# Multidimensional Skill Mismatch*

Fatih Guvenen† Burhan Kuruscu‡ Satoshi Tanaka§ David Wiczer¶

September 2, 2015

**Abstract**

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? In this paper, we propose an empirical measure of skill mismatch for a worker-occupation match, which sheds light on these questions. This measure 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 these skills. In this model, mismatch is central to the career outcomes of workers: it reduces the returns to occupational tenure, and it predicts occupational switching behavior. We construct our empirical analog by combining data from the National Longitudinal Survey of Youth 1979 (NLSY79), the Armed Services Vocational Aptitude Battery (ASVAB) on workers, and the O*NET on occupations. Our empirical results show that the effects of mismatch on wages are large and persistent: mismatch in occupations held early in life has a strong negative effect on wages in future occupations. Skill mismatch also significantly increases the probability of an occupational switch and predicts its direction in the skill space. These results provide fresh evidence on the importance of skill mismatch for the job search process.

**JEL Codes:** E24, J24, J31.
**Keywords:** Skill mismatch; match quality; Mincer regression; ASVAB; O*NET; occupational switching

*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, and Carl Sanders, as well as seminar and conference participants at various institutions. 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.

†University of Minnesota, FRB of Minneapolis, and NBER; guvenen@umn.edu
‡University of Toronto; burhan.kuruscu@utoronto.ca
§University of Queensland; s.tanaka@uq.edu.au
¶Federal Reserve Bank of St. Louis; wiczerd@stls.frb.org

# 1 Introduction

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.

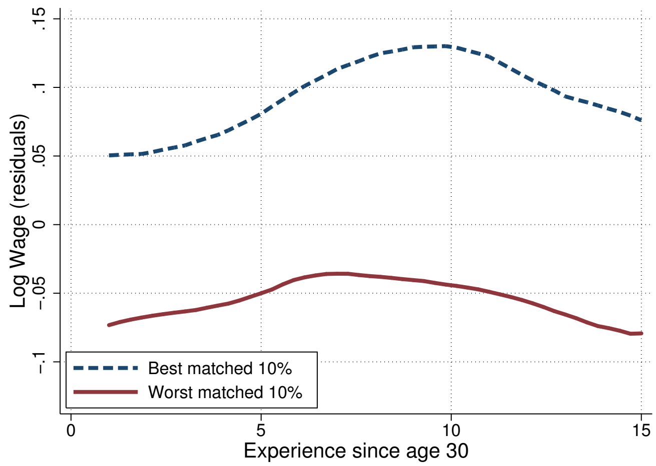
To explain the differences in outcomes of worker-job matches, economists often appeal to the idea of "match quality," that is, some unobservable match-specific factor that determines the productivity of a match after controlling for the observable characteristics of the worker and the job. A long list of papers, going as far back as Jovanovic (1979) and Mortensen and Pissarides (1994), have shown that allowing for such an idiosyncratic match quality can help explain a wide range of labor market phenomena, such as how wages and job separations vary by job tenure, among others (see Rogerson et al. (2005) for a survey of this literature). While theoretically convenient, mapping this abstract notion of match quality onto empirical constructs that can be easily measured has proved elusive. Consequently, in empirical work, match quality is often treated as a residual, whose value is pinned down by making the model fit data on various labor market outcomes.[1]

In this paper, we propose an empirical measure of match quality that can be constructed directly from micro data on workers and their occupations. For reasons that will become clear, it turns out to be convenient to measure the *lack* of match quality, or what we call *skill mismatch*. Rather than interpreting a job as a position in a given firm, we interpret it as a set of tasks to be completed—an occupation. Therefore, our notion of mismatch is based on the discrepancy between the portfolio of skills required by an occupation (for performing the tasks that produce output) and the portfolio of abilities possessed by a worker for learning those skills. If the vector of required skills does not align well with the vector of a worker's abilities, the worker is mismatched, being either overqualified or underqualified along different dimensions of this vector.

Our notion of skill mismatch is multidimensional, as suggested by our title. This viewpoint is motivated by a great deal of psychometric and educational research emphasizing multiple intelligences that can act and develop independently from each other.

[1]For examples, see Miller (1984), Flinn (1986), Jovanovic and Moffitt (1990), Moscarini (2001), and Nagypal (2007).

FIGURE 1 – Wage Gap Between the Best- and Worst-Matched Workers Persists For Several Years.




Note: Workers are grouped by their rank in our mismatch measure after 10 years of labor market experience. Residual wages are obtained by regressing log real wages on demographics, polynomials for occupation tenure, employer tenure, worker experience, a worker's ability measure, an occupational skill requirement measure, and their interactions with occupation tenure, and dummy variables for one-digit-level occupations and industries. See Section 4 for details of those variables. To obtain two lines, we run local polynomial regressions with residual wages on labor market experience for each group of workers, with a rule-of-thumb bandwidth.

Developmental psychologist Howard Gardner, who originally proposed this theory in his 1983 book, *Frames of Mind: The Theory of Multiple Intelligences,* found particular motivation for this idea in the proliferation of occupations. For example, in the preface to the latest edition of his book, Gardner (2011) observes:

> Any complex society has 100–200 distinct occupations at the least; and any university of size offers at least fifty different areas of study. Surely these domains and disciplines are not accidents, nor are the ways they evolve and combine simply random events. The culturally constructed spheres of knowledge must bear some kind of relation to the kinds of brains and minds that human beings have...[2]

Of course, economists are no strangers to the idea of multidimensional skills. After all, a long list of papers have built on the Roy model—which features multiple skills and

[2]Gardner proposed eight types of intelligences: musical-rhythmic, visual-spatial, verbal-linguistic, logical-mathematical, bodily-kinesthetic, interpersonal, intrapersonal, naturalistic, and existential. Of these, we study three in our main analysis and experimented with a fourth, bodily-kinesthetic. We found the latter to have little predictive power for the economic outcomes we studied, so we relegate those results to Appendix D.

comparative advantage—to study wages and occupational choice. Our paper follows this tradition by proposing a measure of mismatch in a world with multiple skills.

Before delving into the details of the paper, we highlight one of the key findings of this paper: workers who are poorly matched with their occupations earn lower wages even many years after they have left the occupation. To show this, in Figure 1, we compute the average of our mismatch measure for each worker over all the occupations held before age 30. Then we group workers who are in the best-matched 10% (blue dashed line) and worst-matched 10% (red solid line) of the population and plot the residual wage[3] of each group over the subsequent 15 years. The well matched group earn wages higher than would be expected based on their characteristics or those of their employer, and the opposite is true of the poorly matched. Notice that the gap is steady, the worst matched earn less to begin and do not close the gap, so that over 15 years they cumulatively have have lost approximately $121,000 (in 2002 dollars).

The empirical measure of skill mismatch we propose naturally emerges from a structural model of occupational choice, multidimensional skill accumulation, and learning about abilities to acquire skills. In this model, output is produced at economic units called "occupations," which combine a vector of distinct skills supplied by their workers. The technology operated by an occupation is given by a vector of *skill requirements*, which specifies the amount of skill *investment* required to be maximally productive in that occupation. Workers who choose occupations with skill requirements below or above their optimal skill investment produce output (and earn a wage) at levels that decline in a concave fashion from the maximum level. Consequently, for each worker there is an optimal amount of investment in each skill type depending on his abilities, and thus an optimal/ideal occupation choice.

How are skills accumulated? Each worker enters the economy possessing a portfolio of skills and accumulates skills of each type by an amount that depends on two factors: (i) his ability to learn that skill and (ii) the occupation he works in. In particular, the same skill requirements that determine current output at the occupation as described above also affect the efficiency of human capital accumulation depending on workers' learning abilities. Workers who are either over- or underqualified accumulate human capital less efficiently than workers who are matched well.[4]

[3]We are controlling for demographics, polynomials for occupation tenure, employer tenure, worker experience, worker's ability measure, an occupational skill requirement measure, and their interactions with occupation tenure, and dummy variables for one-digit-level occupations and industries.

[4]This dual role of a job as producing both output and worker skills is in the spirit of Rosen (1972).

We assume that occupations are distributed continuously in the skill requirements space. Therefore, without any frictions, each worker can choose the occupation that is ideal for him—that is, where he is exactly qualified along all skill dimensions. What prevents this from happening is imperfect information, which arises because workers enter the labor market without full knowledge of their portfolio of abilities (to learn skills). Therefore, a worker may overestimate his ability to learn a certain skill, which will cause him to choose an occupation with skill requirements for that type of skill that are too high relative to his true ability. The opposite occurs when he underestimates his ability. Workers learn about their abilities in an optimal (Bayesian) fashion as they observe changes in their wages from year to year. Each period workers optimally choose a new occupation as they update their beliefs about their true abilities. In this model, we show that skill mismatch is a key determinant of a worker's wages in his current occupation, as well as of his switching behavior across occupations.

In the empirical analysis, we study four key predictions of the model. First, we show that mismatch depresses human capital accumulation and, consequently, reduces both the level and the growth rate of wages with tenure in a Mincer wage regression. Second, current wages also depend negatively on cumulative mismatch in previous occupations. Third, the probability of switching occupation increases with mismatch because each wage observation causes a bigger update of a worker's belief when mismatch is high. Fourth, occupational switches are directional: workers who are overqualified in a skill dimension tend to switch to occupations that are more skill intensive in that dimension. The opposite happens when the worker is underqualified.

In order to test the implications of our framework, we employ the 1979 National Longitudinal Survey of Youth (NLSY79) for information on workers' occupation and wage histories. NLSY79 respondents were also given an occupational placement test—the Armed Services Vocational Aptitude Battery (ASVAB)—that provides detailed measures of occupation-relevant skills and abilities.[5] In addition to this cognitive measure, respondents report several measures of noncognitive skills that we use to describe one's ability for socially interactive work. For comparability with existing work in this area, we focus on the sample of male workers. Turning to the skill requirements of each oc-

[5]We interpret workers' test scores as corresponding to (noisy measures of) abilities in our model. Although it is not obvious whether these scores reflect abilities or accumulated skills, this distinction is not critical because accumulated skills before age 20 are highly correlated with one's abilities to learn those skills: Huggett et al. (2011) estimate that this correlation exceeds 0.85. Since these tests are taken at the beginning of workers' careers, we interpret them as abilities.

cupation, we use data from the U.S. Department of Labor's O*NET project. This data set provides a very detailed picture of the knowledge and skills used in each of the occupations that an NLSY79 respondent might hold. To connect these two data sources, we use the cross-walk provided by the ASVAB project that maps the skills that are tested in ASVAB to the skills measured by the O*NET.[6] Combining these two sources of information allows us to compute both a contemporaneous mismatch measure (in the current occupation) as well as a cumulative mismatch measure (over all past occupations). In the most detailed case, we 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 the Mincer wage regression framework along with flexible interactions with occupational tenure (and a large set of other controls). Consistent with our theory, we find that the coefficient on mismatch is robustly negative and that its interaction with occupational tenure is robustly negative. The estimates imply that the wage rate is 7.4% lower after 10 years of occupation tenure for a worker at the 90th percentile of the mismatch distribution relative to one at the 10th percentile. Even more important, cumulative mismatch also has a significant and negative effect on wages: the implied effect is an 8.9% difference in wages from the top to bottom decile of cumulative mismatch. Our model captures this persistence through the lasting effect of human capital accumulation; it would be missed by theories that postulate that match quality only affects the current match.

Turning to occupational switching behavior, the data reveal patterns consistent with our model. First, estimating a hazard model for occupational switching shows that it is robustly increasing in mismatch. The magnitudes are also fairly large: the switching probability is about 3.5 percentage points higher for a worker at the 90th percentile of the mismatch distribution relative to another worker at the 10th percentile. This gap is about one-fifth of the average switching probability in our sample. Second, we follow workers across occupational transitions to see if they tend to "correct" previous mismatches. Indeed, they do: if a worker is overqualified in his current occupation 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).

[6] The reader might wonder 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, but were only given a fairly wide range; and second, these test scores are themselves noisy measures of individuals' true underlying abilities as discussed further later.

A multidimensional measure of mismatch has important empirical implications. For example, if a worker who is very talented in one type of skill currently works in an occupation requiring another skill intensively, he would be considered mismatched even though both the worker and the occupation might be described as high skill on average. It also allows us to see the potentially different effects of being over- or underqualified along different dimensions, and we show that there are important qualitative differences. For example, mathematical mismatch contributes more to the level of wages, whereas verbal mismatch affects the growth of wages with occupational tenure.

Finally, we extend our wage regressions to distinguish mismatch for overqualified and underqualified workers. We find that both those who were overqualified and those who were underqualified in previous occupations have lower wages today. This implication is consistent with our model, but is inconsistent with a standard Ben-Porath model with multidimensional skills, as we discuss in Section 2.3.

The paper proceeds as follows. In Section 2 we present our model. In Section 3 we describe our data, and Section 4 describes our methodology and how we create our mismatch measures. Section 5 presents the results, and finally, we conclude in Section 6.

## Literature Review

This paper relates to several branches of literature that have emphasized the role of match quality in wage determination and labor market flows. In one such strand, papers such as Altonji and Shakotko (1987), Altonji and Williams (2005), and Topel (1991) emphasized the importance of unobserved employer match quality for the estimated returns to tenure in that job. In this paper, we include our empirical measure of match quality in the wage regression and estimate its interaction with returns to tenure. Relative to these authors, we essentially pull out (a good part of) the match quality component from the residual and directly estimate its importance to wages and returns to tenure, which we find to be significant.

Whereas these earlier papers focused on job match quality, this paper focuses on occupational match quality. Motivating our work is the paper of Kambourov and Manovskii (2009b), which emphasized the importance of returns to occupational tenure over industry or job tenure. We build on this point, showing that not all occupational matches enjoy the same returns to tenure: skill mismatch is a critical factor that determines the returns to occupational tenure. Beyond treating occupational titles as different, we describe occupations by quantifiable skill requirements, similar to other recent papers, notably Ingram and Neumann (2006), Poletaev and Robinson (2008), Gathmann and Schönberg (2010), Bacolod and Blum (2010), and Yamaguchi (2012). These papers have shown that descriptors from the O*NET (as well as from its predecessor, the Dictionary of Occupational Titles (DOT)) have strong explanatory power for wages and career trajectories. We add to this literature by linking this occupation-level information to worker-side information from ASVAB and non-cognitive ability measures in NLSY79 to create a measure of mismatch.

Several papers have studied how workers search for a job that utilizes their comparative advantage as they learn about it (Gibbons et al. (2005), Gervais et al. (2014), Papageorgiou (2014), Sanders (2014), and Antonovics and Golan (2012)). Because different sectors can reward skills differently, as skills are revealed, workers switch toward sectors that maximize their comparative advantage. In a slightly different context, Farber and Gibbons (1996) and Altonji and Pierret (2001) investigate public learning about the workers' quality. Using the Armed Forces Qualification Test (AFQT) score, both papers find evidence that the importance of learning grows with the duration of the match. While our results confirm their findings—the wage effect of worker abilities grows with tenure—we also find that mismatch (both contemporaneous and cumulative) matters greatly even after including worker abilities in the wage regression (see Table IV). This happens because in our model a mismatched worker underaccumulates skills, leading to slow wage growth over time.

A particularly relevant precursor to our paper is Groes et al. (2015). These authors focus on the effects of mismatch on occupational switching in a model in which information frictions drive mismatch and learning ameliorates it. The conclusions of our paper regarding occupational switching are consistent with theirs, but two differences are worth noting. One, these authors define mismatch along a single dimension—an individual's wage deviation from the occupation's average wage—whereas our multidimensional measure enables us to study the effects of mismatch along several distinct types of skill dimensions. Two, because we do not use wages to define mismatch, we are able to include it in Mincer regressions to gauge its effects on wage determination, which is not possible in their approach.

Finally, in contemporaneous work, Lise and Postel-Vinay (2015) study similar features of the data but through quite different methods. Similar to this paper, they use data on workers' skills and occupational requirements to create a measure of match quality, which is intimately tied to wages. Whereas information is the fundamental friction that causes mismatch in our model, in their model search frictions prevent a good match. Both papers highlight the lasting effects of work history, such that mismatch permanently affects wage growth.

# 2 Model

In this section, we present a life-cycle model of occupational choice and human capital accumulation.[7] The structure of the labor market is built upon Rosen (1972), wherein the market for training/learning opportunities is "dual" to the market for jobs. Our model introduces two key features into this framework. First, human capital is multidimensional and workers differ in their learning ability in each of these dimensions, which characterizes the joint choice over human capital accumulation and type of work. Second, learning ability is imperfectly observable, about which individuals have rational beliefs and update these beliefs over time in a Bayesian fashion. We use this framework to study the effects of skill mismatch—between workers' abilities and occupations' skill intensities—on labor market outcomes.

## 2.1 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 $\beta$ is the subjective time discount factor.

**Technology.** There is a continuum of occupations, each using $n$ types of skills, indexed with $j \in \{1, 2, ..., n\}$. Occupations differ in their skill intensity of each skill type, denoted with the vector $\mathbf{r} = (r_1, r_2, ..., r_n) \geq 0$, which remains fixed over time. Each worker is endowed with ability to learn each type of skill, which we denote by the ability vector $\mathbf{A} = (A_1, A_2, ..., A_n)$. The worker enters period $t$ with the vector of human capital of each skill type $\mathbf{h}_t = (h_{1,t}, h_{2,t}, ..., h_{n,t})$. With a slight abuse of notation, let $\mathbf{r}_t = (r_{1,t}, r_{2,t}, ..., r_{n,t})$ denote the occupation chosen by the worker in period $t$. If the worker

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

chooses an occupation indexed by the skill intensity vector $\mathbf{r}_t$, then the amount of skill $j$ that he can effectively utilize in that occupation is assumed to be

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

where $\varepsilon_{j,t}$ is a zero mean noise whose role will become clear once Bayesian learning is introduced. This specification has two key features. First, skill requirement, $r_{j,t}$, enters nonmonotonically—the linear term is thought of as capturing the benefit of an occupation whereas the negative quadratic captures the costs (such as additional training required at high skill jobs). This nonmonotonicity will ensure below that each worker's optimal occupational choice (i.e., choice of $r_{j,t}$) is an interior one; 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 sorting by skill 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.

Workers are paid their marginal products after production takes place. Thus, the wage rate is

$$w_t = \sum_j \alpha_j k_{j,t},$$

where $\alpha_j$'s are weights that are identical across occupations.[8] Note that a worker's wage depends on (the vectors of) his human capital $\mathbf{h}_t$, job choice $\mathbf{r}_t$, and learning ability $\mathbf{A}$. The beginning-of-period human capital in period $t+1$ is given as

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

where $\delta$ is the depreciation rate of human capital, which is assumed to be uniform across skill types and occupations. Thus, $\mathbf{k}_t = (k_{1,t}, ..., k_{n,t})$ determines both the worker's current wage and also the next period's human capital.

We can phrase this market structure in the language of Rosen (1972), where occupations differ not only in the wages they offer but also in the learning opportunities they provide. In our notation, these opportunities are summarized by $\mathbf{r}$, the rate of human

[8] In principle, we can assume that occupations differ in how they weigh each skill type. However, since our intention is to present the simplest model through which we can introduce mismatch, we forgo this possibly reasonable assumption that complicates our model.

capital investment. Crucial to the tradeoff is that wages are net of the cost of this investment: Just as workers sell their labor services, they also "purchase" training from firms. In our model, individuals differ in their learning abilities, $\mathbf{A}$, and so their optimal occupation choice differ too. The heterogeneity in the cost of investment studied in the Rosen model is isomorphic to differences in the return on investment in our model.

**Information Structure.** The underlying friction that generates mismatch is imperfect information about workers' abilities, which is updated over time through a Bayesian process. Specifically, each worker draws ability $A_j$ from a normal distribution at the beginning of his life: $A_j \sim \mathcal{N}(\mu_{A_j}, \sigma^2_{A_j})$. The worker does not observe the true value of $A_j$ but observes a signal given by $\hat{A}_{j,1} = A_j + \eta_j$ where $\eta_j \sim \mathcal{N}(0, \sigma^2_{\eta_j})$, so prior beliefs are unbiased. Thus, the worker starts his career with the prior belief that his ability in skill type $j$ is normally distributed with mean $\hat{A}_{j,1}$ and precision $\lambda_{j,1} \equiv 1/\sigma^2_{\eta_j}$.

We assume that the worker observes $A_j + \varepsilon_{j,t}$ in each period, where the noise is $\varepsilon_{j,t} \sim \mathcal{N}\left(0, \sigma^2_{\varepsilon_j}\right)$. 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 and $\lambda_{j,t}$ be the precision of this distribution at the beginning of period $t$ and $\lambda_{\varepsilon_j}$ 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:

$$\hat{A}_{j,t+1} = \frac{\lambda_{j,t}}{\lambda_{j,t+1}} \hat{A}_{j,t} + \frac{\lambda_{\varepsilon_j}}{\lambda_{j,t+1}} \left(A_j + \varepsilon_{j,t}\right), \tag{3}$$

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

## 2.2 The Worker's Problem

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

$$V_t(\mathbf{h}_t, \hat{\mathbf{A}}_t) = \max_{\{r_{j,t}\}} \mathbb{E}_t \Big[ \sum_j \alpha_j k_j(t) + \beta V_{t+1}(\mathbf{h}_{t+1}, \hat{\mathbf{A}}_{t+1}) \Big],$$

subject to (1), (2), and (3). Since occupations are represented by a vector of skill intensities, this problem yields a choice of occupation in the current period, which 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 $j = 1, ..., n$), given by $\mathcal{N}(\hat{A}_{j,t}, 1/\lambda_{j,t})$, and the distribution of $\varepsilon_{j,t}$, given by $\mathcal{N}(0, \sigma^2_{\varepsilon_j})$.

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

1. *Occupational choice:* $r_{j,t} = \hat{A}_{j,t}$*;*
2. *Value function:*

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

*where* $B_t$ *is a known time-varying function that does not affect the worker's choices.*

Three remarks about this solution are in order. First, since $A_j$'s enter into the worker's objective function linearly, the solution only depends on the worker's expectation of $A_j$, which is $\hat{A}_{j,t}$. Second, 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.

Third, it is also instructive to compare our model with the standard Ben-Porath formulation, from which our model differs in three important ways. One, we introduce multidimensional human capital and abilities. Two, skill accumulation varies not only with a worker's learning abilities ($A_j$) but also with his occupation. Finally, in the Ben-Porath model (assuming perfect information, ignoring multidimensional skills, and interpreting $r_t$ as human capital investment), the current wage is given by $w_t = h_t - r_t^2/2$, and the next period's human capital is given by $h_{t+1} = (1-\delta)(h_t + Ar_t)$. Thus, the choice of $r_t$ that maximizes the current wage is zero, whereas the one that maximizes future human capital is infinite. Thus, there is an intertemporal trade-off between current and future wages. In contrast, in our model, the current wage is $h_t + Ar_t - r_t^2/2$, and the next period's human capital is given by $h_{t+1} = (1-\delta)(h_t + Ar_t - r_t^2/2)$; both equations have the same 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. Thus, the intertemporal trade-off disappears in our model, and the human capital decision essentially becomes a repeated static decision. This model feature is reminiscent of Rosen (1972), in which, if learning ability is constant over time, there is no dynamic tradeoff. We show this result more formally in Appendix B.2.

## 2.3 Skill Mismatch

There is an "ideal" occupation for each worker, which is the occupation that the worker would choose if he had perfect information about his abilities. Denoting this ideal occupation with $\mathbf{r}_t^* = (r_{1,t}^*, ..., r_{n,t}^*)$, it is given as $r_{j,t}^* = A_j$ for all $j$ and $t$. We define the skill *mismatch* in dimension $j$ as $(r_{j,t}^* - r_{j,t})^2$, the deviation of skill-$j$ intensity of the worker's occupation from his ideal occupation's skill-$j$ intensity. Given that $r_{j,t} = \hat{A}_{j,t}$, skill *mismatch* in dimension $j$ can alternatively be written 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 $A_j$'s and his occupation's skill intensities that correspond to $r_{j,t}$'s in order to construct our mismatch measure. By employing the same mismatch measure $(A_j - r_{j,t})^2$, we can rewrite the worker's wage as

$$w_t = \sum_{j=1}^{n} \alpha_j \Big( h_{j,t} + \frac{1}{2}(A_j^2 - (A_j - r_{j,t})^2) \Big) + \sum_{j=1}^{n} \alpha_j r_{j,t} \varepsilon_{j,t},$$

which shows that a worker's wage depends positively on his beginning-of-period human capital $\mathbf{h}_t$ and his ability vector $\mathbf{A}$ and negatively on *mismatch* $(A_j - r_{j,t})^2$. However, note that current human capital depends on past occupational choices and thus past mismatches. In order to see the effect of past mismatches on the current wage, use equations (1) and (2), and repeatedly substitute for human capital. Setting $\delta \equiv 0$ in order to simplify the expression, we obtain an expression that links the current wage to mismatches experienced in all periods:[9]

$$w_t = \sum_j \alpha_j h_{j,1} + \underbrace{\frac{1}{2} \sum_{j=1}^{n} \alpha_j A_j^2 \times t}_{\text{ability} \times \text{experience}} - \underbrace{\frac{1}{2} \sum_{s=1}^{t} \sum_{j=1}^{n} \alpha_j \left( A_j - r_{j,s} \right)^2}_{\text{mismatch}} + \sum_{s=1}^{t} \sum_{j=1}^{n} \alpha_j r_{j,s} \varepsilon_{j,s}. \quad (4)$$

The equation above shows that the current wage is positively related to the worker's ability times his labor market experience and negatively related to the *history* of mismatch values. Notice also that all past mismatch terms have the same effect on the current wage. This is because, first, we have assumed zero depreciation of human capital for simplicity. Otherwise, as shown in Appendix B, mismatches in previous periods would be discounted in the wage expression.

Second, again for simplicity, we have assumed that all occupations put the same weight on all skills; $\alpha_j$'s are the same in all occupations. However, these weights could

[9] In Appendix B, we provide the analogous expression with positive depreciation.

be different in different occupations. As a result, mismatches experienced in different occupations could affect the current wage differently. In order to account for differential impacts of mismatches in different occupations, we separate the mismatch in the current occupation from mismatches in previous occupations in our empirical estimation. We can illustrate this point using the wage equation above. For this purpose, let $t^c$ denote the period in which the worker switched to his current occupation. Thus, $r_{j,s} = r_{j,t^c}$ for all $s \geq t^c$ and the tenure in the current occupation is equal to $t - t^c + 1$. Then, we can rewrite the current wage as

$$\begin{aligned} w_t &= \sum_j \alpha_j h_{j,1} + \underbrace{\frac{1}{2}\sum_j \alpha_j A_j^2 \times t}_{\text{ability}\times\text{experience}} - \underbrace{\frac{1}{2}\sum_j \alpha_j \left(A_j - r_{j,t^c}\right)^2}_{\text{current mismatch}} \times \underbrace{(t - t^c + 1)}_{\text{current tenure}} \qquad (5) \\ &- \underbrace{\frac{1}{2}\sum_{s=1}^{t^c-1}\sum_j \alpha_j \left(A_j - r_{j,s}\right)^2}_{\text{cumulative past mismatch}} + \sum_{s=1}^{t}\sum_j \alpha_j r_{j,s}\varepsilon_{j,s}. \end{aligned}$$

This equation forms the basis for our empirical estimation. It shows that the current wage is negatively related to the current mismatch times the tenure in the current occupation and to the cumulative mismatch in previous occupations.[10]

An important issue in estimating the wage equation above is that the error term is correlated with mismatch measures because the mean of a worker's beliefs about his abilities is correlated with past shocks. By repeatedly substituting equation (3) backward, one can see that beliefs and, therefore, occupational choice and mismatch in each period, depend on all shocks in previous periods. As a result, our estimates of the coefficient on mismatch will be biased. Fortunately, as we state formally in Lemma 1, it turns out that mismatch in a period and past shocks are positively correlated. If we observe a high wage in a period due to positive shocks, we will also observe a high mismatch. Thus, the true effect of mismatch on wages should be stronger than the effect we estimate in our empirical analysis.

[10] If we used the standard Ben-Porath specification with $h_{j,t+1} = (1-\delta)\left(h_{j,t} + A_j r_{j,t}\right)$, then if a worker is employed in an occupation above his ideal skill match today (if the worker is underqualified), he would earn lower wages today but higher wages in the future, since he accumulates more skills in a higher "$r$" occupation. Therefore, in the standard Ben-Porath model negative past cumulative mismatch (worker being underqualified) has a positive effect on current wages and vice versa. In our current set-up, both positive and negative past cumulative mismatches have a negative effect on current wages.

**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,* $Cov(M_{j,t}, \Omega_{j,t}) > 0$. *Therefore, the estimated coefficient of mismatch provides a lower bound for the true effect.*

Another issue that concerns the empirical estimation of the wage equation is that we do not directly observe $A_j$'s. 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 $\widetilde{A}_j \equiv A_j + \nu_j$ where $\nu_j \sim \mathcal{N}(0, \sigma^2_{\nu_j})$ denote the test scores. To see how using $\widetilde{A}_j$ instead of $A_j$ in the estimation affects our results, insert $A_j = \widetilde{A}_j - \nu_j$ into (4), which gives

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

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.

**Lemma 2.** *Let* $\widetilde{\Delta}_{j,t} = \widetilde{A}_j^2 \times t$ *and* $\widetilde{M}_{j,t} \equiv \sum_{s=1}^{t}\left(\widetilde{A}_j - r_{j,s}\right)^2$, $\widetilde{\Omega}_{j,t} \equiv \sum_{s=1}^{t} r_{j,s}\left(\varepsilon_{j,s} - \nu_j\right)$. *Then,*

1. $Cov\left(\widetilde{\Delta}_{j,t}, \widetilde{\Omega}_{j,t}\right) < 0$*: Therefore, the estimated coefficient of ability-experience interaction provides a lower bound for the true effect.*

2. $Cov\left(\widetilde{M}_{j,t}, \widetilde{\Omega}_{j,t}\right) > 0$*: 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 will provide *lower bounds* on the effects of mismatch on wages.

## 2.4 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* $E[(A_j - r_{j,t})^2] = 1/\lambda_{j,t}$*. 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 $m_j > 0$ be a positive number, an occupational switch occurs in period $t$ if $r_{j,t} > r_{j,t-1} + m_j$ or $r_{j,t} < r_{j,t-1} - m_j$ 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 would be 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. Since occupational switch is related to the change in the mean belief, the probability of switching increases with mismatch.[11] Moreover, conditional on mismatch, if the precision of beliefs is higher, the probability of switching occupations will be lower since each 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.

We now turn to the direction of occupational switches. In particular, we can show that occupational switches tend to be in the direction of reducing existing mismatches. That is, workers who are overqualified 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 underqualified. The following proposition formalizes this result.

To establish this, we introduce some notation. Let $\pi^{\text{up}}_{j,t} \equiv \Pr(r_{j,t+1} - r_{j,t} > m_j)$ 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: $\pi^{\text{down}}_{j,t} \equiv \Pr(r_{j,t+1} - r_{j,t} < -m_j)$.

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

**Proposition 3.** **[Direction of Occupational Switches]** *If the worker is overqualified in skill $j$, that is, $r_j^* - r_{j,t} > 0$, then:*

1. *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*

2. *the probability of moving up in skill j increases with the extent of overqualification:* $\frac{\partial \pi_{j,t}^{up}}{\partial (r_j^* - r_{j,t-1})} > 0$.

A worker would be overqualified for his occupation in skill dimension $j$ if he chose an occupation with a lower skill-$j$ intensity than his ideal occupation. 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.

# 3 Data

In this section, we describe the data for our empirical analysis. The main source of data is the NLSY79, which is a nationally representative sample of individuals who were 14 to 22 years of age on January 1, 1979. In addition to the detailed information about earnings and employment, the NLSY79 has three other features that make it suitable for our analysis. First, at the start of the survey all respondents took the ASVAB test, which measures various abilities. Second, respondents were also surveyed about their attitudes that broadly pertain to their social skills (e.g., self-esteem, willingness to engage with others, among others). The ASVAB scores will be used to construct a measure of cognitive abilities, and scores on self-esteem and social interactions will be used to measure social abilities. Third, each individual provided the occupational title for each of their jobs.

We link the ability information on the worker side to the skill requirements information on the occupation side, the latter reported in O*NET (to be explained in detail later), and create a measure of mismatch between a worker and his occupation (taken to be the occupation at his main job). Below, we describe the NLSY79, worker's ability information, occupational skill requirements information, and how we aggregate the ability and skill information into the three components: *verbal*, *math*, and *social*, which are used to create the mismatch measure.

## 3.1 NLSY79

We use the Work History Data File of the NLSY79 to construct yearly panels from 1978 to 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, which includes 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 Appendix C. Our final sample runs from 1978 through 2010 and includes 1,992 individuals and 44,591 individual-year observations.

Descriptive statistics for the sample are reported in Table I. Because of the nature of the survey, which starts with workers when they are young in the workforce, the sample skews younger. As a result, the mean length of employer tenure in our sample is relatively short, although this is a well-understood point about the NLSY79 in the literature.[12] Annual occupational mobility in our sample is 15.94%, comparable to 18.48% reported in Kambourov and Manovskii (2008) who use the Panel Study of Income Dynamics (PSID) for the period 1968–1997.

## 3.2 Data on Workers' Abilities

### ASVAB

Between 1973 and 1975 the U.S. Department of Defense introduced the ASVAB test, designed and maintained by professional psychometricians, to place new recruits into jobs. The version of the ASVAB taken by NLSY79 respondents had 10 component tests.[13] Among those, we focus on the following 4 component tests on verbal and math abilities, which can be linked to skill counterparts: Word Knowledge, Paragraph Comprehension, Arithmetic Reasoning, and Mathematics Knowledge. To process the ASVAB scores, we follow Altonji et al. (2012). In particular, when the test was administered in 1980, the respondents' ages were up to 7 years apart. Because age is likely to have a systematic

[12] Both Parent (2000) and Pavan (2011) report mean employer tenure in the NLSY79 that ranges from 3 to 3.3 years. The corresponding figure in our sample is 3.6 years, which is close.

[13] 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.

Table I – Descriptive Statistics of the Sample, NLSY79, 1978–2010

| Statistics | All Sample | $\leq$ High School | $>$ High School |
|---|---|---|---|
| Total number of observations | 44,591 | 21,618 | 22,973 |
| Total number of individuals | 1,992 | 954 | 1,038 |
| | | | |
| Average age at the time of interview | 33.79 | 32.85 | 34.67 |
| Highest education $<$ high school | 7.01% | 14.45% | - |
| Highest education $=$ high school | 41.48% | 85.55% | - |
| Highest education $>$ high school | 51.51% | - | 100.00% |
| Highest education $\geq$ 4-year college | 31.74% | - | 61.62% |
| Percentage African-American | 10.46% | 13.81% | 7.31% |
| Percentage Hispanic | 6.53% | 6.92% | 6.16% |
| | | | |
| Occupational mobility | 15.94% | 18.10% | 13.90% |
| Occupational tenure (mean) | 6.50 | 6.17 | 6.81 |
| Occupational tenure (median) | 4.00 | 4.00 | 5.00 |
| Employer (job) mobility | 30.39% | 31.97% | 28.90% |
| Employer (job) tenure (mean) | 3.61 | 3.56 | 3.65 |
| Employer (job) tenure (median) | 2.00 | 2.00 | 2.00 |
| Average hours worked within a year | 1983.8 | 1958.8 | 2007.2 |

Note: Occupational mobility is defined as the fraction of individuals who switch occupations in a year. The same definition for employer mobility.

effect on the ASVAB score, we normalize the mean and variance of each test score by their age-specific values.

### Social Ability Scores in NLSY79

The NLSY79 included three attitudinal scales, which describe a respondent's non-cognitive abilities. We focus on two of these measures: the Rotter Locus of Control 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 autonomy in the world, the primacy of his self-determination rather than chance. Heckman et al. (2006) also uses these two scores, and Bowles et al. (2001) review evidence on the influence 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* hereafter.

## 3.3 Occupational Skill Requirements

### O*NET

The U.S. 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, which can be mapped into the 292 occupation categories included in the NLSY79. For each of these occupations, occupation analysts at O*NET give a score for the importance of each of 277 descriptors. These scores are updated periodically using survey data, but we opt for version 4.0, the analysts database, which should yield a more consistent picture across occupations without biases and coding errors of respondents. From these descriptors we will use 26 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. For the complete list, see Table II.[14] O*NET's occupational classification is more detailed than the codes in the NLSY79, which are based on the Three-Digit Census Occupation Codes. We average scores over O*NET occupation codes that map to the same code in the Census Three-Digit Level Occupation Classification.

## 3.4 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 4 ASVAB test categories using the mapping created by the Defense Manpower Data Center (DMDC).[15] The DMDC selected 26 O*NET descriptors that were 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.

[14] For each descriptor, there is both a "level" and an "intensity" score. The ASVAB Career Exploration Program, which we describe below, uses only intensity and so do we.

[15] To increase the ASVAB's general appeal, the ASVAB Career Exploration Program was established by the U.S. 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). The mapping is available at: http://www.asvabprogram.com/downloads/Technical_Chapter_2010.pdf. .

Table II – List of Skills in O*NET

| | | | |
|---|---|---|---|
| Verbal and Math Skills | | | |
| 1. | Oral Comprehension | 2. | Written Comprehension |
| 3. | Deductive Reasoning | 4. | Inductive Reasoning |
| 5. | Information Ordering | 6. | Mathematical Reasoning |
| 7. | Number Facility | 8. | Reading Comprehension |
| 9. | Mathematics Skill | 10. | Science |
| 11. | Technology Design | 12. | Equipment Selection |
| 13. | Installation | 14. | Operation and Control |
| 15. | Equipment Maintenance | 16. | Troubleshooting |
| 17. | Repairing | 18. | Computers and Electronics |
| 19. | Engineering and Technology | 20. | Building and Construction |
| 21. | Mechanical | 22. | Mathematics Knowledge |
| 23. | Physics | 24. | Chemistry |
| 25. | Biology | 26. | English Language |
| Social Skills | | | |
| 1. | Social Perceptiveness | 2. | Coordination |
| 3. | Persuasion | 4. | Negotiation |
| 5. | Instructing | 6. | Service Orientation |

Second, after standardizing each dimension's standard deviation to be one, we reduce these 4 ASVAB categories into 2 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 among occupations.[16]

Likewise, to process the social dimension, we create a single index of social worker ability and another for the occupational skill requirement. 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 one. 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

[16]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 III – Correlations among Ability and Skill Requirement Scores

| | (a) Workers Ability | | | (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: (a) The correlations between each dimension of workers' ability are computed with 1,992 individuals in our sample. (b) The correlation between each dimension of workers' abilities and that of skill requirements in their current occupation are computed using 44,591 observations in our sample.

one, we take the first principal component. Both occupation- and worker-side data are then converted into percentile rank scores.

In Table III, we compute (a) the correlation of workers' verbal, math, and social ability scores for 1,992 individuals in our sample, and (b) the correlation between each dimension of workers' abilities and that of skill requirements in their current occupation for 44,591 observations in our sample. As it turns out in the left panel (a), while the ability scores are correlated to a certain degree, the correlation is not perfect. Between verbal and math ability scores, the correlation is 0.78—positive and high as expected. The correlation between cognitive and social skills is quite a bit lower, which is one of the attractions of using such a measure. In Part (b), we provide a crude look at sorting among workers, the correlation between the occupation's skill requirements and the worker's skill. We see that 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.

# 4 Empirical Methodology

In this section we introduce our main statistic—called *skill mismatch*—designed to measure the lack of fit between the skill portfolio possessed by an individual and the skill requirements of his occupation. We extend this notion creating another statistic—called *cumulative mismatch*—to analyze the persistent effect of past mismatch on current wages. We also present two additional statistics—called *positive* and *negative mismatch*—to analyze the effects of over- and underqualification at a given occupation. All these measures are incorporated into a Mincer regression framework.

## 4.1 An Empirical Measure of Skill Mismatch

The model has made clear the central role of the distance between a worker's abilities and the occupation's requirements. In our empirical measure, we try to operationalize this notion. We have measures of workers' abilities and occupational requirements from the NLSY79 and O*NET, respectively, which we convert into rank scores, as described in Section 3.

**(Contemporaneous) Mismatch.** Specifically, as in Section 2.3, $\tilde{A}_{i,j}$ 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 $q(\tilde{A}_{i,j})$ and $q(\tilde{r}_{c,j})$ 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:

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

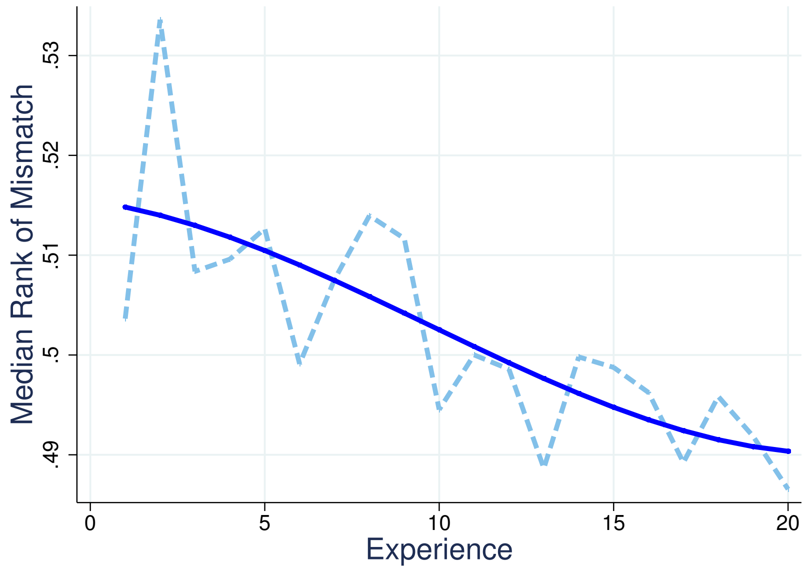
The weights are chosen to be the factor loadings from the first principal component, normalized to sum to 1.[17] Before we include the mismatch measure into our analysis, we rescale it so that its standard deviation is equal to 1.

Figure 2 shows the median of the mismatch measure by labor market experience in our sample. It decreases as workers' experience increases, implying that workers, on average, move to better matches over their life cycle. Table A.1 in Appendix A shows descriptive statistics for the mismatch measure, which 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*

[17] 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 $(\text{verbal}, \text{math}, \text{social}) = (0.43, 0.43, 0.12)$. 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 which sets an equal weight for all dimensions.

Figure 2 – Median Mismatch by Labor Market Experience




Note: Dashed lines show the actual statistic in the data. Solid lines show fitted values by a third-order polynomial. Y-axis show plots the median value of the rank scores of the mismatch measure for each experience group. The rank score for each worker-occupation match is is calculated among all worker-occupation matches observed in our NLSY79 sample.

mismatch as follows. Consider a worker who has worked at $p$ different occupations as of period $t$, whose indices are given by the vector $\{c(1), c(2), \dots, c(p)\}$. The tenure in each of these matches is given by the vector $\left\{\hat{T}_{c(1)}, \hat{T}_{c(2)} \dots, \hat{T}_{c(p-1)}, T_{c(p),t}\right\}$ 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 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:

$$\overline{m}_{i,t} \equiv \frac{m_{i,c(1)}\hat{T}_{c(1)} + m_{i,c(2)}\hat{T}_{c(2)} + \cdots + m_{i,c(p-1)}\hat{T}_{c(p-1)}}{\hat{T}_{c(1)} + \hat{T}_{c(2)} + \cdots + \hat{T}_{c(p-1)}} = \frac{\sum_{s=1}^{p-1} m_{i,c(s)}\hat{T}_{c(s)}}{\sum_{s=1}^{p-1} \hat{T}_{c(s)}}. \tag{6}$$

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 analogue of the cumulative mismatch term in equation (5). This variable represents the lingering effect of previous matches 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 it, as in our model, then poor matches in past occupations can significantly reduce current wages.

**Positive vs. Negative Mismatch.** Equation (5) in Section 2 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

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

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 over qualified (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 4.1.

## 4.2 Empirical Specification of the Wage Equation

We begin with the standard Mincer wage regression and augment it with measures of mismatch to investigate whether current or cumulative mismatch (or both) matters for current wages. If current mismatch matters for the level of wages, then it would lend support to our interpretation of our measure as a proxy for the current occupational match quality, which has been viewed as an unobservable component of the regression residual by much of the extant literature.[18] 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.

### Instrumenting Tenure Variables

As was recognized by Altonji and Shakotko (1987), wage regressions that include a tenure variable are potentially affected by an endogeneity problem that comes from

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

omitted individual- and match-specific factors, which are likely to be correlated with experience and tenure variables. We deal with this issue building on a long list of studies, such as Altonji and Shakotko (1987), Topel (1991), and Altonji and Williams (2005), to instrument for experience and tenure variables.

First, to understand why OLS estimates might be biased, consider the wage equation for individual $i$ who is working with employer $l$ in occupation $c$ at time $t$:

$$\ln w_{i,l,c,t} = X'_{i,t}\beta + \alpha_1 J_{i,l,t} + \alpha_2 T_{i,c,t} + \alpha_3 E_{i,t} + \alpha_4 OJ_{i,t} + \theta_{i,l,c,t}, \tag{7}$$

where $X_{i,t}$ is a vector of worker characteristics, $J_{i,l,t}$ is employer tenure, $T_{i,c,t}$ is occupational tenure, $E_{i,t}$ is labor market experience, and $OJ_{i,t}$ is a dummy variable that indicates a continuing job. The last term $\theta_{i,l,c,t}$ in (7) can be decomposed into

$$\theta_{i,l,c,t} = \nu_i + \mu_{i,l} + \phi_{i,c} + \epsilon_{i,t},$$

where $\nu_i$ is an individual-specific component, $\mu_{i,l}$ is an employer-match component, $\phi_{i,c}$ is an occupational match component, and $\epsilon_{i,t}$ is an orthogonal error.

Specifically, the coefficient on occupational tenure in the model described by Equation (7) could be biased because the duration of the occupational match is endogenous, and could depend on the level of $\phi_{i,c}$. A valid instrument for $T_{i,c,t}$ is given by $\widetilde{T}_{i,c,t} \equiv T_{i,c,t} - \overline{T}_{i,c}$, where $\overline{T}_{i,c}$ is the average tenure of individual $i$ during the spell of working in occupation $c$:

$$\overline{T}_{i,c} \equiv \frac{1}{\hat{T}_c} \sum_{t=1}^{\hat{T}_c} T_{i,c,t}$$

In the above expression, $\hat{T}_c$ is the total length of the spell at occupation $c$. For example, if an individual is observed in an occupation at tenure 1 through 5 years, then $\hat{T}_c$ is 5 years, and $\overline{T}_{i,c}$ is 3 years ($= (1+2+3+4+5)/5$). By construction, $\widetilde{T}_{i,c,t}$ is orthogonal to $\phi_{i,c}$. An appropriate correction for higher order terms is also available: we instrument $(T_{i,c,t})^q$ with $\widetilde{T}^q_{i,c,t} \equiv (T_{i,c,t})^q - \left(\overline{T_{i,c}}\right)^q$ for $q > 1$, where $\left(\overline{T_{i,c}}\right)^q$ is the average of the occupational tenure term raised to power $q$. Because our set of regressors also includes several 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.[19]

[19]Similar to an occupational match component, an employer match component is potentially corre-

For our regressions, we expand on Equation (7) and incorporate our mismatch measure, $m_{i,c}$, and cumulative mismatch measure, $\bar{m}_{i,t}$, and follow the same instrumenting scheme outlined above. We also include the interaction of $m_{i,c}$ with $T_{i,c,t}$, so that mismatch is allowed to have both a fixed level effect on wages as well as an effect that is allowed to change during the occupational tenure. The wage regression then is

$$\begin{aligned}\ln w_{i,l,c,t} =& X'_{i,t}\beta + \gamma_1 m_{i,c} + \gamma_2 \left(m_{i,c} \times T_{i,c,t}\right) + \gamma_3 \overline{m}_{i,t} \\ &+ \gamma_4 \overline{A}_i + \gamma_5 \left(\overline{A}_i \times T_{i,c,t}\right) + \gamma_6 \overline{r}_c + \gamma_7 \left(\overline{r}_c \times T_{i,c,t}\right) \\ &+ \Phi_1 \left(J_{i,l,t}\right) + \Phi_2 \left(T_{i,c,t}\right) + \Phi_3 \left(E_{i,t}\right) + \alpha_4 OJ_{i,t} + \theta_{i,l,c,t}, \end{aligned} \tag{8}$$

where $\Phi_1$, $\Phi_2$, and $\Phi_3$ are polynomials.[20] The vector $X_{i,t}$ includes education and demographics dummies, $\overline{A}_i$ is the ability of worker $i$ averaged across skill dimensions, and $\overline{r}_c$ is the skill requirement of occupation $c$ averaged over skill dimensions.[21] 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. Finally, when estimating Equation (8), we include one-digit level occupation and industry dummies.

## 4.3 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 three reasons for this uncertainty. First, and most important, these tests are clearly not perfect measures of a worker's abilities, but are probably best viewed as noisy signals. The worker is likely to have observed a range of other signals by the time he entered the labor market, and his beliefs at that time are the results of those signals, the test scores being one (but possibly important) component.

Second, even if these test scores were perfect measures of ability, it is important to note that the NLSY respondents were not told their rank in the test, but were rather

lated with employer tenure and the dummy variable of a continuation of a job. An individual-specific component is potentially correlated with labor market experience.

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

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

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 20th and 40th percentiles. Just as in our theoretical model, this is a noisy signal centered around the true mean. As the econometrician, we see the entire NLSY79 sample, so we can compute the worker's precise rank.

Third, and furthermore, as the econometrician, we can process these test scores extract more information than what the respondents could do. For example, we removed age affects 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.

# 5 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 the Mincer regression framework, then study its relationship to switching probability and the direction of switching. We find that mismatch and its interaction with tenure 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.

## 5.1 Mismatch and Wages

Table IV presents the key results from our wage regressions. We present the main coefficients here and the rest are relegated to Appendix A. The first column includes our measure of mismatch into a standard wage regression. The next adds its interaction with occupational tenure. In the third column, 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" and show "OLS" results for robustness.

In column (1) of Table IV, contemporaneous mismatch has an estimated coefficient of –0.027 (and is significant at 1% level), indicating a strong effect on wages. To give a more precise economic interpretation to this coefficient, recall that we have normalized the standard deviation of mismatch to 1, so wages are predicted to be about 5.4% (2.7%

Table IV – Wage Regressions with Mismatch

| | (1) IV | (2) IV | (3) IV | (4) OLS | (5) OLS | (6) OLS |
|---|---|---|---|---|---|---|
| Mismatch | –0.0271** | –0.0145** | –0.0054 | –0.0254** | –0.0214** | –0.0147** |
| Mismatch × Occ Tenure | | –0.0020** | –0.0024** | | –0.0006 | –0.0006 |
| Cumul Mismatch | | | –0.0355** | | | –0.0364** |
| Worker Ability (Mean) | 0.2466** | 0.2475** | 0.3408** | 0.2588** | 0.2585** | 0.3426** |
| Worker Ability × Occ Tenure | 0.0166** | 0.0161** | 0.0140** | 0.0130** | 0.0129** | 0.0127** |
| Occ Reqs (Mean) | 0.1529** | 0.1528** | 0.1576** | 0.2096** | 0.2095** | 0.2224** |
| Occ Reqs × Occ Tenure | 0.0155** | 0.0154** | 0.0161** | 0.0070** | 0.0069** | 0.0061** |
| Observations | 44,591 | 44,591 | 33,072 | 44,591 | 44,591 | 33,072 |
| $R^2$ | 0.355 | 0.355 | 0.313 | 0.371 | 0.371 | 0.332 |

Note: $** \; p < 0.01$, $* \; p < 0.05$, $\dagger \; p < 0.1$. All regressions include a constant, terms for demographics, occupational tenure, employer tenure, work experience, and dummies for one-digit-level occupation and industry. Standard errors are computed as robust Huber-White sandwich estimates. More detailed regression results are in Appendix A.

× 2) lower for workers whose mismatch is one standard deviation above the mean relative to those one standard deviation below it.

In the next column, we introduce a mismatch interaction with occupation tenure. Now the level effect becomes smaller (–0.014 instead of –0.027), partly replaced by a negative tenure effect (also significant at 1% level). Thus, not only does mismatch depress initial wages, it also leads to slower wage growth over the duration of the match. Beyond 7 years, the overall depression in wages due to slower growth rate dominates the losses due to the initial impact.

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, displacing the level effect of current mismatch, which becomes smaller and insignificant. The tenure effect of current mismatch is unaffected however. To help interpret the size of these coefficients, Table V computes the implied wage losses using specification (3). Looking at the effect of current mismatch, we see that the 90th percentile worst-matched workers face 8.8% lower wages after 10 years of occupational tenure compared with a perfectly matched worker. The difference between the 90th percentile and the 10th percentile of *mismatch* is about 4.4% after 5 years of occupational tenure and widens to 7.4% after 10 years. Comparing the 90th percentile to the 10th percentile of *cumulative mismatch*, we see a wage difference of 8.9%.

Table V – Wage Losses from Mismatch & Cumulative Mismatch

| Mismatch Degree (High to Low) | *Mismatch* Effect 5 years | 10 years | 15 years | *Cumul. Mismatch* Effect |
|---|---|---|---|---|
| 90% | –0.052 | –0.088 | –0.124 | –0.116 |
| | (0.010) | (0.014) | (0.024) | (0.012) |
| 70% | –0.034 | –0.057 | –0.080 | –0.081 |
| | (0.007) | (0.009) | (0.015) | (0.008) |
| 50% | –0.023 | –0.039 | –0.054 | –0.062 |
| | (0.004) | (0.006) | (0.010) | (0.006) |
| 30% | –0.015 | –0.026 | –0.036 | –0.046 |
| | (0.003) | (0.004) | (0.007) | (0.005) |
| 10% | –0.008 | –0.014 | –0.020 | –0.027 |
| | (0.002) | (0.002) | (0.004) | (0.003) |

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

Finally, for comparison purposes, the last three columns of Table IV reports the OLS 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 and OLS. As we discussed in Section 4.2, the return to tenure is biased because it is correlated with unobservable match quality. The instruments reduce the return to occupational tenure itself (see Table A.2 in Appendix A) by a factor of about 3, precisely because the OLS estimate on tenure takes some variation from the mismatch times tenure 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.

**Three Dimensions of Skill Mismatch**

In Table VI, 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, $m_{i,c,j} \equiv \left| q(\tilde{A}_{i,j}) - q(\tilde{r}_{c,j}) \right|$. As before, we scale each dimension to have a standard deviation of one.

Looking at math and verbal skills in Table VI, we see a pattern emerge: mismatch in either dimension has a negative effect on wages but with a key difference: 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 persistent growth rate effect. In the most general model of column (3), the interaction term for verbal mismatch has a

Table VI – Wage Regressions with Mismatch by Components

| | (1) IV | (2) IV | (3) IV | (4) OLS | (5) OLS | (6) OLS |
|---|---|---|---|---|---|---|
| Mismatch Verbal | –0.0147** | 0.0030 | 0.0139* | –0.0150** | –0.0053 | 0.0027 |
| Mismatch Math | –0.0130** | –0.0171** | –0.0203** | –0.0109** | –0.0172** | –0.0182** |
| Mismatch Social | –0.0049† | –0.0034 | 0.0067 | –0.0041 | –0.0046 | 0.0017 |
| Mismatch Verbal × Occ Tenure | | –0.0028** | –0.0045** | | –0.0015** | –0.0026** |
| Mismatch Math × Occ Tenure | | 0.0006 | 0.0021* | | 0.0010† | 0.0020** |
| Mismatch Social × Occ Tenure | | -0.0002 | –0.0010 | | 0.0001 | –0.0005 |
| Cumul Mismatch Verbal | | | –0.0123** | | | –0.0107* |
| Cumul Mismatch Math | | | –0.0252** | | | –0.0274** |
| Cumul Mismatch Social | | | –0.0083* | | | –0.0073* |
| Verbal Ability | –0.0440† | –0.0486† | 0.0081 | 0.0112 | 0.0066 | 0.0158 |
| Math Ability | 0.2949** | 0.3001** | 0.3405** | 0.2510** | 0.2547** | 0.3238** |
| Social Ability | 0.0837** | 0.0836** | 0.1017** | 0.0855** | 0.0853** | 0.1137** |
| Verbal Ability × Occ Tenure | 0.0125** | 0.0126** | 0.0088* | 0.0047† | 0.0051* | 0.0066* |
| Math Ability × Occ Tenure | 0.0006 | –0.0003 | 0.0011 | 0.0037 | 0.0031 | 0.0023 |
| Social Ability × Occ Tenure | 0.0072** | 0.0073** | 0.0075** | 0.0076** | 0.0076** | 0.0063** |
| Occ Reqs Verbal | 0.0771 | 0.0757 | 0.0913 | 0.1414* | 0.1482* | 0.1214† |
| Occ Reqs Math | 0.1112† | 0.1075† | 0.1065 | 0.1004† | 0.0917† | 0.1361* |
| Occ Reqs Social | –0.0932** | –0.0894** | –0.0978** | –0.0817** | –0.0803** | –0.0827** |
| Occ Reqs Verbal × Occ Tenure | –0.0071 | –0.0070 | –0.0098 | –0.0229** | –0.0245** | –0.0205* |
| Occ Reqs Math × Occ Tenure | 0.0164† | 0.0172* | 0.0175† | 0.0228** | 0.0249** | 0.0175* |
| Occ Reqs Social × Occ Tenure | 0.0100** | 0.0092** | 0.0131** | 0.0109** | 0.0106** | 0.0135** |
| Observations | 44,591 | 44,591 | 33,072 | 44,591 | 44,591 | 33,072 |
| $R^2$ | 0.358 | 0.358 | 0.317 | 0.374 | 0.375 | 0.335 |

Note: $** \ p < 0.01$, $* \ p < 0.05$, $\dagger \ p < 0.1$. All regressions include a constant, terms for demographics, occupational tenure, employer tenure, work experience, and dummies for one-digit-level occupation and industry. Standard errors are computed as robust Huber-White sandwich estimates. More detailed regression results are in Appendix A.

coefficient of –0.0045 (and highly significant), implying a 9% wage gap after 10 years of tenure between the top and bottom 10% mismatched. Turning to the effects of social mismatch, it has a weaker effect overall, though still negative.

TABLE VII – Wage Regressions with Positive and Negative Mismatch

| | (1) IV | (2) IV | (3) IV | (4) OLS | (5) OLS | (6) OLS |
|---|---|---|---|---|---|---|
| Positive Mismatch | –0.0143** | 0.0031 | 0.0127* | –0.0134** | –0.0066 | 0.0019 |
| Negative Mismatch | 0.0374** | 0.0218** | 0.0253** | 0.0338** | 0.0218** | 0.0272** |
| Pos. Mismatch × Occ Tenure | | –0.0028** | –0.0030** | | –0.0011* | –0.0005 |
| Neg. Mismatch × Occ Tenure | | 0.0025** | 0.0021* | | 0.0019** | 0.0018** |
| Cumul Positive Mismatch | | | –0.0168** | | | –0.0234** |
| Cumul Negative Mismatch | | | 0.0093* | | | 0.0026 |
| Observations | 44,591 | 44,591 | 33,072 | 44,591 | 44,591 | 33,072 |
| $R^2$ | 0.336 | 0.336 | 0.290 | 0.351 | 0.351 | 0.308 |

Note: $**\ p < 0.01$, $*\ p < 0.05$, $\dagger\ p < 0.1$. All regressions include a constant, terms for demographics, occupational tenure, employer tenure, work experience, and dummies for one-digit-level occupation and industry. Standard errors are computed as robust Huber-White sandwich estimates. More detailed regression results are in Appendix A.

Interestingly, the same difference between math and verbal skills is seen in the effects of *ability* on wages (lower panel of Table VI): in the first three columns, math ability has a large level effect (ranging from 29% to 34% across specifications) but little growth effect, whereas verbal ability has little level effect but a robust growth rate effect (ranging from 0.9 to 1.3% per year) on wages. Social skills have an effect broadly similar to that of verbal: the level effect ranges from 8.4% to 10.2%, whereas the growth rate effect is significant and only slightly smaller than that of verbal skills (about 0.7% per year). One interpretation of this difference might be that math skills are easier to observe by employers and the market and so are priced immediately, whereas verbal and social skills capture some more subtle traits that are revealed more slowly over time, leading to a growth rate effect.[22]

Turning to the effects of cumulative mismatch, it is negative in all three dimensions and statistically significant at 5% level or higher. As can be expected however, the individual magnitudes are smaller than in the previous table for total mismatch. Still, the effect for verbal skills is equivalent to about 6 years of mismatch in the current occupation, combining the immediate and tenure effects; cumulative social mismatch is equivalent to about 14 years of mismatch in the current occupation, but this is mainly because the effects of current mismatch are small.

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

**Positive and Negative Mismatch**

Next, in Table VII we investigate the effects of positive and negative mismatch, as defined in Section 4.1, on wages. In Column (1) of Table VII, both positive and negative mismatch reduce wages.[23] However, the effect is not perfectly symmetric; the coefficient on negative mismatch is about 2.5 times larger in Column (1), and when we add the interaction with tenure in Column (2), we see that being overqualified mostly slows wage *growth* rather than having an immediate effect.

Column (3) is especially interesting and also consistent with our model: a history of mismatch, either positive or negative, implies lower wages because mismatch in either direction dampens human capital accumulation. This is not the case under the standard Ben-Porath model as shown in Appendix B.2 and discussed in Footnote 10.

Overall, the results in Tables IV to VII collectively speak to the importance of skill mismatch for the determination of wages. In other words, wages are based not only on the characteristics of the worker and the job separately, but also on the interaction between the two. Further, we see that the tenure effect is especially important. As our model suggested in Equation (5), a worker's wages reflect the history of skill mismatch. Just as the previous literature (e.g., Altonji and Shakotko (1987); Topel (1991)) had suspected that match quality affects the returns to tenure, our results provide direct evidence that it does.

## 5.2 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. What is the effect of mismatch on occupational switching behavior? Does skill mismatch predict the likelihood and direction of occupational switches?

To answer these questions, we estimate a linear probability model for occupational switching on the same set of regressors as in our wage regressions, of which we are chiefly interested in the contribution of mismatch in the current occupation. Table VIII displays our baseline estimates in which we instrument occupational tenure, as we did in the wage regressions. For comparison, again, we also run the regressions by OLS.

[23] Recall that negative mismatch adds all skill dimensions for which the worker is underqualified (so by definition it is a negative number), and the positive estimated coefficients imply that negative mismatch reduces wages. Furthermore, we do not include terms for the level of worker abilities or occupational requirements because in breaking apart the absolute value of the mismatch measure, we would encounter problems of collinearity between the positive and negative mismatch measure and those terms.

Table VIII – Regressions for the Probability of Occupational Switch

| | (1) LPM-IV | (2) LPM-IV | (3) LPM-IV | (4) LPM | (5) LPM | (6) LPM |
|---|---|---|---|---|---|---|
| Mismatch | 0.0135** | | | 0.0066** | | |
| Mismatch Verbal | | 0.0076** | | | 0.0053* | |
| Mismatch Math | | 0.0074** | | | 0.0025 | |
| Mismatch Social | | 0.0007 | | | –0.0012 | |
| Positive Mismatch | | | 0.0134** | | | 0.0087** |
| Negative Mismatch | | | –0.0130** | | | –0.0028 |
| Worker Ability (Mean) | –0.0370** | –0.0370** | - | –0.0208† | –0.0211† | - |
| Worker Ability × Occ Tenure | –0.0003 | –0.0003 | - | 0.0019* | 0.0020* | - |
| Occ Reqs (Mean) | –0.0333* | –0.0334* | - | –0.1225** | –0.1223** | - |
| Occ Reqs × Occ Tenure | –0.0052** | –0.0052** | - | 0.0106** | 0.0106** | - |
| Observations | 41,596 | 41,596 | 41,596 | 41,596 | 41,596 | 41,596 |

Note: ** $p < 0.01$, * $p < 0.05$, † $p < 0.1$. All regressions include a constant, terms for demographics, occupational tenure, employer tenure, work experience, and dummies for one-digit-level occupation and industry. Standard errors are computed as robust Huber-White sandwich estimates. More detailed regression results are in Appendix A.

Notice that the effect of current mismatch on the probability of switching occupations is always positive and significant at the 1% level, with the exception of social mismatch in column (2). To give a better idea about the magnitudes implied by these coefficients, in Table IX we compute the occupational switching probabilities for workers at various percentiles of the mismatch distribution, using the specifications in Columns (1) and (2). A worker who is in the 90th percentile of the mismatch distribution is 3.4 percentage points more likely to switch occupations than an otherwise comparable worker who is in the 10th percentile, a difference corresponding to about 21% of the average switching rate.

Splitting mismatch into components (last three columns of Table IX), we see that the 90th to 10th percentile gap for the switching probabilities are 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, social mismatch seems to only have a modest effect on outcomes once we account for math and verbal skills.

In Column 3 of Table VIII, we see that the effects are roughly symmetric, with increased switching probability similarly associated with positive and negative mismatch. Workers whose skills are worse than their occupations or better are both more likely to switch occupations. These findings are consistent with those of Groes et al. (2015) that

Table IX – Effect of Mismatch on Occupational Switching Probability

| Mismatch Degree (High to Low) | *Mismatch* Effect | Effect by *Component* | | |
|---|---|---|---|---|
| | | Verbal | Math | Social |
| 90% | 0.0407 | 0.0209 | 0.0203 | 0.0020 |
| | (0.0069) | (0.0078) | (0.0076) | (0.0062) |
| 70% | 0.0263 | 0.0129 | 0.0123 | 0.0013 |
| | (0.0045) | (0.0048) | (0.0046) | (0.0038) |
| 50% | 0.0178 | 0.0082 | 0.0076 | 0.0008 |
| | (0.0030) | (0.0030) | (0.0029) | (0.0024) |
| 30% | 0.0119 | 0.0043 | 0.0042 | 0.0004 |
| | (0.0020) | (0.0016) | (0.0016) | (0.0012) |
| 10% | 0.0065 | 0.0014 | 0.0013 | 0.0001 |
| | (0.0011) | (0.0005) | (0.0005) | (0.0004) |

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

Table X – Average Change in Skills When Switching Occupations

| | (a) Fraction of Positive Switch | | | (b) Average Change in Percentile | | |
|---|---|---|---|---|---|---|
| Sample Group | Verbal | Math | Social | Verbal | Math | Social |
| All Workers | 0.56 | 0.55 | 0.55 | 2.44 | 1.94 | 1.56 |
| Less than High School | 0.54 | 0.54 | 0.53 | 1.14 | 1.00 | 0.39 |
| High School | 0.56 | 0.55 | 0.54 | 1.75 | 1.21 | 1.00 |
| Some College | 0.57 | 0.56 | 0.57 | 3.59 | 3.00 | 2.53 |

Note: In Panel (a), we report the fraction of workers who move to an occupation that requires higher skill level in each skill dimension. Panel (b) lists average changes in the percentile rank upon an occupational switch.

workers better sorted into their occupation are less likely to switch out. Of course, the measure of mismatch used in our analysis is based on the portfolio of skills, rather than wages as was done in that paper. But regardless, both papers tell a consistent story.

With our multidimensional measure, we can go one step further and examine if occupational switches show well-defined directions in the skill space. This is what we explore next.

## 5.3 Switch Direction

Not only do mismatched workers switch occupations more frequently, but their switches are also directional as seen in Table X. Workers who are overqualified—their abilities are ranked higher than the skill requirements of their occupations—tend to switch to occupations with higher skill requirements. The converse is true of workers who are underqualified.

In general, switches tend to correct past mismatch. We see this in Panels (a) through (c) of Figure 3. We plot on the vertical axis changes in each occupational skill requirement for every worker who switches occupation and on the horizontal we plot the last positive or negative mismatch in that skill. Here, a change in occupational skill requirement in skill $j$ is defined as the difference between the skill requirement in the last occupation and that in the current one, i.e., $q(\tilde{r}_{c(p),j}) - q(\tilde{r}_{c(p-1),j})$. Positive and negative mismatch *in skill* $j$ is defined as in Section 4.1, but using only one dimension at a time.[24] To give the scatter plots some shape in Figure 3, we run a local polynomial regression for observations that have strictly positive or negative mismatch in skill $j$.

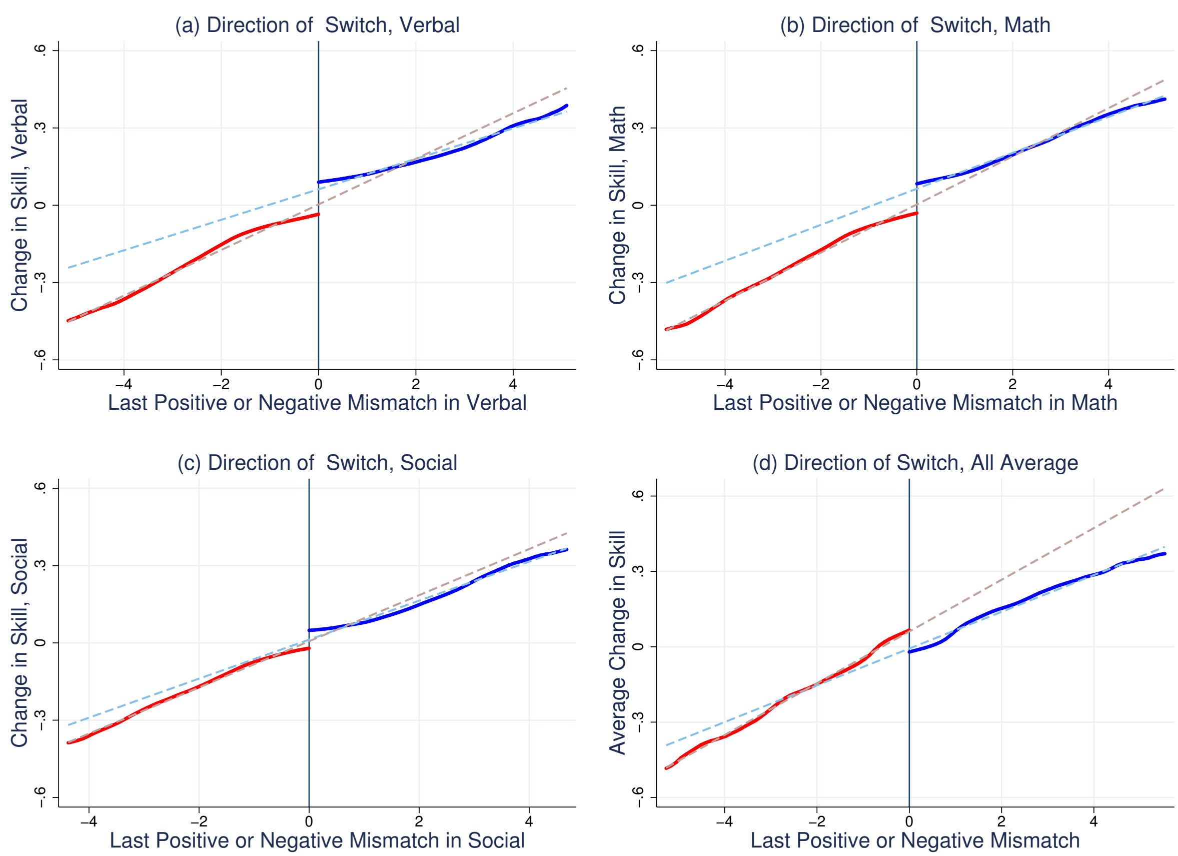
As shown in these panels, the upward-sloping curves on both sides of zero mean that individuals who are overqualified in skill $j$ (the right half of the axes) tend to choose their next occupation with a higher skill requirement, whereas the opposite is true for individuals who are underqualified. The relationship is positive and nearly linear, such that the more mismatched the worker is in the last occupation, the larger the change in occupational requirements of that skill in the next switch. Furthermore, the right branch has a noticeably smaller slope than the left branch in panels (a) and (b), indicating that workers overqualified in verbal and math skills increase the skill requirements in the next occupation by less than the amount under-qualified workers reduce them by. This is not the case for social skills where the two branches are nearly parallel to each other. Finally, panel (d) plots the same relationship by aggregating across all three skill types, which again shows the same patterns.[25]

One drawback of the visual analysis that underlies Figure 3 is that it only documents a univariate relationship—how requirement in one skill dimension changes as a function of current mismatch in the same dimension. To investigate richer dependencies, we turn to a regression framework. Specifically, we regress the change (upon switching occupations) in skill requirement $j$ on positive and negative mismatch in all three skill dimensions

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

[25] 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 mismatch in 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.

Figure 3 – Non-Parametric Plots of Direction of Switch




Note: 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.

for the worker's last occupation.[26] We also include education, demographics, employer tenure, occupational tenure, experience, and the indicator for continuation of job for the last match, and occupation and industry dummies for the current match. The right-hand-side variables are the same as in our wage regressions except that here we omit average worker abilities and occupational requirements, again because of collinearity in wage regressions with positive and negative mismatch.

Columns (1) through (3) of Table XI 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.

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

Table XI – Regressions for Direction of Switch

| Dependent variable → | (1) Verbal | (2) Math | (3) Social | (4) All Average |
|---|---|---|---|---|
| Last Pos. Mismatch, Verbal | *0.0316*** | 0.0097* | 0.0143** | |
| Last Neg. Mismatch, Verbal | *0.0838*** | 0.0536** | 0.0216** | |
| Last Pos. Mismatch, Math | 0.0599** | *0.0898*** | 0.0021 | |
| Last Neg. Mismatch, Math | 0.0558** | *0.0893*** | 0.0076 | |
| Last Pos. Mismatch, Social | 0.0061† | 0.0046 | *0.0774*** | |
| Last Neg. Mismatch, Social | 0.0264** | 0.0166** | *0.1043*** | |
| Last Positive Mismatch | | | | 0.0751** |
| Last Negative Mismatch | | | | 0.1143** |
| Observations | 6,594 | 6,594 | 6,594 | 6,594 |
| $R^2$ | 0.485 | 0.458 | 0.417 | 0.487 |

Note: $**$ $p < 0.01$, $*$ $p < 0.05$, † $p < 0.1$. All regressions include a constant, terms for demographics, occupational tenure before switch, employer tenure before switch, work experience before switch, and dummies for one-digit-level occupation and industry for the last job held. Standard errors are computed as robust Huber-White sandwich estimates. More detailed regression results are in Appendix A.

There are several takeaways from this table. First, the positive coefficients on all regressors confirm the main message of Figure 3: skill change upon switching is an increasing function of current mismatch, so switching works to reduce skill mismatch. The difference here is that this is true even when we consider mismatch along more than one dimension. 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 vs 0.0599). However, if the worker was overqualified in social dimension this has little impact (coefficient of 0.0061) on verbal requirements change.

Second, the other two columns tell a similar story. The change in math skill requirements is responsive to verbal mismatch but much less so to social, whereas change in social is mostly responsive to mismatch in its own dimension. These results echoes the same theme as before that math and verbal skills are distinct, yet closely connected, whereas social skills have more of a life on their own.

Table XII – Effect of Last Mismatch on Change in Skills

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

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

Third, the asymmetry highlighted in Figure 3 also manifests itself here, with the exception of mismatch in math skills. That is, workers who are underqualified move to occupations with an aggressive reduction in that skill requirement, whereas overqualified workers choose a more modest increase in skill requirements in their next occupation.

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 90th, 50th, and 10th percentile rank of each measure in Table XII using (diagonal entries from the) regression results. For example, a highly overqualified worker in the verbal dimension, in the 90th percentile of positively mismatched workers, will choose occupations that require 9.95 percentiles higher verbal skill requirement. A similarly underqualified worker (in the 10th percentile of negative mismatch) reduces his verbal skill requirements by 24.25 percentiles in his next occupation. Similarly, large adjustments are seen for math and social skills in the next two columns. These results show that mismatch is a particularly useful measure for predicting the nature of occupational switching, including both is likelihood and its direction.

# 6 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. Mismatch arises in our model due to workers' imperfect information about their learning abilities, which causes them to choose occupations that are either above or below their optimal level. As workers discover their true abilities over the life cycle, mismatch gradually declines and 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 after controlling for a long list to standard regressors, which includes worker abilities constructed from ASVAB and occupation requirements constructed from 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 highlights a new aspect of occupational switching: workers choose their next occupation 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. Another conclusion we draw is that social skills behave somewhat differently from math and verbal skills. Although social ability appears to matter for wages, mismatch between a worker and an occupation along this dimension does not seem to affect wages too much. The same can be said about switching behavior where mismatch in social skills does not greatly affect the change in other skill requirements upon occupation switches.

These findings should only serve to motivate further work on the mechanisms involved in learning and occupational choice. The empirical evidence we presented suggests a strong link between learning 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. We pursue this approach in separate ongoing work.

## References


**Altonji, Joseph G. and Charles R. Pierret**, "Employer Learning and Statistical Discrimination," *Quarterly Journal of Economics*, 2001, *116* (1), 313–350.

_ **and Nicolas Williams**, "Do Wages Rise with Job Seniority? A Reassessment," *Industrial & Labor Relations Review*, 2005, *58* (3), 370–397.

_ **and Robert A. Shakotko**, "Do Wages Rise with Job Seniority?," *Review of Economic Studies*, 1987, *54* (3), 437–459.

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

**Antonovics, Kate and Limor Golan**, "Experimentation and Job Choice," *Journal of Labor Economics*, 2012, *30* (2), 333–366.

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

**Bowles, Samuel, Herbert Gintis, and Melissa Osborne**, "The Determinants of Earnings: A Behavioral Approach," *Journal of Economic Literature*, 2001, *39* (4), 1137–1176.

**Farber, Henry S. and Robert Gibbons**, "Learning and Wage Dynamics," *Quarterly Journal of Economics*, 1996, *111* (4), 1007–1047.

**Flinn, Christopher J.**, "Wages and Job Mobility of Young Workers," *Journal of Political Economy*, 1986, *94* (3), S88–S110.

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

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

**Gervais, Martin, Nir Jaimovich, Henry E. Siu, and Yaniv Yedid-Levi**, "What Should I Be When I Grow Up? Occupations and Unemployment over the Life Cycle," *NBER Working Paper*, 2014, (20628).

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

**Groes, Fane, Philipp Kircher, and Iourii Manovskii**, "The U-Shapes of Occupational Mobility," *Review of Economic Studies*, 2015, *82* (2), 659–692.

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

**Huggett, Mark, Gustavo Ventura, and Amir Yaron**, "Sources of Lifetime Inequality," *American Economic Review*, 2011, *101* (7), 2923–2954.

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

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

_ **and Robert Moffitt**, "An Estimate of a Sectoral Model of Labor Mobility," *Journal of Political Economy*, 1990, *98* (4), 827–852.

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

_ **and** _ , "Occupational Specificity of Human Capital," *International Economic Review*, 2009b, *50* (1), 63–115.

**Lise, Jeremy and Fabien Postel-Vinay**, "Multidimensional Skills, Sorting, and Human Capital Accumulation," *Working Paper*, 2015. University College London.

**Michaud, Amanda M. and David Wiczer**, "Occupational Hazards and Social Disability Insurance," *Working Paper*, 2014. Federal Reserve Bank of St. Louis.

**Miller, Robert A.**, "Job Matching and Occupational Choice," *Journal of Political Economy*, 1984, *92* (6), 1086–1120.

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

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

**Nagypal, Eva**, "Learning by Doing vs. Learning about Match Quality: Can We Tell Them Apart?," *Review of Economic Studies*, 2007, *74* (2), 537–566.

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

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

**Parent, Daniel**, "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*, 2000, *18* (2), 306–323.

**Pavan, Ronni**, "Career Choice and Wage Growth," *Journal of Labor Economics*, 2011, *29* (3), 549–587.

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

**Rogerson, Richard, Robert Shimer, and Randall Wright**, "Search-Theoretic Models of the Labor Market: A Survey," *Journal of Economic Literature*, 2005, *43* (4), 959–988.

**Rosen, Sherwin**, "Learning and Experience in the Labor Market," *Journal of Human Resources*, 1972, *7* (3), 326–342.

**Sanders, Carl**, "Skill Accumulation, Skill Uncertainty, and Occupational Choice," *mimeo*, 2014.

**Topel, Robert**, "Specific Capital, Mobility, and Wages: Wages Rise with Job Seniority," *Journal of Political Economy*, 1991, *99* (1), 145–176.

**Ware, John E., Mark Kosinski, and Susan D. Keller**, *SF-12: How to Score the SF-12 Physical and Mental Health Summary Scales*, Health Institute, New England Medical Center, 1995.

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

# Supplemental Online Appendix

# NOT FOR PUBLICATION

# Appendix

## A Additional Tables

### A.1 Descriptive Statistics for Mismatch Measure

Table A.1 – Descriptive Statistics for Mismatch Measure

| | Mismatch | |
|---|---|---|
| Group Name | Mean | Std. Dev. |
| *All Observations* | 1.56 | 1 |
| *By Educational Group* | | |
| Less than High School | 1.50 | 1.13 |
| High School | 1.63 | 1.00 |
| Some College | 1.52 | 0.96 |
| *By Race* | | |
| Hispanic | 1.61 | 1.05 |
| Black | 1.52 | 1.02 |
| Non-Black, Non-Hispanic | 1.56 | 0.99 |
| *By Industry* | | |
| Agriculture, Forestry, Fisheries | 1.59 | 1.15 |
| Mining | 1.63 | 1.02 |
| Construction | 1.70 | 1.13 |
| Manufacturing | 1.57 | 1.00 |
| Transportation, Communications, Util. | 1.47 | 0.90 |
| Wholesale and Retail Trade | 1.57 | 0.93 |
| Finance, Insurance and Real Estate | 1.37 | 0.86 |
| Business and Repair Services | 1.63 | 0.99 |
| Personal Services | 1.68 | 1.08 |
| Entertainment and Recreation Services | 1.92 | 1.20 |
| Professional and Related Services | 1.46 | 0.94 |
| Public Administration | 1.49 | 0.97 |

## A.2 Regression Tables

Table A.2 – Wage Regression with Mismatch (Full Results)

| | (1) IV | (2) IV | (3) IV | (4) OLS | (5) OLS | (6) OLS |
|---|---|---|---|---|---|---|
| Mismatch | -0.0271** | -0.0145** | -0.0054 | -0.0254** | -0.0214** | -0.0147** |
| | (0.0027) | (0.0044) | (0.0052) | (0.0027) | (0.0037) | (0.0045) |
| Mismatch × Occ Tenure | | -0.0020** | -0.0024** | | -0.0006 | -0.0006 |
| | | (0.0006) | (0.0007) | | (0.0004) | (0.0005) |
| Cumul Mismatch | | | -0.0355** | | | -0.0364** |
| | | | (0.0035) | | | (0.0035) |
| Worker Ability (Mean) | 0.2466** | 0.2475** | 0.3408** | 0.2588** | 0.2585** | 0.3426** |
| | (0.0189) | (0.0189) | (0.0225) | (0.0164) | (0.0165) | (0.0199) |
| Worker Ability × Occ Tenure | 0.0166** | 0.0161** | 0.0140** | 0.0130** | 0.0129** | 0.0127** |
| | (0.0023) | (0.0023) | (0.0028) | (0.0017) | (0.0017) | (0.0020) |
| Occ Reqs (Mean) | 0.1529** | 0.1528** | 0.1576** | 0.2096** | 0.2095** | 0.2224** |
| | (0.0193) | (0.0193) | (0.0222) | (0.0171) | (0.0171) | (0.0203) |
| Occ Reqs × Occ Tenure | 0.0155** | 0.0154** | 0.0161** | 0.0070** | 0.0069** | 0.0061** |
| | (0.0021) | (0.0021) | (0.0026) | (0.0016) | (0.0016) | (0.0019) |
| Emp Tenure | -0.0136** | -0.0135** | -0.0100† | -0.0013 | -0.0013 | 0.0022 |
| | (0.0040) | (0.0040) | (0.0053) | (0.0032) | (0.0032) | (0.0042) |
| Emp Tenure$^2$ × 100 | 0.0625** | 0.0625** | 0.0621* | -0.0000 | 0.0001 | -0.0122 |
| | (0.0210) | (0.0210) | (0.0298) | (0.0171) | (0.0171) | (0.0244) |
| Occ Tenure | 0.0111* | 0.0147** | 0.0151* | 0.0444** | 0.0456** | 0.0443** |
| | (0.0054) | (0.0056) | (0.0069) | (0.0043) | (0.0044) | (0.0054) |
| Occ Tenure$^2$ × 100 | -0.1705** | -0.1738** | -0.1997** | -0.2563** | -0.2583** | -0.2422** |
| | (0.0452) | (0.0455) | (0.0584) | (0.0383) | (0.0385) | (0.0494) |
| Occ Tenure$^3$ × 100 | 0.0029** | 0.0030** | 0.0043** | 0.0042** | 0.0043** | 0.0041** |
| | (0.0011) | (0.0011) | (0.0015) | (0.0009) | (0.0010) | (0.0013) |
| Experience | 0.0576** | 0.0576** | 0.0569** | 0.0378** | 0.0378** | 0.0343** |
| | (0.0031) | (0.0031) | (0.0051) | (0.0030) | (0.0030) | (0.0048) |
| Experience$^2$ × 100 | -0.1543** | -0.1543** | -0.1449** | -0.0777** | -0.0776** | -0.0600† |
| | (0.0239) | (0.0239) | (0.0347) | (0.0231) | (0.0231) | (0.0330) |
| Experience$^3$ × 100 | 0.0015** | 0.0015** | 0.0013† | 0.0003 | 0.0003 | 0.0000 |
| | (0.0005) | (0.0005) | (0.0007) | (0.0005) | (0.0005) | (0.0007) |
| Old Job | -0.0161† | -0.0162† | -0.0181† | 0.0057 | 0.0056 | 0.0091 |
| | (0.0091) | (0.0091) | (0.0107) | (0.0081) | (0.0081) | (0.0096) |
| < High School | -0.0802** | -0.0807** | -0.0750** | -0.0697** | -0.0699** | -0.0633** |
| | (0.0080) | (0.0080) | (0.0095) | (0.0078) | (0.0078) | (0.0092) |
| 4-Year College | 0.2657** | 0.2660** | 0.2377** | 0.2525** | 0.2525** | 0.2269** |
| | (0.0078) | (0.0078) | (0.0095) | (0.0076) | (0.0076) | (0.0093) |
| Hispanic | 0.0142 | 0.0140 | 0.0014 | 0.0130 | 0.0129 | 0.0031 |
| | (0.0114) | (0.0114) | (0.0131) | (0.0112) | (0.0112) | (0.0130) |
| Black | -0.0674** | -0.0671** | -0.0771** | -0.0641** | -0.0640** | -0.0737** |
| | (0.0088) | (0.0088) | (0.0106) | (0.0086) | (0.0086) | (0.0104) |

| | | | | | | |
|---|---|---|---|---|---|---|
| Constant | 6.4250** | 6.4039** | 6.4416** | 6.3551** | 6.3486** | 6.4252** |
| | (0.0274) | (0.0278) | (0.0359) | (0.0268) | (0.0271) | (0.0349) |
| Observations | 44591 | 44591 | 33072 | 44591 | 44591 | 33072 |
| $R^2$ | 0.355 | 0.355 | 0.313 | 0.371 | 0.371 | 0.332 |

All regressions include occupation and industry dummies.

Robust standard errors in parentheses. $^{\dagger}$ $p < 0.10$, $^{*}$ $p < 0.05$, $^{**}$ $p < 0.01$.

Table A.3 – Wage Regression with Mismatch by Components (Full Results)

| | (1) IV | (2) IV | (3) IV | (4) OLS | (5) OLS | (6) OLS |
|---|---|---|---|---|---|---|
| Mismatch Verbal | -0.0147** | 0.0030 | 0.0139* | -0.0150** | -0.0053 | 0.0027 |
| | (0.0033) | (0.0054) | (0.0064) | (0.0032) | (0.0046) | (0.0056) |
| Mismatch Math | -0.0130** | -0.0171** | -0.0203** | -0.0109** | -0.0172** | -0.0182** |
| | (0.0035) | (0.0056) | (0.0065) | (0.0035) | (0.0049) | (0.0058) |
| Mismatch Social | -0.0049† | -0.0034 | 0.0067 | -0.0041 | -0.0046 | 0.0017 |
| | (0.0027) | (0.0046) | (0.0054) | (0.0027) | (0.0040) | (0.0047) |
| Mismatch Verbal × Occ Tenure | | -0.0028** | -0.0045** | | -0.0015** | -0.0026** |
| | | (0.0007) | (0.0009) | | (0.0005) | (0.0006) |
| Mismatch Math × Occ Tenure | | 0.0006 | 0.0021* | | 0.0010† | 0.0020** |
| | | (0.0008) | (0.0009) | | (0.0006) | (0.0007) |
| Mismatch Social × Occ Tenure | | -0.0002 | -0.0010 | | 0.0001 | -0.0005 |
| | | (0.0006) | (0.0007) | | (0.0005) | (0.0005) |
| Cumul Mismatch Verbal | | | -0.0123** | | | -0.0107* |
| | | | (0.0045) | | | (0.0045) |
| Cumul Mismatch Math | | | -0.0252** | | | -0.0274** |
| | | | (0.0044) | | | (0.0044) |
| Cumul Mismatch Social | | | -0.0083* | | | -0.0073* |
| | | | (0.0034) | | | (0.0034) |
| Verbal Ability | -0.0440† | -0.0486† | 0.0081 | 0.0112 | 0.0066 | 0.0158 |
| | (0.0251) | (0.0252) | (0.0299) | (0.0220) | (0.0220) | (0.0269) |
| Math Ability | 0.2949** | 0.3001** | 0.3405** | 0.2510** | 0.2547** | 0.3238** |
| | (0.0264) | (0.0264) | (0.0314) | (0.0236) | (0.0235) | (0.0288) |
| Social Ability | 0.0837** | 0.0836** | 0.1017** | 0.0855** | 0.0853** | 0.1137** |
| | (0.0160) | (0.0159) | (0.0187) | (0.0137) | (0.0136) | (0.0164) |
| Verbal Ability × Occ Tenure | 0.0125** | 0.0126** | 0.0088* | 0.0047† | 0.0051* | 0.0066* |
| | (0.0034) | (0.0034) | (0.0041) | (0.0025) | (0.0025) | (0.0031) |
| Math Ability × Occ Tenure | 0.0006 | -0.0003 | 0.0011 | 0.0037 | 0.0031 | 0.0023 |
| | (0.0034) | (0.0034) | (0.0040) | (0.0025) | (0.0025) | (0.0030) |
| Social Ability × Occ Tenure | 0.0072** | 0.0073** | 0.0075** | 0.0076** | 0.0076** | 0.0063** |
| | (0.0021) | (0.0021) | (0.0026) | (0.0016) | (0.0015) | (0.0019) |
| Occ Reqs Verbal | 0.0771 | 0.0757 | 0.0913 | 0.1414* | 0.1482* | 0.1214† |
| | (0.0671) | (0.0669) | (0.0822) | (0.0575) | (0.0576) | (0.0733) |
| Occ Reqs Math | 0.1112† | 0.1075† | 0.1065 | 0.1004† | 0.0917† | 0.1361* |
| | (0.0626) | (0.0623) | (0.0756) | (0.0535) | (0.0535) | (0.0679) |
| Occ Reqs Social | -0.0932** | -0.0894** | -0.0978** | -0.0817** | -0.0803** | -0.0827** |
| | (0.0198) | (0.0199) | (0.0238) | (0.0177) | (0.0178) | (0.0211) |
| Occ Reqs Verbal × Occ Tenure | -0.0071 | -0.0070 | -0.0098 | -0.0229** | -0.0245** | -0.0205* |
| | (0.0090) | (0.0089) | (0.0114) | (0.0067) | (0.0067) | (0.0087) |
| Occ Reqs Math × Occ Tenure | 0.0164† | 0.0172* | 0.0175† | 0.0228** | 0.0249** | 0.0175* |
| | (0.0084) | (0.0083) | (0.0103) | (0.0062) | (0.0062) | (0.0079) |
| Occ Reqs Social × Occ Tenure | 0.0100** | 0.0092** | 0.0131** | 0.0109** | 0.0106** | 0.0135** |
| | (0.0024) | (0.0024) | (0.0031) | (0.0018) | (0.0018) | (0.0023) |

| | | | | | | |
|---|---|---|---|---|---|---|
| Emp Tenure | -0.0135** | -0.0135** | -0.0096† | -0.0007 | -0.0008 | 0.0022 |
| | (0.0040) | (0.0040) | (0.0053) | (0.0032) | (0.0032) | (0.0042) |
| Emp Tenure$^2 \times$ 100 | 0.0621** | 0.0626** | 0.0594* | -0.0056 | -0.0049 | -0.0109 |
| | (0.0208) | (0.0208) | (0.0298) | (0.0170) | (0.0170) | (0.0244) |
| Occ Tenure | 0.0066 | 0.0103† | 0.0113 | 0.0389** | 0.0398** | 0.0406** |
| | (0.0055) | (0.0057) | (0.0070) | (0.0043) | (0.0044) | (0.0055) |
| Occ Tenure$^2 \times$ 100 | -0.1660** | -0.1702** | -0.1935** | -0.2382** | -0.2406** | -0.2391** |
| | (0.0451) | (0.0454) | (0.0585) | (0.0381) | (0.0382) | (0.0494) |
| Occ Tenure$^3 \times$ 100 | 0.0028** | 0.0030** | 0.0041** | 0.0038** | 0.0039** | 0.0042** |
| | (0.0011) | (0.0011) | (0.0015) | (0.0009) | (0.0009) | (0.0013) |
| Experience | 0.0588** | 0.0588** | 0.0579** | 0.0393** | 0.0393** | 0.0360** |
| | (0.0031) | (0.0031) | (0.0051) | (0.0030) | (0.0030) | (0.0048) |
| Experience$^2 \times$ 100 | -0.1602** | -0.1599** | -0.1509** | -0.0862** | -0.0860** | -0.0711* |
| | (0.0240) | (0.0240) | (0.0347) | (0.0232) | (0.0232) | (0.0331) |
| Experience$^3 \times$ 100 | 0.0016** | 0.0016** | 0.0015* | 0.0005 | 0.0005 | 0.0002 |
| | (0.0005) | (0.0005) | (0.0007) | (0.0005) | (0.0005) | (0.0007) |
| Old Job | -0.0162† | -0.0164† | -0.0186† | 0.0052 | 0.0052 | 0.0083 |
| | (0.0090) | (0.0090) | (0.0107) | (0.0081) | (0.0081) | (0.0096) |
| < High School | -0.0683** | -0.0686** | -0.0600** | -0.0587** | -0.0587** | -0.0500** |
| | (0.0081) | (0.0081) | (0.0097) | (0.0079) | (0.0079) | (0.0095) |
| 4-Year College | 0.2608** | 0.2616** | 0.2314** | 0.2488** | 0.2492** | 0.2218** |
| | (0.0078) | (0.0078) | (0.0095) | (0.0076) | (0.0076) | (0.0094) |
| Hispanic | 0.0160 | 0.0164 | 0.0075 | 0.0144 | 0.0148 | 0.0081 |
| | (0.0115) | (0.0115) | (0.0132) | (0.0113) | (0.0113) | (0.0130) |
| Black | -0.0596** | -0.0590** | -0.0654** | -0.0580** | -0.0577** | -0.0642** |
| | (0.0090) | (0.0090) | (0.0109) | (0.0088) | (0.0088) | (0.0107) |
| Constant | 6.4174** | 6.3975** | 6.4192** | 6.3423** | 6.3396** | 6.3974** |
| | (0.0288) | (0.0291) | (0.0372) | (0.0280) | (0.0281) | (0.0360) |
| Observations | 44591 | 44591 | 33072 | 44591 | 44591 | 33072 |
| $R^2$ | 0.358 | 0.358 | 0.317 | 0.374 | 0.375 | 0.335 |

All regressions include occupation and industry dummies.
Robust standard errors in parentheses. † $p < 0.10$, * $p < 0.05$, ** $p < 0.01$.

Table A.4 – Wage Regression with Positive and Negative Mismatch (Full Results)

| | (1) IV | (2) IV | (3) IV | (4) OLS | (5) OLS | (6) OLS |
|---|---|---|---|---|---|---|
| Positive Mismatch | -0.0143** | 0.0030 | 0.0120† | -0.0135** | -0.0069 | 0.0014 |
| | (0.0033) | (0.0053) | (0.0064) | (0.0032) | (0.0044) | (0.0056) |
| Negative Negative | 0.0375** | 0.0218** | 0.0259** | 0.0338** | 0.0219** | 0.0278** |
| | (0.0033) | (0.0051) | (0.0063) | (0.0033) | (0.0045) | (0.0054) |
| Positive Mismatch $\times$ Occ Tenure | | -0.0028** | -0.0030** | | -0.0011* | -0.0005 |
| | | (0.0007) | (0.0009) | | (0.0005) | (0.0006) |
| Negative Mismatch $\times$ Occ Tenure | | 0.0025** | 0.0021* | | 0.0019** | 0.0018** |
| | | (0.0007) | (0.0009) | | (0.0005) | (0.0006) |
| Cumul Positive Mismatch | | | -0.0163** | | | -0.0228** |
| | | | (0.0046) | | | (0.0046) |
| Cumul Negative Mismatch | | | 0.0086* | | | 0.0018 |
| | | | (0.0038) | | | (0.0038) |
| Emp Tenure | -0.0133** | -0.0132** | -0.0097† | -0.0012 | -0.0014 | 0.0014 |
| | (0.0041) | (0.0041) | (0.0054) | (0.0032) | (0.0032) | (0.0043) |
| Emp Tenure$^2 \times$ 100 | 0.0575** | 0.0574** | 0.0568† | 0.0003 | 0.0020 | 0.0015 |
| | (0.0211) | (0.0211) | (0.0300) | (0.0173) | (0.0173) | (0.0244) |
| Occ Tenure | 0.0264** | 0.0308** | 0.0295** | 0.0526** | 0.0556** | 0.0551** |
| | (0.0054) | (0.0055) | (0.0068) | (0.0043) | (0.0044) | (0.0054) |
| Occ Tenure$^2 \times$ 100 | -0.1449** | -0.1495** | -0.1802** | -0.2328** | -0.2395** | -0.2505** |
| | (0.0452) | (0.0454) | (0.0586) | (0.0385) | (0.0386) | (0.0497) |
| Occ Tenure$^3 \times$ 100 | 0.0023* | 0.0024* | 0.0038* | 0.0037** | 0.0039** | 0.0045** |
| | (0.0011) | (0.0011) | (0.0015) | (0.0009) | (0.0010) | (0.0013) |
| Experience | 0.0542** | 0.0543** | 0.0549** | 0.0360** | 0.0358** | 0.0342** |
| | (0.0032) | (0.0032) | (0.0051) | (0.0030) | (0.0030) | (0.0048) |
| Experience$^2 \times$ 100 | -0.1342** | -0.1347** | -0.1328** | -0.0684** | -0.0678** | -0.0632† |
| | (0.0243) | (0.0243) | (0.0351) | (0.0235) | (0.0235) | (0.0335) |
| Experience$^3 \times$ 100 | 0.0011* | 0.0011* | 0.0011 | 0.0002 | 0.0002 | 0.0001 |
| | (0.0005) | (0.0005) | (0.0007) | (0.0005) | (0.0005) | (0.0007) |
| Old Job | -0.0167† | -0.0169† | -0.0181† | 0.0064 | 0.0061 | 0.0098 |
| | (0.0092) | (0.0092) | (0.0109) | (0.0082) | (0.0082) | (0.0098) |
| < High School | -0.1397** | -0.1400** | -0.1473** | -0.1314** | -0.1320** | -0.1418** |
| | (0.0078) | (0.0078) | (0.0094) | (0.0078) | (0.0078) | (0.0092) |
| 4-Year College | 0.3446** | 0.3445** | 0.3198** | 0.3295** | 0.3293** | 0.3089** |
| | (0.0076) | (0.0076) | (0.0093) | (0.0074) | (0.0074) | (0.0091) |
| Hispanic | -0.0225† | -0.0224† | -0.0361** | -0.0224* | -0.0229* | -0.0360** |
| | (0.0115) | (0.0115) | (0.0133) | (0.0114) | (0.0114) | (0.0132) |
| Black | -0.1321** | -0.1312** | -0.1514** | -0.1291** | -0.1290** | -0.1523** |
| | (0.0089) | (0.0089) | (0.0108) | (0.0087) | (0.0087) | (0.0106) |
| Constant | 6.7522** | 6.7235** | 6.7828** | 6.7090** | 6.6938** | 6.7905** |
| | (0.0241) | (0.0247) | (0.0338) | (0.0240) | (0.0244) | (0.0330) |
| Observations | 44591 | 44591 | 33072 | 44591 | 44591 | 33072 |

| | | | | | | |
|---|---|---|---|---|---|---|
| $R^2$ | 0.336 | 0.336 | 0.290 | 0.351 | 0.351 | 0.308 |

All regressions include occupation and industry dummies.
Robust standard errors in parentheses. $^{\dagger}$ $p < 0.10$, $^{*}$ $p < 0.05$, $^{**}$ $p < 0.01$.

Table A.5 – Regressions for Probability of Occupational Switch (Full Results)

| | (1) LPM-IV | (2) LPM-IV | (3) LPM-IV | (4) LPM | (5) LPM | (6) LPM |
|---|---|---|---|---|---|---|
| Mismatch | 0.0135** | | | 0.0066** | | |
| | (0.0023) | | | (0.0018) | | |
| Mismatch Verbal | | 0.0076** | | | 0.0053* | |
| | | (0.0028) | | | (0.0022) | |
| Mismatch Math | | 0.0074** | | | 0.0025 | |
| | | (0.0028) | | | (0.0022) | |
| Mismatch Social | | 0.0007 | | | -0.0012 | |
| | | (0.0022) | | | (0.0017) | |
| Positive Mismatch | | | 0.0134** | | | 0.0087** |
| | | | (0.0027) | | | (0.0022) |
| Negative Mismatch | | | -0.0130** | | | -0.0028 |
| | | | (0.0027) | | | (0.0021) |
| Worker Ability (Mean) | -0.0370** | -0.0370** | | -0.0208† | -0.0211† | |
| | (0.0134) | (0.0134) | | (0.0116) | (0.0116) | |
| Worker Ability × Occ Tenure | -0.0003 | -0.0003 | | 0.0019* | 0.0020* | |
| | (0.0018) | (0.0018) | | (0.0009) | (0.0009) | |
| Occ Reqs (Mean) | -0.0333* | -0.0334* | | -0.1225** | -0.1223** | |
| | (0.0146) | (0.0146) | | (0.0125) | (0.0125) | |
| Occ Reqs × Occ Tenure | -0.0052** | -0.0052** | | 0.0106** | 0.0106** | |
| | (0.0018) | (0.0018) | | (0.0009) | (0.0009) | |
| Emp Tenure | 0.0039 | 0.0039 | 0.0040 | -0.0064** | -0.0063** | -0.0071** |
| | (0.0029) | (0.0029) | (0.0029) | (0.0017) | (0.0017) | (0.0017) |
| Emp Tenure$^2$× 100 | -0.0189 | -0.0189 | -0.0192 | 0.0301** | 0.0300** | 0.0327** |
| | (0.0161) | (0.0161) | (0.0161) | (0.0090) | (0.0090) | (0.0090) |
| Occ Tenure | 0.0996** | 0.0996** | 0.0966** | -0.0547** | -0.0547** | -0.0495** |
| | (0.0038) | (0.0038) | (0.0037) | (0.0023) | (0.0023) | (0.0022) |
| Occ Tenure$^2$× 100 | -0.5452** | -0.5453** | -0.5474** | 0.3248** | 0.3248** | 0.3406** |
| | (0.0365) | (0.0365) | (0.0363) | (0.0185) | (0.0185) | (0.0184) |
| Occ Tenure$^3$× 100 | 0.0122** | 0.0122** | 0.0122** | -0.0066** | -0.0066** | -0.0070** |
| | (0.0010) | (0.0010) | (0.0010) | (0.0004) | (0.0004) | (0.0004) |
| Experience | -0.0626** | -0.0625** | -0.0619** | 0.0132** | 0.0133** | 0.0134** |
| | (0.0028) | (0.0028) | (0.0028) | (0.0024) | (0.0024) | (0.0024) |
| Experience$^2$× 100 | 0.2355** | 0.2355** | 0.2312** | -0.1297** | -0.1299** | -0.1312** |
| | (0.0212) | (0.0212) | (0.0211) | (0.0171) | (0.0171) | (0.0172) |
| Experience$^3$× 100 | -0.0036** | -0.0036** | -0.0035** | 0.0027** | 0.0027** | 0.0027** |
| | (0.0005) | (0.0005) | (0.0005) | (0.0004) | (0.0004) | (0.0004) |
| Old Job | 0.1506** | 0.1506** | 0.1508** | 0.0069 | 0.0069 | 0.0063 |
| | (0.0056) | (0.0056) | (0.0056) | (0.0051) | (0.0051) | (0.0051) |
| < High School | 0.0413** | 0.0414** | 0.0505** | 0.0090 | 0.0091 | 0.0180** |
| | (0.0087) | (0.0087) | (0.0085) | (0.0070) | (0.0070) | (0.0069) |
| 4-Year College | -0.0433** | -0.0435** | -0.0551** | -0.0040 | -0.0042 | -0.0093* |
| | (0.0060) | (0.0060) | (0.0059) | (0.0047) | (0.0047) | (0.0046) |

| | | | | | | |
|---|---|---|---|---|---|---|
| Hispanic | 0.0044 | 0.0044 | 0.0104 | 0.0068 | 0.0067 | 0.0092 |
| | (0.0095) | (0.0095) | (0.0094) | (0.0074) | (0.0075) | (0.0074) |
| Black | 0.0108 | 0.0107 | 0.0207** | 0.0029 | 0.0027 | 0.0104† |
| | (0.0080) | (0.0080) | (0.0079) | (0.0064) | (0.0064) | (0.0063) |
| Constant | 0.1474** | 0.1486** | 0.0871** | 0.3702** | 0.3727** | 0.2918** |
| | (0.0250) | (0.0252) | (0.0221) | (0.0206) | (0.0208) | (0.0182) |
| Observations | 41596 | 41596 | 41596 | 41596 | 41596 | 41596 |

All regressions include occupation and industry dummies.

Robust standard errors in parentheses. † $p < 0.10$, * $p < 0.05$, ** $p < 0.01$.

Table A.6 – Regressions for Direction of Occupational Switch (Full Results)

| | (1) Average | (2) Verbal | (3) Math | (4) Social |
|---|---|---|---|---|
| Last Mismatch Positive | 0.0751** | | | |
| | (0.0028) | | | |
| Last Mismatch Negative | 0.1143** | | | |
| | (0.0030) | | | |
| Last Pos. Mismatch, Verbal | | 0.0316** | 0.0097* | 0.0143** |
| | | (0.0043) | (0.0046) | (0.0046) |
| Last Neg. Mismatch, Verbal | | 0.0838** | 0.0536** | 0.0216** |
| | | (0.0053) | (0.0056) | (0.0057) |
| Last Pos. Mismatch, Math | | 0.0599** | 0.0898** | 0.0021 |
| | | (0.0044) | (0.0047) | (0.0048) |
| Last Neg. Mismatch, Math | | 0.0558** | 0.0893** | 0.0076 |
| | | (0.0050) | (0.0053) | (0.0054) |
| Last Pos. Mismatch, Social | | 0.0061† | 0.0046 | 0.0774** |
| | | (0.0033) | (0.0035) | (0.0035) |
| Last Neg. Mismatch, Social | | 0.0264** | 0.0166** | 0.1043** |
| | | (0.0037) | (0.0039) | (0.0039) |
| Employer Tenure | -0.0055 | -0.0044 | -0.0034 | -0.0099† |
| | (0.0046) | (0.0053) | (0.0056) | (0.0057) |
| Employer Tenure$^2\times$ 100 | 0.0278 | 0.0127 | 0.0052 | 0.0603† |
| | (0.0278) | (0.0319) | (0.0339) | (0.0343) |
| Occupational Tenure | 0.0005 | 0.0011 | 0.0007 | -0.0043 |
| | (0.0057) | (0.0065) | (0.0069) | (0.0070) |
| Occupational Tenure$^2\times$ 100 | -0.0067 | -0.0226 | -0.0310 | 0.0924 |
| | (0.0633) | (0.0727) | (0.0772) | (0.0780) |
| Occupational Tenure$^3\times$ 100 | 0.0003 | 0.0008 | 0.0012 | -0.0030 |
| | (0.0019) | (0.0022) | (0.0023) | (0.0023) |
| Experience | 0.0045 | 0.0025 | 0.0038 | 0.0041 |
| | (0.0032) | (0.0037) | (0.0039) | (0.0039) |
| Experience$^2\times$ 100 | -0.0175 | -0.0010 | -0.0001 | -0.0212 |
| | (0.0279) | (0.0321) | (0.0341) | (0.0344) |
| Experience$^3\times$ 100 | 0.0002 | -0.0001 | -0.0003 | 0.0003 |
| | (0.0007) | (0.0008) | (0.0008) | (0.0008) |
| Old Job | -0.0057 | -0.0065 | -0.0071 | 0.0063 |
| | (0.0083) | (0.0095) | (0.0101) | (0.0102) |
| $<$ High School | 0.1163** | 0.1256** | 0.1290** | 0.0764** |
| | (0.0079) | (0.0091) | (0.0096) | (0.0097) |
| 4-Year College | -0.1498** | -0.1607** | -0.1508** | -0.1142** |
| | (0.0075) | (0.0087) | (0.0092) | (0.0093) |
| Hispanic | 0.0538** | 0.0651** | 0.0694** | 0.0096 |
| | (0.0100) | (0.0115) | (0.0122) | (0.0124) |
| Black | 0.1136** | 0.1283** | 0.1340** | 0.0397** |
| | (0.0082) | (0.0094) | (0.0100) | (0.0101) |

| | | | | |
|---|---|---|---|---|
| Constant | 0.2754** | 0.2849** | 0.2395** | 0.3001** |
| | (0.0227) | (0.0262) | (0.0278) | (0.0281) |
| Observations | 6594 | 6594 | 6594 | 6594 |
| $R^2$ | 0.487 | 0.485 | 0.458 | 0.417 |

All regressions include occupation and industry dummies.

Robust standard errors in parentheses. $^{\dagger}$ $p < 0.10$, $^{*}$ $p < 0.05$, $^{**}$ $p < 0.01$.

# B Proofs and Derivations

## B.1 Baseline Model with Depreciation

*Proof.* **Derivation of Human Capital Decision and Wage Equation**

Using

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

we obtain

$$h_{j,t} \;=\; (1-\delta)\left(h_{j,t-1} + \frac{A_j^2}{2} - \frac{(A_j - r_{j,t-1})^2}{2} + r_{j,t-1}\epsilon_{j,t-1}\right)$$

and repeatedly substituting for human capital, we obtain

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

Inserting this expression into the following wage equation

$$w_t \;=\; \sum_j \alpha_j \Big(h_{j,t} + \frac{A_j^2}{2} - \frac{(A_j - r_{j,t})^2}{2}\Big) + \sum_j \alpha_j r_{j,t} \varepsilon_{j,t},$$

gives

$$\begin{aligned}
w_t \;=\;& (1-\delta)^{t-1} \sum_j \alpha_j h_{j,1} \\
&+ \sum_{s=1}^{t-1} (1-\delta)^{t-s} \sum_j \alpha_j \left(\frac{A_j^2}{2} - \frac{(A_j - r_{j,s})^2}{2} + r_{j,s}\epsilon_{j,s}\right) \\
&+ \sum_j \alpha_j \Big(\frac{A_j^2}{2} - \frac{(A_j - r_{j,t})^2}{2}\Big) + \sum_j \alpha_j r_{j,t}\varepsilon_{j,t} \\
=\;& (1-\delta)^{t-1} \sum_j \alpha_j h_{j,1} + \frac{1}{2}\left(\sum_j \alpha_j A_j^2\right)\left(\sum_{s=1}^{t} (1-\delta)^{t-s}\right) \\
&+ \frac{1}{2} \sum_{s=1}^{t} (1-\delta)^{t-s} \sum_j \alpha_j (A_j - r_{j,s})^2 \\
&+ \sum_{s=1}^{t} (1-\delta)^{t-s} \sum_j \alpha_j r_{j,s}\epsilon_{j,s}
\end{aligned}$$

□

Setting the depreciation rate to zero, we would obtain:

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

and

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

*Proof.* (Proposition 1) We solve the worker's problem backwards:

$$\begin{aligned} V_t\left(\mathbf{h}_t,\hat{\mathbf{A}}_t\right) &= \max\nolimits_{\{r_{j,t}\}} E_t \quad \left[\sum_j \alpha_j\left(h_{j,t}+(A_j+\epsilon_{j,t})\,r_{j,t}-\frac{r_{j,t}^2}{2}\right)\right. \\ &+ \qquad \left.\beta V_{t+1}\left(\mathbf{h}_{t+1},\hat{\mathbf{A}}_{t+1}\right)\right] \end{aligned}$$

subject to

$$\hat{A}_{j,t+1}=\frac{\lambda_{j,t}}{\lambda_{j,t+1}}\hat{A}_{j,t}+\frac{\lambda_{\epsilon_j}}{\lambda_{j,t+1}}\,(A_j+\epsilon_{j,t})$$

and

$$h_{j,t+1}=(1-\delta)\left(h_{j,t}+(A_j+\epsilon_{j,t})\,r_{j,t}-\frac{r_{j,t}^2}{2}\right)\text{ for all } j.$$

The worker's problem in the last period of his life is

$$V_T\left(\mathbf{h}_T,\hat{\mathbf{A}}_T\right)=\max_{\{r_{j,T}\}}E_T\left[\sum_j \alpha_j\left(h_{j,T}+(A_j+\epsilon_{j,T})\,r_{j,T}-\frac{r_{j,T}^2}{2}\right)\right],$$

which, due to linearity of the objective in $A_j$'s, can be written as

$$V_T\left(\mathbf{h}_T,\hat{\mathbf{A}}_T\right)=\max_{\{r_j\}}\sum_j \alpha_j\left(h_{j,T}+\hat{A}_{j,T}r_{j,T}-\frac{r_{j,T}^2}{2}\right).$$

The optimal solution is the same as in the previous section

$$r_{j,T}=\hat{A}_{j,T}.$$

Substituting the solution, we obtain

$$V_T\left(\mathbf{h}_T,\hat{\mathbf{A}}_T\right)=\sum_j \alpha_j\left(h_{j,T}+\frac{\hat{A}_{j,T}{}^2}{2}\right).$$

Now look at the problem in period $T-1$:

$$\begin{aligned} V_{T-1}\left(\mathbf{h}_{T-1},\hat{\mathbf{A}}_{T-1}\right) &= \max_{\{r_{j,T-1}\}} E_{T-1}\left[\sum_j \alpha_j \left(h_{j,T-1}+\left(A_j+\epsilon_{j,T-1}\right) r_{j,T-1}-\frac{r_{j,T-1}^2}{2}\right)\right. \\ &+ \left.\beta V_T\left(\mathbf{h_T},\hat{\mathbf{A}}_\mathbf{T}\right)\right] \end{aligned}$$

subject to

$$\hat{A}_{j,T}=\frac{\lambda_{j,T-1}}{\lambda_{j,T}}\hat{A}_{j,T-1}+\frac{\lambda_{\epsilon_j}}{\lambda_{j,T}}\left(A_j+\epsilon_{j,T-1}\right)$$

and

$$h_{j,T}=(1-\delta)\left(h_{j,T-1}+\left(A_j+\epsilon_{j,T-1}\right) r_{j,T-1}-\frac{r_{j,T-1}^2}{2}\right) \text{ for all } j.$$

Substituting the law of motion for human capital, we can write as

$$\begin{aligned} V_{T-1}\left(\mathbf{h}_{T-1},\hat{\mathbf{A}}_{T-1}\right) &= \max_{\{r_{j,T-1}\}} E_{T-1}\left[\sum_j \alpha_j \left(h_{j,T-1}+\left(A_j+\epsilon_{j,T-1}\right) r_{j,T-1}-\frac{r_{j,T-1}^2}{2}\right)\right. \\ &+ \left.\beta\sum_j \alpha_j\left(1-\delta\right)\left(h_{j,T-1}+\left(A_j+\epsilon_{j,T-1}\right) r_{j,T-1}-\frac{r_{j,T-1}^2}{2}\right)\right] \\ &+ E_{T-1}\left[\beta\sum_j \alpha_j \hat{A}_{j,T}{}^2/2\right] \end{aligned}$$

$$\begin{aligned} V_{T-1}\left(\mathbf{h}_{T-1},\hat{\mathbf{A}}_{T-1}\right) &= \max_{\{r_{j,T-1}\}} \left(1+\beta\left(1-\delta\right)\right)\sum_j \alpha_j \left(h_{j,T-1}+\hat{A}_{j,T-1} r_{j,T-1}-\frac{r_{j,T-1}^2}{2}\right) \\ &+ E_{T-1}\left[\beta\sum_j \alpha_j \hat{A}_{j,T}^2/2\right]. \end{aligned}$$

The solution gives $r_{j,T-1}=\hat{A}_{j,T-1}$. And

$$\begin{aligned} V_{T-1}\left(\mathbf{h}_{T-1},\hat{\mathbf{A}}_{T-1}\right) &= \left(1+\beta\left(1-\delta\right)\right)\sum_j \alpha_j \left(h_{j,T-1}+\frac{\hat{A}_{j,T-1}^2}{2}\right) \\ &+ E_{T-1}\left[\beta\sum_j \alpha_j \hat{A}_{j,T}^2/2\right] \end{aligned}$$

Continuing backwards, we obtain

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

Since $B_t$ only involves beliefs and does not depend on $r_j$ this term does not affect the worker's decision rules. $\square$

*Proof.* (Lemma 1) Given the history $(A_j + \epsilon_{j,1}, A_j + \epsilon_{j,2}, ..., A_j + \epsilon_{j,t-1})$, the worker's belief at the beginning of period $t$ is given as

$$\begin{aligned}
\hat{A}_{j,t} &= \frac{\lambda_{j,1}}{\lambda_{j,t}}\hat{A}_{j,1} + \frac{\lambda_{\epsilon_j}}{\lambda_{j,t}}(A_j + \epsilon_{j,1} + A_j + \epsilon_{j,2} + ... + A_j + \epsilon_{j,t-1}) \\
&= \frac{\lambda_{j,1}}{\lambda_{j,t}}\hat{A}_{j,1} + \frac{\lambda_{\epsilon_j}}{\lambda_{j,t}}(t-1)A_j + \frac{\lambda_{\epsilon_j}}{\lambda_{j,t}}(\epsilon_{j,1} + \epsilon_{j,2} + ... + \epsilon_{j,t-1}),
\end{aligned}$$

where

$$\lambda_{j,t} = \lambda_{j,1} + (t-1)\lambda_{\epsilon_j}.$$

Since $\hat{A}_{j,1}$ is normally distributed with $\mathcal{N}\left(A_j, \sigma_{\eta_j}^2 + \sigma_{A_j^2}\right)$. Then, $\hat{A}_{j,t}$ will be normally distributed since

$$\begin{aligned}
\hat{A}_{j,t} &= \frac{\lambda_{j,1}}{\lambda_{j,t}}(A_j + \eta_j) + \frac{\lambda_{\epsilon_j}}{\lambda_{j,t}}(t-1)A_j + \frac{\lambda_{\epsilon_j}}{\lambda_{j,t}}(\epsilon_{j,1} + \epsilon_{j,2} + ... + \epsilon_{j,t-1}) \\
&= A_j + \frac{\lambda_{j,1}}{\lambda_{j,t}}\eta_j + \frac{\lambda_{\epsilon_j}}{\lambda_{j,t}}(\epsilon_{j,1} + \epsilon_{j,2} + ... + \epsilon_{j,t-1})
\end{aligned}$$

From this expression, we obtain

$$\hat{A}_{j,t} - A_j = \frac{\lambda_{j,1}}{\lambda_{j,t}}\eta_j + \frac{\lambda_{\epsilon_j}}{\lambda_{j,t}}(\epsilon_{j,1} + \epsilon_{j,2} + ... + \epsilon_{j,t-1})$$

which implies that $E\left[\hat{A}_{j,t} - A_j\right] = 0$. Inserting $\hat{A}_{j,t} - A_j$ from expression above into $M_{j,t} = \sum_{s=1}^{t} \frac{(A_j - r_{j,s})^2}{2}$, we obtain

$$M_{j,t} = \sum_{s=1}^{t}\left(\frac{\lambda_{j,1}}{\lambda_{j,s}}\eta_j + \frac{\lambda_{\epsilon_j}}{\lambda_{j,s}}(\epsilon_{j,1} + \epsilon_{j,2} + ... + \epsilon_{j,s-1})\right)^2$$

and

$$\Omega_{j,t} = \sum_{s=1}^{t}\left(A_j + \frac{\lambda_{j,1}}{\lambda_{j,s}}\eta_j + \frac{\lambda_{\epsilon_j}}{\lambda_{j,s}}(\epsilon_{j,1} + \epsilon_{j,2} + ... + \epsilon_{j,s-1})\right)\epsilon_{j,s}.$$

Note that $\mathrm{Cov}(M_{j,t}, \Omega_{j,t}) = E(M_{j,t}\Omega_{j,t}) - E(M_{j,t})E(\Omega_{j,t})$. Furthermore, $E(\Omega_{j,t}) = 0$ since $\epsilon_{j,s}$ is uncorrelated with all the terms in $\frac{\lambda_{j,1}}{\lambda_{j,s}}\eta_j + \frac{\lambda_{\epsilon_j}}{\lambda_{j,s}}(\epsilon_{j,1} + \epsilon_{j,2} + ... + \epsilon_{j,s-1})$ and $E(\epsilon_{j,s}) = 0$. Thus, $\mathrm{Cov}(M_{j,t}, \Omega_{j,t}) = E(M_{j,t}\Omega_{j,t})$. An important point to notice is that $M_{j,t}\Omega_{j,t}$ includes multiplication of $\epsilon_{j,s}$, $\epsilon_j^2$, $\eta_j$, and $\eta_j^2$, and ${\epsilon_j}^3$. Note that both $\epsilon_j$ and $\eta_j$ are normal with mean zero. And, for a normal random variable $x$ with mean zero, $E\left(x^n\right)$ is zero if $n$ is an odd number and positive if $n$ is an even number. Thus, we have $E\left({\epsilon_{j,s}}^m \eta_j^n\right) = 0$ if either $m$ or $n$ is an odd number and $E\left({\epsilon_{j,s}}^m \eta_j^n\right) > 0$ if both $m$ and $n$ are even numbers. As a result, only terms that remain are the positive ones. Thus, $E\left(M_{j,t}\Omega_{j,t}\right) > 0$. $\square$

*Proof.* (Lemma 2) Note that $\tilde{\Delta}_{j,t} = \left(A_j^2 + \nu_j^2 + 2A_j\nu_j\right)t$. Using

$$r_{j,s} \quad = \hat{A}_{j,s} = \quad A_j + \frac{\lambda_{j,1}}{\lambda_{j,t}}\eta_j + \frac{\lambda_{\epsilon_j}}{\lambda_{j,t}}\left(\epsilon_{j,1} + \epsilon_{j,2} + ... + \epsilon_{j,s-1}\right),$$

we also obtain

$$\widetilde{M}_{j,t} = \sum_{s=1}^{t}\left(\frac{\lambda_{j,1}}{\lambda_{j,s}}\eta_j + \frac{\lambda_{\epsilon_j}}{\lambda_{j,s}}\left(\epsilon_{j,1} + \epsilon_{j,2} + ... + \epsilon_{j,s-1}\right) - \nu_j\right)^2$$

and

$$\widetilde{\Omega}_{j,t} = \sum_{s=1}^{t}\left(A_j + \frac{\lambda_{j,1}}{\lambda_{j,s}}\eta_j + \frac{\lambda_{\epsilon_j}}{\lambda_{j,s}}\left(\epsilon_{j,1} + \epsilon_{j,2} + ... + \epsilon_{j,s-1}\right)\right)\left(\varepsilon_{j,s} - \nu_j\right).$$

First, note that $\mathrm{Cov}\left(\widetilde{\Delta}_{j,t}, \widetilde{\Omega}_{j,t}\right) = E\left[\widetilde{\Delta}_{j,t}\widetilde{\Omega}_{j,t}\right]$ since $E\left[\widetilde{\Omega}_{j,t}\right] = 0$. Using the fact that odd moments of the normal distribution are zero, we obtain

$$\mathrm{Cov}\left(\widetilde{\Delta}_{j,t}, \widetilde{\Omega}_{j,t}\right) \quad = \quad -A_j^3\nu_j t^2 - 2\left(A_j\nu_j t\right)^2 < 0.$$

Second, similarly $\mathrm{Cov}\left(\widetilde{M}_{j,t}, \widetilde{\Omega}_{j,t}\right) = E\left[\widetilde{M}_{j,t}\widetilde{\Omega}_{j,t}\right]$. Rewrite

$$\begin{aligned}
\widetilde{M}_{j,t} \quad &= \quad \sum_{s=1}^{t}\left(\frac{\lambda_{j,1}}{\lambda_{j,s}}\eta_j + \frac{\lambda_{\epsilon_j}}{\lambda_{j,s}}\left(\epsilon_{j,1} + \epsilon_{j,2} + ... + \epsilon_{j,s-1}\right) - \nu_j\right)^2 \\
&= \quad \sum_{s=1}^{t}\left(\frac{\lambda_{j,1}}{\lambda_{j,s}}\eta_j + \frac{\lambda_{\epsilon_j}}{\lambda_{j,s}}\left(\epsilon_{j,1} + \epsilon_{j,2} + ... + \epsilon_{j,s-1}\right)\right)^2 \\
&- \quad 2\nu_j\sum_{s=1}^{t}\left(\frac{\lambda_{j,1}}{\lambda_{j,s}}\eta_j + \frac{\lambda_{\epsilon_j}}{\lambda_{j,s}}\left(\epsilon_{j,1} + \epsilon_{j,2} + ... + \epsilon_{j,s-1}\right)\right) \\
&+ \quad \nu_j^2 t
\end{aligned}$$

and

$$\begin{aligned}
\widetilde{\Omega}_{j,t} \quad &= \quad \sum_{s=1}^{t}\left(A_j + \frac{\lambda_{j,1}}{\lambda_{j,s}}\eta_j + \frac{\lambda_{\epsilon_j}}{\lambda_{j,s}}\left(\epsilon_{j,1} + \epsilon_{j,2} + ... + \epsilon_{j,s-1}\right)\right)\varepsilon_{j,s} \\
&- \quad \nu_j\sum_{s=1}^{t}\left(A_j + \frac{\lambda_{j,1}}{\lambda_{j,s}}\eta_j + \frac{\lambda_{\epsilon_j}}{\lambda_{j,s}}\left(\epsilon_{j,1} + \epsilon_{j,2} + ... + \epsilon_{j,s-1}\right)\right).
\end{aligned}$$

Noting that

$$M_{j,t} = \sum_{s=1}^{t} \left( \frac{\lambda_{j,1}}{\lambda_{j,s}} \eta_j + \frac{\lambda_{\epsilon_j}}{\lambda_{j,s}} \left( \epsilon_{j,1} + \dots + \epsilon_{j,s-1} \right) - \nu_j \right)^2$$

and

$$\Omega_{j,t} = \sum_{s=1}^{t} \left( A_j + \frac{\lambda_{j,1}}{\lambda_{j,s}} \eta_j + \frac{\lambda_{\epsilon_j}}{\lambda_{j,s}} \left( \epsilon_{j,1} + \dots + \epsilon_{j,s-1} \right) \right) \varepsilon_{j,s},$$

we can write

$$\begin{aligned}
\widetilde{M}_{j,t} \widetilde{\Omega}_{j,t} &= M_{j,t} \Omega_{j,t} \\
&- \nu_j \left( \sum_{s=1}^{t} \left( \frac{\lambda_{j,1}}{\lambda_{j,s}} \eta_j + \frac{\lambda_{\epsilon_j}}{\lambda_{j,s}} \left( \epsilon_{j,1} + \dots + \epsilon_{j,s-1} \right) \right)^2 \right) \left( \sum_{s=1}^{t} \left( A_j + \frac{\lambda_{j,1}}{\lambda_{j,s}} \eta_j + \frac{\lambda_{\epsilon_j}}{\lambda_{j,s}} \left( \epsilon_{j,1} + \dots + \epsilon_{j,s-1} \right) \right) \right) \\
&- 2\nu_j \left( \sum_{s=1}^{t} \left( \frac{\lambda_{j,1}}{\lambda_{j,s}} \eta_j + \frac{\lambda_{\epsilon_j}}{\lambda_{j,s}} \left( \epsilon_{j,1} + \dots + \epsilon_{j,s-1} \right) \right) \right) \left( \sum_{s=1}^{t} \left( A_j + \frac{\lambda_{j,1}}{\lambda_{j,s}} \eta_j + \frac{\lambda_{\epsilon_j}}{\lambda_{j,s}} \left( \epsilon_{j,1} + \dots + \epsilon_{j,s-1} \right) \right) \varepsilon_{j,s} \right) \\
&+ 2\nu_j^2 \left( \sum_{s=1}^{t} \left( \frac{\lambda_{j,1}}{\lambda_{j,s}} \eta_j + \frac{\lambda_{\epsilon_j}}{\lambda_{j,s}} \left( \epsilon_{j,1} + \dots + \epsilon_{j,s-1} \right) \right) \right) \left( \sum_{s=1}^{t} \left( A_j + \frac{\lambda_{j,1}}{\lambda_{j,s}} \eta_j + \frac{\lambda_{\epsilon_j}}{\lambda_{j,s}} \left( \epsilon_{j,1} + \dots + \epsilon_{j,s-1} \right) \right) \right) \\
&= \nu_j^2 t \sum_{s=1}^{t} \left( A_j + \frac{\lambda_{j,1}}{\lambda_{j,s}} \eta_j + \frac{\lambda_{\epsilon_j}}{\lambda_{j,s}} \left( \epsilon_{j,1} + \epsilon_{j,2} + \dots + \epsilon_{j,s-1} \right) \right) \varepsilon_{j,s} \\
&- \nu_j^3 t \sum_{s=1}^{t} \left( A_j + \frac{\lambda_{j,1}}{\lambda_{j,s}} \eta_j + \frac{\lambda_{\epsilon_j}}{\lambda_{j,s}} \left( \epsilon_{j,1} + \epsilon_{j,2} + \dots + \epsilon_{j,s-1} \right) \right).
\end{aligned}$$

Notice that in the expressions above, $\nu_j^n$ is multiplied by other random variables which are not correlated with $\nu_j^n$. Thus, the expectations of all these expressions are zero. Then we are left with

$$E\left[ \widetilde{M}_{j,t} \widetilde{\Omega}_{j,t} \right] = E\left[ M_{j,t} \Omega_{j,t} \right].$$

We already know from the proof of Lemma 1 that $E\left[ M_{j,t} \Omega_{j,t} \right] > 0$. □

*Proof.* (Lemma 3)

$$\begin{aligned}
Var\left[ \hat{A}_{j,t} - A_j \right] &= \frac{\lambda_{j,1}^2}{\lambda_{j,t}{}^2} \sigma_{\eta_j}^2 + \frac{\lambda_{\epsilon_j}^2}{\lambda_{j,t}{}^2} (t-1) \sigma_{\epsilon_j}^2 \\
&= \frac{\lambda_{j,1}}{\lambda_{j,t}{}^2} + \frac{\lambda_{\epsilon_j}}{\lambda_{j,t}{}^2} (t-1) \\
&= \frac{1}{\lambda_{j,t}},
\end{aligned}$$

Since $E\left[ \hat{A}_{j,t} - A_j \right] = 0$, $E\left[ \left( \hat{A}_{j,t} - A_j \right)^2 \right] = Var\left[ \hat{A}_{j,t} - A_j \right] + \left( E\left[ \hat{A}_{j,t} - A_j \right] \right)^2 = Var\left[ \hat{A}_{j,t} - A_j \right] = \frac{1}{\lambda_j(t)}$. Note that $\lambda_{j,t} = \lambda_{j,1} + (t-1)\lambda_{\epsilon_j}$ increases with experience. Thus, $E\left[ \left( \hat{A}_{j,t} - A_j \right)^2 \right]$, which is equal to $E\left[(r_{j,t}-A_j)^2\right]$, declines with age. □

*Proof.* (Proposition 2) Note that the probability of switching occupation in period $t$ is given by

$$\begin{aligned}\Pr\left(\mathbf{r_t}\neq\mathbf{r_{t-1}}\right) &= 1-\prod_j \text{Prob}\left(r_{j,t-1}-m_j\leq r_{j,t}<r_{j,t-1}+m_j\right)\\ &= 1-\prod_j \text{Prob}\left(\hat{A}_{j,t-1}-m_j\leq \hat{A}_{j,t}<\hat{A}_{j,t-1}+m_j\right).\end{aligned}$$

Now look at the term $\text{Prob}\left(\hat{A}_{j,t-1}-m_j\leq \hat{A}_{j,t}<\hat{A}_{j,t-1}+m_j\right)$. Inserting

$$\hat{A}_{j,t}=\frac{\lambda_{j,t-1}}{\lambda_{j,t}}\hat{A}_{j,t-1}+\frac{\lambda_{\epsilon_j}}{\lambda_{j,t}}\left(A_j+\epsilon_{j,t-1}\right)$$

and

$$\lambda_{j,t}=\lambda_{j,t-1}+\lambda_{\epsilon_j},$$

we obtain

$$\begin{aligned}&\text{Prob}\left(\hat{A}_{j,t-1}-m_j\leq \hat{A}_{j,t}<\hat{A}_{j,t-1}+m_j\right)\\ =\ &\text{Prob}\left(-m_j\frac{\lambda_{j,t}}{\lambda_{\epsilon_j}}\leq\left(A_j-\hat{A}_{j,t-1}+\epsilon_{j,t-1}\right)<m_j\frac{\lambda_{j,t}}{\lambda_{\epsilon_j}}\right)\\ =\ &\text{Prob}\left(-m_j\frac{\lambda_{j,t}}{\lambda_{\epsilon_j}}-\left(A_j-\hat{A}_{j,t-1}\right)\leq\epsilon_{j,t-1}<m_j\frac{\lambda_{j,t}}{\lambda_{\epsilon_j}}-\left(A_j-\hat{A}_{j,t-1}\right)\right).\end{aligned}$$

Letting $F$ be the cumulative distribution function of a normal variable with mean zero, and noting that normal distribution with mean zero is symmetric around zero, $F(M+x)-F(-M+x)$ declines with $|x|$. Since $x^2$ and $|x|$ move in the same direction, probability of staying in an occupation declines with mismatch $\left(A_j-\hat{A}_{j,t-1}\right)^2$.

Now evaluate the unconditional probability of staying in an occupation:

$$\begin{aligned}&\text{Prob}\left(\hat{A}_{j,t-1}-m_j\leq \hat{A}_{j,t}<\hat{A}_{j,t-1}+m_j\right)\\ =\ &\text{Prob}\left(-m_j\frac{\lambda_{j,t}}{\lambda_{\epsilon_j}}\leq\left(A_j-\hat{A}_{j,t-1}+\epsilon_{j,t-1}\right)<m_j\frac{\lambda_{j,t}}{\lambda_{\epsilon_j}}\right)\\ =\ &\text{Prob}\left(-m_j\sqrt{\frac{\lambda_{j,t-1}\lambda_{j,t}}{\lambda_{\epsilon_j}}}\leq\frac{A_j-\hat{A}_{j,t-1}+\epsilon_{j,t-1}}{\sqrt{\frac{\lambda_{j,t}}{\lambda_{j,t-1}\lambda_{\epsilon}}}}<m_j\sqrt{\frac{\lambda_{j,t-1}\lambda_{j,t}}{\lambda_{\epsilon_j}}}\right)\end{aligned}$$

Remember that $A_j-\hat{A}_{j,t-1}$ is normally distributed with mean zero and variance $1/\lambda_{j,t-1}$ (see proof of Proposition 3). Thus, $\left(A_j-\hat{A}_{j,t-1}+\epsilon_{j,t-1}\right)$ is normally distributed with mean zero and variance $\frac{1}{\lambda_{j,t-1}}+\frac{1}{\lambda_\epsilon}=\frac{\lambda_{j,t}}{\lambda_{j,t-1}\lambda_\epsilon}$. Thus, $\left(A_j-\hat{A}_{j,t-1}+\epsilon_{j,t-1}\right)\times\sqrt{\frac{\lambda_{j,t-1}\lambda_\epsilon}{\lambda_{j,t}}}$ is normally distributed with mean zero and variance one. Since both $\lambda_{j,t}$ and $\lambda_{j,t-1}$ increases with age, probability of staying in the last period's occupation increases and probability of switching decreases with age. □

*Proof.* (Proposition 3) Probability of switching to an occupation with higher skill-$j$ intensity is

$$\begin{aligned}\pi_{j,t}^{up} &= \text{Prob}\left(\hat{A}_{j,t} > \hat{A}_{j,t-1} + m_j\right) \\ &= \text{Prob}\left(\epsilon_{j,t} > m_j \frac{\lambda_{j,t}}{\lambda_{\epsilon_j}} - \left(A_j - \hat{A}_{j,t-1}\right)\right) \\ &= \text{Prob}\left(\epsilon_{j,t} > m_j \frac{\lambda_{j,t}}{\lambda_{\epsilon_j}} - \left(r_j^* - r_{j,t-1}\right)\right).\end{aligned}$$

Note that probability of switching to an occupation with higher skill-$j$ intensity increases with $\left(r_j^* - r_{j,t-1}\right)$. Thus, to the extent that the worker is overqualified, he will switch to a higher skill occupation. The probability of switching to a lower skill occupation is given by

$$\begin{aligned}\pi_{j,t}^{down} &= \text{Prob}\left(A_{j,t} < A_{j,t-1} - m_j\right) \\ &= \text{Prob}\left(\epsilon_{j,t} < -m_j \frac{\lambda_{j,t}}{\lambda_{\epsilon_j}} - \left(A_j - \hat{A}_{j,t-1}\right)\right) \\ &= \text{Prob}\left(\epsilon_{j,t} < -m_j \frac{\lambda_{j,t}}{\lambda_{\epsilon_j}} - \left(r_j^* - r_{j,t-1}\right)\right).\end{aligned}$$

Using these two equations above, it is easy to observe that $\pi_{j,t}^{up} > \pi_{j,t}^{down}$ when $r_j^* - r_{j,t-1} > 0$. Similar proof can be made for the case $r_{j,t-1} - r_j^* > 0$. But we skip it for the sake of brevity. □

## B.2 Baseline Model vs. Ben-Porath Version

In the baseline model of Section 2, the occupation/human capital choice problem is a static one. To see this most clearly, we consider a simplified version of the model with a single skill and where ability is observed. In this model, the lifetime problem not only reduces to a static one, but the decision rule does not change over time. Next we show that the Ben-Porath version of the same model features an occupation/human capital choice that changes every period, underscoring the dynamic nature of the decision. The following derivations establish these results. These results extend straightforwardly to the case with multiple skills and Bayesian learning at the expense of much more complicated algebra. (These results are available upon request).

### Baseline Model with No Learning

Let us write the problem of the individual sequentially, starting from the last period of life.

$$V_T\left(h_T\right) = \max_{r_T}\left(h_T + Ar_T - \frac{r_T^2}{2}\right).$$

Now insert this into the following problem

$$V_{T-1}\left(h_{T-1}\right) \quad = \quad \max_{r_{T-1}} \left(h_{T-1} + Ar_{T-1} - \frac{r_{T-1}^2}{2}\right) + \beta\left(h_T + Ar_T - \frac{r_T^2}{2}\right)$$

s.t.

$$h_{j,T} = (1-\delta)\left(h_{j,T-1} + Ar_{T-1} - \frac{r_{T-1}^2}{2}\right).$$

Inserting the law of motion into the objective function, we obtain

$$V_{T-1}\left(h_{T-1}\right) \quad = \quad \max_{r_{T-1}} \left(1+\beta(1-\delta)\right)\left(h_{T-1} + Ar_{T-1} - \frac{r_{T-1}^2}{2}\right) + \beta\left(Ar_T - \frac{r_T^2}{2}\right)$$

The problem in period $T-1$ does not depend on any period-$T$ variables or affect any decision in period $T$. This is because the occupation choice/human capital accumulation decision does not depend on the stock of human capital—it only depends on the workers' ability level, which does not change over time. Consequently, the decision rule in period $T-1$ would be the same if we just ignored period $T$ and just solved

$$\max_{r_{T-1}} \left(h_{T-1} + Ar_{T-1} - \frac{r_{T-1}^2}{2}\right).$$

It can be shown that the solution is the same in all periods and given by

$$r_t = A \qquad \forall t. \tag{9}$$

### Ben-Porath with No Learning

In Ben-Porath the wage equation is the same as in our baseline model (with a single skill): $w_t = h_t - \frac{r_t^2}{2}$. The lifetime maximization problem is

$$\begin{aligned} &\max_{\{r_t\}_{t=1}^T} \mathbb{E}_0 \left[\sum_{t=1}^T \beta^{t-1}(h_t - \frac{r_t^2}{2})\right] \\ &\text{s.t.} \\ &h_{t+1} = (1-\delta)\left(h_t + Ar_t\right) \end{aligned}$$

This model corresponds to the standard Ben-Porath formulation with a quadratic cost function. Following the same steps as above, the solution can be shown to be:

$$r_t = \sum_{s=1}^{T-t} \left(\beta(1-\delta)\right)^s A,$$

which unlike the solution in (9) is not constant and changes every period. This is due to the intertemporal trade-off noted above inherent in the Ben-Porath model.

### Mismatch in the Ben-Porath Model

Note that both positive and negative past mismatch reduces current wage in our model. Ben-Porath model on the other hand has different implications for positive and negative past mismatch. In particular, it implies that positive past mismatch (worker being over-qualified in past occupations) reduces the current wage and negative past mismatch (worker being under-qualified in past occupations) increases the current wage. Abstracting from multi-dimensions for simplicity, note that the optimal occupational choice of a worker in period $t$ under perfect information, denoted by $r_t^*$, is given by

$$r_t^* = \sum_{s=1}^{T-t} (\beta(1-\delta))^s A.$$

Under Bayesian Learning, assuming the same information structure as in our model, the worker's optimal choice is given by

$$r_t = \sum_{s=1}^{T-t} (\beta(1-\delta))^s \hat{A}_t,$$

where $\hat{A}_t$ is the mean belief of worker's ability in period $t$. Substituting $h_t = (1-\delta)\left(h_{t-1} + (A + \epsilon_{t-1})\, r_{t-1}\right)$ repeatedly into $w_t = h_t - \frac{r_t^2}{2}$, we obtain

$$w_t = h_1 + A\sum_{s=1}^{t-1} r_s + \sum_{s=1}^{t-1} \epsilon_s r_s - \frac{r_t^2}{2}.$$

By adding and subtracting $A\sum_{s=1}^{t-1} r_s^*$, we obtain

$$w_t = h_1 + A\sum_{s=1}^{t-1} r_s^* - A\sum_{s=1}^{t-1}(r_s^* - r_s) + \sum_{s=1}^{t-1} \epsilon_s r_s - \frac{r_t^2}{2}.$$

# C Data

## C.1 Panel Construction and Sample Selection

To construct annual panel data for our main analysis, we use NLSY79's Work History Data File, which records individuals' employment histories up to five jobs on a weekly basis from 1978 to 2010. Following the approach by Neal (1999) and Pavan (2011), we calculate total hours worked for each job within a year from the information of usual hours worked per week and the number of weeks worked for each job. Then, we define primary jobs for each individual for each year as the one for which an individual spent the most hours worked within the year. We construct panel data with annual frequency (from 1978 to 2010) from a series of observations of primary jobs and out-of-labor-force status for each individual. The annually reported demographic information and detailed information of employment (occupation, industry, and hourly wage) are merged with the panels.[27]

**Occupational Codes** Before we merge the occupation information with the panel data, we clean occupational titles. In NLSY79, every year individuals report their occupations for up to five jobs that they had since their last interviews. Also, NLSY79 provides a mapping between five jobs reported in the current interview and those reported in the last interview if any of them are the same. Using the mapping of jobs across interviews, we first create an employment spell for each job. We then assign, to each employment spell, an occupation code that is most often observed during the spell. This approach is similar to the one by Kambourov and Manovskii (2009b), in which an occupational switch is considered as genuine only when it is accompanied with a job switch. Since the classifications of occupations are not consistent across years, we converted all the occupational codes into the Census 1990 Three-Digit Occupation Code before the cleaning.[28]

**Industry Codes** To clean the industry codes, we take the similar approach as the one for occupation codes. Since industry codes are also reported in the different classifications across years, we use our own crosswalk to convert them into the Census 1970 One-Digit Industry Code.[29] After the conversion into the Census 1970 Code, we clean the industry titles by using job spells. We use those one-digit level industry codes to create industry dummies used in our regression analysis.

**Employer and Occupational Switches** We identify an employer switch when the primary job for an individual is different from the one in the last year. We identify occupational switches when the occupation in his primary job is different from the one in last year's primary job.

[27] More precisely, in NLSY79, the information except weekly employment status is reported annually from 1979 to 1994, and biannually, from 1996 to 2010.

[28] NLSY79 reports workers' occupational titles in the Census 1970 Three-Digit Occupation Code until 2000. After 2000, they are reported in the Census 2000 Three-Digit Occupation Code. All of those codes are converted to the Census 1990 Three-Digit Occupation Code using the crosswalks provided by the Minnesota Population Center (http://usa.ipums.org/usa/volii/occ_ind.shtml).

[29] Industry codes are reported in the Census 1970 Industry Classification Code before 1994, in the Census 1980 Industry Classification Code for the year 1994, and in the Census 2000 Industry Classification Code after 2000. The crosswalk is presented in Appendix C.2.

**Labor Market Experience, Employer and Occupational Tenure** As we will discuss below, we drop the individuals who have already been in labor markets when the survey started. We then set individuals' experience equal to zero when a worker is entering labor markets, and increase it by one every year when the worker reports a job. Employer and occupational tenure increase by one every year when the individual reports a job and are reset to zero when switches happen.

**Wages** Worker's wages are measured by the usual rate of pay for the primary job at the time of interview. All the wages are deflated by the price index for personal consumption expenditures into real term in the 2000 dollars. We drop the observations if their wages are missing. We also drop the top 0.1% and the bottom 0.1% of observations in the wage distribution in each round of the interview. This trimming strategy doesn't affect the regression result.[30]

**Sample Selection** We follow the approach by Farber and Gibbons (1996) for the sample selection. We first limit our sample to the individuals who make their initial long-term transition from school to labor markets during the survey period: that is, we drop those who work more than 1,200 hours in the initial year of the survey. We also focus on the individuals who work more than 1,200 hours at least for two consecutive years during the survey period. The individuals who are in the military service more than two years during the period are also eliminated from our sample. For the individuals who go back to school from the labor force during the survey period, we assume they start their career from the point they reenter labor markets, and drop the observations before that time. Also, the observations after the last time an individual report a job are eliminated. Furthermore, we drop individuals who are weakly attached to the labor force: those who are out of the labor force more than once before they work at least 10 years after they started their career. If an individual is out of the labor force only for one year after he started his career, or if he has worked more than 10 years before he first dropped from the labor force, we only drop those observations. Finally, we restrict our sample to those who have a valid occupation and industry code, who have valid demographics information, who are equal to or above age 16 and not currently enrolled in a school, and who have valid ASVAB scores and valid wage information. The number of the remaining individuals and observations after applying each sample selection criterion are summarized in Table A.7.

## C.2 The Crosswalk of Census Industry Codes

We used the crosswalk in Table A.8 to convert the Census 1970, 1980, and 1990 Three-Digit Industry Code to the Census 1970 One-Digit Industry Code. We use one-digit level industry titles to create industry dummies used in our regression analysis. From the Census 1980 Three-Digit Code to the Census 1970 One-Digit Code, we first aggregate the Census 1980 Three-Digit level into the Census 1980 One-Digit level. Then, we combine Wholesale (500-571) and Retail Trade (580-691) in the Census 1980 One-Digit Code to create the category 6, "Whole Sale and Retail Trade", in the Census 1970 One-Digit Code. For other one-digit-level industry titles, the Census 1970 and 1980 have the same classification.

Unlike the one between the Census 1970 and the Census 1980, the mapping is not straightforward from the Census 2000 to the Census 1970. Sometimes, the same industry titles in three-digit-level are put in different one-digit-level categories. For example, "Newspaper Publishers"

[30]Similarly, Pavan (2011) drops the top ten and the bottom ten observations in the entire sample.

Table A.7 – Sample Selection, NLSY79, 1978 - 2010

| Criterion for Sample Selection | Remaining Individuals | Remaining Observations |
|---|---|---|
| Male Cross-Sectional Sample | 3,003 | 99,099 |
| Started career after the survey started | 2,408 | 79,242 |
| Work more than 1,200 hours for two consecutive periods | 2,311 | 65,041 |
| Not in the military service more than or equal to two years | 2,261 | 63,529 |
| Drop the observations before they go back to school | 2,261 | 63,477 |
| Drop the observations after the last time they worked | 2,261 | 55,406 |
| Drop those who are weakly attached to labor force | 2,095 | 49,154 |
| Valid occupation and industry code | 2,094 | 48,352 |
| Valid demographics information | 2,093 | 48,328 |
| Drop those below age 16 and not enrolled in school | 2,093 | 48,314 |
| Valid ASVAB scores | 1,996 | 46,253 |
| Valid wage information | 1,992 | 44,721 |
| Drop top 0.1% and bottom 0.1% in the wage distribution | 1,992 | 44,591 |

(code number 647 in the Census 2000) is in "Information and Communication" category in the Census 2000, but is put in "Manufacturing" in the Census 1970. Therefore, we check all the three-digit industry titles both in the Census 1970 and 2000 Industry Code, and made necessary changes to create our own mapping. The obtained crosswalk is reported in Table A.8.

Table A.8 – The Crosswalk across the Census 1970, 1980, and 1990 One-Digit Industry Classification Code

| 1970 One-Digit Classification | | 1970 | 1980 | 2000 |
|---|---|---|---|---|
| 1. | Agriculture, Forestry, Fishing, and Hunting | 017-029 | 010-031 | 017-029 |
| 2. | Mining | 047-058 | 040-050 | 037-049 |
| 3. | Construction | 067-078 | 60 | 077 |
| 4. | Manufacturing | 107-398 | 100-392 | 107-399, 647-659, 678-679 |
| 5. | Transportation, Communications, and Other Public Utilities | 407-499 | 400-472 | 57-69, 607-639, 667-669 |
| 6. | Wholesale and Retail Trade | 507-699 | 500-691 | 407-579, 868-869 |
| 7. | Finance, Insurance and Real Estate | 707-719 | 700-712 | 687-719 |
| 8. | Business and Repair Services | 727-767 | 721-760 | 877-879, 887 |
| 9. | Personal Services | 769-799 | 761-791 | 866-867, 888-829 |
| 10. | Entertainment and Recreation Services | 807-817 | 800-802 | 856-859 |
| 11. | Professional and Related Services | 828-899 | 812-892 | 677, 727-779, 786-847 |
| 12. | Public Administration | 907-947 | 900-932 | 937-959 |

# D Physical Skills Dimension

In addition to mathematical, verbal, and social dimensions, one might expect that match quality is affected by the physical requirements of an occupation and a worker's physical abilities. In this appendix we try to incorporate a physical dimension into our measure. Conceptually, it is difficult to measure a worker's physical abilities, as these are going to change quite a bit over his working life and often as a result of the occupations he chooses. This endogeneity makes it quite difficult to identify an underlying ability for physically demanding work, as we had identified in the other ability dimensions. Beyond this direct reverse causality, there is a strong correlation between healthy behaviors and income, whereas most high income jobs have only mild physical requirements.

When we introduce our proxy for worker's physical ability and the occupational physical requirements, as described below, it seems to have little to do with wages. When we use principal components to aggregate dimensions of mismatch, physical gets a very low weight, suggesting its variation is not well related to the rest of the variation in the dataset. While this independence from the other dimensions may have actually been useful, we found that physical mismatch also has little relationship to wages. Generally physical mismatch is insignificant when we include it into a Mincer regression, as we did with the others. All this is not to say that physical match quality is unimportant, but given the measurement hurdles, we were unable to find a solid relationship. Details of the process and findings are given below.

### Health/Physical Scores in NLSY79

Participants in the NLSY79 were asked to take a survey when they turned age 40 to evaluate their health status. In particular, the survey includes questions about how much health limited the respondents' (i) moderate activities; (ii) ability to climb a flight of stairs; and (iii) types of work they can perform; as well as (iv) how participants rated their own health status (often referred to as EVGFP) and (v) whether pain interfered with their daily activities.[31] Each participant was then assigned a composite health score, called PCS-12, by combining their scores on each question. One difficulty in using this health composite score in our analysis is that it is measured after a significant period of working life, so differences across individuals may simply reflect the effects of occupations on workers (see Michaud and Wiczer (2014)).

[31] The survey is conducted by health care survey firm Quality Metric; see Ware et al. (1995) for details.

**Physical Skills in O*NET**

TABLE A.9 – List of Physical Skills in O*NET

| Physical Skills | | | |
|---|---|---|---|
| 1. | Arm-Hand Steadiness | 2. | Manual Dexterity |
| 3. | Finger Dexterity | 4. | Control Precision |
| 5. | Multi-limb Coordination | 6. | Response Orientation |
| 7. | Rate Control | 8. | Reaction Time |
| 9. | Wrist-Finger Speed | 10. | Speed of Limb Movement |
| 11. | Static Strength | 12. | Explosive Strength |
| 13. | Dynamic Strength | 14. | Trunk Strength |
| 15. | Stamina | 16. | Extent Flexibility |
| 17. | Dynamic Flexibility | 18. | Gross Body Coordination |
| 19. | Gross Body Equilibrium | | |

To create a physical measure of an occupation, we again turn to the O*NET, which contains 19 descriptors related to the physical demands of an occupation (e.g., whether it requires strength, coordination, and stamina). To reduce the 19 descriptors to a single index measure of physical skills, we take the first principal component over the 19 descriptors. For the worker's physical ability measure, we use the NLSY's PCS-12 score. Both physical ability and skill scores again are converted into rank scores among individuals or among occupations. Notice that the coefficients to mismatch change relatively little from our previous specification. This is because the loading on physical mismatch is relatively small. Principal components assigns loadings (0.42, 0.42, 0.12, 0.4) when constructing mismatch measure.

**Wage Regression Results with Physical Component**

TABLE A.10 – Four Skills: Wage Regressions with Mismatch

| | (1) IV | (2) IV | (3) IV | (4) OLS | (5) OLS | (6) OLS |
|---|---|---|---|---|---|---|
| Mismatch | -0.0275** | -0.0126** | -0.0048 | -0.0271** | -0.0214** | -0.0158** |
| Mismatch × Occ Tenure | | -0.0023** | -0.0021** | | -0.0009* | -0.0003 |
| Cumul Mismatch | | | -0.0374** | | | -0.0374** |
| Worker Ability (Mean) | 0.2416** | 0.2447** | 0.3445** | 0.2682** | 0.2686** | 0.3414** |
| Worker Ability × Occ Tenure | 0.0177** | 0.0169** | 0.0178** | 0.0116** | 0.0114** | 0.0156** |
| Occ Reqs (Mean) | 0.1728** | 0.1710** | 0.1683** | 0.2202** | 0.2193** | 0.2247** |
| Occ Reqs × Occ Tenure | 0.0123** | 0.0124** | 0.0120** | 0.0050** | 0.0051** | 0.0036* |
| Observations | 37738 | 37738 | 28115 | 37738 | 37738 | 28115 |
| $R^2$ | 0.352 | 0.352 | 0.315 | 0.368 | 0.368 | 0.332 |

Note: $** \ p < 0.01$, $* \ p < 0.05$, $\dagger \ p < 0.1$. All regressions include a constant, terms for demographics, occupational tenure, employer tenure, work experience, and dummies for one-digit-level occupation and industry. Standard errors are computed as robust Huber-White sandwich estimates.

Table A.11 – Four Skills: Wage Regressions with Mismatch by Components

| | (1) IV | (2) IV | (3) IV | (4) OLS | (5) OLS | (6) OLS |
|---|---|---|---|---|---|---|
| Mismatch Verbal | -0.0168** | 0.0040 | 0.0138* | -0.0164** | -0.0029 | 0.0058 |
| Mismatch Math | -0.0110** | -0.0142* | -0.0170* | -0.0112** | -0.0172** | -0.0205** |
| Mismatch Social | -0.0065* | -0.0067 | -0.0005 | -0.0058* | -0.0079† | -0.0039 |
| Mismatch Phys | 0.0005 | 0.0004 | -0.0047 | 0.0031 | 0.0028 | -0.0061 |
| Mismatch Verbal × Occ Tenure | | -0.0032** | -0.0044** | | -0.0021** | -0.0031** |
| Mismatch Math × Occ Tenure | | 0.0004 | 0.0020* | | 0.0009 | 0.0025** |
| Mismatch Social × Occ Tenure | | 0.0001 | -0.0005 | | 0.0003 | -0.0001 |
| Mismatch Phys × Occ Tenure | | 0.0001 | 0.0014† | | 0.0001 | 0.0016** |
| Cumul Mismatch Verbal | | | -0.0137** | | | -0.0117* |
| Cumul Mismatch Math | | | -0.0270** | | | -0.0290** |
| Cumul Mismatch Social | | | -0.0105** | | | -0.0095* |
| Cumul Mismatch Phys | | | -0.0031 | | | 0.0006 |
| Verbal Ability | -0.0804** | -0.0866** | -0.0048 | -0.0264 | -0.0327 | -0.0071 |
| Math Ability | 0.3147** | 0.3230** | 0.3450** | 0.2698** | 0.2751** | 0.3343** |
| Social Ability | 0.0764** | 0.0759** | 0.0860** | 0.0863** | 0.0860** | 0.1084** |
| Phys Ability | 0.0988** | 0.1010** | 0.1330** | 0.1227** | 0.1232** | 0.1130** |
| Verbal Ability × Occ Tenure | 0.0132** | 0.0136** | 0.0090* | 0.0064* | 0.0070** | 0.0082* |
| Math Ability × Occ Tenure | -0.0004 | -0.0019 | 0.0007 | 0.0020 | 0.0010 | 0.0005 |
| Social Ability × Occ Tenure | 0.0082** | 0.0083** | 0.0082** | 0.0070** | 0.0071** | 0.0050* |
| Phys Ability × Occ Tenure | 0.0002 | -0.0001 | 0.0037 | -0.0045** | -0.0046** | 0.0051* |
| Occ Reqs Verbal | 0.0351 | 0.0303 | 0.0631 | 0.0791 | 0.0834 | 0.0738 |
| Occ Reqs Math | 0.1698* | 0.1671* | 0.1637† | 0.1818** | 0.1734** | 0.2047** |
| Occ Reqs Social | -0.0933** | -0.0887** | -0.1045** | -0.0858** | -0.0842** | -0.1022** |
| Occ Reqs Phys | 0.0004 | -0.0013 | -0.0431 | -0.0153 | -0.0175 | -0.0733* |
| Occ Reqs Verbal × Occ Tenure | -0.0031 | -0.0022 | -0.0073 | -0.0165* | -0.0175* | -0.0171† |
| Occ Reqs Math × Occ Tenure | 0.0117 | 0.0121 | 0.0139 | 0.0158* | 0.0175** | 0.0144† |
| Occ Reqs Social × Occ Tenure | 0.0115** | 0.0106** | 0.0124** | 0.0122** | 0.0119** | 0.0142** |
| Occ Reqs Phys × Occ Tenure | 0.0026 | 0.0028 | -0.0009 | 0.0009 | 0.0013 | 0.0016 |
| Observations | 37738 | 37738 | 28115 | 37738 | 37738 | 28115 |
| $R^2$ | 0.356 | 0.356 | 0.320 | 0.372 | 0.372 | 0.337 |

Note: $** \ p < 0.01$, $* \ p < 0.05$, $\dagger \ p < 0.1$. All regressions include a constant, terms for demographics, occupational tenure, employer tenure, work experience, and dummies for one-digit-level occupation and industry. Standard errors are computed as robust Huber-White sandwich estimates.

# E College Graduates

Given that our mismatch measure is based on higher-order cognitive, and social abilities, it is natural that this measure is more relevant to individuals with a higher level of education doing occupations that place greater emphasis on higher-order skills. In this appendix, we restrict our sample to college graduates, and run wage regressions as we did in the main analysis. The results are presented in Table A.12. Most of the coefficients of mismatch, mismatch times tenure, and cumulative mismatch increased in their effect on wages compared to those in Table IV. In particular, the coefficient on the cumulative mismatch in Column (3) almost doubled for the college sample compared to the baseline result.

It is also interesting to see where the increase of the effect is coming from. By breaking down the measures into components in Column (3) of Table A.13, we learn that it is verbal and social components which are particularly strong effects and contribute to the differences in wages among college graduates. In particular, the coefficient on cumulative social mismatch is four times larger than the one in our benchmark result. The results here show that mismatch is a more important wage determinant among college graduates, and that verbal and social components have especially large effects compared to the benchmark case.

Table A.12 – College Graduate: Wage Regressions with Mismatch

| | (1) IV | (2) IV | (3) IV | (4) OLS | (5) OLS | (6) OLS |
|---|---|---|---|---|---|---|
| Mismatch | -0.0393** | -0.0244** | -0.0080 | -0.0377** | -0.0241** | -0.0108 |
| Mismatch × Occ Tenure | | -0.0023* | -0.0014 | | -0.0021** | -0.0004 |
| Cumul Mismatch | | | -0.0671** | | | -0.0736** |
| Worker Ability (Mean) | 0.2373** | 0.2432** | 0.2902** | 0.2584** | 0.2608** | 0.2893** |
| Worker Ability × Occ Tenure | 0.0144** | 0.0128** | 0.0175** | 0.0108** | 0.0097** | 0.0220** |
| Occ Reqs (Mean) | 0.2617** | 0.2791** | 0.3221** | 0.3166** | 0.3339** | 0.3666** |
| Occ Reqs × Occ Tenure | 0.0181** | 0.0157** | 0.0126** | 0.0077** | 0.0052† | 0.0030 |
| Observations | 21908 | 21908 | 15762 | 21908 | 21908 | 15762 |
| $R^2$ | 0.295 | 0.295 | 0.263 | 0.308 | 0.308 | 0.278 |

Note: ** $p < 0.01$, * $p < 0.05$, † $p < 0.1$. All regressions include a constant, terms for demographics, occupational tenure, employer tenure, work experience, and dummies for one-digit-level occupation and industry. Standard errors are computed as robust Huber-White sandwich estimates.

Table A.13 – College Graduate: Wage Regressions with Mismatch by Components

| | (1) IV | (2) IV | (3) IV | (4) OLS | (5) OLS | (6) OLS |
|---|---|---|---|---|---|---|
| Mismatch Verbal | -0.0351** | -0.0130 | -0.0039 | -0.0327** | -0.0202** | -0.0153† |
| Mismatch Math | -0.0004 | -0.0060 | -0.0032 | -0.0015 | 0.0027 | 0.0072 |
| Mismatch Social | -0.0169** | -0.0288** | -0.0085 | -0.0150** | -0.0253** | -0.0095 |
| Mismatch Verbal × Occ Tenure | | -0.0035** | -0.0029* | | -0.0019* | -0.0004 |
| Mismatch Math × Occ Tenure | | 0.0008 | 0.0017 | | -0.0007 | 0.0003 |
| Mismatch Social × Occ Tenure | | 0.0018† | -0.0004 | | 0.0016* | -0.0005 |
| Cumul Mismatch Verbal | | | -0.0322** | | | -0.0334** |
| Cumul Mismatch Math | | | -0.0355** | | | -0.0423** |
| Cumul Mismatch Social | | | -0.0318** | | | -0.0276** |
| Verbal Ability | -0.1062** | -0.0942* | 0.0211 | 0.0030 | 0.0094 | 0.0233 |
| Math Ability | 0.3524** | 0.3523** | 0.3385** | 0.2700** | 0.2681** | 0.3097** |
| Social Ability | 0.0469† | 0.0324 | 0.0170 | 0.0521* | 0.0400† | 0.0358 |
| Verbal Ability × Occ Tenure | 0.0184** | 0.0155** | 0.0107† | 0.0040 | 0.0023 | 0.0116* |
| Math Ability × Occ Tenure | -0.0109* | -0.0112* | -0.0034 | -0.0030 | -0.0031 | 0.0027 |
| Social Ability × Occ Tenure | 0.0121** | 0.0145** | 0.0150** | 0.0127** | 0.0148** | 0.0147** |
| Occ Reqs Verbal | 0.0621 | 0.0708 | 0.0464 | -0.0014 | -0.0090 | -0.0317 |
| Occ Reqs Math | 0.1934* | 0.1813* | 0.2465* | 0.2977** | 0.3114** | 0.3811** |
| Occ Reqs Social | -0.0281 | -0.0179 | -0.0419 | -0.0059 | 0.0035 | -0.0329 |
| Occ Reqs Verbal × Occ Tenure | -0.0150 | -0.0151 | -0.0229 | -0.0135 | -0.0110 | -0.0211 |
| Occ Reqs Math × Occ Tenure | 0.0262* | 0.0278* | 0.0268* | 0.0157† | 0.0129 | 0.0128 |
| Occ Reqs Social × Occ Tenure | 0.0130** | 0.0108** | 0.0195** | 0.0108** | 0.0088** | 0.0184** |
| Observations | 21908 | 21908 | 15762 | 21908 | 21908 | 15762 |
| $R^2$ | 0.298 | 0.298 | 0.268 | 0.311 | 0.312 | 0.283 |

Note: ** $p < 0.01$, * $p < 0.05$, † $p < 0.1$. All regressions include a constant, terms for demographics, occupational tenure, employer tenure, work experience, and dummies for one-digit-level occupation and industry. Standard errors are computed as robust Huber-White sandwich estimates.

# F Effects of Mismatch on Earnings

In the model we presented in Section 2, wages and earnings are identical because we assume worker's labor supply is constant (fixed to 1) over lifecycle. However, in reality, wages and earnings could be significantly different as there is large heterogeneity in individuals' working hours. Therefore, it is worth to see whether our model's implications still hold when we use individuals' earnings data rather than the wage data in regressions. Looking at earnings rather than wages is advantageous in the light of measurement error due to misreporting. As is common in many survey-based datasets, because most workers do not earn an hourly wage, and actual hours are often difficult to recall, earnings are much more precisely reported than hourly wages. Therefore, in this appendix, we check the robustness of our results by using the earnings in place of wages.

In order to create an annual earnings measure, we use two income variables from NLSY79: total income from wage and salary and total income from farm or business. One shortcoming of using these variables is that, after 1994, the information is only available every 2 years. Therefore, we have to reduce the number of observations significantly when we run a regression using the earnings measure. Another important issue to take into account is that these variables pool income from different jobs. Thus, when an individual works for more than one occupation in a year, income from different occupations are pooled in one earnings measure. Therefore, when a worker reports more than one job, we relate a worker's annual earnings to the job in which the worker earned the largest amount of money in the year, which is calculated by the hourly rate of pay of that job times the number of hours the worker spent in that job in the year. Obtained annual earnings measure is deflated by the price index for personal consumption expenditures into real term in the 2000 dollars. Finally, obtained, real annual earnings are put as the left-hand-side variable in a Mincer regression after taking a natural logarithm.

Table A.14 reports results of earnings regressions with mismatch. It is worth to emphasize that those results are very similar to those in previous wage regressions (compare the results with Table IV). In our most preferred specification, (3), the coefficient for mismatch times tenure is slightly larger than the one in the wage regression, and the one for cumulative mismatch is slightly smaller. However, in general, the results are almost same as those in wage regression.

Turning to regressions by components reported in Table A.15, again, the results didn't change from Table VI in general. In the specification (3), the coefficient for cumulative social mismatch obtain a slightly larger value when we use earnings, while that for cumulative verbal mismatch loses its significance. However, over all, the results are in line with those in wage regressions. This fact confirms the robustness of our results even when we use annual earnings as the left-hand-side variable of a Mincer regression.

Table A.14 – Earnings Regressions with Mismatch

| | (1) IV | (2) IV | (3) IV | (4) OLS | (5) OLS | (6) OLS |
|---|---|---|---|---|---|---|
| Mismatch | -0.0215** | -0.0110† | -0.0033 | -0.0187** | -0.0185** | -0.0163* |
| Mismatch × Occ Tenure | | -0.0021* | -0.0027* | | -0.0000 | 0.0006 |
| Cumul Mismatch | | | -0.0305** | | | -0.0290** |
| Worker Ability (Mean) | 0.1887** | 0.1908** | 0.3022** | 0.2412** | 0.2412** | 0.3285** |
| Worker Ability × Occ Tenure | 0.0180** | 0.0171** | 0.0157** | 0.0074** | 0.0074** | 0.0100** |
| Occ Reqs (Mean) | 0.3465** | 0.3460** | 0.3607** | 0.4555** | 0.4555** | 0.5022** |
| Occ Reqs × Occ Tenure | 0.0126** | 0.0128** | 0.0131** | -0.0086** | -0.0086** | -0.0162** |
| Observations | 31351 | 31351 | 21063 | 31351 | 31351 | 21063 |
| $R^2$ | 0.400 | 0.400 | 0.346 | 0.419 | 0.419 | 0.370 |

Note: $** \ p < 0.01$, $* \ p < 0.05$, $\dagger \ p < 0.1$. All regressions include a constant, terms for demographics, occupational tenure, employer tenure, work experience, and dummies for one-digit-level occupation and industry. Standard errors are computed as robust Huber-White sandwich estimates.

Table A.15 – Earnings Regressions with Mismatch by Components

| | (1) IV | (2) IV | (3) IV | (4) OLS | (5) OLS | (6) OLS |
|---|---|---|---|---|---|---|
| Mismatch Verbal | -0.0031 | 0.0116 | 0.0196* | -0.0020 | 0.0109 | 0.0185* |
| Mismatch Math | -0.0243** | -0.0270** | -0.0278** | -0.0223** | -0.0343** | -0.0400** |
| Mismatch Social | 0.0076* | 0.0057 | 0.0149* | 0.0091* | 0.0055 | 0.0115† |
| Mismatch Verbal × Occ Tenure | | -0.0029* | -0.0045** | | -0.0025** | -0.0036** |
| Mismatch Math × Occ Tenure | | 0.0005 | 0.0011 | | 0.0023** | 0.0038** |
| Mismatch Social × Occ Tenure | | 0.0004 | 0.0003 | | 0.0007 | 0.0009 |
| Cumul Mismatch Verbal | | | -0.0046 | | | -0.0032 |
| Cumul Mismatch Math | | | -0.0220** | | | -0.0221** |
| Cumul Mismatch Social | | | -0.0120* | | | -0.0111* |
| Verbal Ability | -0.0699* | -0.0711* | -0.0387 | -0.0140 | -0.0175 | -0.0192 |
| Math Ability | 0.2372** | 0.2415** | 0.2698** | 0.2013** | 0.2046** | 0.2657** |
| Social Ability | 0.0858** | 0.0852** | 0.1572** | 0.1352** | 0.1347** | 0.1807** |
| Verbal Ability × Occ Tenure | 0.0040 | 0.0033 | 0.0041 | -0.0031 | -0.0032 | 0.0004 |
| Math Ability × Occ Tenure | 0.0073 | 0.0064 | 0.0039 | 0.0087* | 0.0085* | 0.0044 |
| Social Ability × Occ Tenure | 0.0122** | 0.0124** | 0.0121** | 0.0039† | 0.0041† | 0.0069* |
| Occ Reqs Verbal | 0.4962** | 0.4963** | 0.3912** | 0.5374** | 0.5542** | 0.4426** |
| Occ Reqs Math | -0.1758* | -0.1803* | -0.1133 | -0.1501* | -0.1689* | -0.0503 |
| Occ Reqs Social | 0.0436 | 0.0458 | 0.1139** | 0.1144** | 0.1146** | 0.1636** |
| Occ Reqs Verbal × Occ Tenure | -0.0231† | -0.0231† | -0.0124 | -0.0386** | -0.0428** | -0.0343* |
| Occ Reqs Math × Occ Tenure | 0.0275* | 0.0287** | 0.0238 | 0.0282** | 0.0330** | 0.0194 |
| Occ Reqs Social × Occ Tenure | 0.0123** | 0.0119** | 0.0057 | 0.0025 | 0.0025 | -0.0003 |
| Observations | 31351 | 31351 | 21063 | 31351 | 31351 | 21063 |
| $R^2$ | 0.402 | 0.402 | 0.349 | 0.421 | 0.421 | 0.372 |

Note: $** \ p < 0.01$, $* \ p < 0.05$, $\dagger \ p < 0.1$. All regressions include a constant, terms for demographics, occupational tenure, employer tenure, work experience, and dummies for one-digit-level occupation and industry. Standard errors are computed as robust Huber-White sandwich estimates.