Occupational Switching and Self-Discovery in the Labor Market

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

February 15, 2015

Preliminary and Incomplete. Comments Welcome.

Abstract

This paper studies workers' occupational switching behavior and how lifetime earnings inequality is affected by the match between workers' ability and the skills required by their occupation. Using Armed Services Vocational Aptitude Battery (ASVAB), O*NET, and National Longitudinal Survey of Youth 1979 (NLSY79), we create empirical measures of the *match quality* between each worker's ability and the skills emphasized by his/her occupation, and analyze their effects on workers' labor market outcomes. We find that low match quality—what we also call "skill mismatch"—between one's skills and required occupational skills reduces wage growth during an occupational tenure. Furthermore there is a persistence across occupations: match quality in occupations held early in life has a strong effect on wages in future occupations. We view these findings within the context of a general equilibrium model of occupational choice and human capital accumulation. We believe that our study sheds light on the importance of (i) occupational match on determination of wages, and (ii) workers' learning on their ability and the skills required by occupations.

*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

Occupation switching is a defining characteristic of the careers of young workers. The average American worker holds about three to four different occupations during his/her twenties, but switching declines rapidly thereafter. This paper aims to answer one question about this behavior: What are the economic gains from occupational switching? Clearly, this question begs for a theory of how workers allocate themselves into any particular occupation.

In this paper, we think of workers who possess a set of differentiated skills and occupations are production units that require different amounts of these skills. Workers, however, are not perfectly aware of how successful they will be in an occupation, but can learn through their labor market experience. The task of choosing an occupation then is to match with an occupation expected to best utilizes one's skills. This gives meaning to "mismatch" in an occupation: a worker is well matched if the skills in which he is relatively strong are those skills used relatively intensively in his occupation, but the match quality may be low due to this informational incompleteness. We quantitatively explore mismatch in occupational skills through a reduced-form empirical and model-based approach.

We construct an empirical measure of mismatch: where mismatch measures the distance between a workers strengths and those required by his occupation. We present evidence that occupational match quality raises (and mismatch lowers) both the level and the growth rate of the wage of a worker. Furthermore, match quality in occupations held early in life has a strong effect on wages in future occupations. This last fact is striking and inconsistent with theories that include only idiosyncratic match quality; we suggest instead that bad matches affect human capital accumulation and hence wages in future occupations. We also bring evidence that occupational switches exhibit specific patterns. A worker with more experience will tend to have a higher match quality. And additional switches tend to improve the match, after conditioning on the initial match quality. Piecing these facts together, occupational switching seems to be an integral part of the process of self-discovery in the labor market, a process that is rewarded with faster wage growth.

This paper uses data from the 1979 National Longitudinal Survey of Youth (NLSY79) for information on worker's occupation and wage histories. Respondents were also given occupational placement tests, the Armed Services Vocational Aptitude Battery

(ASVAB), that provides a detailed snapshot of occupation-relevant skills. In addition to this cognitive measure, respondents report a variety of health-measures that we use to describe one's ability for physical work. To determine the skills used by a particular occupation, we use data from the US Department of Labor's O*NET project. This data gives a very complete 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 are fortunate that the ASVAB project commissioned a group of experts to map the skills that they test to the skills the O*NET measures. Combining these two sources of information allows us to quantify the match quality, and thus quantify the gains from a switch of occupation.

We incorporate our measure of mismatch into the Mincer-regression framework along with flexible interactions with occupational tenure. We find that the coefficient on mismatch is robustly negative and that its interaction with occupational tenure is negative. Even more important, the cumulative effect of past mismatch also has a significant and negative effect on wages. Our work and these results contributes to a large literature on the effects of match quality. Whereas match quality is often left in the residual and dealt with by various techniques, by measuring it and explicitly including it in Mincer regressions, we shed light on its importance to, for example, estimates of the return to occupational tenure.

We can also exploit the panel dimension of our data and find important patterns regarding switching. We estimate the relationship between mismatch and switching probability, which is interestingly non-linear. Higher mismatch is associated with a higher probability of switching, about 2 percentage points annually between the 90th and 10th percentiles of mismatch. However, positive mismatch and negative mismatch also have their own patterns. One step further, we can follow where workers switch. If a worker's skills are higher than the skills required by their current occupation, the next occupation has, on average higher skill.

The paper proceeds as follows: In Section 2 we describe our data on worker skills and occupations' utilization of them and Section 3 describes how we create our match quality measures. In Section 4 we present our reduced-form wage regression results using this data. Section 5 introduces our model of occupational choice, human capital accumulation and learning. Section 6 describes its estimation and our quantitative findings. Finally, we conclude.

Literature Review

[To be completed]

2 Model

In this section, we present a model of multi-dimensional skill, occupational choice, and human capital accumulation in order to discuss mismatch between workers' abilities and occupations' skill requirements. The underlying friction that generates mismatch is imperfect information about workers' abilities. The market learns each worker's abilities over time in a Bayesian manner. In Section 2.1, we start with a reduced form version of the Bayesian Learning model in order to introduce the concept of mismatch. This model captures the basic ideas of the Bayesian learning model and allows us to discuss mismatch without going into the details of Bayesian learning. We discuss the full model in Section 2.2.

2.1 Basics

The economy consists of continuum of occupations that use n types of skills indexed with $i \in \{1, 2, .., n\}$. Occupations differ in their skill intensity of each type, denoted with the vector $\mathbf{r} = (r_1, r_2, ..., r_n)$. 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 (the stock of accumulated skill) of each skill type $\mathbf{h}(\mathbf{t}) = (h_1(t), h_2(t), ..., h_n(t))$. If the worker chooses an occupation with indexed by skill intensity vector $\mathbf{r}$, then the amount of skill i that he can effectively utilize in that occupation is assumed to be

$$k_i(t) \equiv h_i(t) + A_i r_i - \frac{r_i^2}{2}. \tag{1}$$

Worker's wage in a period is given by

$$w(t) = \sum_i \alpha_i k_i(t),$$

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

$$h_i(t+1) = (1-\delta) k_i(t). \tag{2}$$

Thus, $\mathbf{k}(\mathbf{t}) = (k_1(t), ..., k_n(t))$ determines both the worker's current wage and also the next period's human capital. This formulation is similar to the standard Ben-Porath model but it extends the Ben-Porath formulation in several dimensions. First, we introduce multi-dimensional human capital and abilities. Second, skill-accumulation is linked to occupational choice. Third, the negative term $(0.5r_i^2)$ enters in the accumulation of human capital, which it does not in the Ben-Porath model. As we elaborate below, these features imply that there is an ideal occupation for each worker where earnings and human capital accumulation are maximized. And the worker earns lower wages and accumulates less human capital to the extent that his occupational choice is further away from his ideal occupation in skill intensity space. Thus, this framework allows us to formulate the concept mismatch as deviation of worker's occupation from his ideal occupation in skill intensity space.

Information about ability: The worker has imperfect information about his ability $\mathbf{A}$. In this section, we assume that the worker makes his occupational choice in period t based on the belief that his abilities are given by $\hat{\mathbf{A}}(\mathbf{t}) = (\hat{A}_1(t), ..., \hat{A}_n(t))$.

Worker's optimization problem: We assume that the worker discounts future fully. Thus, his objective is to maximize current wage given his belief $\hat{\mathbf{A}}(\mathbf{t})$ by choosing an occupation. Thus, the worker's problem can be written as

$$V_t\left(\mathbf{h}(\mathbf{t}), \hat{\mathbf{A}}(\mathbf{t})\right) = \max_{\{r_i\}} \sum_i \alpha_i \left(h_i(t) + \hat{A}_i(t) r_i - \frac{r_i^2}{2}\right),$$

which yields an optimal solution of the form:

$$r_i(t) = \hat{A}_i(t) \text{ for all } i.$$

[1]In principle, we can assume that each occupation has a different weight for each skill type. However, that would complicate our model. Since our intension is to present the simplest model that we can discuss the concept of mismatch, we forgo this possibly reasonable assumption.

The worker makes his decision based on $\hat{\mathbf{A}}(\mathbf{t})$. However, the realization of his effective human capital and wage also depends on his true ability $\mathbf{A}$:

$$w(t) = \sum_i \alpha_i \left(h_i(t) + A_i \hat{A}_i(t) - \frac{\hat{A}_i(t)^2}{2} \right).$$

Note that the worker's wage is maximized when his belief is correct, i.e. $\hat{\mathbf{A}}(\mathbf{t}) = \mathbf{A}$. Let $\mathbf{r}^* = (r_1^*, ..., r_2^*)$ be the ideal occupation, where

$$r_i^* = A_i \text{ for all } i.$$

To the extent that skill intensity of his current occupation deviates from the skill intensity of his ideal occupation, the worker will earn less wages and accumulate less human capital.

Definition 1. *Skill mismatch between worker and occupation* in skill dimension i in period t is $(A_i - r_i(t))^2$, which is the deviation of the worker's occupation's skill-i intensity from the worker's skill-i ability.

Note that this definition of mismatch is also equivalent to the following two definitions: 1) deviation of worker's occupation's skill intensity from his ideal occupation's skill intensity $(r_i^* - r_i(t))^2$ and 2) deviation of worker's belief from his true ability $(A_i - \hat{A}_i(t))^2$. We use the definition above since, in the empirical analysis, we will use test scores that proxies workers' skills, which correspond to A_i's, and skill intensities of his occupation, which correspond to $r_i(t)$'s.

Using the definition of mismatch above, we can rewrite the worker's wage as

$$w(t) = \sum_i \alpha_i \left(h_i(t) + \frac{A_i^2}{2} - \frac{(A_i - r_i(t))^2}{2} \right),$$

which shows that a worker's wage depends positively on his beginning-of-period human capital $\mathbf{h}(\mathbf{t})$ and his ability $\mathbf{A}$ and negatively on **mismatch** $(A_i - r_i(t))^2$. However, note that the current human capital depends on past occupational choices and thus past mismatches. In order to see the effect of past mismatches on current wage, substitute $k_i(t)$ from equation (1) into equation (2) to obtain

$$h_i(t) = (1-\delta)\left(h_i(t-1) + \frac{A_i^2}{2} - \frac{\left(A_i - \hat{A}_i(t-1)\right)^2}{2}\right).$$

Repeatedly substituting for human capital and setting $\delta \equiv 0$ in order to simplify the expression, we obtain an expression that links current human capital to the path of past mismatches:[2]

$$h_i(t) = h_i(1) + \frac{A_i^2}{2}(t-1) - \sum_{s=1}^{t-1} \frac{\left(A_i - \hat{A}_i(s)\right)^2}{2},$$

and using the definition of mismatch above, the worker's wage at time t is given as[3]

$$w(t) = \sum_i \alpha_i h_i(1) + \underbrace{\sum_i \alpha_i \frac{A_i^2}{2} t}_{\text{ability}\times\text{experience}} - \sum_i \alpha_i \underbrace{\frac{(A_i - r_i(t))^2}{2}}_{\text{current mismatch}} - \underbrace{\sum_{j=1}^{t-1} \sum_i \alpha_i \frac{(A_i - r_i(j))^2}{2}}_{\text{cumulative mismatch}}.$$

2.2 Bayesian Learning

In this section, we expand the framework of the previous section by modeling the learning process of workers' abilities.

[2] If $\delta \neq 0$, the expression for human capital would be

$$h_i(t) = (1-\delta)^{t-1} h_i(1) + (1-\delta)\sum_{s=1}^{t-1}(1-\delta)^{s-1}\left(\frac{A_i^2}{2}\right) + (1-\delta)\sum_{s=1}^{t-1}(1-\delta)^{s-1}\frac{\left(A_i - \hat{A}_i(t-s)\right)^2}{2}.$$

[3] If we used the standard Ben-Porath specification with

$$h_i(t+1) = (1-\delta)\left(h_i(t) + A_i r_i\right),$$

then if somebody is employed in an occupation above his ideal skill match today, he would have lower wages today but on the other hand, he would have higher wages in the future. Therefore, in the standard Ben-Porath, positive past cumulative mismatch (worker being overqualified) has a negative effect on current wage, and vice a versa. In our current set-up, both positive and negative past cumulative mismatches have negative effect on current wages.

2.2.1 Information structure and technology

Each worker draws ability A_i from a normal distribution $N\left(\mu_{A_i}, \sigma_{A_i}^2\right)$ at the beginning of his life. He also receives a signal about A_i, which is given as $\hat{A}_i(1) = A_i + \eta_i$ where $\eta_i \sim N(0, \sigma_{\eta_i}^2)$. Thus, the worker starts his career with the belief that his ability in skill type i is normally distributed with mean $\hat{A}_i(1)$ and precision $\lambda_i(1) = \lambda_i = 1/\sigma_{\eta_i}^2$. The population average of initial beliefs is $E[A_i + \eta_i] = \mu_{A_i}$. Therefore, prior beliefs are unbiased.

If the worker chooses an occupation indexed by skill requirement vector $\mathbf{r}$, then the amount of skill i that he can effectively utilize in that occupation is given by

$$k_i(t) \equiv h_i(t) + (A_i + \epsilon_i(t)) r_i - \frac{r_i^2}{2}, \tag{3}$$

where $\epsilon_i(t) \sim N\left(0, \sigma_{\epsilon_i}^2\right)$. We assume that the worker observes his effective skill of each type, which is equivalent to observing $\tilde{A}_i(t) = A_i + \epsilon_i(t)$ in each period. Given his prior, the worker updates his belief about his ability after observing $\tilde{A}_i(t)$.

The worker's belief at the beginning of each period is normally distributed. Let $\hat{A}_i(t)$ be the mean and $h_i(t)$ be the precision of this distribution. Letting λ_{ϵ_i} be the precision of $\epsilon_i(t)$ and given the history $(\tilde{A}_i(1), \tilde{A}_i(2), ..., \tilde{A}_i(t-1))$, the worker's belief at the beginning of period t will have the mean

$$\hat{A}_i(t) = \frac{\lambda_i}{\lambda_i(t)} \hat{A}_i(1) + \frac{\lambda_{\epsilon_i}}{\lambda_i(t)} \left(\tilde{A}_i(1) + \tilde{A}_i(2) + ... + \tilde{A}_i(t-1)\right)$$

and precision

$$\lambda_i(t) = \lambda_i + (t-1)\lambda_{\epsilon_i}.$$

Alternatively, we can write the updating problem recursively. Given $\hat{A}_i(t-1)$ and $\lambda_i(t-1)$, we have

$$\hat{A}_i(t) = \frac{\lambda_i(t-1)}{\lambda_i(t)} \hat{A}_i(t-1) + \frac{\lambda_{\epsilon_i}}{\lambda_i(t)} \tilde{A}_i(t-1),$$

where

$$\lambda_i(t) = \lambda_i(t-1) + \lambda_{\epsilon_i}.$$

Proposition 1. *(Population Belief Distribution)*

a. $E[\hat{A}_i(t)] = \mu_{A_i}$ *and* $Var[\hat{A}(t)] = \frac{\left(\lambda_i^2 + \lambda_{\epsilon_i}^2 (t-1)^2\right)\sigma_{A_i}^2 + \lambda_i(t)}{\lambda_i(t)^2}$.

b. $E[(A_i - \hat{A}_i(t))^2] = 1/\lambda_i(t)$ *and* $\lim_{t\to\infty} E[(A_i - \hat{A}_i(t))^2] = 0$.

Thus, the average of workers' beliefs are equivalent to the average of A_i's in the population. However, workers over estimate or under estimate some of their abilities. This is given by the positive $E[(A_i - \hat{A}_i(t))^2]$ in part (b) of the proposition. Nevertheless, each individual's belief of his ability becomes more precise and converges to his true ability A_i as labor market experience increases.

2.2.2 Worker's problem:

The law of motion for human capital is the same as in equation 2. At any period t, given that his ability A_i is normally distributed with mean $\hat{A}_i(t)$, the worker maximizes

$$V\left(\mathbf{h}(\mathbf{t}), \hat{\mathbf{A}}(\mathbf{t})\right) = \max_{\{r_i\}} E\left[\sum_i \alpha_i \left(h_i(t) + (A_i + \epsilon_i(t)) r_i - \frac{r_i^2}{2}\right)\right],$$

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

$$V\left(\mathbf{h}(\mathbf{t}), \hat{\mathbf{A}}(\mathbf{t})\right) = \max_{\{r_i\}} \sum_i \alpha_i \left(h_i(t) + \hat{A}_i(t) r_i - \frac{r_i^2}{2}\right).$$

The optimal solution is the same as in the previous section

$$r_i(t) = \hat{A}_i(t).$$

Thus, the worker's wage is given as

$$w(t) = \sum_i \alpha_i \left(h_i(t) + (A_i + \epsilon_i(t)) \hat{A}_i(t) - \frac{\hat{A}_i(t)^2}{2}\right)$$

and his beginning-of-next-period human capital is

$$h_i(t+1) = (1-\delta)\left(h_i(t) + (A_i + \epsilon_i(t)) \hat{A}_i(t) - \frac{\hat{A}_i(t)^2}{2}\right)$$

Note that the worker's **expected** wage is maximized when his belief is correct, i.e. $\hat{\mathbf{A}}(\mathbf{t}) = \mathbf{A}$. Using the same definition of **mismatch**, we can rewrite the workers wage as

$$w(t) = \sum_i \alpha_i \left(h_i(t) + \frac{A_i^2}{2} - \frac{(A_i - r_i(t))^2}{2}\right) + \sum_i \alpha_i r_i(t) \epsilon_i(t)$$

Note that the only difference between the wage above and the one in the previous section is the term $\sum_i \alpha_i r_i(t)\epsilon_i(t)$. Since $\epsilon_i(t)$ and $r_i(t)$ are uncorrelated, the worker's expected wage is the same as in the previous section. Using the law of motion for human capital and setting $\delta = 0$ (in order to simplify the expression), we can obtain the current wage as a function of current and past mismatches as in the previous section:[4]

$$w(t) = \sum_i \alpha_i h_i(1) + \underbrace{\sum_i \alpha_i \frac{A_i^2}{2} t}_{\text{ability}\times\text{experience}} - \sum_i \alpha_i \underbrace{\frac{(A_i - r_i(t))^2}{2}}_{\text{current mismatch}} - \underbrace{\sum_{j=1}^{t-1}\sum_i \alpha_i \frac{(A_i - r_i(j))^2}{2}}_{\text{cumulative mismatch}} + \sum_{s=1}^{t-1}\sum_i \alpha_i r_i(s)\epsilon_i$$

The following proposition is a direct consequence of proposition 1 and the fact that $r_i(t) = \hat{A}_i(t)$.

Proposition 2. *(Mismatch by Age) Average mismatch is given by* $E[(A_i - r_i(t))^2] = 1/\lambda_i(t)$. *Since the precision* $\lambda_i(t)$ *increases with age, average mismatch declines by age. Moreover, it converges to zero as labor market experience reaches infinity, i.e.* $\lim_{t\to\infty} E[(A_i - r_i(t))]^2 = 0$.

2.2.3 Occupational Switching

In this section, we discuss the worker's occupational switching decision and how it depends on mismatch. We assume that occupational switch occurs when a worker

[4] Using

$$h_i(t) = (1-\delta)\left(h_i(t-1) + \frac{A_i^2}{2} - \frac{\left(A_i - \hat{A}_i(t-1)\right)^2}{2}\right) + \hat{A}_i(t-1)\epsilon_i(t-1)$$

and repeatedly substituting for human capital, we obtain

$$\begin{aligned} h_i(t) &= (1-\delta)^{t-1} h_i(1) + (1-\delta)\sum_{s=1}^{t-1}(1-\delta)^{s-1}\left(\frac{A_i^2}{2}\right) + (1-\delta)\sum_{s=1}^{t-1}(1-\delta)^{s-1}\frac{\left(A_i - \hat{A}_i(t-s)\right)^2}{2} \\ &+ (1-\delta)\sum_{s=1}^{t-1}(1-\delta)^{s-1}\hat{A}_i(t-s)\epsilon_i(t-s). \end{aligned}$$

S $\delta \equiv 0$, we obtain an expression that links current human capital to the path of past mismatches:

$$h_i(t) = h_i(1) + \frac{A_i^2}{2}(t-1) - \sum_{s=1}^{t-1}\frac{\left(A_i - \hat{A}_i(s)\right)^2}{2} + \sum_{s=1}^{t-1}\hat{A}_i(s)\epsilon_i(s)$$

.

choses 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_i be some neighborhood of skill-i intensity, an occupation switch occurs in period t if $r_i(t) > r_i(t-1) + m_i$ or $r_i(t) < r_i(t-1) - m_i$ for some i. Thus, the probability of switching occupation in period t is given by

$$\Pr(\mathbf{r_t} \neq \mathbf{r_{t-1}}) = 1 - \prod_i \text{Prob}\left(r_i(t-1) - m_i \leq r_i(t) < r_i(t-1) + m_i\right).$$

Proposition 3. *(Occupational Switching by Age) Probability of switching occupation declines with age.*

This proposition is a direct consequence of the fact that the precision of beliefs increase with age and as a result, workers' occupational choices converge to their ideal choices and mismatch declines with age.

Proposition 4. *(Mismatch vs. Occupational Switch) Probability of switching to another occupation in period t increases with the mismatch in the occupation in period $t-1$.*

Whether the worker is over-qualified or underqualified, the probability of switching to another occupation increases with mismatch.

Proposition 5. *(Direction of Occupational Switches) This proposition has two components:*

a. Upward mobility: If $r_i^* - r_i(t-1) > 0$, i.e. if the worker is overqualified in dimension i, then the probability of switching to an occupation with a higher (lower) skill-i intensity is larger than the probability of switching to an occupation with a lower skill-i intensity. Moreover, the probability of switching to an occupation with a higher skill-i intensity increases with $r_i^* - r_i(t-1)$.

b. Downward mobility: If $r_i(t-1) - r_i^* > 0$, i.e. if the worker is underqualified in dimension i, then the probability of switching to an occupation with a lower skill-i intensity is larger than the probability of switching to an occupation with a higher skill-i intensity. Moreover, the probability of switching to an occupation with a lower skill-i intensity increases with $r_i(t-1) - r_i^*$.

3 Data

The main source of data for this paper is the NLSY79, which is a nationally representative sample of individuals who were 14 to 21 years of age on January 1, 1979. Each individual is tracked from 1979 to 2010, providing up to 32 years of labor market information. In addition to detailed information about earnings and employment, NLSY79 has three other features that makes it suitable for our analysis. First, in the first year of the survey all respondents took the ASVAB, which measures various abilities. Second, respondents were also surveyed about their health status, physical conditions, as well as relevant family health background. These questions provide detailed information about various aspects of each individual's ability to perform jobs at different points in the life cycle. In particular, ASVAB test scores will be used to measure cognitive abilities, where health/physical scores will be used to construct measures of physical skills.

Third, each individual also provided information about the occupational title they hold for each job. In order to create measures of match quality and mismatch between workers' skills measured and the skills required by occupations, we link these individual test scores to the information about the skills required for their occupations reported in O*NET (to be described in detail below). Below, we describe (i) the NLSY79, (ii) the ASVAB, and health scores (iii) the O*NET, and (4) the mapping from individual abilities to O*NET skill categories.

3.1 NLSY79

We use data on males from the nationally representative sample of the NLSY79, which includes 3,003 individuals. Individuals in the sample report their main job during the year (although they also provide information about additional jobs at higher frequency). We use the weekly wages at the main employer as our measure of wages for a given year and convert nominal to real using the CPI index with 2000 as the base year.[5] The Census codes for occupations and industries often change over time, so we use the code correspondence provided by the NLSY79 to obtain consistent codes over time. The final coding results in 292 occupations that are tracked consistently over time. Finally, we consolidate industry codes to the 1-digit level. We also correct for spurious changes by ignoring switches that immediately revert.

[5]This approach is consistent with others in the literature; see, e.g., Pavan (2011). Further details are in Appendix Section 6.

We exclude individuals who were already working when the sample began so as to avoid the left-truncation in their employment history. As explained later, such truncation would pose problems for our empirical measures which require the complete work history to be recorded for each individual. We also restrict our sample to workers that are reasonably well attached to the labor force by including individuals who work at least 1,200 hours for two consecutive years. These assumption as consistent with earlier work, such as Farber and Gibbons (1996) and Pavan (2011). Our final sample runs from 1979 through 2010 and includes 1,992 individuals and 44,589 individual-year observations. For some regressions, the number of observations could be lower because further variables used in a particular regression might be missing.

The descriptive statistics for the sample is reported in Table I. Because of the nature of the survey, starting with workers as they are young in the workforce, the survey generally skews younger. As a result, the mean length of employer tenure in our sample is relatively short,[6] although this is a well understood point about the NLSY79.[7] Annual occupational mobility in our sample is 16.10%, comparable to that reported in Kambourov and Manovskii (2008) who use PSID for the period, 1968–1997.

3.2 ASVAB

Between 1973 and 1975 all major branches of the U.S. military introduced the ASVAB test, designed and maintained by professional psychometrics, to place new recruits into jobs. The version of the ASVAB taken by NLSY79 respondents had 10 component tests: Word Knowledge, Arithmetic Reasoning, Mechanical Comprehension, Automotive and Shop Information, Electronics Information, Mathematics Knowledge, General Science, Paragraph Comprehension, Numerical Operations, and Coding Speed. To process the ASVAB scores, we followed a simplified version of the process described in Altonji *et al.* (2008). 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 effect on ASVAB score, we normalize the mean and variance of each test score by their age-specific values.

To increase the ASVAB's general appeal, the ASVAB Career Exploration Program was established to provide career guidance to high school students. For this purpose, they created a mapping between ASVAB test scores and O*NET occupation requirements

[6]Using PSID, Kambourov and Manovskii (2009b) report the mean employer tenure as 9.87 years.

[7]The mean employer tenure reported in Pavan (2011) is 3.3 years. Parent (2000) also calculates the mean employer tenure as 3.0 years using NLSY79.

Table I – Descriptive Statistics of the Sample, NLSY79, 1979–2010

Statistics	All Sample	$\leq$ High School	$>$ High School
Total number of the observations	44,589	21,616	22,973
Total number of the individuals	1,992	954	1,038
Average age at the time of interview	33.79	32.85	34.67
Highest education $<$ high school	7.83 %	16.34%	-
Highest education $=$ high school	40.06 %	83.76%	-
Highest education $>$ high school	52.11 %	-	100.00%
Highest education $\geq$ 4-year college	31.92 %	-	37.98%
Percentage of black people	11.40 %	15.09 %	8.00 %
Percentage of Hispanic people	6.78 %	7.34 %	6.26 %
Occupational mobility	16.10 %	18.27 %	14.05 %
Occupational tenure (mean)	6.41	6.10	6.69
Occupational tenure (median)	4.00	4.00	5.00
Employer (job) mobility	30.54 %	32.14 %	29.05 %
Employer (job) tenure (mean)	3.57	3.53	3.61
Employer (job) tenure (median)	2.00	2.00	2.00
Average hours worked within a year	1983	1958	2007

(described below). Later, we will describe how we use this information.

In this ASVAB and O*NET matching, they focused on only 7 tests, and from these 7 tests we reduce them into 2 categories, math and verbal, where the weights are set by principal components. The math component includes math knowledge and arithmetic reasoning and the verbal component combines word knowledge and paragraph comprehension. That these test actually measure different skills will be crucial to our study. It is well known, that many times individuals perform similarly on a battery of presumably dissimilar tests. Scores are certainly correlated between tests, between about 62% and 85%. Within math and verbal composites, the scores are very similar, but across verbal and math the correlation is only about 75%.

3.3 O*NET

Th 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, O*NET surveys workers in that occupation to give a score for the importance of each of 277 descriptors. From these descriptors we will use 26 descriptors that are most related to the ASVAB component tests and another 19 descriptors related to the physical demands. For the complete list, see Table II.[8] O*NET's occupational classification is more detailed than the codes we find associated with the NLSY workers, which are based on 3-digit census occupation codes. We average scores over O*NET occupation codes that map to the same 3-digit level occupation.[9]

TABLE II – List of Skills in O*NET

Skill's Name	
1.Inductive Reasoning	2.Written Comprehension
3.Oral Comprehension	4.English Language
5.Reading Comprehension	6.Deductive Reasoning
7.Inductive Reasoning	8.Technology Design
9.Number Facility	10.Mathematical Reasoning
11.Mathematics Information	12.Ordering
13.Science	14.Biology
15.Chemistry	16.Computers and Electronics
17.Physics	18.Engineering and Technology
19.Installation	20.Building and Construction
21.Troubleshooting	22.Operation and Control
23.Repairing Equipment	24.Maintenance
25.Mechanical	26.Equipment Selection

3.4 Mapping between ASVAB and O*NET

To match the O*NET information with ASVAB, we refer to a study by the Defense Manpower Data Center (DMDC) for their OCCU-Find program[10] that mapped ASVAB

[8]For each descriptor, there is both a "level" and an "intensity" score. Following the advice of the experts at the ASVAB Career Exploration Program, we only use the intensity score.

[9]Earlier papers, e.g. Autor *et al.* (2003) and Autor *et al.* (2003), aggregated DOT occupation codes into the relatively coarse Census occupation codes by using 1971 April CPS data which categorizes workers with both census and DOT occupation codes, which allows them to take a weighted average of DOT codes to the census code. For each DOT code, they can construct the fraction of CPS sampling weight within its a census code. We are using the O*NET, which, to our knowledge, does not have a CPS sample which categorizes workers by both O*NET and census codes which would allow us to create a similar weight scheme. Instead, we use a simple average.

[10]http://www.asvabprogram.com/downloads/Technical_Chapter_2010.pdf

scores to occupation information in the O*NET to help non-military high schoolers use their scores for career guidance.[11]

To explain the procedure, first let s_i^j denote the test score of individual i in test component j.[12] For each ASVAB category test, we create an O*NET analog by summing the 26 descriptors and weighting them by their "relatedness score" as judged by these experts. The result is that each occupation gets a score in each of the 7 categories, which are weighted averages of the 26 original O*NET descriptors. To put this mathematically, we use the experts' relatedness score to create a weighting matrix Ω, the rows of which all sum to one. We will work with a vector of scores for each occupation k, $\mathbf{r}_k = \Omega O_k^*$, where O_k^* is the vector of raw O*NET scores. The occupations' scores for these 7 O*NET categories, $\mathbf{r}_k$, are then comparable to the NLSY79 respondents' scores, $\mathbf{s}_i = \{s_i^1, s_i^2, ..., s_i^7\}$, a vector of 7 ASVAB category tests. The resulting O*NET descriptors, now in 7 dimensions to correspond to the 7 ASVAB tests, are highly correlated across occupations.

3.5 Health and Physical Requirements

Participants in the NLSY79 were asked to take a survey when they turned age 40 to evaluate their health status. In particular, the survey included 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 (EVGFP) and (v) whether pain interfered with their daily activities.[13] 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)). To circumvent this problem, we regress the PCS-12 score onto an array of health indicators available in the NLSY79 that are measured at the beginning of an individuals working life. In particular, we use parents' health status, body mass index (BMI) at first year of survey participation, height, and AFQT, along with several fixed demographics.

[11] The DMDC commissioned 14 expert judges, 12 PhDs, and 2 MAs to perform the mapping analysis.

[12] These components refer to the 10 categories of questions described earlier.

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

TABLE III – First stage regression results of current health on various work-life-independent statistics.

Dependent variable: Health index			
Regressors	Estimates	Regressors	Estimate
BMI>25	87.535***	AFQT	31.916***
	(26.482)		(4.740)
BMI<25	62.062***	Dad Alive	–46.714
	(18.891)		(32.316)
BMI^2	–1.806***	Dad Death Age	–1.472***
	(0.441)		(0.514)
Height	–6.066***	Mom Alive	258.068***
	(1.330)		(45.196)
Year of birth	–5.504***	Mom Death Age	3.720***
	(1.729)		(0.742)
Observations	28,144	R-squared	0.054

Note: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Because BMI could have a non-linear relationship with the health score, we segment it into below average and above average. As for the AFQT score, even though it is not a measure of health, it is likely to be correlated with other unobservable factors that contribute to one's general physical well-being.

This first stage regression is summarized in Table III. The F-statistic of the joint test that all coefficients are jointly zero is 116, with a *p-value* of zero (to the fifth digit).[14] In our later wage regressions, we will include these variables in the list of instruments to purge the PCS-12 of its effects from earnings and working. Because PCS-12 has a somewhat arbitrary scale, we rescale our health index to have a minimum at zero and standard deviation of one, as with all of our worker ability measures.

As mentioned earlier, to measure the physical requirements 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). We provide the full list of these descriptors in Table II. To reduce the 19 descriptors to a single index measure, we take the first principal component and then normalize it to have

[14] The statistic has a distribution F(14, 28129) with a p-value of 0.000. Further, all coefficients are individually significant at 1% level, except for race and whether the birth father is alive.

a zero minimum and standard deviation of one. As it turns out, physical and cognitive skills are not particularly highly correlated, either as provided by workers or demanded by their occupations. The correlation matrix between occupations' physical demands, the workers' physical abilities and the first principal component of cognitive skills is given in Table IV

Table IV – Correlations between skill dimensions

	Worker Verb	Worker Math	Worker Phys	Occ Verb	Occ Math	Occ Phys
Worker Verb	1.00					
Worker Math	0.80	1.00				
Worker Phys	-0.10	-0.11	1.00			
Occ Verb	0.35	0.42	-0.11	1.00		
Occ Math	0.31	0.39	-0.09	0.95	1.00	
Occ Phys	-0.33	-0.38	0.09	-0.49	-0.45	1.00

4 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.[15] To implement this measure, we first create a three-dimensional skill index in which we take Verbal and Math skills from the ASVAB, and Physical abilities from the instrumented PCS-12 health score. In the next section, these measures will be used in Mincer-type wage regressions to estimate the (potential) wage gains from a better occupational match. Then, we will study how match quality evolves over the life cycle.

4.1 (Skill) Mismatch Measure

If there were only a single type of skill then we could rank all workers and occupations by how able the workers were at that skill and how intensively the occupation uses that skill. In a perfectly-sorted world, as if labor allocations were frictionless, the worker with the 99th percentile score would take the job with the 99th percentile highest intensity

[15] We have also considered a second measure that seems plausible. This measure—called *match quality*—is based on the correlation between the vector of skills possessed by the worker and the vector of (corresponding) skill requirements at his occupation. The results turned out to be quite similar to the baseline case, so we omit a discussion of those results in the interest of space. These results are available from the authors upon request.

TABLE V – Summary Statistics of the Distribution of Mismatch and Match Quality Measures.

	Mismatch measure		
	Mean	St Dev	Kelley
VMP	0.26	0.16	0.32
CP	0.28	0.15	0.14

VMP is our prefered measure, using skills in verbal, mathematical and physical dimensions. For comparisons' sake, we also incldue distributional features when we aggregate verbal and math to have cognitive and physical dimensions.

using this skill. In this one dimensional world, then, a natural definition of mismatch would be the distance between the worker's rank and the occupation's rank. Because in reality skills are multi-dimensional, we extend this notion by taking a weighted sum of the distances in each skill dimension.

To this end, we begin by converting every worker's skill test score to a quantile rank, $q(s_i^j)$. For every occupation's O*NET descriptors, we transform them to correspond to tests by using the mapping between ASVAB and O*NET described above, and then use the occupation's rank in each of the skill categories, $q(r_k^j)$. Then, we take the difference in each dimension and define weights using factor loadings from the first principal component and normalize these weights to sum to 1. We take the absolute value of these differences and sum them using our weights, ω^j, to obtain:[16]

$$m_{i,k} \equiv \sum_{j=1}^{J} \left(\omega^j \times |q(r_k^j) - q(s_i^j)| \right).$$

This measure is intended to capture the deviation from some perfect sorting. However, the weight on each type of skill was derived using principal components analysis. Thus, they do not correspond to the magnitude of the welfare loss due to mismatch, i.e. giving more weight to skills in which mismatch contribute. Rather, we give a higher weight to dimensions that capture higher interpersonal variation in match quality. This means that our measure captures the highest possible information from all skill dimensions, but does not weight to emphasize the productivity impact of mismatch.

The distribution of $m_{i,k}$ has a strong right skew to it, which is interesting because neither the density of $\bar{s}_i$ nor $\bar{r}_k$ is significantly skewed—in fact are both slightly left

[16]It turns out not to matter whether we use factors associated with the PCA of the distance or the absolute distances.

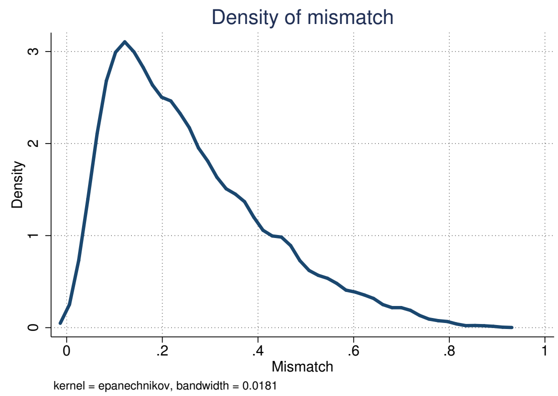

FIGURE 1 – Empirical Density of Mismatch

skewed. Table V gives summary statistics for this distribution, and Figure 1 displays the kernel estimate of its density.

4.2 The Effect of Past Matches

A key idea that we will explore in this paper is whether a poor match between a worker and his current occupation can have persistent effects *that last beyond the current job.* To this end, we construct a measure of *cumulative* mismatch. If an individual's prior occupations are indexed by $p = 1, \ldots, k$, then

$$\overline{m}_{i,k} \equiv \frac{\sum_{p<k} \tau_{i,p} m_{i,p}}{\sum_{p<k} \tau_{i,p}}$$

is the cumulative mismatch measure up to (but excluding) the current occupation k. Notice that each past match measure is weighted by the total tenure in that occupation, $\tau_{i,p}$. We then divide the summed measure by the total number of years worked prior to the current occupation.

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 would be the case, for example, in a model in which occupational match-quality is completely idiosyncratic) 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.

Table VI – Mismatch Across Education Groups

	Average	Std. Dev.
Less than HS	0.24	0.19
High school	0.28	0.17
College	0.23	0.15

4.3 The Distributions of Match Quality and Mismatch

To get a sense for the properties of the mismatch variable, we will discuss briefly how it varies across demographic groups. First, we divide the population into three education groups: high school dropouts, high school graduates, and college graduates. While the means across education groups are different, statistically speaking, they account for only 0.6% of the variance in cumulative mismatch. Further, only 1.8% of the variance in mismatch and 0.9% of the variance in cumulative mismatch is between education groups.

Table VI presents summary statistics for mismatch by education groups. Mismatch is lowest for the most educated, but the highest level is actually found among the high school educated, not those without a diploma. Although our work will not directly address this observation—our model will be agnostic as to formal education and we will attempt to control for education differences in our empirical work—it is consistent with our notion that "learning" plays a role in match quality. In this case, it may be that learning is happening outside of the labor market during education. For example, workers can gain a great deal of experience with different fields of work by sampling the corresponding fields of study.

How does mismatch vary across industry groups? We find that most of the dispersion is within industries, with only about 1.3% of mismatch variance found between industries. Table VII presents summary statistics for mismatch within industries along with the real average weekly wage rate. Notice that industries with higher wages generally have higher match quality.[17]

[17] However, when we look only at two dimensions, combining math and verbal skills into one cognitive skill measure, there is little relationship between high wage and low mismatch industries. The two highest wage sectors also have the highest match quality, the financial sector (Finance, Insurance, Real Estate) and professional services and notice that they also have the least dispersion in match quality.

Table VII – Mismatch Across Industries

	Average	Std. Dev.	Avg. Hourly Wage
FIRE	0.23	0.14	26.86
Prof Services	0.24	0.15	23.00
Trans Comm Util	0.24	0.15	20.80
Public Admin	0.25	0.16	20.74
Mfg	0.25	0.16	19.22
Agriculture	0.26	0.18	11.74
Retail Trade	0.26	0.16	15.43
Mining	0.27	0.18	16.78
Business Svcs	0.27	0.17	16.28
Construction	0.28	0.18	16.67
Personal Svcs	0.29	0.19	13.84
Entertainment Rec	0.32	0.20	21.72

4.4 Econometric Model of Wages

We begin with a basic Mincer-type wage regression, and then 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 wage growth turns out the matter for current wages, then this would provide evidence that match quality affects human capital accumulation and life-cycle wage dynamics.

In the basic wage regression, log real wages are linked to a variety of factors: education, labor market experience, and occupational tenure. Our addition is a measure of occupational match quality and its interaction with occupational tenure. As was recognized as far back as Altonji and Shakotko (1987), wage regressions that include a tenure variable are potentially affected by an endogeneity problem that comes from omitted individual- and match-specific factors, which are likely to be correlated with tenure. We deal with this issue in two separate ways. First, the fact that we are including a measure of match quality as a regressor ameliorates this omitted variables problem, by taking out

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

an important factor that was previously treated as unobservable (and hence part of the residual) into the set of observable regressors.

Second, and further, we build on a long list of studies, such as Altonji and Shakotko (1987), Topel (1991) and Altonji and Williams (2005) to instrument for occupational tenure as explained in more detailed next.

Instrumental Variables

First to understand why a straight OLS might be biased, consider the wage equation for individual i in (occupation) match k at time t:

$$\log w_{i,k,t} = x'_{i,t}\beta_1 + y'_{k,t}\beta_2 + \alpha a_{i,t} + \gamma\tau_{i,k,t} + \epsilon_{i,k,t}, \tag{4}$$

where $x_{i,t}$ and $y_{k,t}$ are vectors of observable worker and job characteristics, respectively; $a_{i,t}$ is labor market experience, and $\tau_{i,k,t}$ is the tenure in that occupation. The error term $\epsilon_{i,k,t}$ can be decomposed into

$$\epsilon_{i,k,t} = \mu_i + \theta_k + \phi_{i,k} + \eta_{i,k,t},$$

an individual and an occupation-specific effect plus an orthogonal error.

Coefficients on tenure in the model described by Equation 4 could be biased because the duration of the match is endogenous, so the levels of $\phi_{i,k}$ and μ_i could affect $\tau_{i,k,t}$. To address this issue, Altonji and Shakotko (1987) proposed to instrument $\tau_{i,k,t}$ with $\tilde{\tau}_{i,k,t} = \tau_{i,k,t} - \overline{\tau}_{i,k}$. The term $\overline{\tau}_{i,k}$ is the average match duration for agent i over the job spell at occupation k:

$$\overline{\tau}_{i,k} \equiv \frac{1}{t^*}\Sigma_{t=1}^{t^*}\tau_{i,k,t},$$

where t^* is the total length of the spell.[19] For example, if an individual is observed in an occupation at tenure 1 through 5 years, then $\overline{\tau}_{i,k}$ is simply 3 years (15/5). The normalized residual, $\tilde{\tau}_{i,k,t}$, is orthogonal to μ_i and $\eta_{i,k}$ because it looks only within the duration of the match. An appropriate correction for higher order polynomial terms is also available: $\tilde{\tau}^p_{i,k,t} = \tau^p_{i,k,t} - \overline{\tau^p}_{i,k}$ for $p > 1$.[20]

Further, our set of regressors also include several variables that interact with tenure and, hence, inherit the potential endogeneity problem with tenure. For each of these

[19] Talk about right truncation at the end of the sample...

[20] There is still a great deal of discussion about how to properly estimate the returns to tenure, see e.g. Buchinsky *et al.* (2010).

interaction terms, we create an instrument replacing tenure with its instrument.[21] Also, because our indicator for a continuing job, "OJ," is a function of tenure, we instrument this too. This indicator separates the effect of the first year of a match from the rest of the path from tenure, and hence its effect is relative to the length of tenure. We instrument that with OJ - mean(OJ), computed just as in our $\tau_{i,k,t} - \bar{\tau}_{i,k}$.

As noted earlier, the physical skill measure might also suffer from endogeneity and needs to be instrumented. We do this slightly differently. Because health score enters the mismatch variable in a nonlinear fashion, we cannot directly use the PCS-12 composite health score to compute mismatch and then include our health instruments in the wage regression because that would be a linear instrumenting scheme whereas PCS-12 enters nonlinearly. Instead, we run a first stage regression directly on the health variable, and then use this projected value to construct mismatch, cumulative mismatch and its interactions. These exogenous health variables, are then included in the vector of exogenous variables we use for later regressions as with all of our other instruments.

For each of our endogenous variables, the first stage regressions are quite strong, as expected given our instrumenting scheme. For example, the correlation between $\tau_{i,t,k}$ and $\bar{\tau}_{i,k}$ is 0.70. The *p-value* of the F-test for joint significance is approximately zero in every instance.

For our regressions, we expand on Equation 4 and incorporate our mismatch measure, $m_{i,k}$, and follow the same instrumenting scheme outlined above, as in Altonji and Shakotko (1987), Parent (2000), and Kambourov and Manovskii (2009b). We also include the interaction of $m_{i,k}$ with $\tau_{i,k,t}$, 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

$$\log w_{i,k,t} = x'_{i,t}\beta_1 + y'_{k,t}\beta_2 + A(a_{i,t}) + B(\tau_{i,k,t}) + \beta_m \times \overline{m}_{i,k} + \alpha \times \tau_{i,k,t} \times m_{i,k} + \epsilon_{i,k,t}, \quad (5)$$

where A and B are both polynomials.[22] The vector $x_{i,t}$ includes the usual suspects, such as education dummies, but also a measure of general intelligence. This is an important variable to include because we might worry that our match quality measures are just proxies for an individual effect. Cawley *et al.* (1996) incorporated “cognitive ability” or

[21] For example, we treat $\bar{s}_i \times \tau_{i,t,k}$ as an endogenous variable and instrument with $\bar{s}_i \times \tilde{\tau}_{i,t,k}$.

[22] It is possible to include a higher polynomial term in tenure to interact with mismatch, i.e., $\sum_{p=0}^{P} \alpha_p \times \tau^p_{i,k,t} \times m_{i,k}$. We have considered up to a fifth order polynomial for C, though, it turns out that the higher-power interactions are unnecessary, so we report the results for the linear case only.

general intelligence into a wage regression in much the same way as we incorporate match quality. They build their measure from the first principal component of ASVAB scores while we include the rank of the average of ASVAB scores.

5 Empirical Results

5.1 Mismatch and Wages

Table VIII presents the key results from our wage regressions incorporating mismatch in two dimensions: cognitive and physical skills. The OLS estimates change relatively little, shown in IX. The full regression results are presented in the Appendix in Table XIII along with OLS results. The first column includes into the a standard wage regression our measure of mismatch and its interaction with occupational tenure (reported in the first two rows). Notice that cumulative past mismatch has a strong effect even when the effect of the current match is included. We have also controlled for individual-specific characteristics such as average O*NET score, average worker-side skills, their interactions with occupational tenure and the AFQT score. Cumulative mismatch has an effect beyond all of these descriptors.

Table VIII – IV Wage Regressions with Mismatch

Regressors	(1)	(2)	(3)	(4)	(5)
Mismatch	–0.101***	–0.060*			
Mismatch × Occ Tenure	–0.015***	–0.013**			
Mismatch Negative Components			0.153**		
Mismatch Positive Components			0.033		
Mismatch Negative × Occ Tenure			–0.009		
Mismatch Positive × Occ Tenure			–0.029***		
Mismatch Verbal				0.006	0.041
Mismatch Math				–0.080**	–0.086**
Mismatch Phys.				–0.054**	–0.033
Mismatch Verbal × Occ Tenure				–0.016***	–0.023***
Mismatch Math × Occ Tenure				–0.000	0.008
Mismatch Phys. × Occ Tenure				0.005*	0.010***
Cumul. Mismatch		–0.243***			
Cumul Mismatch, Neg.			–0.152***		
Cumul Mismatch, Pos.			–0.455***		
Cumul Mismatch, Verbal					–0.007
Cumul Mismatch, Math					–0.200***
Cumul Mismatch, Phys.					–0.113***
Mean Worker Ability	–0.003	0.172***	0.235***		
Worker Ability × Occ Tenure	0.024***	0.024***	0.043***		
Worker Ability Verbal				–0.228***	–0.111***
Worker Ability Math				0.220***	0.281***
Worker Ability Phys.				0.045**	0.038*
Worker Ability Verbal × Occ Tenure				0.016***	0.009**
Worker Ability Math × Occ Tenure				–0.003	0.001
Worker Ability Phys. × Occ Tenure				0.001	0.006**
Mean Occ Reqs	0.236***	0.265***	0.353***		
Occ Reqs × Occ Tenure	0.018***	0.018***	–0.008		
Occ Reqs Verbal				–0.098	–0.109
Occ Reqs Math				0.242***	0.282***
Occ Reqs Phys				0.061**	0.032
Occ Reqs Verbal × Occ Tenure				0.011	0.009
Occ Reqs Math × Occ Tenure				0.003	0.006
Occ Reqs Phys × Occ Tenure				–0.003	–0.007**
Observations	35,778	26,066	26,395	35,778	26,066
R–squared	0.356	0.321	0.327	0.359	0.326

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 occupation and industry. Standard errors are computed as robust Huber-White sandwich estimates.

Table IX – OLS Wage Regressions with Mismatch

Regressors	(1)	(2)	(3)	(4)	(5)
Mismatch	–0.135***	–0.115***			
Mismatch × Occ Tenure	–0.008***	–0.002			
Mismatch Negative Components			0.180***		
Mismatch Positive Components			–0.041		
Mismatch Negative × Occ Tenure			–0.014**		
Mismatch Positive × Occ Tenure			–0.017***		
Mismatch Verbal				–0.011	0.006
Mismatch Math				–0.095***	–0.100***
Mismatch Phys				–0.054***	–0.061**
Mismatch Verbal × Occ Tenure				–0.013***	–0.017***
Mismatch Math × Occ Tenure				0.003	0.011***
Mismatch Phys × Occ Tenure				0.008***	0.014***
Cumul Mismatch		–0.257***			
Cumul Mismatch, Neg			–0.207***		
Cumul Mismatch, Pos			–0.494***		
Cumul Mismatch, Verbal					–0.006
Cumul Mismatch, Math					–0.219***
Cumul Mismatch, Phys					–0.092***
Mean Worker Ability	0.062*	0.169***	0.272***		
Worker Ability × Occ Tenure	0.012***	0.021***	0.039***		
Worker Ability Verbal				–0.172***	–0.103***
Worker Ability Math				0.179***	0.263***
Worker Ability Phys				0.067***	0.023
Worker Ability Verbal × Occ Tenure				0.010***	0.008**
Worker Ability Math × Occ Tenure				–0.001	0.002
Worker Ability Phys × Occ Tenure				–0.002	0.008***
Mean Occ Reqs	0.314***	0.332***	0.393***		
Occ Reqs × Occ Tenure	0.005*	0.005	–0.017***		
Occ Reqs Verbal				–0.019	–0.100
Occ Reqs Math				0.237***	0.326***
Occ Reqs Phys				0.044**	0.006
Occ Reqs Verbal × Occ Tenure				–0.007	0.000
Occ Reqs Math × Occ Tenure				0.010	0.006
Occ Reqs Phys × Occ Tenure				–0.005**	–0.007***
Observations	35,778	26,066	26,395	35,778	26,066
R–squared	0.371	0.337	0.343	0.374	0.343

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 occupation and industry. Standard errors are computed as robust Huber-White sandwich estimates.

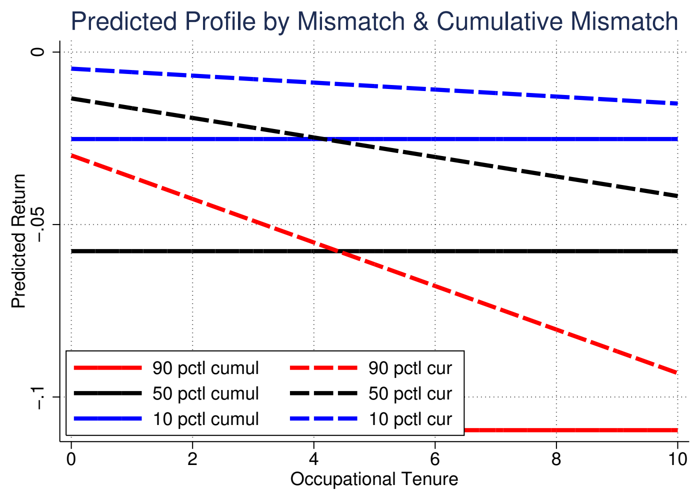

FIGURE 2 – For verbal and mathematical dimensions, predicted effect of cumulative mismatch at various quantiles of the distribution (solid) and the predicted effect of mismatch in the current occupation (dotted).

The cumulative effect of past mismatch has a persistent and significant effect on current wages. Current mismatch also has a strong effect in levels and this gets worse over the course of the match. To help interpret the meaning of these coefficients Figure 2 plots the implied effect on wages as a function of tenure in the occupation match. Looking at the effect of current mismatch, the difference between the 90th percentile worst matched and 10th percentile of mismatch is small at first, about 2.5%. It widens to 7.7% after 10 years. Comparing the 90th percentile to 10th percentile of cumulative mismatch, we see a wage difference of 8.6%.

Looking at math and verbal skills a pattern emerges: mismatch in either dimension has a negative effect on wages but the verbal dimension has a more dynamic effect within a match. It strongly increases over the course of match, whereas the mismatch effect in math is immediate and then tapers slightly. We see this in Figure 3, where we have plotted the predicted effects of mismatch. Overall the effect of math-side mismatch is stronger until about 7 years of tenure have been accrued. It is also notable that cumulative mismatch is much stronger in the math dimension than verbal. While it's difficult to perfectly explain this finding, it may be a result of occupational switching. In our data, the median occupational tenure is only 4 years (it is skewed significantly right). If the effect of verbal mismatch materializes only after 7 years, then workers might not have time to accumulate the effect.

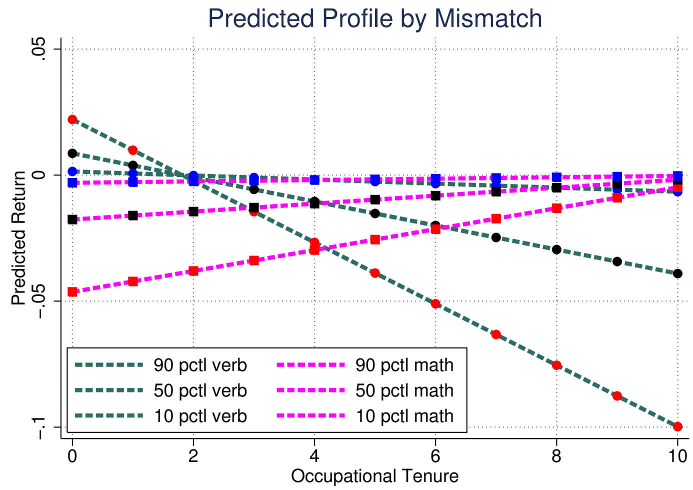

FIGURE 3 – For verbal and mathematical dimensions, predicted effect of mismatch in the current occupation for the Verbal (emerald) and Mathematical (magenta) dimensions. The implied effect of physical mismatch is suppressed for visual clarity.

5.2 Mismatch and Occupational Switching

Just as worse matches are related to lower wages, they also correspond to a greater likelihood of switching occupations. To study this, we estimate a probit model for occupational switching on various demographic characteristics and mismatch. On the left hand side of this probit we have the switch outcome in the next period and on the right hand side, we are chiefly interested in the contribution from mismatch in the current occupation. Marginal effects are presented in Table X, where the dependent on each column is an occupation switch. We leave off the effects of various demographic controls, AFQT score and occupational and industry dummies. Notice that mismatch is strongly related to the probability of switching occupations. From the 90% most mismatched to the 10% least mismatched reduced one's switching probability by 2.1 percentage points, a fall of about a 14.2% of the average switching rate.

Previous literature, e.g. Groes *et al.* (2010), had suggested that both over- and under-matched workers were more likely to switch occupations and we see a version of this result. We split out the negative and positive components of mismatch, i.e. for each individual, we create two variables, one containing the negative components and the other containing the positive components. In the probit, if the negative components are more negative, the workers skills are worse than the occupations, then the chance of switching is higher. Similarly, positive components also predict a higher chance of

TABLE X – Probit regression results with mismatch. All regressions include terms for demographics, AFQT and dummies for occupation and industry.

Regressors	(1)	(2)	(3)	(4)
Mismatch	0.034***			
Mismatch Verbal			0.024**	
Mismatch Math			0.007	
Mismatch Phys			0.008	
Mismatch Negative Components		-0.008		
Mismatch Positive Components		0.062***		
Negative Mismatch Verbal				0.067***
Positive Mismatch Verbal				0.116***
Negative Mismatch Math				-0.053**
Positive Mismatch Math				-0.015
Negative Mismatch Phys				0.016
Positive Mismatch Phys				0.112
AFQT	-0.002	-0.005	-0.002	-0.006
Mean Worker Ability	-0.037**	-0.058***	-0.036**	-0.073***
Mean Occ Reqs	-0.055***	-0.017	-0.056***	
Occ Tenure	-0.012***	-0.012***	-0.012***	-0.012***
Experience	0.005**	0.005**	0.005**	0.005**
Experience$^2 \times$ 100	-0.064***	-0.063***	-0.064***	-0.062***
Experience$^3 \times$ 100	0.001***	0.001***	0.001***	0.001***
Observations	33,544	33,544	33,544	33,544

Note: *** p<0.01, ** p<0.05, * p<0.1

switching. These coefficients are in Column 2 of Table X. Column 3 tries to break apart even further the switching probability result. In each dimension, we see that this pattern maintains.

In Figures 4 and 5 we present the non-parametric relationship between switching probability and mismatch. In each figure we present raw correlations computed locally across. On the right panel we present the local correlation of mismatch and switching probability after each has been projected onto an array of demographics and occupation-industry dummies as we had included in our probit regression above. In Figure 4 we see how mismatch is uniformly associated with a higher probability of switching occupations. At lower levels of mismatch, some of the raw correlation is explained by other

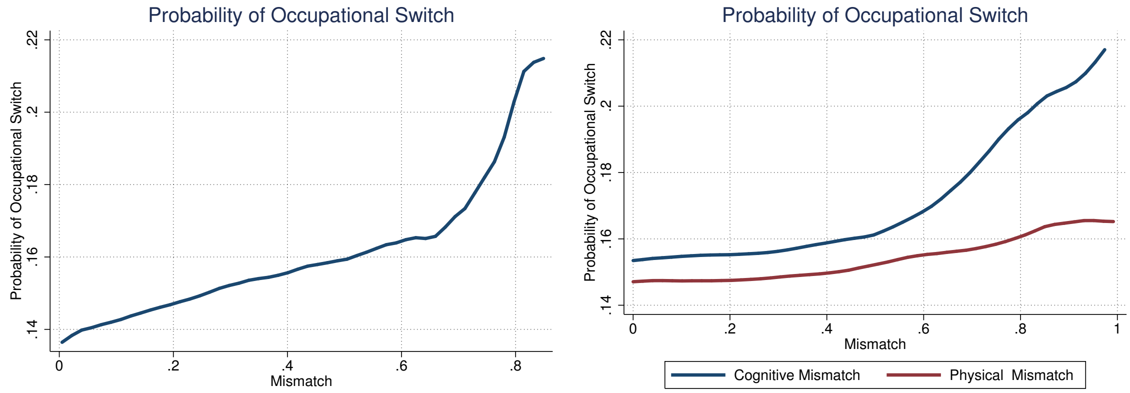

FIGURE 4 – The probability of switching occupations as a non-parametric function of mismatch. The probability of switching increases both in the unconditional relationship (left) and conditional on age, demographics, AFQT, average workers' ability and occupation/industry dummies (Table X). The effect of cognitive mismatch (right) is stronger than physical mismatch.

demographics, but overall the conditional correlation is stronger than the unconditional.

Then, in Figure 5, we see that negative mismatch has a considerably larger effect on switching probability than positive mismatch. This is especially true when we look at the raw correlation in the left panel. On the right panel, we see positive mismatch associated with a slightly higher probability of switching. The difference between unconditional and conditional correlations comes in part from the relationship between mismatch and overall skills, proxied by education and AFQT. Higher skill workers have lower probabilities of switching and higher skill workers are more likely to have skills greater than their occupations' requirements, positively mismatched. Once we condition on the effect of AFQT, mean worker's ability, education and the average for the occupation-industry, we see that marginally overqualified workers are more likely to switch occupations.

As in the wage regressions, we split the cognitive measure of mismatch to look within verbal and mathematical components. The results of this probit are in Table XI, where we suppress coefficients on the other demographic controls and dummies for occupation and industry. The magnitude on the marginal effect of overall mismatch is approximately comparable. To help understand the coefficient, the difference in switch probability between the 90th 10th deciles of mismatch are 2.2%, which is 14.5% of the average switching rate.

Again, physical mismatch has a smaller effect than cognitive mismatch, but it seems that verbal mismatch is more significant than mathematical. Notice also that in this case,

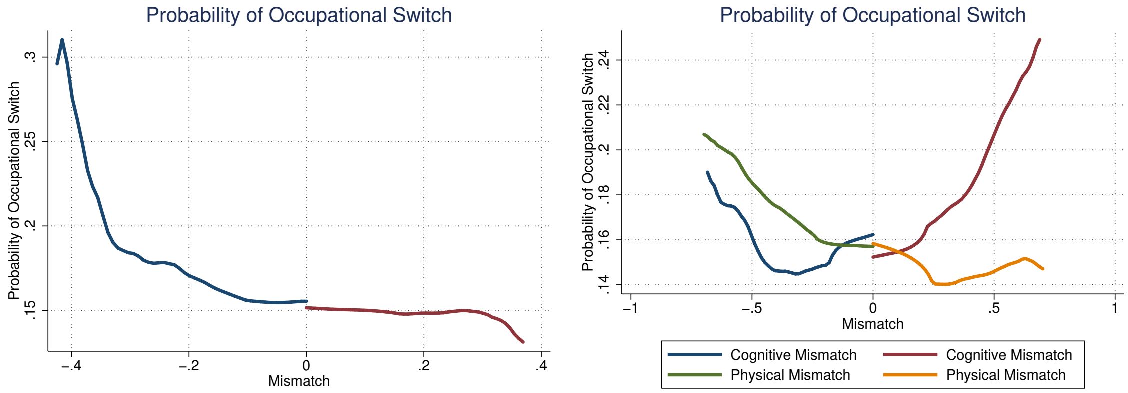

FIGURE 5 – The probability of switching occupations as a non-parametric function of mismatch split between positive and negative dimensions. The probability of switching is quite high for the negative mismatch, both in the unconditional relationship (Left) and conditional on age, demographics, AFQT and occupation/industry dummies (Table X). However, only when we condition on these other aspects, does positive mismatch predict occupational switching. Separating out cognitive and physical dimensions (Right), notice that cognitive mismatch is associated with increased switching for both positive and negative mismatch, whereas physical mismatch only affects switching probability for workers with negative mismatch.

the positive mismatch has a stronger effect than before. This is because the marginal effect of the positive verbal and mathematical components actually go in different directions, as we can see in Column (3). Being over qualified in mathematical fields actually has little to do with switching, but slightly reduces it whereas for verbal skills positive mismatch is associated with higher occupational mobility.

Visualizing the relationship in Figures 7 and 6 we see that again mismatch is positively associated with switching, both conditional on our battery of controls and unconditionally. Different this time, the effect of negative mismatch terms disappears when we condition on observable characteristics in our regressions. Evidently, under-matched workers also have other characteristics that predict higher rates of occupational mobility. In particular, workers whose skills are too low for their occupation tend to have lower observable measures of skill like education and AFQT which correlates with higher switching rates and picks up this effect. On the other hand, positive mismatch still corresponds to a higher probability of switching, especially at high values.

5.3 Switch Direction

Beyond that mismatched workers switch more frequently, their switches are also directional. Workers who are lacking in a skill relative to their occupation will, on

Table XI – Probit regression results with within-cognitive mismatch. All regressions include terms for demographics, AFQT, average workers' ability, and dummies for occupation and industry.

VARIABLES	(1)	(2)	(3)	(4)
Mismatch	0.049***			
Mismatch Verbal			0.028**	
Mismatch Math			0.015	
Mismatch Phys			0.016*	
Mismatch Negative Components		–0.020		
Mismatch Positive Components		0.317***		
Negative Mismatch Verbal				0.092***
Positive Mismatch Verbal				0.148***
Negative Mismatch Math				–0.094***
Positive Mismatch Math				–0.029
Negative Mismatch Phys				–0.029
Positive Mismatch Phys				0.201*
AFQT	–0.006*	–0.010***	–0.007**	–0.010**
Mean Worker Ability	–0.040**	–0.066***	–0.038**	–0.079***
Age	–0.004**	–0.004*	–0.004**	–0.004*
Age^2	–0.006*	–0.006**	–0.006*	–0.006*
Observations	33,544	33,544	33,544	33,544

*** p<0.01, ** p<0.05, * p<0.1

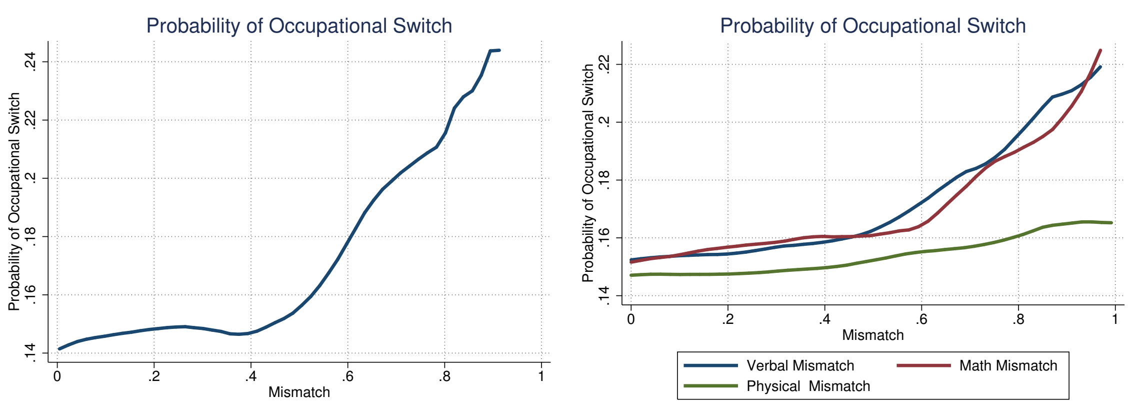

Figure 6 – The probability of switching occupations as a non-parametric function of mismatch including the within-cognitive terms. The probability of switching increases both in the unconditional relationship (Left) and in separating the dimensions of mismatch (Right) we see that the cognitive dimensions have a stronger effect than physical.

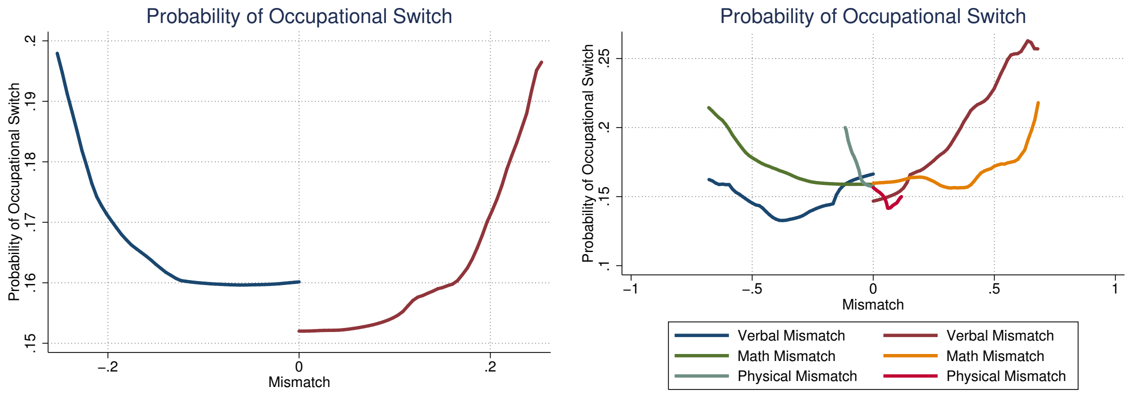

FIGURE 7 – The probability of switching occupations as a non-parametric function mismatch separating within-cognitive dimensions and split between positive and negative. Looking at the overall mismatch measure (Left) probability of switching is quite high for the negative mismatch, but only for the unconditional relationship (Table XI). When we split out across dimensions, again physical mismatch has little effect, positive verbal mismatch is quite strongly associated with switching and mathematical mismatch has a moderate relationship that is somewhat U-shaped.

average, switch to an occupation using that skill less intensively. Those with more of a skill than their occupation requires tend to switch into occupations requiring more. These patterns also hold when we look at average skills, so that a worker whose skills are, on average, better than required for his occupation will switch to an occupation with greater skill requirements, on average.

To help visualize this result, Figure 8 plots "signed mismatch" along the horizontal axis, we look at negative terms and separate them from positive terms of

$$m^+ \equiv \sum_{j=1}^{J} 1\{q(r_k^j) - q(s_i^j) > 0\} \times \omega^j \left(q(r_k^j) - q(s_i^j)\right)$$

for the prior match. On the vertical axis, we have the change in occupational skill requirements from the prior to next match. The upward sloping line means that workers whose own skills were low compared to their occupation will tend to switch into a lower skilled occupation and those whose skills were relative high move up the skill ladder. The picture looks much the same when we first condition on a list of demographics and measures of workers' ability.

Table XII includes results from the regression of the change in skill on the left hand side and the direction of mismatch on the right hand side. We also condition on the same variables as in our switching probability regressions, education, age, demographics, AFQT, mean ability and occupation and industry dummies. Notice that all of the coefficients on mismatch are significant and positive. Hence, in columns (2) and (3) the positive coefficient on negative mismatch means that workers whose skill is lower ranked than their occupations in a particular dimension tend to move to occupations with a lower skill rank and the term on positive mismatch implies that workers with higher skills than their occupation will move to more skill intensive occupations in that dimension. In column (1), the negative terms are associated with an average change in skill downwards and positive terms are associated with movements towards higher skill jobs.

As we have seen, there is a distinct direction to switching. Positive mismatch terms have an association with switches to higher skill occupations and the converse is true for negative mismatch terms. Both verbal and mathematical mismatch predict changes in skill requirements in these directions. The median positive verbal mismatch brings an increase in verbal skill requirements of 23 percentage points and a median negative mismatch brings a decline in skills of 20 percentage points. In the mathematical dimensions, the median positive mismatch corresponds to a change in skill of 22 percentage points and the median negative mismatch switches down in mathematical skills by 19 points. Because we are running the regression for physical mismatch on its own, we get the exact same implied skills change as before, down by 27 or up by 24 percentage points though the coefficient differs because of different scaling when we use the individual components of two or three dimensions of mismatch.

Both mathematical and verbal skill requirements tend to increase with age and with each switch. The average increase in verbal skills with each switch is 2.2 percentage points and it is 1.8 points for math. For both, the median worker with positive mismatch moves up in the skill distribution by 3 points more than that the median worker with negative mismatch moves down. Again, both math and verbal skills are increasing on average over the life-cycle and this result shows that some of that is associated with the larger moves of workers who are too well qualified.

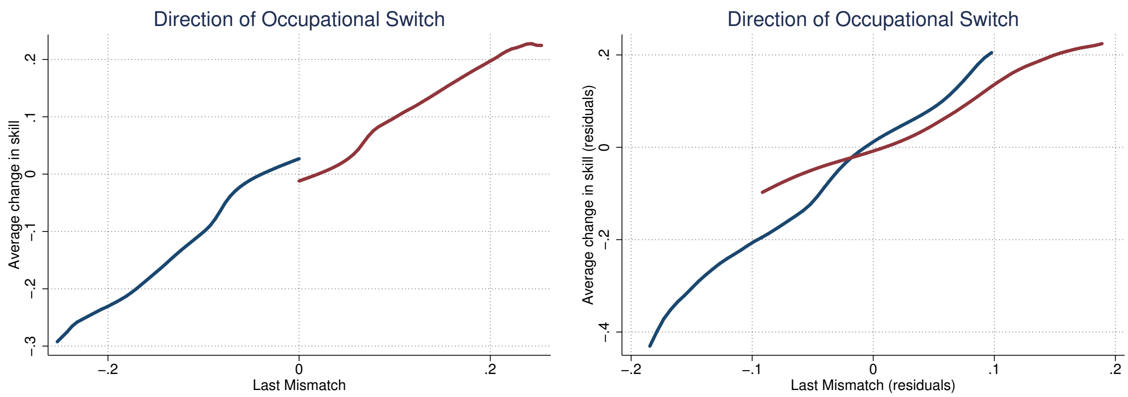

FIGURE 8 – The relationship between a worker's mismatch and the change in occupational skill requirements after a switch. (Left) is the non-parametric function of change in average skill from one occupation to another against the signed mismatch and (Right) is the same relationship but as a partial correlation, after conditioning on various demographics, including a quadratic in age, eduction, standardized AFQT and the worker's average skill level.

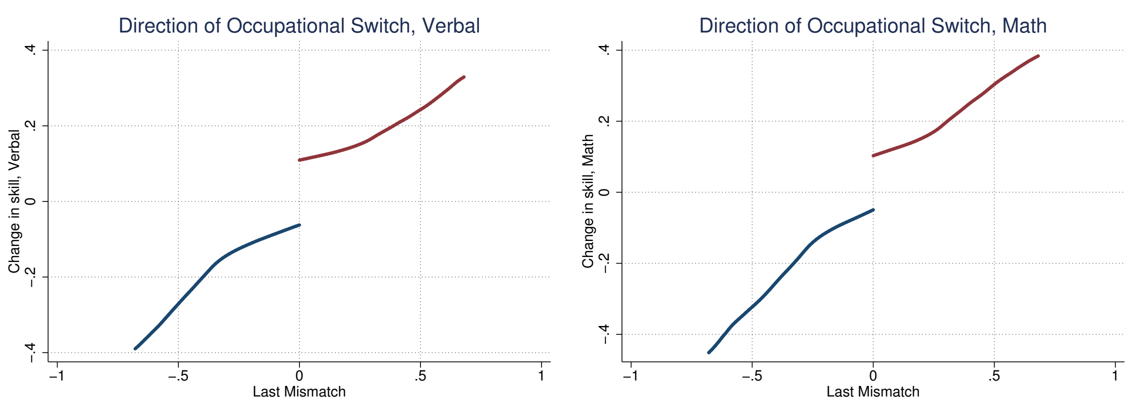

FIGURE 9 – The relationship between a worker's mismatch and the change in occupational skill requirements after a switch. (Left) is the change in the cognitive dimension as a function of verbal mismatch. (Right) is the same relationship for the mathematical dimension.

Table XII – Regression results of a worker's mismatch and the average change in skill required by an occupation. All regressions include terms for demographics, AFQT, average workers' ability, and dummies for occupation and industry.

Variables	(1)	(2)	(3)	(4)
Last Mismatch, Pos	1.854***			
Last Mismatch, Neg	2.140***			
Last Mismatch Pos, Verbal		1.336***		
Last Mismatch Pos, Math			1.344***	
Last Mismatch Pos, Phys				8.423***
Last Mismatch Neg, Verbal		1.354***		
Last Mismatch Neg, Math			1.344***	
Last Mismatch Neg, Phys				8.418***
Worker Ability Verbal		–0.972***		
Worker Ability Math			-0.827***	
Worker Ability Phys				-0.986***
Mean Worker Ability	–0.309***			
Age	0.002	0.007**	0.007**	-0.002
Age^2	–0.002	–0.007	–0.007	0.001
AFQT	–0.049***	0.027***	0.000	–0.011***
< High School	0.004	0.002	0.010	–0.010
4-Year College	–0.013**	0.006	–0.007	–0.032***
Hispanic	0.015*	–0.007	–0.008	0.001
Black	–0.000	–0.016*	–0.004	–0.016**
Constant	0.807***	0.284***	0.516***	0.463***
Observations	5,053	5,053	5,053	5,053
R-squared	0.472	0.655	0.622	0.721

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

6 Conclusion

In this paper, we provide new evidence on the importance of the match between workers' ability and the skills required by occupations on worker's wages and occupational switching behavior. Our novel approach is that we create measures of the match quality between worker's ability and the skills emphasized by his/her occupation by using data from ASVAB, O*NET, and NLSY79, and analyze their effects on worker's wages and switching patterns. Our finding is striking; the mismatch between one's ability and required occupational skills decreases the wage growth over occupational tenure while a high correlation between one's ability and occupational requirements increases the returns to tenure. Furthermore, occupational switches are directed in the sense that they improve match quality. Even more interesting from the perspective of choosing the correct model, we see a very strong effect from previous occupations' match quality. We believe that this evidence suggests that self-discovery is an important aspect of a worker's experience in the labor market and switching occupations is part of this process. It seems that mismatch with occupations affects the accumulation of human capital and that this has a non-trivial effect on a worker's entire career.

We build a quantitative model in which agents are heterogeneous in their ability to acquire different types of human capital. These types of human capital are required in different proportions by different occupations and so working in an occupation that highly values the type of human capital the worker is good at accumulating is important for his wages. Complicating the occupational matching process, however, workers do not immediately know their ability and must learn it over time. The model introduced an interesting general equilibrium factor through a non-linear prices for human capital.

We have not yet completed our quantitative analysis using this model.

For future research, the next concern will be how the importance of the match quality changes over time. Kambourov and Manovskii (2008) show that the occupational mobility in the US has been rising during the period 1968 to 1997. Kambourov and Manovskii (2009a) argue that the rise of the occupational mobility is related to the increased with-in group wage inequality. We are planning to test their hypothesis directly by using the NLSY97 and the set of measures we provided in this paper.

References

Altonji, J., Bharadwaj, P. and Lange, F. (2008). Changes in the characteristics of american youth: Implications for adult outcomes. *NBER Working Paper*.

— and Shakotko, R. (1987). Do wages rise with job seniority? *The Review of Economic Studies*, **54** (3), 437–459.

— and Williams, N. (2005). Do wages rise with job seniority? a reassessment. *Industrial & Labor Relations Review*, **58** (3), 4.

Autor, D. H., Levy, F. and Murnane, R. J. (2003). The Skill Content Of Recent Technological Change: An Empirical Exploration. *The Quarterly Journal of Economics*, **118** (4), 1279–1333.

Buchinsky, M., Fougere, D., Kramarz, F. and Tchernis, R. (2010). Interfirm Mobility, Wages and the Returns to Seniority and Experience in the United States. *Review of Economic Studies*, **77** (3), 972–1001.

Cawley, J., Conneely, K., Heckman, J. and Vytlacil, E. (1996). *Cognitive Ability, Wages, and Meritocracy*. NBER Working Papers 5645, National Bureau of Economic Research, Inc.

Farber, H. S. and Gibbons, R. (1996). Learning and wage dynamics*. *The Quarterly journal of economics*, **111** (4), 1007–1047.

Groes, F., Kircher, P. and Manovskii, I. (2010). *The U-Shapes of Occupational Mobility*. Workingpaper <importmodel: Workingpaperimportmodel>, Department of Economics, Copenhagen Business School.

Kambourov, G. and Manovskii, I. (2008). Rising occupational and industry mobility in the united states: 1968–97*. *International Economic Review*, **49** (1), 41–79.

— and — (2009a). Occupational mobility and wage inequality. *The Review of Economic Studies*, **76** (2), 731–759.

— and — (2009b). Occupational specificity of human capital. *International Economic Review*, **50** (1), 63–115.

Michaud, A. M. and Wiczer, D. (2014). *Occupational hazards and social disability insurance*. Working Papers 2014-24, Federal Reserve Bank of St. Louis.

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

Pavan, R. (2011). Career choice and wage growth. *Journal of Labor Economics*, **29** (3), 549–587.

Topel, R. (1991). Specific capital, mobility, and wages: Wages rise with job seniority. *Journal of Political Economy*, **99** (1).

Ware, J., Kosinski, M. and Keller, S. (1995). Sf-12: How to score the sf-12 physical and mental health summary scales, 2nd edition. *Boston: The Health Institute, New England Medical Center*.

Appendix

A. Additional Tables and Figures

A.1. Wage Regressions

Table XIII – Additional Regressors Not Reported in Table VIII

Regressors	(1)	(2)	(3)	(4)	(5)
Emp Tenure	–0.012***	–0.007	–0.008	–0.013***	–0.008
Occ Tenure	0.013*	0.009	0.013	0.019***	0.015*
Emp Tenure$^2\times$ 100	0.055**	0.049	0.055*	0.063***	0.054
Occ Tenure$^2\times$ 100	–0.193***	–0.183***	–0.187***	–0.201***	–0.185***
Occ Tenure$^3\times$ 100	0.003***	0.004**	0.004**	0.004***	0.004**
Experience	0.061***	0.056***	0.055***	0.063***	0.058***
Experience$^2\times$ 100	–0.172***	–0.143***	–0.144***	–0.183***	–0.155***
Experience$^3\times$ 100	0.002***	0.001*	0.001*	0.002***	0.002**
Old Job	–0.020*	–0.021*	–0.020*	–0.020*	–0.021*
BMI–25 | BMI>25	–0.078***	–0.103***	–0.098***	–0.094***	–0.126***
BMI–25 | BMI<25	–0.038**	–0.062***	–0.059***	–0.051***	–0.078***
BMI^2	0.001***	0.002***	0.001***	0.001***	0.002***
height	–0.000	0.002	0.003**	–0.001	0.002*
Year of birth	–0.007***	0.003*	0.005***	–0.008***	0.002
AFQT	0.088***	0.072***	0.097***	0.093***	0.066***
Dad Alive	0.009	–0.005	0.005	0.017	–0.005
Dad Death Age	–0.000	–0.000	–0.000	–0.000	–0.000
Mom Alive	–0.120***	–0.118***	–0.071*	–0.108***	–0.106***
Mom Death Age	–0.003***	–0.003***	–0.002***	–0.003***	–0.003***
< High School	–0.023**	–0.019*	–0.017	–0.022**	–0.013
4–Year College	0.280***	0.259***	0.252***	0.270***	0.239***
Hispanic	0.011	0.032**	0.032**	0.012	0.035**
Black	–0.010	–0.026*	–0.022	–0.016	–0.038***
Observations	35,778	26,066	26,395	35,778	26,066
R–squared	0.356	0.321	0.327	0.359	0.326

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 occupation and industry. Standard errors are computed as robust Huber-White sandwich estimates.

Table XIV – Additional Regressors Not Reported in Table IX

Regressors	(1)	(2)	(3)	(4)	(5)
Emp Tenure	-0.001	0.003	0.005	-0.003	0.001
Occ Tenure	0.047***	0.037***	0.038***	0.050***	0.039***
Emp Tenure$^2\times$ 100	-0.007	-0.023	-0.034	-0.001	-0.010
Occ Tenure$^2\times$ 100	-0.250***	-0.205***	-0.193***	-0.255***	-0.224***
Occ Tenure$^3\times$ 100	0.004***	0.003**	0.003**	0.004***	0.003**
Experience	0.042***	0.037***	0.036***	0.044***	0.039***
Experience$^2\times$ 100	-0.095***	-0.078**	-0.080**	-0.107***	-0.089**
Experience$^3\times$ 100	0.000	0.000	0.000	0.001	0.001
Old Job	-0.000	0.004	0.004	0.000	0.005
BMI-25 $\mid$ BMI$>$25	-0.085***	-0.091***	-0.082***	-0.101***	-0.111***
BMI-25 $\mid$ BMI$<$25	-0.042***	-0.051***	-0.046***	-0.054***	-0.065***
BMI^2	0.001***	0.001***	0.001***	0.001***	0.002***
height	0.000	0.003**	0.004***	0.000	0.003**
Year of birth	-0.005***	0.003	0.006***	-0.005***	0.002
AFQT	0.085***	0.072***	0.101***	0.091***	0.066***
Dad Alive	0.013	0.002	0.007	0.023	0.007
Dad Death Age	-0.000	-0.000	-0.000	0.000	-0.000
Mom Alive	-0.121***	-0.130***	-0.073*	-0.111***	-0.107***
Mom Death Age	-0.003***	-0.003***	-0.002***	-0.003***	-0.003***
$<$ High School	-0.015	-0.013	-0.012	-0.015	-0.009
4-Year College	0.268***	0.249***	0.241***	0.259***	0.229***
Hispanic	0.010	0.031**	0.030*	0.007	0.030*
Black	-0.009	-0.026**	-0.020	-0.017	-0.038***
Observations	35,778	26,066	26,395	35,778	26,066
R-squared	0.370	0.337	0.343	0.374	0.342

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 occupation and industry. Standard errors are computed as robust Huber-White sandwich estimates.

B. Data

B.1. Panel construction and sample selection

To construct annual panel data for the main analysis, we use Work History Data File which records individuals' employment histories up to five jobs on weekly basis from 1978 to 2010. Following the approach by 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 spend the most hours worked within the year. We construct panel data with annual frequency (from 1979 to 2010) from a series of observations of primary jobs and out-of-labor-force status for each individuals. Finally, the annually-reported[23] demographic information and detailed information of employment (occupation, industry, and hourly wage) is merged to the panels.

Occupational Codes Before we merge the occupation information to the panel data, we clean occupational titles. In NLSY79, every year individuals report their occupations for up-to five jobs which they worked for 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 that mapping of jobs across interviews, we create an employment spell for each job. We then assign, to each employment spell, an occupation code which is most often observed during the spell.[24] 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.[25]

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,[26] we use our own crosswalk to convert them into the Census 1970 One-Digit

[23]In NLSY79, the information except weekly employment status is reported annually from 1979 to 1994, and biannually, from 1996 to 2010.

[24]We explain this cleaning procedure of occupation title more in details in the appendix.

[25]We used the Census 1970 Three-Digit Occupation Code to specify worker's occupational until 2000. After 2000, we used the Census 2000 Three-Digit Occupation Code. All of those codes are converted to the Census 1990 Three-Digit Occupation Code using the crosswalks, which are down loadable at the Minnesota Population Center's website (http://usa.ipums.org/usa/volii/occ_ind.shtml).

[26]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.

Industry Code.[27] We use those one-digit level industry codes to create industry dummies used in our analysis. After the conversion, we clean the codes by using job spells as described in the case of occupation codes.

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. To clean some potentially spurious switches, we do not count those that immediately revert, instead treating the year in question as an improper categorization.

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

Wages Worker's wages are measured by the usual rate of pay for the primary occupation at the time of interview. All the wages are deflated by the consumer price index, and converted 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 is similar to Pavan (2011), and doesn't affect the regression result.[28]

Sample Selection We follow the approach by Farber and Gibbons (1996) and Pavan (2011) for the sample selection. We, first, limit the sample to individuals who make their initial long-term transition from school to labor markets during the survey period. Thus, we drop the individuals who work more than 1,200 hours in the initial year of the survey. We also drop the individuals who don't 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 the sample. For the individuals who go back to school, we assume they start their career from the point they re-enter labor markets, and drop the observations before that. At this stage, we also drop the observations which are after the last time an individual is in the labor

[27] The crosswalk is shown in the appendix.

[28] Pavan (2011) drop the top ten and the bottom ten observations in the entire sample.

Table XV – 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 period	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 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	48,350
Drop those below age 16 and not enrolled in school	2,093	48,312
Valid ASVAB scores	1,996	46,251
Valid wage information	1,993	44,719
Drop top 0.1% and bottom 0.1% in the wage distribution	1,992	44,589

force. Furthermore, we drop individuals who are weakly attached to labor force; those who are out of labor force more than once before they work at least 10 years after they started their career. If an individual is out of 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 the sample to those who have valid occupation and industry code, who are equal to or above age 16 and not currently enrolled in a school, and who have valid ASVAB scores. The number of the remained individuals and observations after applying each sample selection criterion are summarized in Table XV.

B.2. The Crosswalk of Census Industry Codes

We used the crosswalk in Table XVI to convert the Census 1970, 1980, and 1990 Three-Digit Industry Code to the Census 1970 One-Digit Industry Code. 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, and 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. Otherwise, the Census 1970 and 1980 have the same classification in one-digit level.

It is not so straight-forward 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" (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 Table XVI.

TABLE XVI – The Crosswalk across the Census 1970, 1980, and 1990 Digit Industry 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

C. Proofs

Proof. (Proposition 1)

a. Note that

$$\begin{aligned}
\hat{A}_i(t) &= \frac{\lambda_i}{\lambda_i(t)}\hat{A}_i(1) + \frac{\lambda_{\epsilon_i}}{\lambda_i(t)}\left(\tilde{A}_i(1) + \tilde{A}_i(2) + \ldots + \tilde{A}_i(t-1)\right) \\
&= \frac{\lambda_i}{\lambda_i(t)}\hat{A}_i(1) + \frac{\lambda_{\epsilon_i}}{\lambda_i(t)}\left(A_i + \epsilon_i(1) + A_i + \epsilon_i(2) + \ldots + A_i + \epsilon_i(t-1)\right) \\
&= \frac{\lambda_i}{\lambda_i(t)}\hat{A}_i(1) + \frac{\lambda_{\epsilon_i}}{\lambda_i(t)}(t-1)A_i + \frac{\lambda_{\epsilon_i}}{\lambda_i(t)}\left(\epsilon_i(1) + \epsilon_i(2) + \ldots + \epsilon_i(t-1)\right)
\end{aligned}$$

$\hat{A}_i(1)$ is normally distributed with $N\left(A_i, \sigma^2_{\eta_i} + \sigma_{A_i^2}\right)$. Then, $\hat{A}_i(t)$ is sum of normal random variables as given

$$\begin{aligned}\hat{A}_i(t) &= \frac{\lambda_i}{\lambda_i(t)}(A_i+\eta_i)+\frac{\lambda_{\epsilon_i}}{\lambda_i(t)}(t-1)A_i+\frac{\lambda_{\epsilon_i}}{\lambda_i(t)}(\epsilon_i(1)+\epsilon_i(2)+...+\epsilon_i(t-1)) \\ &= A_i+\frac{\lambda_i}{\lambda_i(t)}\eta_i+\frac{\lambda_{\epsilon_i}}{\lambda_i(t)}(\epsilon_i(1)+\epsilon_i(2)+...+\epsilon_i(t-1))\end{aligned}$$

it will be normally distributed. From this expression, we see that

$$E\left[\hat{A}_i(t)\right]=A_i$$

and

$$\begin{aligned}Var\left(\hat{A}_i(t)\right) &= \frac{\lambda_i^2\left(\sigma_{\eta_i}^2+\sigma_{A_i}^2\right)+\lambda_{\epsilon_i}^2(t-1)^2\sigma_{A_i}^2+\lambda_{\epsilon_i}^2(t-1)\sigma_{\epsilon_i}^2}{\lambda_i(t)^2} \\ &= \frac{\left(\lambda_i^2+\lambda_{\epsilon_i}^2(t-1)^2\right)\sigma_{A_i}^2+\lambda_i+\lambda_{\epsilon_i}(t-1)}{\lambda_i(t)^2} \\ &= \frac{\left(\lambda_i^2+\lambda_{\epsilon_i}^2(t-1)^2\right)\sigma_{A_i}^2+\lambda_i(t)}{\lambda_i(t)^2}.\end{aligned}$$

Thus, the distribution of beliefs will have mean A_i and variance $\frac{\left(\lambda_i^2+\lambda_{\epsilon_i}^2(t-1)^2\right)\sigma_{A_i}^2+\lambda_i(t)}{\lambda_i(t)^2}$.

b. Using

$$\begin{aligned}\hat{A}_i(t) &= \frac{\lambda_i}{\lambda_i(t)}(A_i+\eta_i)+\frac{\lambda_{\epsilon_i}}{\lambda_i(t)}(t-1)A_i+\frac{\lambda_{\epsilon_i}}{\lambda_i(t)}(\epsilon_i(1)+\epsilon_i(2)+...+\epsilon_i(t-1)) \\ &= A_i+\frac{\lambda_i}{\lambda_i(t)}\eta_i+\frac{\lambda_{\epsilon_i}}{\lambda_i(t)}(\epsilon_i(1)+\epsilon_i(2)+...+\epsilon_i(t-1))\end{aligned}$$

we obtain

$$\begin{aligned}\hat{A}_i(t)-A_i &= \frac{\lambda_i}{\lambda_i(t)}\eta_i+\frac{\lambda_{\epsilon_i}}{\lambda_i(t)}(\epsilon_i(1)+\epsilon_i(2)+...+\epsilon_i(t-1)) \\ Var\left[\hat{A}_i(t)-A_i\right] &= \frac{\lambda_i^2}{\lambda_i(t)^2}\sigma_{\eta_i}^2+\frac{\lambda_{\epsilon_i}^2}{\lambda_i(t)^2}(t-1)\sigma_{\epsilon_i}^2 \\ &= \frac{\lambda_i}{\lambda_i(t)^2}+\frac{\lambda_{\epsilon_i}}{\lambda_i(t)^2}(t-1) \\ &= \frac{1}{\lambda_i(t)}.\end{aligned}$$

Since $E\left[\hat{A}_i(t) - A_i\right] = 0$,

$$E\left[\left(\hat{A}_i(t) - A_i\right)^2\right] = Var\left[\hat{A}_i(t) - A_i\right] + \left(E\left[\hat{A}_i(t) - A_i\right]\right)^2 = Var\left[\hat{A}_i(t) - A_i\right]$$

Since $\lambda_i(t) = \lambda_i + (t-1)\lambda_{\epsilon_i}$, it goes the infinity as t goes to infinity. Thus, the variance $\lim_{t\to\infty} Var\left[\hat{A}_i(t) - A_i\right] = 0$ and $\lim_{t\to\infty} E\left[\left(\hat{A}_i(t) - A_i\right)^2\right] = 0$. □

Proof. (Proposition 3) Note that probability of switchnig is given by

$$1 - \prod_i \text{Prob}\left(r_i(t-1) - m_i \leq r_i(t) < r_i(t-1) + m_i\right) = 1 - \prod_i \text{Prob}\left(A_i(t-1) - m_i \leq A_i(t) < A_i(t-1)\right.$$

Now look at the term $\text{Prob}\left(A_i(t-1) - m_i \leq A_i(t) < A_i(t-1) + m_i\right)$. Inserting

$$\hat{A}_i(t) = \frac{\lambda_i(t-1)}{\lambda_i(t)}\hat{A}_i(t-1) + \frac{\lambda_{\epsilon_i}}{\lambda_i(t)}\tilde{A}_i(t-1)$$

and

$$\lambda_i(t) = \lambda_i(t-1) + \lambda_{\epsilon_i},$$

we obtain

$$\begin{aligned} & \text{Prob}\left(A_i(t-1) - m_i \leq A_i(t) < A_i(t-1) + m_i\right) \\ = & \text{Prob}\left(-m_i \leq \frac{\lambda_{\epsilon_i}}{\lambda_i(t)}\left(A_i - \hat{A}_i(t-1) + \epsilon_i(t)\right) < m_i\right) \\ = & \text{Prob}\left(-m_i\frac{\lambda_i(t)}{\lambda_{\epsilon_i}} \leq \left(A_i - \hat{A}_i(t-1) + \epsilon_i(t)\right) < m_i\frac{\lambda_i(t)}{\lambda_{\epsilon_i}}\right). \end{aligned}$$

Given that $\hat{A}_i(t-1)$ is normally distributed with mean A_i and variance $\frac{\left(\lambda_i^2 + \lambda_{\epsilon_i}^2(t-2)^2\right)\sigma_{A_i}^2 + \lambda_i(t-1)}{\lambda_i(t-1)^2}$. Since $\frac{\left(\lambda_i^2 + \lambda_{\epsilon_i}^2(t-2)^2\right)\sigma_{A_i}^2 + \lambda_i(t-1)}{\lambda_i(t-1)^2}$ declines with age, probability of staying in the last period's occupation increases and probability of switching decreases with age. □

Proof. (Proposition 4) Given $\hat{A}_i(t-1)$, we will evaluate

$$\begin{aligned}
& \text{Prob}\left(A_i(t-1) - m_i \leq A_i(t) < A_i(t-1) + m_i\right) \\
= \; & \text{Prob}\left(-m_i \frac{\lambda_i(t)}{\lambda_{\epsilon_i}} \leq \left(A_i - \hat{A}_i(t-1) + \epsilon_i(t)\right) < m_i \frac{\lambda_i(t)}{\lambda_{\epsilon_i}}\right) \\
= \; & \text{Prob}\left(-m_i \frac{\lambda_i(t)}{\lambda_{\epsilon_i}} - \left(A_i - \hat{A}_i(t-1)\right) \leq \epsilon_i(t) < m_i \frac{\lambda_i(t)}{\lambda_{\epsilon_i}} - \left(A_i - \hat{A}_i(t-1)\right)\right).
\end{aligned}$$

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 $\left(A_i - \hat{A}_i(t-1)\right)^2$. □

Proof. (Proposition 5) Probability of switching to an occupation with higher skill-i intensity is

$$\begin{aligned}
\text{Prob}\left(A_i(t) > A_i(t-1) + m_i\right) &= \text{Prob}\left(\epsilon_i(t) > m_i \frac{\lambda_i(t)}{\lambda_{\epsilon_i}} - \left(A_i - \hat{A}_i(t-1)\right)\right) \\
&= \text{Prob}\left(\epsilon_i(t) > m_i \frac{\lambda_i(t)}{\lambda_{\epsilon_i}} - \left(r_i^* - r_i(t-1)\right)\right).
\end{aligned}$$

Note that probability of switching to an occupation with higher skill-i intensity increases with $(r_i^* - r_i(t-1))$. 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

$$\begin{aligned}
\text{Prob}\left(A_i(t) < A_i(t-1) - m_i\right) &= \text{Prob}\left(\epsilon_i(t) < -m_i \frac{\lambda_i(t)}{\lambda_{\epsilon_i}} - \left(A_i - \hat{A}_i(t-1)\right)\right) \\
&= \text{Prob}\left(\epsilon_i(t) < -m_i \frac{\lambda_i(t)}{\lambda_{\epsilon_i}} - \left(r_i^* - r_i(t-1)\right)\right).
\end{aligned}$$

Using these two equations above, it is easy to see that the probability of switching to a higher skill occupation is higher than the probability of switching to a lower skill occupation when $r_i^* - r_i(t-1) > 0$. Similar proof can be made for the case $r_i(t-1) - r_i^* > 0$. But we skip it for the sake of brevity. □