are fewer observations because estimating cumulative mismatch requires that the worker held at least one previous occupation.

In column 1 of Table 2, contemporaneous mismatch has an estimated coefficient of -0.0205 (and is significant at the 1 percent level), indicating a strong effect on wages. To give an economic interpretation to this coefficient, recall that we have normalized the standard deviation of mismatch to 1, so wages are predicted to be about 4.1 percent ($2.05\% \times 2$) lower for workers whose mismatch is 1 standard deviation above the mean relative to those 1 standard deviation below it.

In the next column, we introduce a mismatch interaction with occupational tenure. Now part of the level effect is replaced by a negative tenure effect (also significant at the 1 percent level). Mismatch depresses initial wages and also leads to slower wage growth over the duration of the match. After five years, the overall depression in wages due to slower growth exceeds the losses due to the initial impact.

TABLE 2—WAGE REGRESSIONS WITH MISMATCH

	IV-GLS			GLS		
	(1)	(2)	(3)	(4)	(5)	(6)
Mismatch	−0.0205	−0.0091	−0.0054	−0.0196	−0.0136	−0.0128
	(−6.69)	(−1.76)	(−0.88)	(−6.44)	(−3.27)	(−2.56)
Mismatch × Occ Tenure		−0.0020	−0.0022		−0.0010	−0.0007
		(−2.75)	(−2.28)		(−2.08)	(−1.13)
Cumulative Mismatch			−0.0374			−0.0383
			(−8.73)			(−9.10)
Worker Ability (mean)	0.3015	0.3044	0.4095	0.3094	0.3101	0.3987
	(11.95)	(12.04)	(13.50)	(14.90)	(14.94)	(15.83)
Worker Ability × Occ Tenure	0.0115	0.0109	0.0072	0.0090	0.0088	0.0081
	(4.16)	(3.93)	(2.00)	(4.62)	(4.50)	(3.17)
Occ Reqs (mean)	0.0977	0.0962	0.0878	0.1383	0.1374	0.1426
	(4.50)	(4.43)	(3.46)	(7.35)	(7.29)	(6.45)
Occ Reqs × Occ Tenure	0.0144	0.0144	0.0182	0.0084	0.0084	0.0092
	(5.41)	(5.41)	(5.37)	(4.51)	(4.53)	(3.89)
Observations	41,596	41,596	30,599	41,596	41,596	30,599

Notes: All regressions include (*standard controls*) a constant, terms for demographics, occupational tenure, employer tenure, work experience, and dummies for one-digit-level occupation and industry; t-statistics are in parentheses and calculated with standard errors that are estimated via GLS assuming AR(1) autocorrelation. More detailed regression results are in Table B.2 in online Appendix B.

TABLE 3—WAGE LOSSES FROM MISMATCH AND CUMULATIVE MISMATCH

Mismatch degree	*Mismatch* effect			*Cumulative mismatch* effect
(High to low)	5 years	10 years	15 years	
90 percent	−0.049	−0.081	−0.114	−0.123
	(0.011)	(0.018)	(0.030)	(0.014)
50 percent	−0.021	−0.036	−0.050	−0.065
	(0.005)	(0.008)	(0.013)	(0.007)
10 percent	−0.008	−0.013	−0.018	−0.029
	(0.002)	(0.003)	(0.005)	(0.003)

Notes: Wage losses (relative to the mean wage) are computed for each percentile of each measure in the above table. Standard errors are in parentheses.

In column 3, we introduce cumulative mismatch while keeping all the regressors from column 2. Cumulative mismatch has a significant and negative effect on wages. The tenure effect of mismatch in the current match is unaffected, though the initial level effect becomes smaller and insignificant. To help interpret the size of these coefficients, Table 3 computes the implied wage losses using specification (3). Looking at the effect of current mismatch, we see that the ninetieth percentile worst-matched workers face 8.1 percent lower wages after 10 years of occupational tenure compared with a perfectly matched worker. The difference between the ninetieth percentile and the tenth percentile of *mismatch* is about 4.1 percent after 5 years of occupational tenure and widens to 6.8 percent after 10 years. Comparing the ninetieth percentile to the tenth percentile of *cumulative mismatch*, we see a wage difference of 9.4 percent.

For comparison purposes, the last three columns of Table 2 report the GLS estimates of the same specifications in the first three columns. Notice that the coefficient on the mismatch and tenure interaction is quite different between IV-GLS and GLS. As we discussed in Section IIIB, the return to tenure is biased because it is correlated with unobservable match quality. The instruments reduce the return to occupational tenure itself by a factor of about 3 (see Table B.2 in online Appendix B), precisely because the OLS estimate on tenure takes some variation from the mismatch-tenure interaction term. When we instrument tenure, we purge its correlation with match quality so it is instead ascribed to the interaction between mismatch and occupational tenure, making its coefficient larger.

In online Appendix G, we split the sample by education (college and non-college) and reestimate our benchmark wage regression separately for each group. We find that the negative effects of skill mismatch are larger for college graduates. This is true especially of the cumulative mismatch measure, which nearly doubles in magnitude. Among the dimensions, social and verbal have a particularly pronounced effect among this subsample.

Finally, a recent paper by Addison, Chen, and Ozturk (2017) constructs mismatch for the women in the NLSY following the methodology proposed in this paper. While they confirm our key findings, they also highlight some interesting differences between women and men and between the 1979 and 1997 cohorts of women. In particular, highly educated women are more mismatched than men, which explains part of the gender wage gap. They also find that these gender differences in mismatch are lower in the later cohort.

Cost of Mismatch.—We now estimate the overall cost of mismatch. For this purpose, we compute the predicted wages in our full regression and the counterfactual predicted wages if all mismatch terms were set to zero, and we compute the average wages for each case. We find that the overall cost of mismatch is substantial: the average wages would have been 11 percent higher if mismatch could be eliminated. The largest effect on wages comes from cumulative mismatch, followed by *Mismatch* $\times$ *Tenure* interaction; current mismatch has the smallest effect.[31] We should emphasize here that this estimate does not reveal why mismatch exists, but it provides evidence for the lingering effects of mismatch through human capital accumulation.

Table 4 illustrates the cost of mismatch at different percentiles of the predicted and counterfactual wage distributions. The first two columns are in 2002 dollars, and the third is the percent difference. Overall, the cost of mismatch is typically higher at higher percentiles of the wage distribution.

We have also computed the overall cost of mismatch separately for younger and older workers. Young workers are defined as those who are less than 35 years old. For young workers, the average predicted wage is 8.5 log points lower than without

[31] We can provide a back-of-the-envelope calculation to illustrate this point. From column 3 of Table 2, we have the following coefficients: Mismatch: -0.0054; Mismatch $\times$ Occ Tenure: -0.0022; and Cumulative Mismatch: -0.0374. The average value of Mismatch is 1.56 in Table B.1. The average Cumulative Mismatch is 1.9, and the average of Mismatch $\times$ Occ Tenure is 10. Based on these, we can compute the wage loss due to each component: Mismatch $= -0.0054 \times 1.56 = -0.0084$; Mismatch $\times$ Occ Tenure $= -0.0022 \times 10 = -0.0220$; and Cumulative Mismatch $= -0.0374 \times 1.9 = -0.0711$. The total effects is -0.1015.

Table 4—Cost of Mismatch over the Wage Distribution

	No mismatch	Actual	Difference (%)
10 percent	10.43	9.63	8.3
50 percent	16.20	14.51	11.7
90 percent	26.97	24.61	9.7

Note: The table reports wages at different percentiles.

mismatch, compared with a 12.5 log points gap for older workers. The higher cost for the latter group comes mostly from the larger effect of cumulative mismatch for that group, again confirming the important effect of cumulative mismatch for wage loss. The regression results are in online Appendix H.

We could also try to summarize the overall contribution of mismatch to wage growth in a calculation akin to Topel and Ward (1992), which ascribes 40 percent of cumulative wage growth during the first 10 years to job-to-job transitions. We could further break these transitions into those that reduce occupational mismatch and those that do not. We first replicated their exercise in the NLSY 79, using their definition of excess wage growth:

$$\Delta w_E = w_{t+1} - w_{t-2} - \mathbb{E}\left[w_{t+1} - w_t \,|\, \cdot\right] - \mathbb{E}\left[w_{t-1} - w_{t-2} \,|\, \cdot\right],$$

where a job transition occurred between t and $t-1$. The expectations terms, $\mathbb{E}\left[w_{t+1} - w_t\right], \mathbb{E}\left[w_{t-1} - w_{t-2}\right]$, for wage growth are taken with the predicted values from our baseline regression excluding mismatch, skill, and occupational requirements terms. Essentially, excess wage growth at time t is the residual wage growth above that was predicted by the trend.

Overall, our data show essentially the same total excess wage growth as Topel and Ward (1992). Job changes bring an average of 8.5 percent excess wage growth and account for 35 percent of growth during the first 10 years. When we look only at job transitions that also include an occupation switch and divide these switches that increase mismatch from those that decrease mismatch, we find that occupational transitions that decrease mismatch have average excess wage growth of about 12 percent. These switches that reduce mismatch account for 14 percent of the cumulative wage growth in the first 10 years of experience. To put this another way, of the 35 percent of wage growth attributable to all job changes, reductions in mismatch account for 40 percent of that growth.

Three Dimensions of Skill Mismatch.—When presenting the theoretical results of our model, we aggregated over skills, which helped keep it tractable. However, we could create an analogous wage equation by splitting dimensions apart and allowing them to enter equation (7) separately. Here, we consider this notion empirically and look for wage implications dimension-wise. In Table 5, we report the results when we include each component mismatch measure in our regressions. The component mismatch measure in skill j is defined as the difference in the rank scores of ability and occupational requirement, $m_{i,c,j} \equiv \left|q\left(\tilde{A}_{i,j}\right) - q\left(\tilde{r}_{c,j}\right)\right|$. As before, we scale each dimension to have a standard deviation of 1 so that they are comparable.

Table 5—Wage Regressions with Mismatch by Components

	IV-GLS			GLS		
	(1)	(2)	(3)	(4)	(5)	(6)
Mismatch Verbal	−0.0103	0.0082	0.0133	−0.0104	0.0032	0.0051
	(−2.58)	(1.20)	(1.63)	(−2.63)	(0.57)	(0.76)
Mismatch Math	−0.0108	−0.0172	−0.0193	−0.0096	−0.0178	−0.0193
	(−2.71)	(−2.52)	(−2.39)	(−2.46)	(−3.22)	(−2.92)
Mismatch Social	−0.0027	0.0005	0.0055	−0.0021	−0.0012	0.0034
	(−0.86)	(0.10)	(0.87)	(−0.70)	(−0.29)	(0.65)
Mismatch Verbal × Occ Tenure		−0.0031	−0.0043		−0.0022	−0.0028
		(−3.35)	(−3.67)		(−3.46)	(−3.37)
Mismatch Math × Occ Tenure		0.0010	0.0020		0.0013	0.0021
		(1.06)	(1.67)		(2.01)	(2.54)
Mismatch Social × Occ Tenure		−0.0005	−0.0007		−0.0001	−0.0006
		(−0.67)	(−0.75)		(−0.27)	(−0.89)
Cumulative Mismatch Verbal			−0.0148			−0.0137
			(−2.60)			(−2.44)
Cumulative Mismatch Math			−0.0241			−0.0260
			(−4.38)			(−4.81)
Cumulative Mismatch Social			−0.0088			−0.0078
			(−2.05)			(−1.86)
Verbal Ability	−0.0622	−0.0652	0.0029	−0.0132	−0.0181	−0.0006
	(−1.69)	(−1.78)	(0.07)	(−0.44)	(−0.61)	(−0.02)
Math Ability	0.3660	0.3711	0.4260	0.3209	0.3254	0.4027
	(9.86)	(9.99)	(9.63)	(10.60)	(10.74)	(11.09)
Social Ability	0.0997	0.1003	0.1078	0.1046	0.1046	0.1230
	(4.45)	(4.47)	(4.03)	(5.78)	(5.77)	(5.64)
Verbal Ability × Occ Tenure	0.0128	0.0129	0.0083	0.0066	0.0070	0.0083
	(2.96)	(2.97)	(1.51)	(2.10)	(2.23)	(2.08)
Math Ability × Occ Tenure	−0.0051	−0.0060	−0.0066	−0.0016	−0.0024	−0.0045
	(−1.15)	(−1.35)	(−1.22)	(−0.52)	(−0.76)	(−1.13)
Social Ability × Occ Tenure	0.0061	0.0060	0.0071	0.0059	0.0060	0.0056
	(2.30)	(2.25)	(2.09)	(3.12)	(3.12)	(2.27)
Occ Reqs Verbal	0.0190	0.0209	−0.0420	0.0761	0.0832	0.0211
	(0.26)	(0.28)	(−0.46)	(1.25)	(1.36)	(0.28)
Occ Reqs Math	0.1313	0.1244	0.1772	0.1176	0.1080	0.1731
	(1.93)	(1.81)	(2.12)	(2.11)	(1.93)	(2.53)
Occ Reqs Social	−0.1115	−0.1084	−0.1014	−0.1064	−0.1047	−0.0966
	(−4.50)	(−4.37)	(−3.44)	(−5.12)	(−5.03)	(−3.89)
Occ Reqs Verbal × Occ Tenure	−0.0055	−0.0061	0.0034	−0.0184	−0.0202	−0.0123
	(−0.55)	(−0.59)	(0.25)	(−2.64)	(−2.86)	(−1.28)
Occ Reqs Math × Occ Tenure	0.0138	0.0153	0.0077	0.0189	0.0213	0.0120
	(1.50)	(1.62)	(0.62)	(2.96)	(3.26)	(1.38)
Occ Reqs Social × Occ Tenure	0.0095	0.0088	0.0115	0.0112	0.0108	0.0134
	(2.97)	(2.74)	(2.71)	(5.05)	(4.86)	(4.49)
Observations	41,596	41,596	30,599	41,596	41,596	30,599

Notes: All regressions include the same set of standard controls defined in the notes to Table 2; *t*-statistics are in parentheses and calculated with standard errors that are estimated via GLS assuming AR(1) autocorrelation. More detailed regression results are in Table B.3 in online Appendix B.

Looking at math and verbal skills in Table 5, we see a pattern emerge: mismatch in either dimension has a negative effect on wage, but math mismatch reduces the level of wages without a significant growth rate effect, whereas the opposite is true for verbal, which has a small level effect but a strong growth rate effect. The interaction term for verbal mismatch implies a 13 percent wage gap after 10 years of tenure between the top and bottom deciles of mismatch. Math mismatch implies a 5 percent initial wage difference between ninetieth and tenth percentiles that attenuates over time. Social mismatch has a weaker effect overall, though still negative.

Interestingly, the same difference between math and verbal skills is seen in the effects of *ability* on wages (lower panel of Table 5): in the first three columns, math ability has a large level effect (ranging from 37–43 percent across specifications) but little tenure effect, whereas verbal ability has a higher tenure effect (ranging from 0.8–1.3 percent per year) on wages, though the effect of verbal ability on wages is insignificant in our most general specification (column 3). Social ability has both significant level and tenure effects, with the tenure effect exceeding the level effect in two years. One interpretation of these results might be that math ability is easier to observe by employers and the market and so is priced immediately, whereas verbal and social abilities capture some more subtle traits that are revealed more slowly over time, leading to a growth rate effect.[32]

The effect of cumulative mismatch is negative in all three dimensions and statistically significant at the 5 percent level or higher. The cumulative effect for verbal skills is equivalent to about seven years of mismatch in the current occupation, combining the immediate and tenure effects; cumulative social mismatch is equivalent to about 20 years of mismatch in the current occupation, mainly because the effects of current mismatch are small.

Positive and Negative Mismatch.—Next, we investigate the effects of positive and negative mismatch, as defined in Section IIIA, on wages. Recall that, as we discussed in Section ID, in our model both workers who are overqualified and under-qualified experience lower skill accumulation and wages. We also noted that this is not the case in the standard Ben-Porath model, where a worker who is under-qualified at his current occupation (i.e., negative mismatch) earns a lower wage today but a higher wage in the future since his human capital accumulation is still faster than what what would be at his optimal occupation.

To examine these different implications, we rerun our benchmark wage regressions but now split the mismatch term into the positive and negative components. Column 1 of Table 6 presents the simplest version—the analog of column 1 in Table 2: positive and negative mismatch both reduce wages.[33] However, the effect is not symmetric as predicted by our model; the coefficient on negative mismatch is about four times larger than that of positive mismatch. In column 2, we add the interaction terms with

[32] This view is consistent with Altonji and Pierret's (2001) interpretation of public learning about unobserved abilities.

[33] Recall that negative mismatch adds all skill dimensions for which the worker is under-qualified (so by definition it is a negative number), and the positive estimated coefficients imply that negative mismatch reduces wages. Note that we do not include the level of worker abilities or occupational requirements in this regression because once mismatch is split into positive and negative parts, they become highly correlated with those levels.

Table 6—Wage Regressions with Positive and Negative Mismatch

	IV-GLS			GLS		
	(1)	(2)	(3)	(4)	(5)	(6)
Positive Mismatch	−0.0072	0.0087	0.0141	−0.0074	−0.0015	0.0010
	(−5.05)	(−2.05)	(1.49)	(−5.19)	(−3.11)	(0.25)
Negative Mismatch	0.0305	0.0172	0.0231	0.0276	0.0137	0.0195
	(0.93)	(−0.50)	(2.23)	(0.57)	(−0.22)	(2.28)
Positive Mismatch × Occ Tenure		−0.0028	−0.0032		−0.0011	−0.0005
		(−1.24)	(−1.47)		(−0.93)	(0.53)
Negative Mismatch × Occ Tenure		0.0023	0.0014		0.0024	0.0022
		(1.47)	(0.73)		(1.02)	(1.69)
Cumulative Positive Mismatch			−0.0145			−0.0220
			(−13.60)			(−15.37)
Cumulative Negative Mismatch			0.0159			0.0085
			(−2.51)			(−4.35)
Observations	41,596	41,596	30,599	41,596	41,596	30,599

Notes: All regressions include the same set of standard controls defined in the notes to Table 2; t-statistics are in parentheses and calculated with standard errors that are estimated via GLS assuming AR(1) autocorrelation. More detailed regression results are in Table B.4 in online Appendix B.

tenure, which are statistically significant for both types of mismatch. Both positive and negative mismatch-tenure interaction terms have signs supporting our specification.

Finally, in column 3, we add cumulative positive and negative mismatches, which also speaks to the distinction between our model and the Ben-Porath specification. Two points are worth noting. First, both types of cumulative mismatch reduce wages as predicted by our model (and in contrast to the Ben-Porath version). Further, the magnitudes of the two coefficients are quite close to each other (−0.0145 for positive and 0.0159 for negative), which is consistent with the symmetric loss assumed in our model. Second, the effects of contemporaneous positive and negative mismatch are consistent with the model: they both reduce wages when the level and tenure effects are combined. The point estimate of the level of positive mismatch is positive, but it is barely significant (at 10 percent only). Even with that positive coefficient (0.0141), the negative estimated tenure effect (−0.0032) implies that after 5 years of tenure the overall effect turns negative. Overall, the signs and statistical significance (four out of six in column 3) of the estimated coefficients are consistent with the predictions of our model and provide empirical support for the technology for skill accumulation and wages at different occupations as we specified in our model (see equation (1)).

To summarize, the results in Tables 2–6 collectively speak to the importance of skill mismatch for current wages as well as for generating long-lasting scarring effects on future wages. We now turn to address various issues about the sensitivity of these results to various assumptions we made in the benchmark analysis.

B. *Robustness of the Wage Regression Results*

We have conducted various robustness checks. First, to eliminate any concern about errors being serially correlated, we estimated panel-robust standard errors, clustering at the individual level. As we discuss in further detail in online

Appendix D, all of our mismatch coefficients remain still highly significant despite larger standard errors than if, as in our benchmark, we assume errors have an AR(1) structure as in GLS. Second, one might worry that our estimates merely reflect fixed differences across people: perhaps some workers are simultaneously not good at identifying the right occupation and also not productive at work, so they have higher mismatch on average and lower wages without a direct relationship between the two. To address this, we reran all our wage regressions by including worker fixed effects and found quantitatively similar estimates for the effects of mismatch (reported in online Appendix E). Third, we have added a fourth type of skill—physical or manual—and reconstructed skill mismatch in this expanded state space. We were somewhat surprised to find that adding physical skills only had a marginal effect on the mismatch measure's predictive power for wages. We have also experimented with using earnings instead of wages (see online Appendix I), using the logarithm of mismatch rather than its level, and including higher-order terms for ability, skill requirements, and ability-experience interaction terms (see online Appendix J for the latter three). Overall, our substantive results were robust to these variations.

Mismatch and Occupational Switching.—So far we have focused on the impact of mismatch on wages. We now turn to the second key question we raised in the introduction and implied by the model. What is the effect of mismatch on the probability of occupational switching? In the model, highly mismatched workers have larger mistakes in their beliefs. This also leads to a higher probability of an occupational switch because Bayesian learning implies a larger expected correction to their beliefs. Proposition 2 formalized this logic, and we now look for this potential effect in the data.

We estimate a linear probability model for occupational switching on the same set of regressors as in our wage regressions, again instrumenting for occupational tenure. We are most interested in the coefficient on mismatch in the current occupation. Table 7 displays our baseline estimates.

Notice that the effect of current mismatch on the probability of switching occupations is always positive and significant at the 1 percent level, with the exception of social mismatch in column 2. To give a better idea about the magnitudes implied by these coefficients, in Table 8 we compute the occupational switching probabilities across the mismatch distribution using the specifications in columns 1 and 2. A worker who is in the ninetieth percentile of the mismatch distribution is 3.4 percentage points more likely to switch occupations than an otherwise comparable worker in the tenth percentile, a difference corresponding to about 21 percent of the average switching rate.

Splitting mismatch into components (last three columns of Table 8), we see that the gap in switching probability between the ninetieth and tenth percentiles is approximately 2 percentage points for verbal and math skills but is close to zero for social skills. Thus, consistent with what we found for wages, contemporaneous social mismatch seems to only have a modest effect on outcomes once we account for math and verbal skills.

In column 3 of Table 7, we see that the effects are roughly symmetric, with increased switching probability similarly associated with positive and negative

Table 7—Regressions for the Probability of Occupational Switch

	LPM-IV			LPM		
	(1)	(2)	(3)	(4)	(5)	(6)
Mismatch	0.0135			0.0066		
	(5.87)			(3.64)		
Mismatch Verbal		0.0076			0.0053	
		(2.68)			(2.38)	
Mismatch Math		0.0074			0.0025	
		(2.66)			(1.13)	
Mismatch Social		0.0007			−0.0012	
		(0.33)			(−0.72)	
Positive Mismatch			0.0134			0.0087
			(4.87)			(3.98)
Negative Mismatch			−0.0130			−0.0028
			(−4.79)			(−1.32)
Worker Ability (mean)	−0.0370	−0.0370	—	−0.0208	−0.0211	—
	(−2.76)	(−2.75)	—	(−1.79)	(−1.82)	—
Worker Ability × Occ Tenure	−0.0003	−0.0003	—	0.0019	0.0020	—
	(−0.15)	(−0.15)	—	(2.24)	(2.28)	—
Occ Reqs (mean)	−0.0333	−0.0334	—	−0.1225	−0.1223	—
	(−2.28)	(−2.29)	—	(−9.81)	(−9.80)	—
Occ Reqs × Occ Tenure	−0.0052	−0.0052	—	0.0106	0.0106	—
	(−2.99)	(−2.99)	—	(12.34)	(12.31)	—
Observations	41,596	41,596	41,596	41,596	41,596	41,596

Notes: All regressions include the same set of standard controls defined in the notes to Table 2; *t*-statistics are in parentheses and calculated with standard errors that are estimated as robust Huber-White sandwich estimates. More detailed regression results are in Table B.5 in online Appendix B.

Table 8—Effect of Mismatch on Occupational Switching Probability

Mismatch degree (high to low)	*Mismatch* effect	Effect by *Component*		
		Verbal	Math	Social
90 percent	0.0407	0.0209	0.0203	0.0020
	(0.0069)	(0.0078)	(0.0076)	(0.0062)
50 percent	0.0178	0.0082	0.0076	0.0008
	(0.0030)	(0.0030)	(0.0029)	(0.0024)
10 percent	0.0065	0.0014	0.0013	0.0001
	(0.0011)	(0.0005)	(0.0005)	(0.0004)

Notes: Each cell reports the change in the probability of switching occupations. Standard errors are in parentheses.

mismatch. Workers whose skills are worse than their occupations or better are both more likely to switch occupations. This is consistent with the "U shape" of Groes, Kircher, and Manovskii (2015), that workers either under- or overqualified for their occupation are more likely to switch.

C. *Switch Direction*

Not only do mismatched workers switch occupations more frequently, but their switches are also directional. They switch to improve their match quality, and the

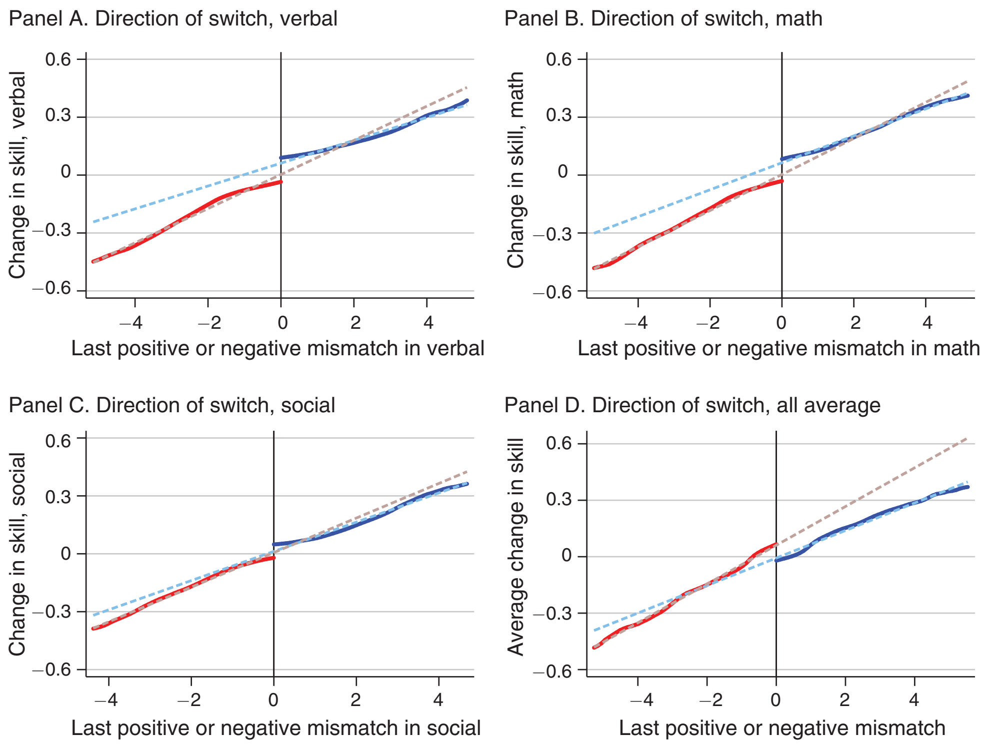

Figure 1. Nonparametric Plots of Direction of Switch

Notes: We run local polynomial regressions with a simple rule-of-thumb bandwidth (solid lines). On the x-axis, we have the value of the last positive or negative mismatch measure. On the y-axis, a change in a skill is computed as the difference in the rank score of the skill in the current occupation and the one in the last occupation. An average change is computed as the mean of the changes in the rank scores in all skills.

magnitude of their mismatch predicts the magnitude of the correction. Workers who are overqualified tend to switch to occupations with higher skill requirements and the converse for under-qualified workers.

As Lemma 4 lays out, our model of learning suggests that switches ought to be directed because as workers learn, their beliefs tend toward their true ability. Furthermore, Proposition 3 shows that the magnitude of the change is related to the magnitude of mismatch. This is a prediction of our model that contrasts with search models of mismatch. If search frictions induced mismatch, workers will switch toward better matches, but the size of the switch will be independent of the size of mismatch.

Our parametric and nonparametric evidence shows that switches tend to correct past mismatch and the magnitude of the correction depends on the magnitude of mismatch. Nonparametrically, we see this in panels A through C of Figure 1. We plot on the vertical axis the change in skill requirement for every worker who switches occupation, and on the horizontal we plot the last positive or negative mismatch in that skill. Specifically, the vertical axis is the difference between the jth skill requirement in the last occupation and that in the current one, that is, $q(\tilde{r}_{c(p),j}) - q(\tilde{r}_{c(p-1),j})$. The horizontal axis, positive and negative mismatch *in skill j*, is defined as in Section IIIA, but using only one dimension at a time.[34] To give the scatter plots some shape in Figure 1, we fit local polynomial regressions for positive and negative mismatch.

As shown in these panels, the upward-sloping curves on both sides of zero mean that individuals who are overqualified in skill j (the right half of the axes) tend to choose their next occupation with a higher skill requirement, whereas the opposite is true for individuals who are under-qualified. The positive relationship means that the more mismatched the worker was in the last occupation, the larger the change in requirements of that skill is in the next switch. Panel D plots the same relationship by aggregating across all three skill types, which again shows the same patterns.[35]

Figure 1 only documents a univariate relationship between changes in requirements as a function of current mismatch in the same skill dimension. To investigate richer dependencies, we regress the change (upon switching occupations) in skill requirement j on positive and negative mismatch in all three skill dimensions for the worker's last occupation.[36] We also include all of the worker characteristics from our wage regression on positive and negative mismatch, for example, demographics, employer, and occupational tenure.

Columns 1–3 of Table 9 report the coefficient estimates from this regression. Column 4 reports the case where the average change in skills is regressed on positive and negative mismatch.

There are several takeaways from this table. Most importantly, the positive coefficients on all regressors confirm the main message of Figure 1: the skill change upon switching is an increasing function of current mismatch, so switching works to reduce skill mismatch. This is true even when we control for mismatch along other dimensions. For example, column 1 tells us that a worker will choose his next occupation to have a higher verbal skill requirement if he is currently overqualified in verbal dimension (first row), but even more so if he is currently overqualified in math skills (coefficients of 0.0316 versus 0.0599). Mismatch in the social dimension has little impact (coefficient of 0.0061) on verbal requirements, but a strong effect on changes in social skill requirements. These results echo the same theme as before that math and verbal skills are distinct, yet closely connected, whereas social skills have their own dynamics. The coefficients also show an asymmetry that was apparent in Figure 1: with the exception of math, workers who are under-qualified in a dimension reduce that skill requirement by more than an overqualified worker increases his skill requirement.

To provide some interpretation of the estimated coefficients, we compute the effect of positive and negative mismatch in skill j on the change in that skill for ninetieth, fiftieth, and tenth percentile rank of each measure in Table 10 using (diagonal entries from the) regression results. For example, a worker with high positive verbal mismatch, in the ninetieth percentile, will choose an occupation 9.95 percentiles

[34] $m^{+}_{i,c(p-1),j} \equiv \max\left[q\left(\tilde{A}_{i,j}\right) - q\left(\tilde{r}_{c(p-1),j}\right), 0\right]$ and $m^{-}_{i,c(p-1),j} \equiv \min\left[q\left(\tilde{A}_{i,j}\right) - q\left(\tilde{r}_{c(p-1),j}\right), 0\right]$.

[35] Again, we restrict our observations to those who have strictly positive mismatch (to the right of the axis) and those who have strictly negative mismatch (to the left of the axis). Unlike positive or negative mismatchin skill j, observations don't split into either the positive or negative side in this case. That is, a number of observations show up on both sides.

[36] As before, skill requirement is measured in terms of percentile rank.

TABLE 9—REGRESSIONS FOR DIRECTION OF SWITCH

Dependent variable →	Verbal (1)	Math (2)	Social (3)	All average (4)
Last positive mismatch, verbal	*0.0316*	0.0097	0.0143	
	(7.30)	(2.11)	(3.08)	
Last negative mismatch, verbal	*0.0838*	0.0536	0.0216	
	(15.86)	(9.55)	(3.80)	
Last positive mismatch, math	0.0599	*0.0898*	0.0021	
	(13.47)	(19.02)	(0.44)	
Last negative mismatch, math	0.0558	*0.0893*	0.0076	
	(11.10)	(16.71)	(1.41)	
Last positive mismatch, social	0.0061	0.0046	*0.0774*	
	(1.85)	(1.31)	(22.01)	
Last negative mismatch, social	0.0264	0.0166	*0.1043*	
	(7.17)	(4.25)	(26.42)	
Last positive mismatch				0.0751
				(26.60)
Last negative mismatch				0.1143
				(37.75)
Observations	6,594	6,594	6,594	6,594

Notes: All regressions include the same set of standard controls defined in the notes to Table 2; t-statistics are in parentheses and calculated with standard errors that are estimated as robust Huber-White sandwich estimates. More detailed regression results are in Table B.6 in online Appendix B.

TABLE 10—EFFECT OF LAST MISMATCH ON CHANGE IN SKILLS

	Predicted percentile change in skill j		
Last mismatch percentile in skill j	Verbal	Math	Social
Positive			
90 percent	9.95	27.91	22.87
	(1.36)	(1.47)	(1.04)
50 percent	4.03	10.70	9.71
	(0.55)	(0.56)	(0.44)
10 percent	0.67	1.88	1.65
	(0.09)	(0.10)	(0.07)
Negative			
90 percent	−1.59	−1.77	−1.82
	(0.10)	(0.11)	(0.07)
50 percent	−9.07	−10.02	−10.31
	(0.57)	(0.60)	(0.39)
10 percent	−24.25	−26.35	−29.59
	(1.53)	(1.58)	(1.12)

Notes: These values are changes in percentile rank scores in each skill dimension. Standard errors are in parentheses.

higher in verbal skill requirements. A similarly under-qualified worker in the tenth percentile of negative mismatch reduces his verbal skill requirements by 24.25 percentiles when switching.

V. Conclusion

In this paper, we propose an empirical measure of multidimensional skill mismatch that is implied by a dynamic model of skill acquisition and occupational choice. In this model, workers self-select based on comparative advantage, but here they face dynamic consequences: long-term wage loss due to mismatch's effect on human capital accumulation. Mismatch arises in our model due to workers' imperfect information about their learning abilities that causes them to choose occupations that are either above or below their optimal level. As workers discover their true abilities, workers better allocate themselves toward their optimal careers.

Our empirical findings provide support to the notion of mismatch proposed in this paper. In particular, we find that mismatch predicts wages even with a long list of controls, including worker fixed effects and abilities and occupation requirements constructed from ASVAB and O*NET. Furthermore, mismatch has a long-lasting impact on workers' wages, depressing them even in subsequent occupations. This latter finding is consistent with the human capital channel that is embedded in our theoretical model.

A second set of findings shows workers choose to switch occupations so as to reduce their skill mismatch. This is true even when we split mismatch into its components. The magnitudes involved are also quite large, revealing large adjustments for workers in the skill space upon switching. These findings are consistent with our model of learning. Comparing skill dimensions, we find social skills behave somewhat differently than math and verbal skills. Although social ability appears to matter for wages, mismatch between a worker and an occupation along this dimension has a weaker relationship to wages than math or verbal mismatch.

These findings should only serve to motivate further work on the mechanisms involved in learning, human capital accumulation, and occupational choice. The empirical evidence we presented suggests a strong link between human capital accumulation and lifetime earnings, but fully quantifying its effects will require a structural quantitative model. Such a model will also allow us to conduct policy experiments and quantify their impact on lifetime welfare.

REFERENCES

Addison, John T., Liwen Chen, and Orgul D. Ozturk. 2017. "Occupational Match Quality and Gender over Two Cohorts." IZA Discussion Paper 11114.

Akcigit, Ufuk, Santiago Caicedo, Ernest Miguelez, Stefanie Stantcheva, and Valerio Sterzi. 2018. "Dancing with the Stars: Innovation through Interactions." NBER Working Paper 24466.

Altonji, Joseph G., Prashant Bharadwaj, and Fabian Lange. 2012. "Changes in the Characteristics of American Youth: Implications for Adult Outcomes." *Journal of Labor Economics* 30 (4): 783–828.

Altonji, Joseph G., and Charles R. Pierret. 2001. "Employer Learning and Statistical Discrimination." *Quarterly Journal of Economics* 116 (1): 313–50.

Altonji, Joseph G., and Robert A. Shakotko. 1987. "Do Wages Rise with Job Seniority?" *Review of Economic Studies* 54 (3): 437–59.

Altonji, Joseph G., and Nicolas Williams. 2005. "Do Wages Rise with Job Seniority? A Reassessment." *Industrial and Labor Relations Review* 58 (3): 370–97.

Antonovics, Kate, and Limor Golan. 2012. "Experimentation and Job Choice." *Journal of Labor Economics* 30 (2): 333–66.

Arellano, M. 1987. "Computing Robust Standard Errors for Within-Groups Estimators." *Oxford Bulletin of Economics and Statistics* 49 (4): 431–34.

Bacolod, Marigee P., and Bernardo S. Blum. 2010. "Two Sides of the Same Coin: U.S. 'Residual' Inequality and the Gender Gap." *Journal of Human Resources* 45 (1): 197–242.

Ben-Porath, Yoram. 1967. "The Production of Human Capital and the Life Cycle of Earnings." *Journal of Political Economy* 75 (4): 352–65.

Bowles, Samuel, Herbert Gintis, and Melissa Osborne. 2001. "The Determinants of Earnings: A Behavioral Approach." *Journal of Economic Literature* 39 (4): 1137–76.

Cameron, A. Colin, and Douglas L. Miller. 2015. "A Practitioner's Guide to Cluster-Robust Inference." *Journal of Human Resources* 50 (2): 317–72.

Cameron, A. Colin, and Pravin K. Trivedi. 2005. *Microeconometrics: Methods and Applications*. Cambridge, UK: Cambridge University Press.

Carrillo-Tudela, Carlos, and Ludo Visschers. 2014. "Unemployment and Endogenous Reallocation over the Business Cycle." http://www.tatsuro-senga.net/e1macro/wp-content/uploads/2017/03/CarrilloTudela.pdf.

Farber, Henry S., and Robert Gibbons. 1996. "Learning and Wage Dynamics." *Quarterly Journal of Economics* 111 (4): 1007–47.

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): 3303–38.

Gardner, Howard. 2011. *Frames of Mind: The Theory of Multiple Intelligences*. 3rd ed. New York: Basic Books.

Gathmann, Christina, and Uta Schönberg. 2010. "How General Is Human Capital? A Task-Based Approach." *Journal of Labor Economics* 28 (1): 1–49.

Gervais, Martin, Nir Jaimovich, Henry E. Siu, and Yaniv Yedid-Levi. 2016. "What Should I Be When I Grow Up? Occupations and Unemployment over the Life Cycle." *Journal of Monetary Economics* 83: 54–70.

Gibbons, Robert, Lawrence F. Katz, Thomas Lemieux, and Daniel Parent. 2005. "Comparative Advantage, Learning, and Sectoral Wage Determination." *Journal of Labor Economics* 23 (4): 681–724.

Groes, Fane, Philipp Kircher, and Iourii Manovskii. 2015. "The U-Shapes of Occupational Mobility." *Review of Economic Studies* 82 (2): 659–92.

Guvenen, Fatih. 2007. "Learning Your Earning: Are Labor Income Shocks Really Very Persistent?" *American Economic Review* 97 (3): 687–712.

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): 411–82.

Herkenhoff, Kyle, Jeremy Lise, Guido Menzio, and Gordon M. Phillips. 2018. "Production and Learning in Teams." NBER Working Paper 25179.

Huggett, Mark, Gustavo Ventura, and Amir Yaron. 2011. "Sources of Lifetime Inequality." *American Economic Review* 101 (7): 2923–54.

Ingram, Beth F., and George R. Neumann. 2006. "The Returns to Skill." *Labour Economics* 13 (1): 35–59.

Jarosch, Gregor, Ezra Oberfield, and Esteban Rossi-Hansberg. 2018. "Learning from Coworkers." https://www.princeton.edu/~erossi/LC.pdf.

Jovanovic, Boyan. 1979. "Job Matching and the Theory of Turnover." *Journal of Political Economy* 87 (5, Part 1): 972–90.

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

Kambourov, Gueorgui, and Iourii Manovskii. 2009. "Occupational Specificity of Human Capital." *International Economic Review* 50 (1): 63–115.

Lazear, Edward P. 2009. "Firm-Specific Human Capital: A Skill-Weights Approach." *Journal of Political Economy* 117 (5): 914–40.

Lise, Jeremy, and Fabien Postel-Vinay. 2016. "Multidimensional Skills, Sorting, and Human Capital Accumulation." https://econ.ucalgary.ca/sites/econ.ucalgary.ca/files/lise_postel-vinay.pdf.

Michaud, Amanda, and David Wiczer. 2016. "Occupational Hazards and Social Disability Insurance." Federal Reserve Bank of St. Louis Working Paper 2014-024B.

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

Neal, Derek. 1999. "The Complexity of Job Mobility among Young Men." *Journal of Labor Economics* 17 (2): 237–61.

Papageorgiou, Theodore. 2014. "Learning Your Comparative Advantages." *Review of Economic Studies* 81 (3): 1263–95.

Parent, Daniel. 2000. "Industry-Specific Capital and the Wage Profile: Evidence from the National Longitudinal Survey of Youth and the Panel Study of Income Dynamics." *Journal of Labor Economics* 18 (2): 306–23.

Pavan, Ronni. 2011. "Career Choice and Wage Growth." *Journal of Labor Economics* 29 (3): 549–87.

Perry, Anja, Simon Wiederhold, and Daniela Ackermann-Piek. 2014. "How Can Skill Mismatch Be Measured? New Approaches with PIAAC." *Methods, Data, Analyses* 8 (2): 137–74.

Poletaev, Maxim, and Chris Robinson. 2008. "Human Capital Specificity: Evidence from the Dictionary of Occupational Titles and Displaced Worker Surveys, 1984–2000." *Journal of Labor Economics* 26 (3): 387–420.

Rosen, Sherwin. 1972. "Learning and Experience in the Labor Market." *Journal of Human Resources* 7 (3): 326–42.

Roy, A.D. 1951. "Some Thoughts on the Distribution of Earnings." *Oxford Economic Papers* 3 (2): 135–46.

Sanders, Carl. 2014. "Skill Accumulation, Skill Uncertainty, and Occupational Choice." https://economics.wustl.edu/files/economics/imce/sandersjobmarketpaper2010.pdf.

Sanders, Carl, and Christopher Taber. 2012. "Life-Cycle Wage Growth and Heterogeneous Human Capital." *Annual Review of Economics* 4 (1): 399–425.

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

Topel, Robert. 1991. "Specific Capital, Mobility, and Wages: Wages Rise with Job Seniority." *Journal of Political Economy* 99 (1): 145–76.

Topel, Robert H., and Michael P. Ward. 1992. "Job Mobility and the Careers of Young Men." *Quarterly Journal of Economics* 107 (2): 439–79.

Ware, John E., Mark Kosinski, and Susan D. Keller. 1995. *SF-12: How to Score the SF-12 Physical and Mental Health Summary Scales*. Boston, MA: Health Institute, New England Medical Center. https://www.researchgate.net/publication/242636950_SF-12_How_to_Score_the_SF-12_Physical_and_Mental_Health_Summary_Scales.

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