Multidimensional Skills, Wealth, and Occupational Choices

Kyungmin Kang*

School of Economics, Shanghai University of Finance and Economics

Jong Jae Lee†

Economics and Management School, Wuhan University

[Preliminary and Incomplete. Do not cite without authors' consent.]

March, 2022

ABSTRACT: We provide a framework to study the effect of wealth and the role of learning under uncertainty and risk aversion on occupational choices. Our model allows wealth to play two roles–make workers seek high-skill-requirement jobs to reduce uncertainty going forward and make them seek low-skill-requirement jobs to reduce immediate uncertainty regarding wages. Our analysis suggests that workers are uncertain of their skills throughout their career and workers continuously update their beliefs using wages as signals of their skills. Moreover, we provide a tractable method to estimate the theoretical model that requires the evaluation of multiple first-order conditions with unknown states.

*kyungmin.kang@outlook.com.

†piuslee83@gmail.com.

1 Introduction

The initial occupation of a person is found to have a lasting impact on the worker's wage and occupation choices throughout the lifecycle **?**. The match quality between workers and jobs has received particular attention, as good matches facilitate skill development and income growth. Differences in occupational choice across demographic groups are concerning if they reflect sub-optimal investments in human capital investment that exacerbate income and wealth equality.

In this paper, we study the role of uncertainty in occupational choice with a focus on the dynamic effect of wealth on earnings over the life cycle. We first develop a model of occupational choice where a risk-averse agent makes an occupational choice under uncertainty surrounding the worker's skill set. Workers can choose occupations with higher skill-level requirements than are optimal under certainty to reduce skills uncertainty in the long run. On the other hand, the workers can choose occupations with lower skill-level requirements to reduce wage uncertainty in the short run. We show that these two forces act in the opposite direction and the effect of wealth on occupation is thus theoretically ambiguous.

We provide evidence of the main assumption of the theoretical model that workers are uncertain about skills and wages provide signals of unknown skills by using wage residuals as a proxy of signals of unknown skills. We show using OLS and FE regressions that wage residuals, which we interpret as the unexpected portion of the wage that captures the difference between the true skill level and current belief. Using this strategy, we show that workers with high unexpected shocks move to occupations with a higher skill requirement. If no uncertainty exists or learning does not occur, wage residuals would have no effect on subsequent occupational choices.

Next, we document that, consistent with the other main assumption of the model, the precision of the information about the skills depends on the skill requirements of the current job. By using the interaction of the wage residuals and the level of skill requirements, we show that jobs with higher skill-level requirements provide a more precise signal about the workers' true skills. That is, occupations that require high cognitive skills, such as economists or engineers, contain more information about their workers' cognitive skill levels than occupations that require lower cognitive skills.

The large state space and the intricate link between state variables and the choice variables make the estimation of the model using the traditional structural method infeasible. We provide a tractable estimation method in the spirit of **?** and **?**. We first approximate the true decision rules of the individuals by general polynomial functions of the state variables, extending the method developed in **?** for discrete dynamic choice problems to continuous dynamic choice. We then use the Kalman filter and simulated the maximum likelihood to recover the distribution of the hidden true skills and to identify the key parameters of the theoretical model. We further discuss the identification of the coefficient of risk aversion.

The paper adds to three strands of literature. First, we add to the literature on learning and labor market transitions. There is extensive literature on skill uncertainty and job choice following the seminal work by **?**. In Jovanovic's model, both workers and employers are uncertain about workers' skills, and information is revealed only if they are matched. We add to the literature by introducing risk-aversion and thus by allowing wealth to affect the occupational choice of workers.

A rapidly growing literature considers task-specific approaches (**?** is among the early and more influential work; **?** provides good theoretical references). **?** is among the first empirical papers that consider occupational choices given multidimensional skills and skill requirements. This paper is most closely related to **?** who considers the effect of wealth on occupational choice through the myopic channel we explained above. We add to the literature by providing a theoretical framework and empirical methods to study the long- and short-term role of wealth–'probing', or the incentive to get high-skill-requirement jobs, and 'caution', or the incentive to get low-skill-requirement jobs.

Section 2 introduces the model. Section 3 discusses the mechanisms of the model and its theoretical implications. Section 4 describe the data use used in the project and document some basic patterns in the data that provides evidence supporting assumptions made in the model. Section 5 discuss identification and method of estimation. Section 6 concludes.

2 Model

2.1 Environment

We consider a dynamic model of an individual worker's consumption and occupational choices over a lifetime of $T < \infty$ periods.

Preferences Let $c_t \in \mathbb{R}_+$ denote a worker's consumption in period t. The worker's preferences are defined over consumption vectors $(c_0, c_1, \dots, c_{T-1})$ and represented by a lifetime utility function that is additively separable across time with an exponential discounting factor $\beta \in (0, 1)$ and the per-period utility function $u(\cdot)$. Then, the worker's expected lifetime utility is

$$\mathbb{E}_0 \sum_{t=0}^{T-1} \beta^t u(c_t) \tag{1}$$

We assume that the worker's per-period utility function takes the CRRA (Constant Relative Risk Aversion) form with a coefficient of relative risk aversion $\rho > 0$, that is, $u(c_t) = c_t^{1-\rho}/(1-\rho)$ if $\rho \neq 1$, and $u(c_t) = \ln(c_t)$ if $\rho = 1$.

Skill and Occupation A worker's skills are classified into $K \in \mathbb{N}$ different types. The worker's skill portfolio at the beginning of period t is denoted by a column vector $\mathbf{s_t} = (s_{1t}, s_{2t}, \ldots, s_{Kt}) \in \mathbb{R}_+^K$. An occupation is identified with the relative skill requirement in the occupation and therefore denoted by a column vector $\mathbf{x_t} = (x_{1t}, \ldots, x_{kt}, \ldots, x_{Kt}) \in (0,1)^K$ in which x_{kt} is the proportion of skill j's component required for the occupation.

Skill Accumulation The worker begins each period t with the skill portfolio $\mathbf{s_t}$ and accumulates her skills in the current occupation $\mathbf{x_t}$ according to the following technology:

$$\mathbf{s_{t+1}} = \mathbf{s_t} + \mathbf{A_1}\mathbf{x_t} + \mathbf{A_2}(\mathbf{x_t} \circ \mathbf{x_t}) + \mathbf{A_3}\mathbf{Z} + \eta_{\mathbf{t+1}}, \tag{2}$$

where $\mathbf{x_t} \circ \mathbf{x_t} = (x_{1t}^2, \ldots, x_{Kt}^2)$ denotes the Hadamard product of $\mathbf{x_t}$ with itself, Z is L-dimensional column vector of the worker's characteristics (such as gender and race), and $\eta_{t+1} = (\eta_{1,t+1}, \ldots, \eta_{k,t+1}, \ldots, \eta_{K,t+1}) \sim N(\mathbf{0}, \mathbf{Q})$ is a K-dimensional column vector of skill portfolio shock that is independently and identically distributed over time according to the normal distribution with a K-dimensional zero mean vector and a $K \times K$ variance-covariance matrix $\mathbf{Q}$. Note that A_1 and A_2 are $K \times K$ diagonal matrices and A_3 is a $K \times L$ matrix.

Wage Equation The worker supplies one unit of labor inelastically in a competitive labor market. Then, the worker's wage w_t equals her productivity, which in turn depends on the skill portfolio $\mathbf{s_t}$, the current occupation $\mathbf{x_t}$, tenure t, and productivity shock ϵ_t in the following way:

$$w_t = B_0 + \mathbf{B_1}\mathbf{s_t} - \mathbf{B_2}(\mathbf{x_t} \circ \mathbf{x_t}) + \mathbf{B_3}(\mathbf{x_t} \circ \mathbf{s_t}) + B_4\, t + \mathbf{B_5}\mathbf{Z} + \epsilon_t \tag{3}$$

where $\epsilon_t \sim N(0, \sigma_\epsilon^2)$ is a productivity shock that is independently and identically distributed over time according to the normal distribution with mean zero and variance $\sigma_\epsilon^2 > 0$. Note that B_0 and B_4 are scalars, $\mathbf{B_1}$, $\mathbf{B_2}$, and $\mathbf{B_3}$ are $1 \times K$ vectors, $\mathbf{B_5}$ is a $1 \times L$ vector.

Our wage equation in the above is consistent with the production function in Altonji

(2005) and therefore features the following: (i) Skill is in general valuable. The more skilled the worker is (regardless of which type of skill and which occupation), the higher the wage (a higher s_{kt} leads to a higher wage for any type-k skill), (ii) The wage is particularly higher when the current occupation appreciates the skills that the worker has a comparative advantage (a higher $\mathbf{x_t} \circ \mathbf{s_t}$ implies a higher wage), (iii) the overhead costs rises with the occupation (a higher $\mathbf{x_t} \circ \mathbf{x_t}$ negatively affects an output and hence the wage).

Asset Accumulation Let m_t denote the total monetary holding available for a worker at the beginning of period t. The worker consumes c_t and invest the rest $a_t = m_t - c_t$ in a financial asset that yields the gross return $R > 1$. In addition to the financial income Ra_t, the worker also earns the labor income (the wage) w_t. The total income evolves as follows:

$$m_{t+1} = R(m_t - c_t) + w_t.^1 \tag{4}$$

2.2 Skill Uncertainty and Learning

Each worker has imperfect knowledge of her own skill portfolio. Assume that the worker's prior belief about her initial skill $\mathbf{s_0}$ (at the beginning of period 0) follows a normal distribution: $\mathbf{s_0} \sim \mathcal{N}(\mathbf{\hat{s}_0}, \mathbf{\Sigma_0})$ where $\mathbf{\Sigma_0}$ is a diagonal matrix. In each period t, the worker may learn about her own skill portfolio s_t by observing the wage in the current occupation. To describe this learning process, consider period t. The worker's belief about s_t (at the beginning of period t) follows a normal distribution with mean vector $\mathbf{\hat{s}_t}$ and variance-covariance matrix $\mathbf{\Sigma_t}$.[2] The worker observes wage w_t as a noisy signal about s_t and updates her beliefs. Particularly, we define the signaling component of the wage as

$$g_t = w_t - B_0 + \mathbf{B_2}(\mathbf{x_t} \circ \mathbf{x_t}) - B_4 t - \mathbf{B_5 Z} = (\mathbf{B_1} + \mathbf{B_3} \circ \mathbf{x_t^\intercal})\mathbf{s_t} + \epsilon_t = \mathbf{H_t s_t} + \epsilon_t \tag{5}$$

where $\mathbf{H_t} := \mathbf{B_1} + \mathbf{B_3} \circ \mathbf{x_t^\intercal}$ and $\mathbf{x_t^\intercal}$ is a transpose of $\mathbf{x_t}$. By Bayes' rule, the posterior distribution of s_t is normal with mean vector and variance-covariance matrix

$$\begin{aligned} &\mathrm{E}(\mathbf{s_t}|\mathbf{\Omega_{t+1}}) = (\mathbf{I} - \mathbf{K_t H_t})\mathbf{\hat{s}_t} + \mathbf{K_t} g_t, \quad \mathbf{K_t} = \mathbf{\Sigma_t H_t^\intercal}(\mathbf{H_t \Sigma_t H_t^\intercal} + \sigma_\epsilon^2)^{-1}, \\ &\mathrm{Var}(\mathbf{s_t}|\mathbf{\Omega_{t+1}}) = (\mathbf{I} - \mathbf{K_t H_t})\mathbf{\Sigma_t} \end{aligned} \tag{6}$$

where $\mathbf{K_t}$ is a $K \times 1$ vector of the precision weights on the signal z_t, and $\mathbf{\Omega_{t+1}}$ denotes the information at the beginning of period $t+1$ that contains the wage profile $(w_1, \ldots, w_t)$ and

[1] We assume that the wage w_t is paid after the worker's consumption decision. Therefore, the wage does not affect the asset holding $a_t = m_t - c_t$ in period t although it affects a_{t+1} through m_{t+1} in period $t+1$.

[2] Note that $\mathbf{\Sigma_t}$ for $t \neq 0$ need not be a diagonal matrix.

$(\mathbf{\hat{s}_0}, \mathbf{\Sigma_0})$.

To understand how an individual worker learns about her skill portfolio, note that she chooses her occupation $\mathbf{x_t}$ while anticipating the wage w_t based on the information set $\mathbf{\Omega_t}$. In other words, the occupational choice $\mathbf{x_t}$ depends on $\hat{w}_t := \mathrm{E}[w_t|\mathbf{\Omega_t}]$, or equivalently,

$$\hat{w}_t = B_0 - \mathbf{B_2}(\mathbf{x_t} \circ \mathbf{x_t}) + B_4\, t + \mathbf{B_5 Z} + \mathbf{H_t \hat{s}_t}.$$

After being paid w_t, the worker updates her belief regarding the skill-based component $\mathbf{H_t \hat{s}_t}$ in the wage payment.

Facing uncertainty about the skill component $\mathbf{H_t s_t}$ of wage, a worker has two imperfect measures, namely $\mathbf{H_t \hat{s}_t}$ and $g_t = \mathbf{H_t s_t} + \epsilon_t$. The updated belief $\mathbf{H_t}\mathrm{E}(\mathbf{s_t}|\mathbf{\Omega_{t+1}})$, which is the posterior belief about $\mathbf{H_t s_t}$, is the average of these two imperfect measures weighed by their relative precisions:

$$\mathbf{H_t}\mathrm{E}(\mathbf{s_t}|\mathbf{\Omega_{t+1}}) = \lambda g_t + (1-\lambda)\mathbf{H_t \hat{s}_t} \tag{7}$$

where $\lambda := \mathbf{H_t K_t} \in (0,1) \subset \mathbb{R}$ is the precision weight placed on the signal g_t, that is,

$$\mathbf{H_t K_t} = \frac{\mathbf{H_t \Sigma_t H_t^\intercal}}{\mathbf{H_t \Sigma_t H_t^\intercal} + \sigma_\epsilon^2} = \frac{\frac{1}{\sigma_\epsilon^2}}{\frac{1}{\mathbf{H_t \Sigma_t H_t^\intercal}} + \frac{1}{\sigma_\epsilon^2}}$$

The posterior distribution leads to the next period's prior distribution $s_{t+1} \sim \mathcal{N}(\mathbf{\hat{s}_{t+1}}, \mathbf{\Sigma_{t+1}})$ through the skill evolution, Equation (**??**), as follows:

$$\begin{aligned} \mathbf{\hat{s}_{t+1}} &:= \mathrm{E}(\mathbf{s_{t+1}}|\mathbf{\Omega_{t+1}}) = \mathrm{E}(\mathbf{s_t}|\mathbf{\Omega_{t+1}}) + \mathbf{A_1 x_t} + \mathbf{A_2}(\mathbf{x_t} \circ \mathbf{x_t}) + \mathbf{A_3 Z}, \\ \mathbf{\Sigma_{t+1}} &:= \mathrm{Var}(\mathbf{s_{t+1}}|\mathbf{\Omega_{t+1}}) = \mathrm{Var}(\mathbf{s_t}|\mathbf{\Omega_{t+1}}) + \mathbf{Q} \end{aligned} \tag{8}$$

2.3 Worker's Problem

The worker begins each period $t = 0, 1, \ldots, T$ with the monetary holding m_t and the prior belief about her current skill portfolio $\mathbf{s_t} \sim \mathcal{N}(\mathbf{\hat{s}_t}, \mathbf{\Sigma_t})$. Let $V(m_t, \mathbf{\hat{s}_t}, \mathbf{\Sigma_t})$ be the worker's value from t and onward. Then, the worker's problem can be stated recursively as the following Bellman equation:

$$\begin{aligned} V_t(m_t, \mathbf{\hat{s}_t}, \mathbf{\Sigma_t}; \mathbf{Z}) = &\max_{c_t, \mathbf{x_t}} \; u(c_t) + \beta \mathrm{E}\left[V_{t+1}(m_{t+1}, \mathbf{\hat{s}_{t+1}}, \mathbf{\Sigma_{t+1}})|\mathbf{\Omega_t}\right] \\ &\text{s.t. } (\mathbf{??}), (\mathbf{??}), (\mathbf{??}), (\mathbf{??}), (\mathbf{??}). \end{aligned} \tag{P}$$

The relevant necessary conditions for the optimality in Problem (**??**) are the following.

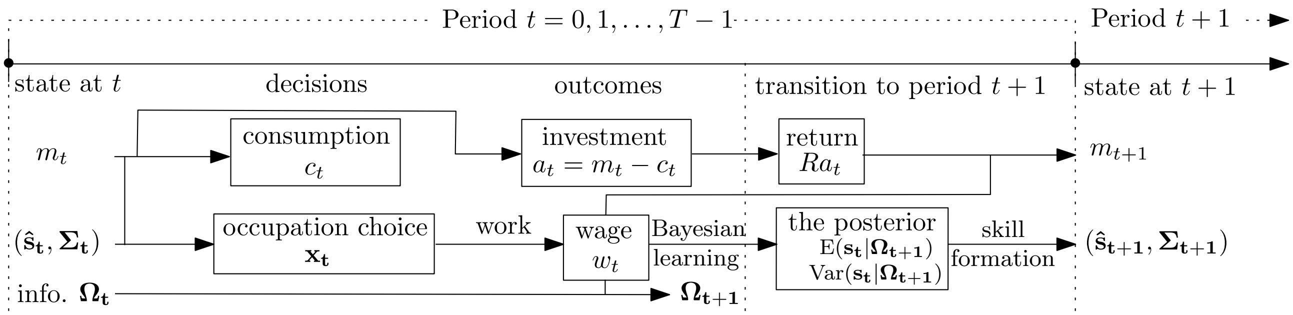


Figure 1: Worker's Problem in Period t

For $k = 1, 2, \ldots, K$, the envelop condition with respect to $\hat{s}_{kt}$ is

$$\begin{aligned}\frac{\partial V_t}{\partial \hat{s}_{kt}} &= \beta \mathrm{E_t}\left[\frac{\partial V_{t+1}}{\partial m_{t+1}}\frac{\partial m_{t+1}}{\partial w_t}\frac{\partial w_t}{\partial \hat{s}_{kt}} + \sum_{l=1}^{K}\frac{\partial V_{t+1}}{\partial \hat{s}_{l,t+1}}\frac{\partial \hat{s}_{l,t+1}}{\partial \hat{s}_{kt}} + \sum_{l,n}\frac{\partial V_{t+1}}{\partial \sigma_{ln}(t+1)}\frac{\partial \sigma_{ln}(t+1)}{\partial \hat{s}_{kt}}\right] \\ &= \beta \mathrm{E_t}\left[\sum_{l=1}^{K}\frac{\partial V_{t+1}}{\partial \hat{s}_{lt+1}}(\mathbf{I} - \mathbf{K_t H_t})(l,k)\right],\end{aligned} \tag{9}$$

where $(\mathbf{I} - \mathbf{K_t H_t})(l,k)$ is the (l,k)-th component of matrix $(\mathbf{I} - \mathbf{K_t H_t})$, that is,

$$(\mathbf{I} - \mathbf{K_t H_t})(j,k) = \begin{cases} 1 - \frac{\left(\sum_{k'}\sigma_{kk'}(t)H_{k't}\right)H_{kt}}{\mathbf{H_t \Sigma_t H_t^\intercal} + \sigma_\epsilon^2} & \text{if } j = k \\ -\frac{\left(\sum_{k'}\sigma_{jk'}(t)H_{k't}\right)H_{kt}}{\mathbf{H_t \Sigma_t H_t^\intercal} + \sigma_\epsilon^2} & \text{if } j \neq k \end{cases}$$

For $j, k = 1, 2, \ldots, K$, the corresponding envelope condition with respect to $\sigma_{jk}(t)$ is

$$\begin{aligned}\frac{\partial V_t}{\partial \sigma_{jk}(t)} &= \beta \mathrm{E_t}\left[\frac{\partial V_{t+1}}{\partial m_{t+1}}\frac{\partial m_{t+1}}{\partial w_t}\frac{\partial w_t}{\partial \sigma_{jk}(t)} + \sum_{l=1}^{K}\frac{\partial V_{t+1}}{\partial \hat{s}_{lt+1}}\frac{\partial \hat{s}_{lt+1}}{\partial \sigma_{jk}(t)} + \sum_{l,n}\frac{\partial V_{t+1}}{\partial \sigma_{ln}(t+1)}\frac{\partial \sigma_{ln}(t+1)}{\partial \sigma_{jk}(t)}\right] \\ &= \beta \mathrm{E_t}\left[\sum_{l,n}\frac{\partial V_{t+1}}{\partial \sigma_{ln}(t+1)}\left[(\mathbf{I} - \mathbf{K_t H_t})(l,j)\right]\left[(\mathbf{I} - \mathbf{K_t H_t})(n,k)\right]\right]\end{aligned} \tag{10}$$

Lastly, the envelope condition with respect to m_t is

$$\begin{aligned}\frac{\partial V_t}{\partial m_t} &= \beta \mathrm{E}\left[\frac{\partial V_{t+1}}{\partial m_{t+1}}\frac{\partial m_{t+1}}{\partial m_t} + \sum_{l=1}^{K}\frac{\partial V_{t+1}}{\partial \hat{s}_{lt+1}}\frac{\partial \hat{s}_{lt+1}}{\partial m_t} + \sum_{l,n}\frac{\partial V_{t+1}}{\partial \sigma_{ln}(t+1)}\frac{\partial \sigma_{ln}(t+1)}{\partial m_t}\right] \\ &= \beta \mathrm{E}\left[\frac{\partial V_{t+1}}{\partial m_{t+1}}R\right]\end{aligned} \tag{11}$$

The first-order condition with respect to c_t is simply expressed as

$$c_t : \quad u'(c_t) = \beta \mathrm{E_t}\left[\frac{\partial V_{t+1}}{\partial m_{t+1}}\frac{\partial m_{t+1}}{\partial c_t}\right] = \beta \mathrm{E_t}\left[\frac{\partial V_{t+1}}{\partial m_{t+1}} R\right]. \tag{12}$$

Lastly, the first-order condition with respect to x_{kt} for $k = 1, 2, \ldots, K$ is

$$x_{kt} : \quad \beta \mathrm{E_t}\left[\frac{\partial V_{t+1}}{\partial m_{t+1}}\frac{\partial m_{t+1}}{\partial w_t}\frac{\partial w_t}{\partial x_{kt}} + \sum_{l=1}^{K}\frac{\partial V_{t+1}}{\partial \hat{s}_{lt+1}}\frac{\partial \hat{s}_{lt+1}}{\partial x_{kt}} + \sum_{l,n}\frac{\partial V_{t+1}}{\partial \sigma_{ln}(t+1)}\frac{\partial \sigma_{ln}(t+1)}{\partial x_{kt}}\right] = 0 \tag{13}$$

where

$$\frac{\partial w_t}{\partial x_{kt}} = B_{3k}s_{kt} - 2B_{2k}x_{kt},$$

$$\begin{aligned}\frac{\partial \hat{s}_{l,t+1}}{\partial x_{kt}} &= \frac{\partial \mathbf{K_t}(l)}{\partial x_{kt}}(g_t - \mathbf{H_t}\mathbf{\hat{s}_t}) + \mathbf{K_t}(l)\frac{\partial (g_t - \mathbf{H_t}\mathbf{\hat{s}_t})}{\partial x_{kt}} + \mathbf{A_1}(l,k) + 2\mathbf{A_2}(l,k)x_{kt} \\ &= B_{3k}\left[\frac{\sigma_{lk}(t)}{\mathrm{Var_t}(w_t)} - 2\mathbf{K_t}(l)\mathbf{K_t}(k)\right](g_t - \mathbf{H_t}\mathbf{\hat{s}_t}) + \mathbf{K_t}(l)B_{3k}(s_{kt} - \hat{s}_{kt}) + \mathbf{A_1}(l,k) + 2\mathbf{A_2}(l,k)x_{kt}\end{aligned}$$

and

$$\begin{aligned}\frac{\partial \sigma_{ln}(t+1)}{\partial x_{kt}} &= \sum_{k'}\left[\frac{\partial \mathbf{K_t}(l)}{\partial x_{kt}}H_{k't} + \mathbf{K_t}(l)\frac{\partial H_{k't}}{\partial x_{kt}}\right]\sigma_{k'n}(t) \\ &= B_{3k}\mathbf{K_t}(n)\left[\sigma_{lk}(t) - 2\mathbf{K_t}(l)\mathbf{K_t}(k)\mathrm{Var_t}(w_t)\right] + \mathbf{K_t}(l)B_{3k}\sigma_{kn}(t)\end{aligned}$$

3 Data

In this section, we discuss the data sets used in the project and conduct a preliminary data analysis of occupation choices and skills uncertainty. Section **??** introduces the Occupational Information Network (O*NET) database and the construction of task intensity measures following **?**. Section **??** introduces the National Longitudinal Survey of Youth (NLSY). Section **??** establishes that wage residuals are informative, i.e., that they are predictive of occupation choices in the subsequent period. We provide preliminary evidence that consistent with our model in section **??**, task intensities and skill signals proxied by wage residuals both affects future task choice.

3.1 O*NET data

The Occupational Information Network (O*NET) is a database containing detailed occupation-level information, including skill and knowledge requirements for 974 different occupations. The ratings come from two sources: a survey of workers who are asked to rate their occupation and a survey of occupation analysts who are asked to rate other descriptors in the O*NET dataset. The measures of task intensity we use are coded from the analysts' ratings.

The dataset and its predecessor, the Dictionary of Occupational Titles (DOT), are used by many authors studying job characteristics to create continuous, multidimensional task intensity measures. A series of papers including **?** and **?** proposes abstract, manual, and routine task measures constructed using O*NET. **?** proposes verbal, math, and social components to study the cost of occupational mismatch. **?** and **?** uses two-dimensional measures, cognitive and manual for the first and cognitive and social for the latter.

We follow task measures proposed by **?** as these measures can be related to the ASVAB component test scores to study occupational sorting by task intensities and skill levels. While **?** task measures are more widely adopted in the literature, measures of skills matching those set of tasks are harder to create using the NLSY dataset. Most of the other proposed measures are percentile scores of task complexity or ranks of the jobs and are thus harder to interpret within our model. **?** task measures that provide levels, not ranks, of task requirements at each occupation give a better match for our model where skills are also measured in levels.

The task measures **?** provide cognitive, manual, and interpersonal task intensity indices in unit intervals. To create these indices, the authors first run Principal Component Analysis (PCA) on 277 O*NET database variables relating to skills, abilities, knowledge, work activities, and work contexts. The first three principal components are recombined with the following exclusion restrictions: mathematics, mechanical knowledge, and social perceptiveness score only reflects cognitive, manual, and interpersonal task intensities, respectively. These indices are then rescaled using linear transformation so that the final task intensity indices lie in unit intervals.

3.2 The NLSY Data

The NLSY is an ongoing, nationally representative survey of the individuals born between 1957 and 1964 in the United States. The NLSY data contain rich information on respondents' work history as well as wage and asset evolution. The data set also contains measures of cognitive and noncognitive skills, including the subsection scores of the Armed Services Vocational Aptitude Battery (ASVAB) scores, Rotter locus-of-control scale and the Rosenberg self-esteem tests, and measures of health and antisocial behavior.

We reduce the dimension of the skill measures to the initial cognitive, manual, and interpersonal skill levels following **?**. The skill indices are created using Principal Component Analysis (PCA) on the test scores and using the exlusion restriction that the ASVAB math score, automotive and shop information score, and Rosenberg self-esteem score only reflects cognitive, manual, and interpersonal skills. The indices are then scaled so that each skill indices fall in the unit interval where 0 denote the lowest and 1 denote the highest level of initial skills.

The main analytic sample is restricted to 2161 male respondents who had or have strong ties to the labor market for whom the three skill indices can be created. We drop all work history until the first occurence of a non-employment spell of 18 months or more. Because we consider annual work history of these workers, the analytic sample consists of 27312 worker-year observations. We drop observations for which the workers have worked for more than 15 years. The occupation characteristics and wage patterns in the data from year 16 onward show cyclical behavior that is different from continuous increase in wage prior to year 15.

Table **??** summarizes the workers who compose the analytic sample. Column (1) does so for the full sample and columns (2)-(3) do so separately by workers' educational attainment. About 24% of workers in the sample completed a four-year degree. College graduates on average have higher initial cognitive, manual, and interpersonal skills compared to those who do not have four-year college degrees. Stark differences interpersonal skills are consistent with **?** that externalizing behavior, which is linked negatively with interpersonal skills, reduces schooling. Difference in initial cognitive skills by education is also consistent with studies that find strong role of cognitive abilities on schooling (**?**). While Blacks and Hispanics are slightly over-represented compared to the national average, the educational attainment observed is consistent with that observed in other studies.

Consistent with patterns in initial skills, college graduates work in occupations that require high cognitive and interpersonal skills. Interestingly, average manual task intensities is higher for some college or less group even though the average initial manual skill levels are lower. This could be due to workers selecting into occupations that they have comparative advantage over: initial manual skill, on average, is the highest for the less-educated and is the lowest for the more-educated. Also, this pattern suggests that task choices are influenced by the levels of all three skills, not only by the levels of the skill that is directly related to the task. That is, choice of cognitive task intensity is affected by manual and interpersonal skill levels on top of cognitive skill levels, and the statement holds for all other tasks.

Figure **??** show change in the wages and occupational task intensities observed over time. Panel (a) reports average wages by labor market tenure along with the standard deviation of the mean wages. Panel (b), (c), and (d) report the means and standard deviations for cognitive, manual, and interpersonal tasks, respectively. In all four panels, we observe an upward trend in the means. For wages, we observe an increase in wage variations over time, which is consistent with the findings in the literature. For the task intensities, the spread of the distributions stay relatively constant over time. Exact values for the means and standard deviations depicted in Figure **??** are shown in Table **??**.

For cognitive choices, we observe a relatively fast increase in early years of life from 0.34 in year 1 to to 0.39 in year 7 followed periods with minimal changes in task intensities. For manual tasks, the changes are more linear with the exception of slightly lower mean value in the first year. Interpersonal task choices follow similar trend with that of manual task.

3.3 Key Patterns in the Data

3.3.1 Job Transitions

A key pattern in the data that is that while most workers transit to similar jobs across periods, they frequently move to different occupations. The transition matrices reported in Table **??** document the frequency of such moves. Between 60 to 90 percent of workers stay in the same task intensity quintile across two years, and the people in 2nd or 3rd quintiles tend to move the most often. Bigger jumps in occupation tasks are less likely compared to smaller jumps, but the probability of those jumps do not seem to depend on the direction of the move. That is, The probability of moving from a 1st quintile task choice to a 5th quantile task is as likely to happen as a move from a 5th quantile to 1st.

3.3.2 Wage Residuals Are Predictive

Understanding evidence of skill learning, both learning about uncertain skills and learning-by-doing, is a precursor to credibly identifying the effect of learning. The difficulty, however, is that search models where workers endogenously switch jobs until they find a good match are virtually indistinguishable empirically from a learning model. **?** provides a comprehensive review of the literature. Because of this, structural models often take a stance on which of the two stories to capture.

We show that wage residuals, or portion of wages that are presumably unanticipated by workers, affect career transition following empirical strategy adopted in **?**. Our model in Section **??** predicts that if learning occurs, a positive shock in wage will lead the worker to form higher expectation the following period. If learning story holds, wage residuals would be predictive in future task choices to the degree that expected skill levels affect task choices. Further, we allow the predictiveness of wage shocks on skill levels may differ by task intentisies as in our model.

We begin by estimating log wage equations of the form

$$log\ w_{it} = \mathbf{s_{i0}}\beta_1 + \mathbf{x_{it}}\beta_2 + \mathbf{s_{i0}} \times \mathbf{x_{it}} + Z_i\gamma + u_{it}. \tag{14}$$

The regression model mimics closely the wage equation (**??**) with three key departures: one, linear task intensities are added to address potential omitted variable problem coming from the use of higher order terms of these variables. Two, measures of initial skills are used in place of true skill levels in each time period. Third, demgraphic variables, Z_i are used as additional controls.

Table **??** reports OLS estimates in column (1) and worker fixed effects estimates in column (2). Task intensity measures are predictive of wages. Interestingly, cognitive task intensities interacted with manual and interpersonal skills both show highly negative coefficients, indicating that wages are higher on jobs specializing in cognitive skills than jobs that also require high intensity of other skills. This is consistent with the pattern observed in Table **??** that workers with less than college education sort into jobs with higher manual task intensities than college graduates do even though their skill manual skill levels are lower.

Contrary to the model setup in **?**, squared term in task intensities show positive coefficients. The sign of these coefficnts suggest that productivity increases with increase in task intensities, or instead of overhead costs rising with the occupation. This could also reflect measurement error in skill levels coming from the fact that we are proxying true skill levels in each period with measures of initial skills. The negative interaction terms and positive squared terms together highlights reward from choosing specialized occupations rather than taking well-rounded jobs.

We show that wage residuals obtained from above regressions affect next period's occupational choice by estimating equations of the form:

$$x_{j,t+1} = \mathbf{x_t}\beta_1, + u_{k,t}\beta_2 + u_{k,t} \times x_{j,t}\beta_3 + \nu_{j,t+1},\ j \in \{C, M, I\},\ k \in \{OLS, FE\}. \tag{15}$$

As described earlier, positive marginal effect of u_k on next period's task choice is consistent with our learning model in that unexpected wage shock incentivize workers to try occupations with higher skill requirements. Further, the coefficient β_2 test whether the effect of these wage shocks differ by task intensities.

Table **??** reports these estimates that are consistent with our theoretical analysis in Section **??**. Columns (1)-(3) shows the OLS estimates and columns (4)-(6) show the fixed effects estimates. Coefficients β_3 s are highly significant and positive. Treating wage residuals as rough measures of skill signals, this indicates that precision of skill signals are is inversely related to task intensities. While coefficients on u are highly significant and negative across all specifications, marginal effect of u, $\beta_2 + x_{j,t}\beta_3$, are positive for most values of x.

3.4 Wealth Effects Are Small

Prevalence of wealth inequality is well documented. **?** develops and estimates a structural model of occupational choices where risk aversion leads to suboptimal choices by workers with lower wealth. As we have shown in Section **??**, under uncertainty and risk aversion, two opposite forces determine the net effect of wealth on task choices. OLS estimates show that effect of wealth on task choices are small: a one percent increase in wealth increases task choices indices by about 0.001, where the indices range from 0 to 1. These results are presented in Table **??**. For the reasons we outlined in Section **??**, this does not mean that wealth effect does not exist, but should be interpreted as the two opposing effects are about equal in size.

4 Empirical Strategy

This section describes the estimation strategy used in the paper. I start by describing how I approximate workers' occupation choice decision rules with policy functions. I also discuss the benefits and costs of this approach. I then discuss how I recover the distribution of unknown skils. I show under a set of plausible assumptions the skills distributions are nonparametrically identified. I then describe the simulated maximum likelihood method I use to estimate the empirical model.

4.1 Estimation Strategy

The empirical strategy involves approximating occupational decisions with policy functions and estimating these jointly with evolution of skills, financial wealth, and wages. This method can be seen as an extension to **?** in which a polynomial of state variables is used to approximate value function for each choices in each periods in a discrete dynamic choice setting. The alternative approach would be to fully estimate the dynamic model described in Section **??** by backward induction. However, that would require computing and approximating value functions and first derivatives of the value function in all periods.

As explained in Section **??**, the workers' occupational cohice in time t will depend on the whole state space he faces at the start of period t (Ω_t) and on his beliefs on the evolution of the state space. Note that this means we can write policy functions for the worker's choices as:

$$x_{j,t} = f_t^j(\mathbf{m_t}, \hat{\mathbf{s}}_\mathbf{t}, \mathbf{\Sigma_t}; Z) + \nu_t^j, \ j \in \{C, M, I\} \tag{16}$$

where ν_t^js capture shocks to the worker's decision. Standard method in estimating the function f_t^j is first choosing fixed set of points in the state space then solving by backward induction the policy functions. Some interpolation methods are used to infer function values at the points that were not chosen initially.

Instead of using backward induction we make parametric assumptions on the functions f_t^j and we estimate these reduced-form parameters of function f_t^j. This method has two clear advantages. First, it allows solving dynamic structural models without using dynamic programming. This brings the computing cost of solving the models to that of a static model. Second, one could specify f_t^j in a general functional form to minimize interpolation and appoximation error coming from computing multiple derivatives of value functions and taking expectations.

The main disadvantage of this empirical strategy is that it does not allow me to recover the risk aversion parameter without further assumptions. We will discuss assumptions needed to recover the utility parameter.

4.1.1 Linear Policy Functions

Ideally I would like to estimate the policy functions nonparametrically. However, given the large state space and highly interdepndent structure of the state transition rules, number of parameters to estimate increase exponentially as we introduce higher order terms to the policy functions. For computation and identification reasons, we assume that the policy functions are linear-in-parameters. We do allow the parameters of the function f_t^j over time following a quadratic function of time.

Even if we assume simplistic form of policy functions, number of parameters exceed 50 per each period, leaving the total number of parameters to estimate in excess of 700. Given the number of observations per period are around 2,000 estimating such a large number of parameters could also cause identification issues. Our setup thus calls for methods to reduce number of parameters for computational and identification reasons.

We make two assumptions to make estimation feasible and also make model tractable. First, we assume that the policy functions are linear-in-parameters. Second, we assume that the parameters across time periods follow a quadratic function in time. This reduces the number of parameters to around 120. While higher-older polynomial functions are desirable to capture the full and rich interactions among each variables in our state space, a function with third-order polynomials with interaction terms would require estimation of over 300 parameters per policy function.

As a result, the policy functions for worker's occupation choice can be described by:

$$\mathbf{x_t} = \pi_{0,t} + \Pi_{1,t}\mathbf{\hat{s}_t} + \Pi_{2,t}\tilde{\mathbf{\Sigma}}_\mathbf{t} + \Pi_{3,t}m_t + \nu_t, \tag{17}$$

where $\tilde{\mathbf{\Sigma}}_\mathbf{t}$ is a column vector of all the unique elements in $\mathbf{\Sigma_t}$.

4.1.2 Kalman Filter

A precursor to estimating the policy functions as above is correctly estimating the distribution of unobserved state variables–means, variances, and covariances of skills. We use the Kalman filter to estimate these variables and calculate the likelihood for estimation.

The Kalman filter is an algorithm used to estimate the distribution of unobserved state variables from observed noisy signals given parameter values. In its original form, the Kalman filter only allows for signals with conditionally normal distributions. There are extensions that allow for non-normal error terms using approximation techniques through first order Taylor expansion, but at an added cost. For this reason and also from the observation that occupational movement appears to be a symmetric, unimodal distribution with decreasing densities as one move away from the center, we deem normality assumption to be appropriate.

This means that nobserved state variables, the distribution of skills, can be recovered from observed occupation choices and wages using the Kalman filter. Adopting common notation of the Kalman filter, mean and variance-covariance matrix of the posterior and the next period's prior distribution of skills in equations (**??**) and (**??**) can be relabeled as

$$\begin{aligned} \mathbf{\hat{s}_{t|t-1}} &:= \mathrm{E}(\mathbf{s_t}|\mathbf{\Omega_t}) \\ \mathbf{\Sigma_{t|t-1}} &:= \mathrm{Var}(\mathbf{s_t}|\mathbf{\Omega_t}) \\ \mathbf{\hat{s}_{t|t}} &:= \mathbf{\hat{s}_{t+1}} = \mathrm{E}(\mathbf{s_{t+1}}|\mathbf{\Omega_t}) \\ \mathbf{\Sigma_{t|t}} &:= \mathrm{Var}(\mathbf{s_{t+1}}|\mathbf{\Omega_t}). \end{aligned} \tag{18}$$

Given the policy rule in equation (**??**) the conditional mean and variance of $\mathbf{x_t}$ are

$$\begin{aligned} \mathrm{E}(\mathbf{x_t}|\mathbf{\Omega_t}) &= \pi_{0,t} + \Pi_{1,t}\mathbf{\hat{s}_t} + \Pi_{2,t}\tilde{\mathbf{\Sigma}}_\mathbf{t} + \Pi_{3,t}m_t \\ \mathrm{Var}(\mathbf{x_t}|\mathbf{\Omega_t}) &= \Pi_{1,t}\mathbf{\Sigma_{t|t-1}}\Pi'_{1,t} + \Sigma_\nu \end{aligned} \tag{19}$$

From this we can update the conditional distribution of skills using $\mathbf{x_t}$ so that

While the Kalman filter assumes knowledge of the parameter values, by normality assumption the it provides a natural way to calculate the likelihood function. Given observation of wages, task intensities, and wealth, we can estimate the parameters using maximum likelihood estimation.[3]

4.1.3 Wealth

We use lognormal approximation of consumption Euler equation widely used in consumption literature together with income evolution policy (equation **??**) and linear-in-parameter policy function approximation to track wealth evolution. Note that the usual consumption Euler equation

$$u'(c_t|\Omega_t) = \mathrm{E}(u'(c_{t+1})|\Omega_t) \tag{20}$$

holds. Rearranging, we have

$$\mathrm{E}(\frac{u'(c_{t+1})|\Omega_t)}{u'(c_t|\Omega_t)}) = E((\frac{c_{t+1}}{c_t})^{-\rho}|\Omega_t) = 1 \implies log(c_t) = log(\mathrm{E}(c_{t+1})|\Omega_t)). \tag{21}$$

From this we have consumption transition rule:

$$log(c_{t+1}) = log(c_t) + \epsilon_t^C,\ \epsilon_t^C \sim N(0, \sigma_C^2). \tag{22}$$

We estimate consumption rule $c_t(m_t, \mathbf{\hat{s}_t}, \mathbf{\Sigma_t})$ using linear policy function approximation together with equations **??** and **??**.

4.2 Simulated Maximum Likelihood

The econometric model is described in Section **??**. We collect the parameters to be estimated into a vector denoted Ξ:

$$\Xi = \{\mathbf{A_1}, \mathbf{A_2}, \mathbf{A_3}, \mathbf{Q}, B_0, \mathbf{B_1}, \mathbf{B_2}, \mathbf{B_3}, B_4, \mathbf{B_5}, \sigma_\epsilon^2, \pi_0, \Pi_1, \Pi_2, \Pi_3, \sigma_C^2\}. \tag{23}$$

We estimate Ξ using simulated maximum likelihood (**?**). In the inner loop of the estimation algorithm, we compute the likelihood for a particular set of candidate parameters, which are indexed by (g) and denoted $\Xi^{(g)}$. To calculate the log likelihood for a given set of candidate

[3] The derivation of the likelihood is similar to **?** because in both models wages and task choices serve as measures for unobserved skills and assume normality of error terms. One key difference is that Yamaguchi constructs roy model of occupation choice with risk-neutral agents and thus consumption behavior need not be modeled. This yielded Yamaguchi to obtain linear policy functions and to identify all structural parameters of the model. State spaces of the two models also differ: as we focus on the importance of learning, variance and covariance terms are part of state space.

parameters $\Xi^{(g)}$, we first fill in missing wage and asset observations for each individual i for each time period t by random draws from the distribution given by $\Xi_{(g)}$. We denote this dataset as θ^{g}[4]. For each θ^{g}, we use distributional assumptions to calculate the likelihood contribution for each person. We first calculate the likelihood contribution from occupation choice, $l(\mathbf{x_t}|\Omega_t)$. Next, we calculate the likelihood contribution from wage, $l(log\ w_t|\Omega_t, x_t)$. Using consumption policy function we then calculate $l(m_{t+1}|\Omega_t, x_t)$ for $t < T_i$ where T_i is the last observed period for individual i. Using these components, we calculate the value of the log likelihood for each individual i in $\theta^{(g)}$ as:

$$l_i^g = l_i^g(\Xi^{(g)}|\{m_t, \mathbf{\hat{s}_t}\}_{t=1}^{T_i}), \mathbf{\Sigma_t}) = log(\prod_{t=1}^{T_i} l(\mathbf{x_t}|\Omega_t)(log\ w_t|\Omega_t, x_t)l(m_{t+1}|\Omega_t, x_t)). \tag{24}$$

Finally, we sum over all N individuals to obtain the log-likelihood:

$$l^{(g)} = \sum_{i=1}^{N} l_i^{(g)}. \tag{25}$$

In the outer loop, we repeat the inner loop for different sets of candidate parameters until the log likelihood function is maximized.

4.3 Identification

There are three levels of identification needed. First, the hidden states need to be identified. Next, the policy functions–task and consumption choices–need be identified. Lastly, the risk-aversion parameter needs to be identified. We summarize the main sources of information and assumptions of the model that identify the components of our model.

The distribution of unobserved skills, $\mathrm{E}(\mathbf{s_t}|\mathbf{\Omega_t}), \mathrm{Var}(\mathbf{s_t}|\mathbf{\Omega_t})$ are identified by conditional independence and linearity assumption upto normalization. Observed wages in periods $t-1, t$, and $t+1$, as well as and occupation choices each period act as instruments as they are conditionally independent given unobserved skills and occupations. Thus the minimum data requirement for non-parametric identification is satisfied (**?**, **?**). Normalization is done by using initial skill level constructed in Section **??**. While normal error terms aid in the accuracy of Kalman filter and the speed of the estimation, normality is not required for

[4] We draw a black matrix of size $N \times T \times 2$ from a standard normal distribution once, where N denote the number of individuals in the sample, T is the number of years we follow, then one column to draw shocks for missing wages and the second column to draw shocks for missing asset observations. This helps to avoid the so-called "chattering" effect, which can lead to different values of the likelihood function given the same parameters due to differences in random draws at each parameter set.

identification.

Identification of the wage parameter follows from conditional indepenence of wage shocks once the distribution of $\mathbf{s_t}$ are recovered. Given that all state space vectors are observed and wage shock is exogenous, wage parameters can be estimated, for example, by least squares method. Thus, the variation in level of wages conditional on the condition set $\mathbf{Z}$ identify the wage parameters.

Identification of the reduced-form parameters of the task policy functions are then identified from comparisons of the set of job types that are chosen by workers with different skills and wealth distribution. Identification of consumption functions follow from comparison of asset evolution path across different workers using equations **??** and **??**.

5 Conclusion

We provide a framework to study the effect of wealth and the role of learning under uncertainty and risk aversion on occupational choices. Our model allows wealth to play two roles–make workers seek high-skill-requirement jobs to reduce uncertainty going forward and make them seek low-skill-requirement jobs to reduce immediate uncertainty regarding wages. Our analysis suggests that workers are uncertain of their skills throughout their career and workers continuously update their beliefs using wages as signals of their skills. Moreover, we provide a tractable method to estimate the theoretical model that requires the evaluation of multiple first-order conditions with unknown states.

Table 1: Summary Statistics

	(1) All Sample	(2) Some College or Less	(3) 4-Yr Degree or More
Initial Cognitive skill	0.526	0.454	0.750
	(0.208)	(0.174)	(0.128)
Initial Manual skill	0.538	0.518	0.599
	(0.163)	(0.167)	(0.134)
Initial Interpersonal skill	0.500	0.416	0.764
	(0.205)	(0.149)	(0.107)
Cognitive skill requirement (y_C)	0.327	0.278	0.457
	(0.180)	(0.147)	(0.194)
Manual skill requriement (y_M)	0.489	0.513	0.426
	(0.191)	(0.190)	(0.180)
Interpersonal skill requirement (y_I)	0.327	0.273	0.472
	(0.170)	(0.131)	(0.179)
Demographics			
Hispanics	0.191	0.210	0.129
	(0.393)	(0.408)	(0.336)
Blacks	0.267	0.297	0.175
	(0.442)	(0.457)	(0.380)
Whites	0.542	0.493	0.696
	(0.498)	(0.500)	(0.460)
Highest Grade Completed (Years)	13.435	12.262	17.082
	(2.497)	(1.391)	(1.425)
Less than HS	0.094	0.125	0.000
	(0.292)	(0.331)	(0.000)
HS Diploma	0.436	0.576	0.000
	(0.496)	(0.494)	(0.000)
Some College	0.226	0.299	0.000
	(0.419)	(0.458)	(0.000)
4-Year College or More	0.243	0.000	1.000
	(0.429)	(0.000)	(0.000)
Observations	2161	1635	526

Notes: This table presents means of variables where workers are the unit of analysis. Standard deviations are reported in parentheses. HS denotes high school. Initial skills and skill requirements are on a 0-1 scale.

Table 2: Transition Matrices of Task Choices

	Task Intensity Quintile				
	1st	2nd	3rd	4th	5th
			y_C		
1st Quintile	74.33	14.39	6.98	3.06	1.24
2nd Quintile	15.40	66.17	12.30	4.48	1.65
3rd Quintile	5.21	14.27	67.89	10.06	2.57
4th Quintile	1.98	3.42	11.15	76.33	7.11
5th Quintile	0.77	1.09	1.48	6.70	89.95
			y_M		
1st Quintile	81.93	9.63	4.45	2.44	1.55
2nd Quintile	12.03	73.34	8.45	4.25	1.92
3rd Quintile	3.60	11.39	69.02	10.54	5.46
4th Quintile	2.07	4.07	13.26	69.04	11.55
5th Quintile	1.19	2.25	4.59	13.43	78.54
			y_I		
1st Quintile	74.42	15.10	7.25	2.24	0.99
2nd Quintile	14.64	68.46	11.97	3.61	1.31
3rd Quintile	5.93	11.50	70.45	9.12	3.01
4th Quintile	1.85	2.46	8.15	78.97	8.57
5th Quintile	0.79	0.93	1.81	7.83	88.65

Notes: Each entry reports the percentage of observations that fall in the particular period t-period $t + 1$ task intensity quintile.

Table 3: Effect of Wage Signals on Quality of Next Job

	OLS			FE		
	(1)	(2)	(3)	(4)	(5)	(6)
	$y_{C,t+1}$	$y_{M,t+1}$	$y_{I,t+1}$	$y_{C,t+1}$	$y_{M,t+1}$	$y_{I,t+1}$
$y_{C,t}$	0.925***	-0.013**	0.038***	0.560***	-0.016	-0.009
	(0.005)	(0.006)	(0.005)	(0.009)	(0.011)	(0.009)
$y_{M,t}$	-0.067***	0.834***	-0.045***	-0.029***	0.537***	-0.012*
	(0.004)	(0.006)	(0.004)	(0.006)	(0.008)	(0.006)
$y_{I,t}$	-0.044***	-0.056***	0.867***	-0.019**	-0.025**	0.579***
	(0.005)	(0.006)	(0.006)	(0.008)	(0.010)	(0.008)
$tenure$	-0.001**	0.001	-0.001	0.004***	0.003***	0.003***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$tenure^2$	0.000	-0.000	0.000	-0.000***	-0.000***	-0.000*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$u_{OLS,t}$	-0.014***	-0.046***	-0.013***			
	(0.004)	(0.005)	(0.003)			
$y_C \times u_{OLS}$	0.064***					
	(0.011)					
$y_M \times u_{OLS}$		0.097***				
		(0.010)				
$y_I \times u_{OLS}$			0.060***			
			(0.008)			
$u_{FE,t}$				-0.012***	-0.041***	-0.016***
				(0.003)	(0.005)	(0.003)
$y_C \times u_{FE}$				0.042***		
				(0.008)		
$y_M \times u_{FE}$					0.079***	
					(0.008)	
$y_I \times u_{FE}$						0.059***
						(0.007)
Constant	0.091***	0.113***	0.067***	0.174***	0.250***	0.149***
	(0.003)	(0.004)	(0.003)	(0.004)	(0.005)	(0.004)
N	27312	27312	27312	27312	27312	27312
Adjusted R^2	0.78	0.73	0.81	0.27	0.23	0.30

Notes: *$p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. $y_{C,t}$, $y_{M,t}$, and $y_{I,t}$ denote the cognitive, manual, and interpersonal tasks at period t respectively. Tenure indicates years of work experience. u_{OLS} and u_{FE} are residuals from the OLS and fixed effects regressions. The dependent variable in columns (1) and (4) is cognitive task intensity in the subsequent period on a 0-1 scale. The dependent variable in columns (2) and (5) is manual task intensity in the following period, and the depdendent variable in columns (3) and (6) is interpersonal task intensity in the following period. Parentheses contain standard errors that are robust to clustering at the individual level.

Table 4: Effect of Wealth on Quality of Next Job

	OLS			FE		
	(1)	(2)	(3)	(4)	(5)	(6)
	$y_{C,t+1}$	$y_{M,t+1}$	$y_{I,t+1}$	$y_{C,t+1}$	$y_{M,t+1}$	$y_{I,t+1}$
$y_{C,t}$	0.898***	0.012	-0.027***	0.492***	-0.009	-0.022*
	(0.007)	(0.007)	(0.007)	(0.014)	(0.015)	(0.013)
$y_{M,t}$	-0.036***	0.859***	-0.014***	-0.023**	0.469***	-0.001
	(0.005)	(0.007)	(0.005)	(0.010)	(0.011)	(0.010)
$y_{I,t}$	-0.059***	-0.043***	0.873***	-0.014	-0.026*	0.528***
	(0.006)	(0.007)	(0.007)	(0.012)	(0.014)	(0.012)
$log(wage_t)$	0.009***	-0.001	0.009***	0.004	-0.002	0.006***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)
$log(wealth_t)$	0.001***	0.001**	0.001	0.001*	0.001*	0.001
	(0.000)	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)
N	14158	14158	14158	14158	14158	14158
Adjusted R^2	0.82	0.81	0.85	0.08	0.08	0.14

Notes: *$p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. $y_{C,t}$, $y_{M,t}$, and $y_{I,t}$ denote the cognitive, manual, and interpersonal tasks at period t respectively. Tenure indicates years of work experience. u_{OLS} and u_{FE} are residuals from the OLS and fixed effects regressions. The dependent variable in columns (1) and (4) is cognitive task intensity in the subsequent period on a 0-1 scale. The dependent variable in columns (2) and (5) is manual task intensity in the following period, and the depdendent variable in columns (3) and (6) is interpersonal task intensity in the following period. OLS regression also controls for a squared term in years of exerience, highest graded completed, as well as indicators for worker race, highschool graduation, and college completion. Parentheses contain standard errors that are robust to clustering at the individual level.

Table 5: Wage and Learning Parameter Estimates, Base

	Estimate	SE
$A_{1,c}$	-0.037***	(0.014)
$A_{1,m}$	0.021	(0.020)
$A_{1,i}$	0.105***	(0.005)
$A_{2,c}$	-0.028	(0.028)
$A_{2,m}$	-0.044	(0.029)
$A_{2,i}$	-0.095***	(0.007)
$\sigma^2_{\eta,c}$	0.076***	(0.008)
$\sigma^2_{\eta,m}$	0.183***	(0.042)
$\sigma^2_{\eta,i}$	0.009***	(0.000)
B_0	5.039***	(0.025)
$B_{1,c}$	-0.947***	(0.047)
$B_{1,m}$	-0.519***	(0.060)
$B_{1,i}$	1.154***	(0.031)
$B_{2,c}$	0.302***	(0.053)
$B_{2,m}$	0.244***	(0.056)
$B_{2,i}$	0.168***	(0.051)
$B_{3,c}$	1.333***	(0.071)
$B_{3,m}$	1.024***	(0.117)
$B_{3,i}$	0.536***	(0.063)
σ^2_{ϵ}	0.019***	(0.000)

XXX

Table 6: Wage and Learning Parameter Estimates, Base

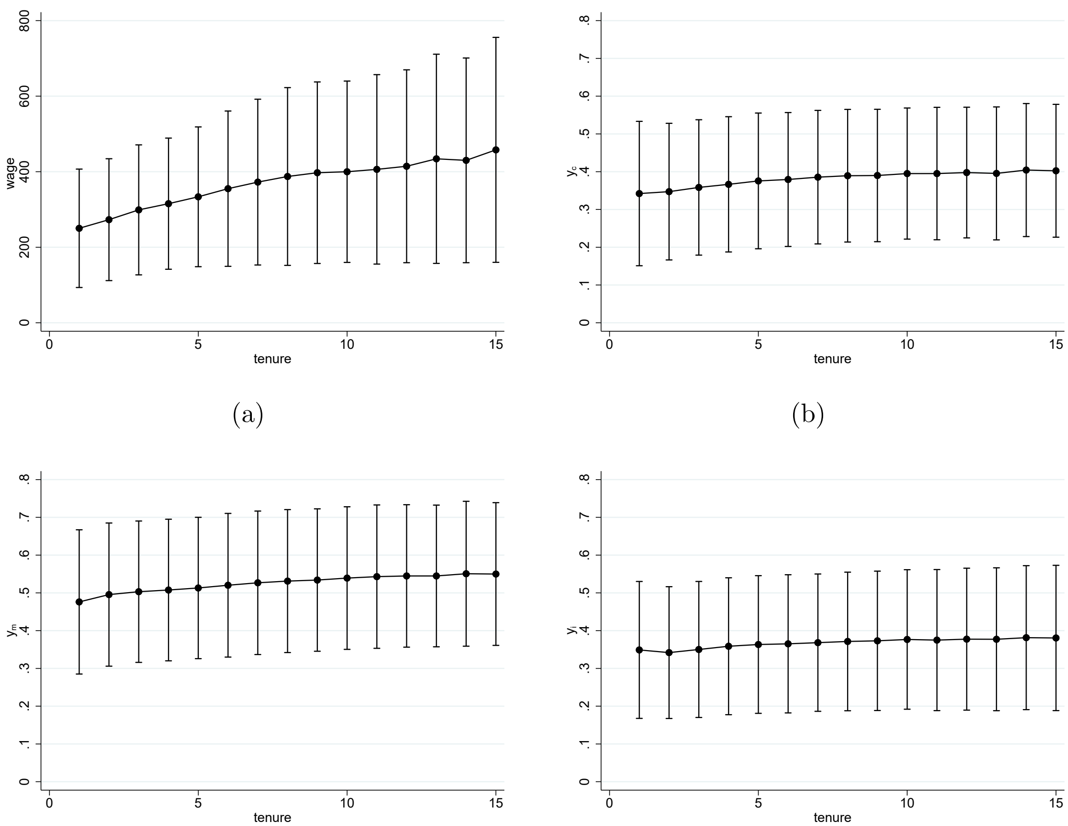
	Estimate	SE
$A_{1,c}$	-0.020	(0.036)
$A_{1,m}$	0.034	(0.054)
$A_{1,i}$	0.101***	(0.012)
$A_{2,c}$	-0.037	(0.048)
$A_{2,m}$	-0.054	(0.053)
$A_{2,i}$	-0.100***	(0.012)
$\sigma^2_{\eta,c}$	0.074***	(0.008)
$\sigma^2_{\eta,m}$	0.167***	(0.040)
$\sigma^2_{\eta,i}$	0.009***	(0.001)
B_0	5.046***	(0.027)
$B_{1,c}$	-0.951***	(0.051)
$B_{1,m}$	-0.543***	(0.064)
$B_{1,i}$	1.154***	(0.031)
$B_{2,c}$	0.262***	(0.059)
$B_{2,m}$	0.279***	(0.064)
$B_{2,i}$	0.138***	(0.052)
$B_{3,c}$	1.335***	(0.077)
$B_{3,m}$	1.067***	(0.125)
$B_{3,i}$	0.485***	(0.066)
σ^2_{ϵ}	0.019***	(0.000)
$A_{3,1,c}$ (Highest Grade)	-0.001*	(0.001)
$A_{3,2,c}$ (College or More)	0.010*	(0.005)
$A_{3,3,c}$ (Race–Black)	0.010***	(0.004)
$A_{3,4,c}$ (Race–Hispanic)	0.002	(0.004)
$A_{3,1,m}$ (Highest Grade)	-0.000	(0.001)
$A_{3,2,m}$ (College or More)	-0.010	(0.008)
$A_{3,3,m}$ (Race–Black)	0.021***	(0.005)
$A_{3,4,m}$ (Race–Hispanic)	0.006	(0.006)
$A_{3,1,i}$ (Highest Grade)	0.000	(0.000)
$A_{3,2,i}$ (College or More)	0.006***	(0.001)
$A_{3,3,i}$ (Race–Black)	-0.004***	(0.001)
$A_{3,4,i}$ (Race–Hispanic)	-0.003**	(0.001)

XXX

Table 7: Wage and Learning Parameter Estimates, Base

	Estimate	SE
$A_{1,c}$	-0.010***	(0.000+0.000i)
$A_{1,m}$	-0.298***	(0.000+0.003i)
$A_{1,i}$	0.157***	(0.000+0.008i)
$A_{2,c}$	0.048***	(0.000+0.000i)
$A_{2,m}$	0.317***	(0.000+0.004i)
$A_{2,i}$	0.321***	(0.000+0.018i)
$\sigma^2_{\eta,c}$	0.004***	(0.000+0.000i)
$\sigma^2_{\eta,m}$	0.105***	(0.000+0.003i)
$\sigma^2_{\eta,i}$	1.957***	(0.000+0.006i)
B_0	4.823***	(0.004)
$B_{1,c}$	0.650***	(0.000+0.002i)
$B_{1,m}$	-0.885***	(0.000+0.002i)
$B_{1,i}$	0.209***	(0.000)
$B_{2,c}$	0.523***	(0.000+0.005i)
$B_{2,m}$	0.205***	(0.000+0.005i)
$B_{2,i}$	-0.609***	(0.000+0.004i)
$B_{3,c}$	1.082***	(0.000+0.004i)
$B_{3,m}$	1.440***	(0.000+0.006i)
$B_{3,i}$	-0.139***	(0.000+0.000i)
$B_{4,t}$	0.088***	(0.000+0.003i)
σ^2_{ϵ}	0.005***	(0.000+0.000i)

XXX



(a) (b)

(c) (d)

Figure 2: Wage and Occupation Choices by Years of Experience Panel **??** plots the average wages and standard deviations by years of experience. Panel **??**, **??**, and **??** plot average cognitive, manual, and interpersonal task intensities by years of labor market experience, respectively.

Appendix A Additional Tables and Figures

Table A1: Sample Mean by Years of Experience–Wage and Skill Requirements

	(1) Wage			(2) Cognitive Task			(3) Manual Task			(4) Interpersonal Task		
	mean	sd	count	mean	sd	count	mean	sd	count	mean	sd	count
1	250.15	156.74	1757	0.34	0.19	1745	0.48	0.19	1745	0.35	0.18	1745
2	273.00	161.34	2370	0.35	0.18	2374	0.50	0.19	2374	0.34	0.17	2374
3	298.92	172.08	2425	0.36	0.18	2427	0.50	0.19	2427	0.35	0.18	2427
4	315.35	173.61	2459	0.37	0.18	2461	0.51	0.19	2461	0.36	0.18	2461
5	333.52	185.05	2464	0.38	0.18	2474	0.51	0.19	2474	0.36	0.18	2474
6	355.06	205.60	2469	0.38	0.18	2484	0.52	0.19	2484	0.37	0.18	2484
7	372.45	219.49	2463	0.39	0.18	2469	0.53	0.19	2469	0.37	0.18	2469
8	387.34	235.36	2450	0.39	0.18	2464	0.53	0.19	2464	0.37	0.18	2464
9	397.33	240.33	2360	0.39	0.18	2379	0.53	0.19	2379	0.37	0.18	2379
10	399.92	240.08	2306	0.40	0.17	2327	0.54	0.19	2327	0.38	0.18	2327
11	406.20	250.80	2189	0.40	0.18	2211	0.54	0.19	2211	0.37	0.19	2211
12	414.38	255.36	2117	0.40	0.17	2137	0.54	0.19	2137	0.38	0.19	2137
13	434.14	276.91	1938	0.40	0.18	1962	0.54	0.19	1962	0.38	0.19	1962
14	430.05	271.00	1803	0.40	0.18	1824	0.55	0.19	1824	0.38	0.19	1824
15	457.83	297.79	1601	0.40	0.18	1616	0.55	0.19	1616	0.38	0.19	1616
Total	365.61	230.99	33171	0.38	0.18	33354	0.52	0.19	33354	0.37	0.18	33354

Notes: This table presents means, standard deviations, and number of observations of variables by years of experience.

Table A2: Log Wage Regression

	(1) OLS	(2) FE
$y_{C,t}$	1.796***	1.291***
	(0.273)	(0.167)
$y_{M,t}$	0.272	1.074***
	(0.284)	(0.164)
$y_{I,t}$	-0.040	-0.107
	(0.346)	(0.207)
$y_{C,t} * y_{M,t}$	-4.303***	-3.603***
	(0.433)	(0.305)
$y_{C,t} * y_{I,t}$	-1.861***	-1.249***
	(0.503)	(0.359)
$y_{M,t} * y_{I,t}$	1.742***	1.294***
	(0.426)	(0.292)
$y^2_{C,t}$	2.276***	1.864***
	(0.370)	(0.229)
$y^2_{M,t}$	0.808***	0.352**
	(0.225)	(0.145)
$y^2_{I,t}$	0.182	0.711***
	(0.324)	(0.211)
$s_{C,0}$	0.680***	
	(0.172)	
$s_{M,0}$	-0.241	
	(0.187)	
$s_{I,0}$	-0.231	
	(0.307)	
$y_{C,t} * s_{C,0}$	-0.258	
	(0.260)	
$y_{M,t} * s_{M,0}$	0.633***	
	(0.241)	
$y_{I,t} * s_{I,0}$	0.554**	
	(0.276)	
N	32585	32585
Adjusted R^2	0.42	0.03

Notes: $*p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$. The dependent variable is log wage. Parentheses contain standard errors that are robust to clustering at the individual level. Estimates in column (1) and (2) are from OLS and individual fixed effects regressions, respectively. OLS regression also controls for a squared term in years of exerience, highest graded completed, as well as indicators for worker race, highschool graduation, and college completion.