HITOTSUBASHI UNIVERSITY

Global COE Hi-Stat Discussion Paper Series 088

Research Unit for Statistical and Empirical Analysis in Social Sciences (Hi-Stat)

Formation of Heterogeneous Skills and Wage Growth

Shintaro Yamaguchi

September 2009

Hi-Stat Discussion Paper

Hi-Stat
Institute of Economic Research
Hitotsubashi University
2-1 Naka, Kunitatchi Tokyo, 186-8601 Japan
http://gcoe.ier.hit-u.ac.jp

Formation of Heterogeneous Skills and Wage Growth*

Shintaro Yamaguchi†

September 20, 2009

Abstract

This paper examines how primitive skills associated with occupations are formed and rewarded in the labor market over the careers of men. The objective task complexity measurement from the Dictionary of Occupational Titles enables a more direct look into the primitive skills of workers. I show that the optimal choice of task complexity is a linear function of unobserved skills, worker characteristics, and preference shocks, which implies that the observed task complexity is a noisy signal of underlying skills. Using career histories from the NLSY79, the growth of cognitive and motor skills as well as structural parameters are estimated by the Kalman filter. The results indicate that both cognitive and motor skills account for a considerable amount of cross-sectional wage variation. I also find that cognitive skills grow over careers and are the main source of wage growth; this pattern is particularly pronounced for the highly educated. In contrast, motor skills grow and contribute to wage growth substantially only for high school dropouts.

1 Introduction

Heterogeneity of the type of worker skills is a central feature of labor economics. One approach to understanding skill heterogeneity is to assume homogeneous skills for workers classified by such criteria as race, sex, and education. Another approach focuses on the sector affiliation of workers. This approach not only defines economically more meaningful skill categories, but is also empirically successful and several papers[1] find evidence of sector specific skills. While these

*Supplementary appendix is available from http://socserv.mcmaster.ca/yamtaro.

†Department of Economics, McMaster University, 1280 Main St. W., Hamilton, ON. Canada L8S 4M4. Email: yamtaro@mcmaster.ca. I am grateful to comments from Chris Ferrall, Hiro Kasahara, John Kennan, Salvador Navarro, Ronni Pavan, Chris Robinson, Yuya Takahashi, Chris Worswick and other participants of the 14th meeting of Society of Labor Economists, The 43rd Annual Conference of the Canadian Economics Association, Society for Economic Dynamics Annual Meeting, and seminar participants at Alberta, Bureau of Labor Statistics, Hitotsubashi, Osaka (ISER), and Western Ontario. The use of SHARCNET computational facilities is gratefully acknowledged.

[1]Papers in this area include Heckman and Sedlacek (1985), Neal (1995), Keane and Wolpin (1997), Parent (2000), Pavan (2006), Kambourov and Manovskii (2009), Yamaguchi (forthcoming), and Sullivan (2008, forthcoming)

papers highlight the importance of sector affiliation in understanding worker skills, the theory of specific human capital ignores apparent similarity in skills between sectors. If skills are assumed idiosyncratically different across finely defined sectors, it is hard to develop a widely applicable theory of skills. When only a small number of sectors are considered, skill specificity is more plausible and interpretation is simplified, but plausible heterogeneity in skills within the broadly defined sector is ignored.

This paper departs from the previous contributions by taking a more direct look at the primitive skills of workers and their tasks. The objective task complexity measures from the Dictionary of Occupational Titles (DOT) makes this approach possible. In the model, individuals and jobs are characterized in continuous and multidimensional spaces of skills and task complexity, respectively. Individuals are heterogeneous in their endowment of primitive skills such as cognitive and motor skills. They synthesize these different skills in order to perform their tasks. Similarly, jobs are heterogeneous in cognitive and motor task complexity. Individuals are engaged in both cognitive and motor tasks in any job, but the complexity of each task varies across jobs. This skill and task complexity space approach allows for a clearer interpretation of heterogeneous workers and jobs from the viewpoint of a few fundamental skills and tasks. Moreover, this approach also enables the model to account for heterogeneity in hundreds of occupations without suffering from the curse of dimensionality, because neither the state variables nor the parameters increase with the number of occupations in the model.

Some recent papers attempt to look into the skills associated with jobs by using information about job tasks. Ingram and Neumann (2006) and Bacolod and Blum (forthcoming) include task complexity measures in wage regressions and study the resulting implications for wage inequality. Autor, Levy, and Murnane (2003) find that workplace computerization has replaced routine cognitive and manual tasks, which results in a labor demand shift in favor of the educated. These examples demonstrate the usefulness of job task information in understanding worker skills. An important challenge for most papers in the task-based approach literature is a lack of an explicit distinction between skills that are possessed by workers and skills that are required for a job.[2] This lack of distinction implicitly assumes that workers in the same occupation have identical skills, and thus, the returns to skills are confused with the returns to tasks. This paper departs from these previous contributions by distinguishing between skills and task complexity explicitly.

Heterogeneous jobs affect individual welfare in three different ways according to task complexity. First, skills are better rewarded when the relevant task is complex. For example, those who are endowed with cognitive skills are paid better in a job with a complex cognitive task. Sec-

[2] An important exception is Poletaev and Robinson (2008). Using the Displaced Worker Surveys and the Dictionary of Occupational Titles, they examine wage changes following displacement. Poletaev and Robinson (2008) find a greater wage loss for those who have moved to jobs that require a different skill portfolio than their pre-displacement jobs.

ond, individuals learn more skills when the relevant task is complex. Skills are acquired through learning-by-doing, with the amount of learning increasing in the intensity of the task. Individuals who have spent many years in a motor skill intensive job will have accumulated a large amount of motor skills. Individuals also directly receive utility from the job characteristics. Skilled individuals tend to like complex tasks, while unskilled individuals find them unpleasant.

Individuals enter the labor market with a heterogeneous initial skill endowment. They are also heterogeneous in preference and learning ability. In each period, until the terminal period, individuals choose an optimal job. Unlike many other papers that consider the occupational choice problem as a discrete choice problem, this paper models it as a continuous choice problem in a task complexity space.

The main theoretical finding is that the optimal policy function for occupational choice is a linear function of unobserved skills, worker characteristics, and preference shocks under the set of my assumptions. This policy function is remarkably simple relative to other structural dynamic models of occupational choice such as Keane and Wolpin (1997). This analytical solution allows me to include many state variables including observed worker characteristics without major computational burden. This is not the case for other structural models that must be solved numerically using discrete state space. The linear policy function implies that the observed task complexity can be interpreted as a noisy signal of unobserved skills. Hence, the growth of unobserved skills as well as the structural parameters can be estimated by the Kalman filter.

The data used to construct wage and occupational choice histories of men are drawn from the NLSY79. They are merged with a task complexity measurement from the DOT. Using the detailed task complexity information from the DOT, I construct two broadly defined task categories: cognitive and motor tasks. The estimation results are intuitive and indicate that the returns to skills and the amount of skill learning increases with task complexity. I also find that both cognitive and motor skills account for a considerable amount of cross-sectional logwage variation, although cognitive skills are more important than motor skills in explaining the wage distribution. Finally, I examine the growth of unobserved skills and how the skills contribute to wage growth for each level of education. The average cognitive skill grows over time and they are the main source of wage growth for all education levels. The amount of cognitive skill growth increases with the level of education. In contrast, average motor skills grow only for high school dropouts. The growth of motor skills account for about 30% of the wage growth of high school dropouts during the first 10-20 years of their careers.

The rest of the paper is organized as follows: Section 2 lays out the structural model and shows the optimal policy function. Section 3 explains the estimation strategy of the model by the Kalman filter. The identification issue is also discussed here. Section 4 describes the data set. The estimation results are presented and discussed in Section 5. Section 6 concludes.

2 Model

In this section I describe a dynamic model of occupational choice and skill formation. In the model, each individual who made a long term transition to the full-time labor market has a finite decision horizon ending in an exogenously fixed retirement age. There is no distinction between jobs and occupations. In other words, jobs in the same occupation are homogeneous in terms of task complexity. In each year t, an individual chooses an occupation that lies in a K-dimensional continuous space of task complexity x_t. Sufficiently many occupations exist so that an individual can choose any occupation in the task complexity space. Skills in year t are denoted by a K-dimensional vector s_t.

2.1 Wage Function

Skills are differently rewarded across occupations according to task complexity. Let $p(x_t)$ be a K-dimensional vector of the marginal rate of returns to skills when the task complexity of the job in year t is x_t. Wages depend on skill quantity and its returns;

$$\ln w_t \quad = \quad p_0 + p^{'}(x_t)s_t + \eta_t, \tag{1}$$

where p_0 is a constant term and η_t is an i.i.d. error term that follows the normal distribution with mean zero and variance σ_η^2. This error term η_t can be interpreted either as a productivity shock unrelated to worker skills s_t or as a measurement error. The return to skills $p(x_t)$ increases with task complexity; $\partial p^j(x_t)/\partial x_t^j > 0$, where j is a superscript for skill dimension. For example, cognitive skills are better rewarded in a job where the cognitive task is more complex. When task complexity is low, worker skills have little effect on the productivity of a job. A low-skill individual can perform the tasks of less skill demanding job such as house keeping satisfactorily. In addition, a high-skill individual is unlikely to far outperform a low-skill individual in such a simple task. In contrast, the productivity of a skill demanding job such as a managerial task is sensitive to worker skills. Because the quality of a manager affects the productivity of her subordinates, a small difference in managerial skills can translate into a large productivity difference. A low-skill individual performs managerial tasks poorly and produces little output relative to a high-skill individual.

I parametrize the wage equation as

$$\ln w_t \quad = \quad p_0 + [p_1 + P_2^{'} x_t]^{'} s_t + \eta_t, \tag{2}$$

where p_1 is a K-dimensional vector and P_2 is a $K \times K$ positive definite diagonal matrix of the parameters.

This wage function[3] also provides an interpretation about job match quality introduced by Jovanovic (1979). While many empirical papers[4] find that the wage gain from job search is substantial for young workers, their models do not explain why match quality varies across jobs explicitly. In this model, a worker receives a wage gain when he moves to an occupation in which the returns to skills are high. The size of this wage gain depends on the skill endowment of the worker. Consider a worker with high motor skills and little cognitive skills. This worker can expect a large wage gain by taking a job with complex motor tasks, but cannot expect a significant wage gain by taking a job with complex cognitive tasks. Desirable occupations from the viewpoint of wage gain vary across individuals, according to his skill endowment. The job search process can be interpreted as a process to reach an occupation that offers a higher return to skills that a worker has abundantly.

2.2 Skill Formation

The amount of skill formation varies across individuals in three ways. First, it depends on tasks in which a worker is engaged; a worker learns more skills by using them more intensely. For example, individuals accumulate more motor skills by taking on motor skill intensive tasks. This approach is a natural extension of learning-by-doing to a model in which jobs are heterogeneous. Second, individuals differ in learning ability. For example, education affects not only the initial skill endowment at the entry to the labor market, but also learning ability. Third and lastly, an i.i.d. skill shock influences skill growth.

The skill transition is defined by a linear function of task complexity, worker characteristics, and the current skill level. Let d be a L-dimensional vector of individual characteristics fixed at labor market entry such as race and education. A vector of skill shocks ε_t is normal, independent and identically distributed with mean zero and variance Σ_ε: $\varepsilon_t \sim N(0, \Sigma_\varepsilon)$. Skills grow from year t to year $t+1$ according to the following skill transition equation

$$s_{t+1} = Ds_t + a_0 + A_1 x_t + A_2 d + \varepsilon_{t+1}, \tag{3}$$

where D is a K-dimensional diagonal matrix for skill depreciation, a_0 is a K-dimensional vector of parameters, A_1 is a $K \times K$ diagonal matrix of the marginal effects of task complexity on learning, A_2 is a $K \times L$ dimensional matrix that represents heterogeneous learning ability.

[3]Lazear (2004) and Gibbons, Katz, Lemieux, and Parent (2005) present wage functions similar to this model in that they allow for returns to skills to vary across jobs depending on tasks.

[4]The contributions in this area include Topel and Ward (1992), Schönberg (2007), Neal (1999), Pavan (2006), Yamaguchi (forthcoming), and Sullivan (forthcoming).

An alternative to this learning-by-doing assumption is an on-the-job training (or skill investment) model such as that proposed by Ben-Porath (1967). However, Heckman, Lochner, and Cossa (2002) find it hard to distinguish learning-by-doing from on-the-job training when using features observed in the data. Moreover, Altonji and Spletzer (1991) find that a skill demanding job offers more skill training. The key property that skills grow more in skill demanding jobs holds true regardless of whether I assume that skills accumulate through learning-by-doing or on-the-job training.

Individuals start their careers with different amounts of initial skills s_1 in both observable and unobservable ways. The initial skill endowment is given by

$$s_1 = h_0 + Hd + \varepsilon_1 \tag{4}$$

where h_0 is a K-dimensional vector, H is a $K \times L$ matrix of parameters, d is observed individual characteristics at labor market entry, and ε_1 is an unobserved component of initial skills which is distributed as $\varepsilon_1 \sim N(0, \Sigma_{s1})$.

2.3 Job Preference

Job preference, as well as skill endowment, can rationalize the observed occupational choices. The following quadratic function of task complexity determines utility from work,

$$\begin{aligned} v_t &= v(x_t, \bar{x}_t, s_t, \tilde{v}_t; d) && (5) \\ &= (g_0 + G_1 d + G_2 s_t + \tilde{v}_t)' x_t + x_t' G_3 x_t + (x_t - \bar{x}_t)' G_4 (x_t - \bar{x}_t), && (6) \end{aligned}$$

where g_0 is a K-dimensional vector of preference parameters, G_1 is a $K \times L$ matrix of preference parameters, $\tilde{v}_t$ is a K-dimensional vector of preference shocks with zero mean, G_2, G_3, G_4 are $K \times K$ diagonal matrices, and $\bar{x}_t$ is a K-dimensional vector of work habits.

The utility from job characteristics varies across individuals according to individual characteristics d, the skill levels s_t, and a preference shock $\tilde{v}_t$. Skilled workers prefer complex tasks if the parameter matrix G_2 is positive definite. I assume that the matrix G_3 is negative definite. This restriction implies that, for a very high value of x_t, the marginal utility from job characteristics is negative; this is the cost of entering an occupation with complex tasks. The parameters g_0, G_1, and G_2 are unrestricted. The last term in the above equation captures the effect of work habits on utility. The work habits of individuals are measured by the weighted average of the task complexity of previous occupations the individual has held. Individuals may have difficulty in adjusting themselves to a new work environment. The mental and physical costs of this adjustment are high when an individual enters into an occupation that is very different from the past occupations held. This effect can be interpreted as a sort of search friction as it prevents workers from reaching a best-paying job immediately. This work habit effect is introduced to approximate the fact that workers do not change occupations every year. It is true that the model can predict more realistic worker mobility patterns by introducing a fixed entry cost to a new occupation or by assuming that workers do not receive job offers every period, but the proposed specification is necessary to derive the linear policy function shown below. I argue that the benefit from this approximation exceeds its loss of realism.

This utility function allows for a rich form of heterogeneity in job preferences: it varies across individuals according to skills, work habits, other observed worker characteristics, and unobserved preference shocks. Consider the utility change of a worker being promoted to a job with complex tasks. If he is unskilled, this promotion may decrease the utility from job characteristics because he may not like the complex tasks and will suffer from adjusting himself to the new tasks. If he is skilled, in contrast, this promotion may increase his utility from job characteristics despite the work habit effect, because a skilled worker likes a complex task.

Individuals form their work habits by the following transition equation

$$\bar{x}_{t+1} = A_3\bar{x}_t + (I - A_3)x_t \tag{7}$$

where A_3 is a K-dimensional diagonal matrix of which elements take values between zero and one and I is a K-dimensional identity matrix. Hence, work habits $\bar{x}$ are a weighted average of the task complexity of the past jobs. When $A_3 = 0$, only the tasks in the last occupation affect work disutility. In contrast, when $A_3 = I$, work habits remain constant at the initial value $\bar{x}_1$. For all other cases where the elements of A_3 are between 0 and 1, the tasks of all past jobs affect work disutility.

Individuals experience part-time jobs and/or are engaged with other activities in and out of school before they transit to the full-time labor market. These experiences outside the full-time labor market form individuals' initial work habits as well as the initial skill endowment. The initial condition for $\bar{x}_t$ varies across individuals according to initial observed characteristics d such that

$$\bar{x}_1 = \bar{x}_{1,0} + Xd, \tag{8}$$

where $\bar{x}_{1,0}$ is a K-dimensional vector of parameters and X is a $K \times L$ matrix of parameters.

2.4 Bellman Equation

The Bellman equation for an individual is given by

$$
\begin{aligned}
V_t(s_t,\bar{x}_t,\tilde{\nu}_t,\eta_t;d) &= \max_{x_t} \ln w(x_t,s_t,\eta_t) + v(x_t,\bar{x}_t,s_t,\tilde{\nu}_t;d) + \beta E V_{t+1}(s_{t+1},\bar{x}_{t+1},\tilde{\nu}_{t+1},\eta_{t+1};d) && (9)\\
\text{s.t.} & \\
\ln w_t &= p_0 + [p_1 + P_2^{'} x_t]^{'} s_t + \eta_t && (10)\\
v_t &= (g_0 + G_1 d + G_2 s_t + \tilde{\nu}_t)^{'} x_t + x_t^{'} G_3 x_t + (x_t - \bar{x}_t)^{'} G_4 (x_t - \bar{x}_t) && (11)\\
s_{t+1} &= D s_t + a_0 + A_1 x_t + A_2 d + \varepsilon_{t+1} && (12)\\
\bar{x}_{t+1} &= A_3 \bar{x}_t + (I - A_3) x_t && (13)\\
s_1 &\sim N(h_0 + Hd,\ \Sigma_{s1}) && (14)\\
\bar{x}_1 &= \bar{x}_{1,0} + Xd. && (15)
\end{aligned}
$$

Because this is a stochastic optimal linear regulator problem, the optimal policy function is a linear function of skills, work habits, individual characteristics, and preference shocks. It can be expressed as

$$
x_t^* = c_{0,t} + C_{1,t} d + C_{2,t} s_t + C_{3,t} \bar{x}_t + \nu_t, \qquad (16)
$$

where $c_{0,t}$ is a K-dimensional vector, $C_{1,t}$ is a $K \times L$ matrix, $C_{2,t}$ and $C_{3,t}$ are K-dimensional diagonal matrices, and ν_t is a K-dimensional vector of rescaled preference shocks (i.e., I can write $\nu_t = M_t \tilde{\nu}_t$ where M_t is a K-dimensional diagonal matrix). The proof is shown in Appendix C. The rescaled preference shocks ν_t are normal, independent, and identically distributed random variables with zero mean and variance matrix Σ_ν. The parameters $c_{0,t}$, $C_{1,t}$, $C_{2,t}$, and $C_{3,t}$ are functions of structural parameters and are not estimated as free parameters. Because the problem has a finite horizon, I solve the value function and the policy function by backward recursion.[5]

This analytical solution allows me to include many state variables including observed worker characteristics without major computational burden. Most structural dynamic models of occupational choice are formulated as a dynamic discrete choice problem. In these models, allowing for rich worker heterogeneity is computationally intractable, because the value function is calculated at each point in the discrete state space. The derived policy function also provides with a useful interpretation of the observed occupational choice: it is a noisy signal of underlying skills. Worker skills are not homogeneous within occupation, but task complexity of the occupation contains useful information about worker skills. Skills and tasks are often not explicitly distinguished in

[5] Many other methods are available for infinite horizon problems. See Anderson, Hansen, McGrattan, and Sargent (1996) for a survey.

empirical literature, but this result characterizes the relationship between them.

3 Estimation Strategy

3.1 Identification Restrictions

I estimate the wage equation (2) and the policy function (16) jointly, given the transition equations (3) and (7) and the initial conditions defined by Equations (4) and (8), under the parameter constraints imposed on the policy function coefficients C's by the structural model. Notice that this model is an application of a state-space model. Like other state-space models (or factor models), there is an identification issue related to unobserved skills. The scale parameters of skill are not identified because observed variables (i.e. wage and task complexity) are the product of unobserved skills and unknown parameters such as the returns to skills. Either a large amount of skills or high returns can rationalize an observed high wage. The location parameters of skills are identified by the nonlinearity of the model except for special cases in which $D = 0$ or $D = I$, but no credible source of identification exists. Hence, I specify the unconditional mean and standard deviation of initial skills as .50 and .10, respectively. Specifically, it is imposed that

$$E(s_1) = h_0 + HE(d) = \begin{pmatrix} .5 \\ .5 \end{pmatrix} \tag{17}$$

$$diag[Var(s_1)] = diag[HE(dd')H' + \Sigma_{s1}] = \begin{pmatrix} .01 \\ .01 \end{pmatrix} \tag{18}$$

where $diag$ is an operator that converts a matrix into a vector that consists of diagonal elements of the matrix. In the estimation, h_0 and $diag[\Sigma_{s1}]$ are not estimated as free parameters, but recovered from these restrictions. The off-diagonal elements of Σ_{s1} are unrestricted and estimated as free parameters. Under these restrictions, skills are essentially always positive, and thus, results are interpretable. For a robustness check, I also estimate the location parameters.[6] The results are very similar to my preferred specification.

For the distinction between unobserved skills (signal) and work disutility shocks (noise), the time dimension of the data is useful. Notice that the skills are serially correlated while work disutility shocks are not. Hence, the unobserved skills are identified by the persistent component of the residuals. Another source of identification of unobserved skills comes from the correlation between wages and task complexity. Heckman, Stixrud, and Urzua (2006) and Cunha and Heckman (2008) use multiple skill measures such as test scores and identify an underlying skill from the cor-

[6] See supplementary appendix for the estimation results.

relation between the measures. Although this paper considers only one task complexity measure for each skill, these authors' identification strategy can still be used because the wage, as well as the task complexity, is affected by unobserved skills.

Not all parameters for the utility from job characteristics g_0, G_1, G_2, G_3, and G_4 are identified, because the number of these parameters is greater than the number of the parameters in the policy function. To see this unidentifiability, consider the optimal choice of occupation at the terminal period T. The optimal task complexity x_T^* is given by

$$\begin{aligned} x_T^* &= -\frac{1}{2}(G_3+G_4)^{-1}[g_0+G_1 d+(P_2+G_2)s_T-2G_4\bar{x}_T+\tilde{v}_T] && (19)\\ &\equiv c_{0,T}+C_{1,T}d+C_{2,T}s_T+C_{3,T}\bar{x}_T+v_T. && (20)\end{aligned}$$

Here, the number of structural parameters to be estimated is $(4+L)K$, while the number of parameters in the policy function is $(3+L)K$. To proceed, K parameters have to be fixed. Specifically, it is imposed that G_3 be the negative of an identity matrix. This normalization is innocuous as long as G_3 is negative definite, because the policy function parameters are consistently estimated regardless of how I normalize G_3. The structural parameters G_1, G_2, and G_4 are identified by the variations of worker characteristics d, skills s_t, and work habit $\bar{x}_t$, respectively. The parameter g_0 is identified by the constant term of the policy function. The exact relationships between the structural parameters and the policy function parameters are outlined in Appendix C.

3.2 Kalman Filter

I use the Kalman filter to calculate the likelihood. The Kalman filter is an algorithm used to estimate recursively the distribution of unobserved state variables (i.e. skills) from observed noisy signals (i.e. the task complexity of occupation and wages).

Suppose that skills are normally distributed given task complexity x_t and wages w_t up to year $t-1$ and the initial work habit $\bar{x}_1$ and the fixed worker characteristics d. The conditional mean and variance of skills are

$$\begin{aligned} E(s_t|x_1,w_1,\cdots,x_{t-1},w_{t-1};\bar{x}_1,d) &\equiv E(s_t|Y_{t-1}) && (21)\\ &\equiv \hat{s}_{t|t-1} && (22)\\ Var(s_t|x_1,w_1,\cdots,x_{t-1},w_{t-1};\bar{x}_1,d) &\equiv Var(s_t|Y_{t-1}) && (23)\\ &\equiv \Sigma^s_{t|t-1} && (24)\end{aligned}$$

where Y_{t-1} summarizes all the information up to year $t-1$. The optimal choice of task complexity is also normally distributed, because the policy function (see Equation 16) is linear in skills, the work habit (i.e. the weighted average of the task complexity of the past occupations), and preference shocks $\tilde{v}_t$. The conditional mean and variance of x_t given Y_t are

$$\begin{aligned} E(x_t|Y_{t-1}) &= c_{0,t} + C_{1,t}d + C_{2,t}\hat{s}_{t|t-1} + C_{3,t}\bar{x}_t &(25)\\ Var(x_t|Y_{t-1}) &= C_{2,t}\Sigma^s_{t|t-1}C^{'}_{2,t} + \Sigma_v. &(26)\end{aligned}$$

I then update the conditional distribution of skills using task complexity in the current period x_t so that

$$\begin{aligned} E(s_t|Y_{t-1},x_t) &= \hat{s}_{t|t-1} + \Sigma^s_{t|t-1}C^{'}_{2,t}(C_{2,t}\Sigma^s_{t|t-1}C^{'}_{2,t} + \Sigma_v)^{-1}\hat{v}_t &(27)\\ Var(s_t|Y_{t-1},x_t) &= \Sigma^s_{t|t-1} - \Sigma^s_{t|t-1}C^{'}_{2,t}(C_{2,t}\Sigma^s_{t|t-1}C^{'}_{2,t} + \Sigma_v)^{-1}C_{2,t}\Sigma^s_{t|t-1}, &(28)\end{aligned}$$

where $\hat{v}_t$ is a vector of residuals and $\hat{v}_t = x_t - E(x_t|Y_{t-1})$. Notice that the logwage is a linear function of normal random variables given information up to $t-1$ and the current occupational characteristics x_t. Thus, the logwage is also normally distributed given Y_{t-1} and x_t. The conditional mean and variance of the logwage are

$$\begin{aligned} E(\ln w_t|Y_{t-1},x_t) &= p_0 + [p_1 + P^{'}_2 x_t]^{'}E(s_t|Y_{t-1},x_t) &(29)\\ Var(\ln w_t|Y_{t-1},x_t) &= [p_1 + P^{'}_2 x_t]^{'}Var(s_t|Y_{t-1},x_t)[p_1 + P^{'}_2 x_t] + \sigma^2_\eta. &(30)\end{aligned}$$

Again, I then update the conditional distribution of skills using the information obtained in the current period,

$$\begin{aligned} &E(s_t|Y_{t-1},x_t,w_t)\\ =\ &E(s_t|Y_{t-1},x_t) + Var(s_t|Y_{t-1},x_t)[p_1 + P^{'}_2 x_t][Var(\ln w_t|Y_{t-1}x_t)]^{-1}\hat{\eta}_t &(31)\\ &Var(s_t|Y_{t-1},x_t,w_t)\\ =\ &Var(s_t|Y_{t-1},x_t) -\\ &Var(s_t|Y_{t-1},x_t)[p_1 + P^{'}_2 x_t][Var(\ln w_t|Y_{t-1}x_t)]^{-1}[p_1 + P^{'}_2 x_t]^{'}Var(s_t|Y_{t-1},x_t), &(32)\end{aligned}$$

where $\hat{\eta}_t$ is a logwage residual and $\hat{\eta}_t = \ln w_t - E(\ln w_t|Y_{t-1},x_t)$. Finally, I calculate the conditional distribution of skills in year $t+1$ given information up to year t using the skill transition equation (see Equation 3). Because skills in year $t+1$ are linear in current skills and task complexity, they are also normally distributed with mean and variance,

$$\begin{aligned} \hat{s}_{t+1|t} &= DE(s_t|Y_{t-1},x_t,w_t) + a_0 + A_1 x_t + A_2 d &(33)\\ \Sigma^s_{t+1|t} &= DVar(s_t|Y_{t-1},x_t,w_t)D + \Sigma_\varepsilon. &(34)\end{aligned}$$

This algorithm allows me to calculate the conditional distribution of skills, wages, and occupational characteristics sequentially from the first period $t = 1$ to the last period $t = T$, because the initial skills are normally distributed by assumption. More specifically, using the Kalman filter I calculate the likelihood contribution of each individual as a product of the conditional likelihoods. I have observations of wage and task complexity measures of occupations for each individual $(w_{i1}, x_{i1}, \cdots, w_{iT_i}, x_{iT_i})$,where i is an index for individual and T_i is the last period in the sample for individual i. The likelihood contribution of individual i is

$$\begin{aligned} & l(w_{i1}, x_{i1}, \cdots, w_{iT_i}, x_{iT_i} | \bar{x}_{i1}, d_i) \\ = \; & l(x_{i1} | \bar{x}_{i1}, d_i) l(w_{i1} | x_{i1}; \bar{x}_{i1}, d_i) \times \cdots \times l(x_{iT_i} | Y_{iT_i - 1}) l(w_{i1} | Y_{iT_i - 1}, x_{iT_i}). \end{aligned} \tag{35}$$

The likelihood for the whole sample consisting of N individuals is given by

$$l = \Pi_{i=1}^{N} l(w_{i1}, x_{i1}, \cdots, w_{iT_i}, x_{iT_i} | \bar{x}_{i1}, d_i). \tag{36}$$

4 Data

4.1 Dictionary of Occupational Titles

The DOT contains information on 12,099 occupations that are defined on the basis of the tasks performed. It is constructed by the U.S. Department of Labor to provide standardized occupational information for an employment service matching job applicants with job openings. The information included in the DOT is based on on-site observation of jobs as they are performed in diverse business establishments and, for jobs that are difficult to be observed, on information obtained from professional and trade associations.[7] On this basis, in the revised fourth edition of the DOT, analysts rate each occupation with respect to 62 characteristics that include the aptitudes, temperaments, and interests necessary for adequate performance; the training time necessary to prepare for an occupation; the physical demands of the occupation; and the working conditions under which the occupation typically occurs.

To facilitate interpretation of the data, I summarize the detailed information in the DOT by constructing a low dimensional vector of occupational characteristics by Principal Component Analysis (PCA). Previous studies take two major different approaches. The first approach, which

[7] One might be concerned that task complexity cannot be correctly measured by observing jobs performed, because what analysts observe is a realized combination of job tasks and worker skills in equilibrium: even when the task is simple, an analyst might consider it complex if the worker is skilled. This confusion should be at least partially avoided because the job information is obtained from other sources as well (e.g. interviewing incumbents and supervisors.) It is also worth noting that the DOT explicitly states that each occupation is defined on the basis of the tasks performed. See Miller, Treiman, Cain, and Roos (1980) for a critical review of the DOT.

is employed by Bacolod and Blum (forthcoming), Yamaguchi (2008), and this paper, assumes that a subset of DOT variables measures complexity of a single task. For example, it is assumed that three General Educational Development (reasoning, mathematics, and language) variables measure cognitive task complexity, but does not measure other types of skills such as motor skills. This approach requires a priori knowledge about which variable measures which type of skill. The second approach assumes that a DOT variable contains information about several underlying skills that are orthogonally distributed. For example, unlike the first approach, it is assumed that the General Educational Development variables above contain information about both cognitive and motor task complexity. The second approach does not require a priori knowledge mentioned above, but does impose task complexity be orthogonally distributed. Ingram and Neumann (2006) employ the method similar to this.

It seems impossible to determine which approach is better than the other in general, because these two approaches impose different restrictions. Nevertheless, I find the first approach is more suitable for this paper given the data, because the constructed task complexity vector has a clear interpretation. This is not necessarily the case under the orthogonality assumption in the second approach, particularly when two seemingly unrelated DOT variables have high loadings on the same factor. For example, it is hard to interpret a factor when a variable that is related to worker intelligence and a variable related to physical strength have high loadings for that factor. In the first approach, this possibility is excluded by construction: a variable related to worker intelligence is not used to construct a physical strength measure.

In light of the job analysis literature as well as previous economics papers that use the DOT, this paper assumes that tasks are broadly categorized into either cognitive tasks or motor tasks. While I assume these two task categories a priori, the PCA with the orthogonality assumption also yields a similar set of factors.[8] Autor, Levy, and Murnane (2003) consider different task groups including routine and non-routine manual tasks to understand the role of technological change in the labor market, but these tasks capture different aspects of the motor tasks needed to perform a job. More narrowly defined task categories could be considered. For example, cognitive tasks could be separated into intelligent and communication tasks. However, these two are highly correlated with each other and do not seem to provide additional insights into underlying worker skills. Breaking down tasks increases the number of free parameters and thus, complicates the source of identification.[9]

[8]See supplementary appendix for the robustness check on this issue.

[9]Although I believe that cognitive and motor skills are most suitable for my sample of men, different skill categorization may be more useful for a different demographic group. For example, if one is interested in the highly educated, considering numerical skills and verbal skills may make more sense than cognitive and motor skills, because the highly educated are unlikely to be rewarded for their physical abilities. Skill categories should be chosen depending the demographic group.

By examining the textual definitions of the DOT variables, I assume what variables measure which type of task complexity. The following choices seem to be reasonable and I confirm that the constructed task complexity index is robust to the choice of the DOT variables.[10] The DOT variables that measure cognitive task complexity consists of 2 worker function variables (Data and People), 3 General Educational Development variables (reasoning, mathematical, and language), 3 aptitude variables (Intelligence, Verbal, and Numerical), and 3 adaptability variables (influencing people, accepting responsibility for direction and dealing with people). Motor task complexity is measured by 1 worker function variable (Things), Motor Coordination, Finger Dexterity, and Manual Dexterity, Eye-hand-foot Coordination, Spatial Perception, Form Perception, Color Discrimination, Setting Limits, Tolerance or Standards, and 20 physical demand variables.

In the PCA, factor loadings are calculated so that variation of the data explained by the constructed variables is maximized (or minimize the information loss, equivalently.)[11] For this purpose, I take a sample of occupational characteristics from the April 1971 CPS augmented by the fourth edition of the DOT.[12]This augmented CPS file contains the 1970 census occupation code, the DOT occupation code, and the DOT variables. I update occupational characteristics using the revised fourth edition by matching the DOT occupation code.[13] Most previous studies including Ingram and Neumann (2006), Bacolod and Blum (forthcoming), and Poletaev and Robinson (2008) use the 1995 DOT alone that do not contain the number of workers holding each occupation. Those authors are forced to assume the weights of each DOT occupation are equal, despite the fact that, in the real economy, some occupations may have few workers and some occupations may have thousands. This limitation may bias the task complexity index in an unpredictable direction. I avoid this problem because the augmented CPS contains the number of workers in each DOT occupation. The results of the PCA in Table 1 indicate that the estimated factor loadings are quite intuitive. Following Autor, Levy, and Murnane (2003), I further convert these first principal components into percentile scores using the weights from the augmented CPS file. I also estimate the model using the first principal components without converting them into percentile scores. The results are almost unchanged (see supplementary appendix for detail.) This calculation by the PCA generates the task complexity indexes for each DOT occupation, not the 1970 census occupation.

To combine the task complexity indexes with career histories from the NLSY, task complexity indexes need to be constructed for each census occupation. The indexes for the census occupations are constructed by taking the average values of the constructed indexes over individuals in a census occupation using the augmented CPS. Notice that the number of DOT occupations contained in

[10]See supplementary appendix for the robustness check on this issue.

[11]See supplementary appendix for a technical discussion on PCA.

[12]This is compiled by the Committee on Occupational Classification and Analysis at the National Academy of Sciences and available from ICPSR (http://www.icpsr.umich.edu.)

[13]The DOT occupational code crosswalk between the fourth and the revised editions are used for matching.

each of the 1970 census occupation is available in the augmented CPS. If an occupational crosswalk between the census code and the DOT code is used like other papers, one has to assume that all DOT occupations included in the same census occupation have the same weight. Again, this is an advantage of using the augmented CPS over using the DOT alone.

To see if the constructed variables characterize occupations reasonably, I report the mean and standard deviation of the task complexity measures for each census 1-digit occupation in Table 2. The cognitive tasks of professionals are most complex, followed by those of managers. Laborers and household service workers use the lowest level of cognitive skills. This cognitive task complexity measure largely matches the conventional one-dimensional notion of skill found in the empirical literature (Gibbons, Katz, Lemieux, and Parent (2005), for example). However, this index alone is not rich enough to describe heterogeneous tasks across occupations. For example, cognitive task complexity is similar between sales and craft occupations, although the complete nature of tasks differs very much between the two. Motor task complexity more clearly characterizes the difference between sales workers and craft workers. Motor tasks of craftsmen such as automobile mechanics and carpenters are most complex. Tasks of sales workers, household service workers, and managers require little motor skills. These features are quite intuitive and the proposed measurement is a useful description of the heterogeneity of occupations. Another important finding here is that task complexity varies considerably *within* 1-digit occupation. The reported standard deviation is large in all 1-digit occupations. To assess the extent of heterogeneity within occupation more formally, I decompose the total task complexity variance into the within-occupation variance and the between-occupation variance.[14] For cognitive task complexity, about 59% of the total variance is explained by the within-occupation variance. For motor task complexity, the within-occupation variance explains a larger fraction of the total variance at 75%. This variance decomposition indicates that tasks are greatly heterogeneous within 1-digit occupation and that quite a large part of task complexity variation would be lost if one relies on the 1-digit occupation code. Dealing with occupations at the 3-digit level is essentially important in accounting for heterogeneity of jobs thoroughly.

4.2 National Longitudinal Survey of Youth 1979

4.2.1 Sampling Criteria

The National Longitudinal Survey of Youth (NLSY) 1979 is particularly suitable for this study because it contains detailed individual career histories and focuses on young individuals who change

[14] The law of total variance states that $Var(X) = E(Var(X|O)) + Var(E(X|O))$ where X is the task complexity index and O is the occupational affiliation at 1-digit level. The first term is the within-occupation variance and the second term is the between-occupation variance.

occupations more frequently than older individuals. I concentrate on male workers who make a long-term transition to the full-time labor market during the period between 1979 and 2000. Observations after 2000 are not included in the sample, because in surveys later than 2002 occupation code is not compatible with that until 2000.[15] I define a long-term transition to occur when an individual spends three consecutive years working 30 hours per week or more. In the NLSY cross-section sample, 3,003 men are included. I dropped 294 individuals who served actively in the armed forces during the sample period. Out of 2,709 individuals, 149 individuals did not make a long-term transition to the full-time labor market until 2000. I also excluded 11 individuals who made the long-term transition at age 16 or younger, because they are likely to be mismeasured. I also dropped 131 out of 2,549 remaining individuals whose AFQT scores are missing. Finally, 132 individuals are excluded because there are less than 3 years of observations for these individuals. Hourly wages are deflated by the 2005 CPI. If the recorded hourly wage is greater than $100 or less than one dollar, they are regarded as missing because they are likely to be mismeasured. After imposing all sample restrictions, the sample contains the career histories of 2,417 men, and contains 32,849 person-year observations of occupational choices and 27,063 person-year observations of wages. I change these sampling criteria and estimate the model in order to check the robustness of the parameter estimates. Specifically, I relax the restriction of the full-time work and include part-time jobs in the sample. This change does not affect the main results greatly, because male workers take part-time jobs only in the beginning of their careers. In another exercise, I also exclude self-employed workers. This additional restriction greatly decreases the number of individuals in the sample, but it does not change the main results substantially. I conclude from these exercises that the sampling criteria are not crucial for the main results of the paper.[16] The sample mean AFQT score is 0.50 and the sample standard deviation is 0.29. The sample mean years of education is 13.22 and the sample standard deviation is 2.45. In the sample, 10% of individuals are black and 7% are Hispanic.

Previous empirical papers, including Neal (1999) and Sullivan (2009), report that the occupation codes in the NLSY are often misclassified. One possible way to correct these errors is to assume that all occupation changes within the same employer are false. Neal (1999), Pavan (2006), and Yamaguchi (forthcoming) take this approach to identify their broadly defined occupation changes. However, this edit is likely to result in a downward bias in the mean task complexity, because many occupation code changes within the same employer are promotions to managers.

[15]Since 2002 survey, the NLSY adopts the 2000 census three-digit occupation code and it is quite different from the 1970 census code. Although IPUMS-CPS provides an occupational crosswalk, some occupations in the 2000 census code do not exist in the 1970 census code and thus, the DOT occupational characteristics are missing. O*NET, which is the successor of the DOT, contains occupational characteristics for these recent occupations, but the O*NET variables are not compatible with those of the DOT. Due to these limitations, I drop the post-2000 observations in the NLSY.

[16]The detailed results for the robustness check are available in supplementary appendix.

Another editing method assumes that cycles of occupation code are false. If an individual's occupation code changes from A to B, and then comes back to A in the next year, I edit the code so that he remains in occupation A in all of these three years. I also edit missing occupation code similarly. If I find the same occupation codes in the years bracketing a year in which the occupation code is missing, the missing code is replaced with that found in the bracketing years.

To minimize missing values in the sample, I apply this occupation code correction method after I exclude those who served active armed forces and before I check for a long-term transition to the labor market. This correction method edits 21,995 cases out of 26,923 apparent occupation changes and reduces the annual occupation change rate of the pooled male sample at the 3-digit level from 49% to 40%. This rate is still high, but consistent with the rate reported by Moscarini and Vella (2003) who use the NLSY.[17]

4.3 Career Progression Patterns

The time profiles of the average wage and occupational task complexity are presented in Figure 1. At the point of long-term transition to the labor market, the average cognitive task complexity index of men is 0.41. Individuals take on more and more cognitive skill demanding tasks over time; the cognitive task complexity index reaches .52 in 10 years and .55 in 20 years. The mean motor task complexity slightly decreases over time: at the entry to the labor market, it is .53 and decreases to .52 in 10 years and .51 in 20 years. The mean wages grow by 48% in 10 years and by 64% in 20 years measured by logwage differences.

These profiles are also presented for three different schooling levels: high school dropouts (education less than 12 years), high school graduates (12 years of education), and college graduates (education more than 12 years). The cognitive task complexity indexes at labor market entry are substantially different across education groups and more educated workers take cognitive skill demanding jobs. They are .28 for dropouts, .33 for high school graduates, and .53 for college graduates. But, for all education groups, the cognitive task complexity increases over time. It grows to .36 in 10 years and .44 in 20 years for dropouts and .41 in 10 years and .50 in 20 years for high school graduates. The profile for college graduates is concave. The cognitive task complexity of college graduates grows to .67 in 10 years, but the speed of the growth slows down and it reaches .69 in 20 years.

The profiles of motor task complexity are also very different across education groups. Less educated workers take more motor skill demanding jobs. The motor task complexity at labor market entry is .57 for dropouts, .55 for high school graduates, and .50 for college graduates. Unlike

[17] This is not a problem only with the NLSY. Kambourov and Manovskii (2008) and Moscarini and Thomsson (2008) find occupational classification errors in the PSID and the CPS, respectively.

cognitive task complexity, motor task complexity does not increase over time for all education levels. For high school dropouts, it monotonically grows to .63 in 10 years and .64 in 20 years. The profile for high school graduates is hump-shaped. It peaks at .60 in 6 years and decreases to .53 in 20 years. The motor task complexity for college graduates monotonically decreases to .45 in 10 years and .44 in 20 years.

5 Estimation Results

5.1 Model Fit

Using the estimated parameters, I calculate the predicted paths of the mean wage and the mean task complexity of occupations through simulation. I simulate each individual in the sample 300 times, from his entrance to the full-time labor market until the last year when he is seen in the data. To account for potential attrition problems, if an observation is missing in the data, I treat the corresponding simulation outcomes as missing.

Figure 1 compares the actual and predicted profiles of mean task complexity and hourly logwage over time. The predicted profiles for all men are very close to the actual profiles from the data. For high school dropouts, the model fit to the profiles of motor task complexity and logwage is good, but the predicted cognitive task complexity is lower than the actual profile. For high school graduates, cognitive task complexity is slightly over-predicted, but the level of the predicted motor task complexity and logwage is close to that of the data. Finally, for college graduates, the model fit is very well for all of three dimensions. All in all, the model shows an ability to fit these interesting features of the data.

5.2 Parameter Estimates

5.2.1 Wage Equation

Table 3 presents the parameter estimates of the wage equation (see Equation 2) and their standard errors. The returns to skills significantly increase with task complexity. To see how the wages of an identical worker would vary across occupations solely due to differences in returns to skills, I calculate the difference of two potential wages for some selected workers; the wage that worker would receive in the occupation offering the lowest returns to skills and the wage in the occupation offering the highest returns to skills. First consider an average new labor market entrant who has .50 units of cognitive and motor skills by construction. The returns to cognitive skills range between 5.83 and 5.98 across occupations, because the cognitive task complexity index ranges between 0 and 1. These estimates imply that the wages of this worker can vary by about 8%[18] across occupations due to differences in the returns to cognitive skills alone. Similarly, I find that his wages can vary by about 9% due to differences in returns to motor skills across occupations. The wage change is greater for those endowed with more skills. As shown below, college graduates tend to accumulate more cognitive skills than other individuals, while high school dropouts tend to accumulate more motor skills. Consider an average college graduate with 20 years of experience. He has .69 units of cognitive skills. The wage of this worker can vary by about 10% across occupations due to differences in returns to cognitive skills alone. An average high school dropout with 20 years of experience has .53 units of motor skills. The wage of this worker can vary by about 9% solely due to differences in returns to motor skills.

To assess the importance of each skill in explaining the variation of wages, I decompose the logwage variance. Because the wage consists of two skill components and a white noise term, the logwage variance is the sum of the variance of cognitive skill component, the variance of motor skill component, the covariance of the two, and the variance of the white noise.[19] Table 7 presents the results for different education levels for years 1, 10, and 20. Both the cognitive and motor skill components of wage have a considerable variance for all groups in all years. Although the variance of the cognitive skill component is larger than that of the motor skill component, the results indicate that motor skills account for a considerable amount of the variation in wages. I also find that the relative importance of the cognitive skill component grows over time. In year 20, the variance of the cognitive skill component is more than double of that of the motor skill component. Another interesting finding is that cognitive and motor skills are strongly negatively correlated with each other. The correlation coefficient is between about -.75 and -.79. This negative correlation seems reasonable, because college graduates tend to occupy cognitive skill demanding jobs while high school dropouts tend to occupy motor skill demanding jobs. Lastly, the contribution from an i.i.d. shock is 0.01, which is tiny. This exercise shows that the wage variance is mostly explained by unobserved skills, not by white noise, and that both types of skills account for considerable amount of the cross-sectional wage variation.

[18]This is measured by the logwage difference between the wage that this worker would received in the occupation offering the lowest returns to cognitive skills and the wage that this worker would receive in the occupation offering the highest returns to cognitive skills in the data. The returns to motor skills are held constant between these two occupations.

[19]The wage can be written as

$$\ln w_t = p_0 + p(x_t)^C s_t^C + p(x_t)^M s_t^M + \eta_t$$

where superscripts C is for cognitive skills and M is for motor skills. For example, $p(x_t)^C$ is the returns to cognitive skills and s_t^C is quantity of cognitive skills. I call the second term, $p(x_t)^C s_t^C$, as the cognitive skill component and the third term, $p(x_t)^M s_t^M$, as the motor skill component.

5.2.2 Skill Transition

Parameter estimates for the skill transition equation (see Equation 3) and the initial skill endowment (see Equation 4) are reported in Table 4. The annual skill depreciation rates for cognitive and motor skills are 7-8%. I believe these estimates are in reasonable scale, although other papers do not estimate a parameter that can be directly compared with mine. Other papers that test if human capital depreciates under the assumptions of homogeneous skills and jobs extract the variation of time out of work.[20] This paper cannot apply this identification strategy, because all individuals in the model work full time. Nevertheless, skill depreciation parameters are still identified by the extent to which the task complexity of previous jobs affects the current wage and the current occupational choice.

The amount of learning increases with both cognitive and motor task complexity, although only the coefficient for the cognitive task is significant. To see how learning can vary across jobs, consider a white male high school graduate whose AFQT score is at the median (i.e. .50.) The amount of cognitive skill learning of this individual ranges from .0436 to .0502 across occupations due to differences in task complexity, because the task complexity index lies between 0 and 1 in the data. Similarly, the amount of motor skill learning ranges from .0353 to .0413 across occupations. These estimates indicate that skill learning opportunity diverge considerably across occupations. Skill growth is also different across individuals depending on observed characteristics. In particular, education and AFQT scores are positively correlated with cognitive skill growth, while they are negatively correlated with motor skill growth. I find no significant difference in skill growth across race. The cognitive and motor skill shocks are negatively correlated and the correlation coefficient is about -.77, which is consistent with the distribution of the skills reported above.

The initial skill endowment is significantly different across individuals. Given that the initial skill variances are normalized to 0.01 and the variances of the unobserved components of cognitive and motor skills are 0.0078 and 0.0091, only about 10-20% of the initial skill variation is explained by observed characteristics, which implies that accounting for unobserved skills is important for consistent estimation of the skill growth processes. I also find that the unobserved component of the initial cognitive and motor skills are negatively correlated with the correlation coefficient being -.37. High AFQT score holders tend to have more initial cognitive skills and less initial motor skills. Education has little correlation with initial cognitive skills and is positively correlated with initial motor skills. But, these relationships are not statistically significant. I also do not find a significant difference in initial skill endowment across race. However, these observed characteristics are jointly significant.

[20] See, for example, Mincer and Polacheck (1974), Mincer and Ofek (1982), Albrecht, Edin, Sundström, and Vroman (1999), Edin and Gustavsson (2008), and Gorlich and de Grip (2009).

5.2.3 Job Preference

The parameter estimates for job preference (see Equation 6) are reported in Table 5. Although no single demographic variables are significantly correlated with job preference, they are jointly significant. The utility from job characteristics for skilled workers significantly increases with task complexity, which implies that skilled workers prefer complex tasks. Work habit also strongly affects utility. The transition parameter for work habit is significantly positive for cognitive tasks at .3188, but not significantly different from zero for motor tasks at .0009. In both skill dimensions, the magnitude of the parameters is small, which implies that characteristics from jobs held more than a year ago do not strongly influence occupational choice through job preference.

Initial work habit varies considerably across individuals. High AFQT score holders have better initial conditions for both task dimensions. The educated have a better initial condition for cognitive task, but not for motor tasks. Blacks tend to enter more motor skill demanding and less cognitive skill demanding jobs than whites, while Hispanics tend to enter more cognitive and motor skill demanding jobs than whites.

5.2.4 Policy Function

Table 6 presents the parameter estimates for the policy function (see Equation 16) in year 1. Although the policy function parameters change over time, I report only those in year 1 because they are not substantially different, at least during the sample period. AFQT scores are positively correlated with both cognitive and motor task complexity, although they are not statistically significant. Education is significantly and positively correlated with the optimal cognitive task complexity, but it is insignificantly and negatively correlated with the optimal motor task complexity. Blacks tend to occupy less cognitive and motor skill demanding occupations than whites, while Hispanics tend to occupy more cognitive and less motor skill demanding jobs, but the differences across race are insignificant. The coefficients $C_{2,t}$ for skills are positive and significant, which implies that skilled workers enter more skill demanding jobs. Although skills are similarly scaled, the parameter size is quite different between cognitive and motor tasks. The coefficients $C_{3,t}$ for work habit are positive and significant, and this parameter is larger for cognitive tasks than motor tasks. These coefficients $C_{2,t}$ and $C_{3,t}$ suggest that work habit plays a more important role than skills in the choice of cognitive task complexity, while the opposite is true in the choice of motor task complexity.

To assess the performance of the model in explaining the variation of task complexity, I decompose the variance of task complexity into the component explained by the model and that explained by the preference shocks ν. Table 8 presents the results for the task complexity variance decomposition for each education level for years 1, 10, and 20. For cognitive task complexity, the model accounts for about 40% of the variance for all men in year 1, although it explains only a modest fraction of the within group variance. However, in years 10 and 20, the model accounts for more than half of the variation of cognitive task complexity. The model performs better in explaining the variance of motor task complexity. In year 1, about the half of the variance is explained by the model. In years 10 and 20, the model accounts for about 70% of the variance.

5.3 The Growth of Skills and Wages

Using the parameter estimates, I calculate the time profiles of unobserved skills and their contributions to wage growth. Figure 2 illustrates the growth of average skills by education group. While cognitive skills grow over time for all education levels, the pace of the growth is considerably different and increases with the level of education. At the entry to the labor market, mean cognitive skills are .47, .49, and .52 for high school dropouts, high school graduates, and college graduates, respectively. They grow to .49, .55, and .63 in year 10 and to .51, .58, and .69 in year 20 for each education group. Given that the standard deviation of initial skills is normalized to .10, the cognitive skill differences in later years are substantial. Motor skills grow only for high school drop outs, while they decrease over time for the other education groups. At labor market entry, mean motor skills are .51, .50, and .49 for high school dropouts, high school graduates, and college graduates, respectively. High school dropouts' motor skills grow to .52 in year 10 and to .53 in year 20. Motor skills of high school graduates are almost unchanged (but slightly decreasing) throughout their careers. College graduates experience a modest decline in their motor skills. They decrease to .47 in year 10 and to .46 in year 20. Generally, the changes of motor skills are modest relative to those of cognitive skills.

How does the growth of these skills translate into wages? To answer this question, I decompose wage growth into contributions from cognitive skills and contributions from motor skills. The results for the wage growth decomposition are graphically presented in Figure 3. Cognitive skills are the main source of wage growth for all education groups, with the importance increasing in education level. Cognitive skills increase wages by about 15%, 38%, and 70% in 10 years and by 23%, 57%, and 105% in 20 years for high school dropouts, high school graduates, and college graduates, respectively. For high school and college graduates, wages decrease due to the deterioration of motor skills by 1% and 11% in 10 years and by 2% and 17% in 20 years. However, for high school dropouts, the growth of motor skills is an important source of wage growth. Their wages grow by 7% in 10 years and 10% in 20 years, which implies that motor skills account for about 30% of dropouts' wage growth during the first 10 and 20 years in the labor market. This exercise shows that cognitive skills are the main source of wage growth and the amount of skill growth is considerably different across education levels.

6 Conclusion

This paper constructs and estimates a structural dynamic model of occupational choice. A key feature of the model is that worker skills and job tasks are explicitly distinguished; this is often not done in previous empirical papers. Another important feature of the model is that workers and occupations are characterized in a multidimensional space of skills and task complexity, respectively. This approach has its merit in ease of interpretation and computational simplicity when analyzing heterogeneity of skills and occupational tasks. I also show that the optimal policy function is linear in unobserved skills, preference shocks, and other worker characteristics. This analytical solution allows me to include many worker characteristics in the model without major computational burden. It also provides with an interpretation of the observed task complexity as a noisy signal of underlying skills.

The model is estimated by the Kalman filter using the data from the DOT and NLSY. The empirical results indicate that the returns to skills and the amount of learning increase with the task complexity of a job. By decomposing logwage variance, I show that both cognitive and motor skills account for a considerable amount of cross-sectional wage variation. I also find that the skill growth patterns are significantly different across levels of education. Cognitive skills grow and drive wage growth for all education levels and the amount of skill growth increases with education. Motor skills grow only for high school dropouts and about 30% of their wage growth is accounted for by the motor skill growth.

This paper demonstrates that the proposed model is useful for empirical research of human capital. Given its analytical simplicity, the current model can be easily extended to answer various questions. For example, one might want to allow for part-time work and non participation, which are important aspects of the labor market, particularly for female labor supply. Another possible extension is to endogenize pre-labor market skill investment. The current model is unable to answer questions regarding the effect that schooling may have on individual cognitive and motor skills. These important and interesting issues are to be addressed in future research.

References

ALBRECHT, J. W., P.-A. EDIN, M. SUNDSTRÖM, AND S. B. VROMAN (1999): "Career Interruptions and Subsequent Earnings: A Reexamination Using Swedish Data," *Journal of Human Resources*, 34(2), 294–311.

ALTONJI, J. G., AND J. R. SPLETZER (1991): "Worker characteristics, job characteristics, and the receipt of on-the-job training," *Industrial and Labor Relations Review*, 45(1), 58–79.

ANDERSON, E. W., L. P. HANSEN, E. R. MCGRATTAN, AND T. J. SARGENT (1996): "Mechanics of Forming and Estimating Dynamic Linear Economies," in *Handbook of Computational Economics*, ed. by H. M. Amman, D. A. Kendrick, and J. Rust. Elsevier Science Publisher (North-Holland), vol. 1, pp. 171–252.

AUTOR, D. H., F. LEVY, AND R. J. MURNANE (2003): "The Skill Content of Recent Technological Change: An Empirical Exploration," *Quarterly Journal of Economics*, 118(4), 1279–1333.

BACOLOD, M., AND B. S. BLUM (forthcoming): "Two Sides of the Same Coin: U.S. "Residual" Inequality and the Gender Gap," *Journal of Human Resources*.

BEN-PORATH, Y. (1967): "The Production of Human Capital and the Life Cycle of Earnings," *Journal of Political Economy*, 75, 352–65.

CUNHA, F., AND J. J. HECKMAN (2008): "Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation," *Journal of Human Resources*, 43(4), 738–782.

EDIN, P.-A., AND M. GUSTAVSSON (2008): "Time Out of Work and Skill Depreciation," *Industrial and Labor Relations Review*, 61(2), 163–180.

GIBBONS, R., L. F. KATZ, T. LEMIEUX, AND D. PARENT (2005): "Comparative Advantage, Learning and Sectoral Wage Determination," *Journal of Labor Economics*, 23(4), 681–724.

GORLICH, D., AND A. DE GRIP (2009): "Human Capital Depreciation during Hometime," *Oxford Economic Papers*, 61(S1), 98–121.

HECKMAN, J. J., L. LOCHNER, AND R. COSSA (2002): "Learning-By-Doing Vs. On-the-Job Training: Using Variation Induced by the EITC to Distinguish Between Models of Skill Formation," University of Chicago.

HECKMAN, J. J., AND G. SEDLACEK (1985): "Heterogeneity, Aggregation, and Market Wage Functions: An Empirical Model of Self-Selection in the Labor Market," *Journal of Political Economy*, 93(6), 1077–1125.

HECKMAN, J. J., J. STIXRUD, AND S. URZUA (2006): "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior," *Journal of Labor Economics*, 24(3), 411–482.

INGRAM, B. F., AND G. R. NEUMANN (2006): "The Returns to Skill," *Labour Economics*, 13, 35–59.

JOVANOVIC, B. (1979): "Job Matching and the Theory of Turnover," *Journal of Political Economy*, 87(5), 972–90.

KAMBOUROV, G., AND I. MANOVSKII (2008): "Rising Occupational and Industry Mobility in the United States: 1968-1997," *International Economic Review*, 49(1), 41–79.

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

KEANE, M. P., AND K. I. WOLPIN (1997): "The Career Decisions of Young Men," *Journal of Political Economy*, 105(3), 473–522.

LAZEAR, E. P. (2004): "Firm-Specific Human Capital: A Skill-Weights Approach," Stanford University.

MILLER, A. R., D. J. TREIMAN, P. S. CAIN, AND P. A. ROOS (eds.) (1980): *Work, Jobs, and Occupations: A Critical Review of the Dictionary of Occupational Titles*. National Academy Press.

MINCER, J., AND H. OFEK (1982): "Interrupted Work Careers: Depreciation and Restoration of Human Capital," *Journal of Human Resources*, 27(1), 3–24.

MINCER, J., AND S. POLACHECK (1974): "Family Investments in Human Capital: Earnings of Women," *Journal of Political Economy*, 82(2), 76–108.

MOSCARINI, G., AND K. THOMSSON (2008): "Occupational and Job Mobility in the US," *Scandinavian Journal of Economics*, 109(4), 807–36.

MOSCARINI, G., AND F. VELLA (2003): "Occupational Mobility and Employment Reallocation: Evidence From The NLSY79," Yale University.

NEAL, D. (1995): "Industry-Specific Human Capital: Evidence from Displaced Workers," *Journal of Labor Economics*, 13, 653–77.

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

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, 306–21.

PAVAN, R. (2006): "Career Choice and Wage Growth," University of Rochester.

POLETAEV, M., AND C. ROBINSON (2008): "Human Capital Specificity: Evidence from the Dictionary of Occupational Titles and Displaced Worker Surveys 1984-2000," *Journal of Labor Economics*, 26(4), 387–420.

SCHÖNBERG, U. (2007): "Wage Growth due to Human Capital Accumulation and Job Search: A Comparison between the United States and Germany," *Industrial and Labor Relations Review*, 60, 562–586.

SULLIVAN, P. J. (2008): "Empirical Evidence on Occupation and Industry Specific Human Capital," University of Michigan.

——— (2009): "Estimation of an Occupational Choice Model when Occupations are Misclassified," *Journal of Human Resources*, 44(2), 495–535.

——— (forthcoming): "A Dynamic Analysis of Educational Attainment, Occupational Choices, and Job Search," *International Economic Review*.

TOPEL, R., AND M. WARD (1992): "Job mobility and the careers of young men," *Quarterly Journal of Economics*, 107(2), 439–79.

YAMAGUCHI, S. (2008): "Career Progression and Comparative Advantage," McMaster University.

——— (forthcoming): "The Effect of Match Quality and Specific Experience on Career Decisions and Wage Growth," *Labour Economics*.

A Tables

Table 1: Factor Loadings

Variables	Cognitive	Motor
data	0.536	
people	0.484	
GED reasoning	0.305	
GED mathematics	0.317	
GED language	0.364	
intelligence	0.213	
verbal	0.231	
numerical	0.196	
influencing people	0.041	
dealing w/ people	0.082	
making directions	0.087	
things		0.873
motor coordination		0.151
finger dexterity		0.152
manual dexterity		0.135
eye-hand-foot coordination		0.066
color descrimination		0.086
form perception		0.139
spacial perception		0.151
setting limits		0.102
strength		0.054
climbing		0.041
balancing		0.026
stooping		0.027
kneeling		0.049
crouching		0.037
crawling		0.014
reaching		0.058
handling		0.045
fingering		0.085
feeling		0.058
talking		−0.136
hearing		−0.111
tasting/smelling		0.006
near acuity		0.083
far acuity		0.005
depth perception		0.161
accomodation		0.116
color vision		0.082
field of vision		0.001

Source: The 1971 April CPS augmented with job characteristics variables from the Revised Fourth Edition of the DOT (1991).

Table 2: Task Complexity (Percentile Scores) by Occupation at 1-Digit Classification

	Cognitive Task		Motor Task		Nobs.
	Mean	Std.Dev.	Mean	Std.Dev.	
Professional	0.85	0.14	0.45	0.33	7522
Manager	0.79	0.15	0.21	0.21	5538
Sales	0.57	0.17	0.23	0.15	3748
Clerical	0.49	0.16	0.56	0.22	9270
Craftsmen	0.52	0.20	0.82	0.20	6557
Operatives	0.20	0.18	0.58	0.20	5824
Transport	0.28	0.15	0.63	0.10	1774
Laborer	0.15	0.16	0.46	0.13	2818
Farmer	0.68	0.19	0.78	0.14	1117
Farm Laborer	0.18	0.19	0.53	0.16	882
Service	0.32	0.22	0.44	0.24	6834
Household Service	0.20	0.11	0.24	0.23	1469
All Occupations	0.49	0.29	0.50	0.29	53353

Source: The 1971 April CPS augmented with job characteristics variables from the Revised Fourth Edition of the DOT (1991).

Table 3: Wage Equation

Notation	Estimates	Std. Error
p_0	-2.9176	1.0071
$p_1(1)$	5.8292	1.9961
$p_1(2)$	4.5970	1.5307
$P_2(1,1)$	0.1503	0.0219
$P_2(2,2)$	0.1753	0.0348
σ_η^2	0.0116	0.0002

Note: The wage equation is $\ln w_t = p_0 + (p_1' + P_2 x_t)' s_t + \eta_t$ where $\eta_t \sim N(0, \sigma_\eta^2)$. The first element of vectors and (1,1) element of matrices are for cognitive skills and the second element of vectors and (2,2) element of matrices are motor skills.

Table 4: Skill Transition

Notation	Estimates	Std. Error
$D(1,1)$	0.9261	0.0023
$D(2,2)$	0.9183	0.0089
$a_0(1)$	0.0223	0.0049
$a_0(2)$	0.0415	0.0117
$A_1(1,1)$	0.0066	0.0023
$A_1(2,2)$	0.0060	0.0043
$A_2(1,1)$, AFQT	0.0114	0.0135
$A_2(1,2)$, Edu	0.0013	0.0008
$A_2(1,3)$, Black	−0.0021	0.0120
$A_2(1,4)$, Hispanic	−0.0055	0.0111
$A_2(2,1)$, AFQT	−0.0099	0.0164
$A_2(2,2)$, Edu	−0.0001	0.0013
$A_2(2,3)$, Black	0.0008	0.0166
$A_2(2,4)$, Hispanic	0.0071	0.0146
$\Sigma_\varepsilon(1,1)$	0.0018	0.0012
$\Sigma_\varepsilon(2,1)$	−0.0014	0.0004
$\Sigma_\varepsilon(2,2)$	0.0017	0.0009
$\Sigma_{s1}(1,1)$	0.0078	0.0052
$\Sigma_{s1}(2,1)$	−0.0063	0.0018
$\Sigma_{s1}(2,2)$	0.0091	0.0050
$H(1,1)$, AFQT	0.1126	0.1788
$H(1,2)$, Edu	−0.0016	0.0126
$H(1,3)$, Black	0.0063	0.1628
$H(1,4)$, Hispanic	−0.0474	0.1528
$H(2,1)$, AFQT	−0.0744	0.2054
$H(2,2)$, Edu	0.0037	0.0145
$H(2,3)$, Black	−0.0173	0.2042
$H(2,4)$, Hispanic	0.0704	0.1825

Note: The skill transition equation is $s_{t+1} = Ds_t + a_0 + A_1x_t + A_2d + \varepsilon_{t+1}$. The conditional initial mean and variance of skills are given by $E(s_1|d) = h + Hd$ and $Var(s_1|d) = \Sigma_{s1}$ where d is a vector of AFQT score, education, and dummy variables for race (whites are the reference group). The unconditional mean and variance of initial skills are normalized to .50 and .01, respectively. The variance matrix of iid skill shocks is denoted by Σ_ε. The first element of vectors and (1,1) element of matrices are for cognitive skills and the second element of vectors and (2,2) element of matrices are motor skills.

Table 5: Job Preference

Notation	Estimates	Std. Error
$g_0(1)$	-3.1584	0.4626
$g_0(2)$	-0.8270	0.4279
$G_1(1,1)$, AFQT	0.2642	0.2041
$G_1(1,2)$, Edu	0.0672	0.0148
$G_1(1,3)$, Black	-0.1083	0.1906
$G_1(1,4)$, Hispanic	0.0738	0.1774
$G_1(2,1)$, AFQT	0.3031	0.8230
$G_1(2,2)$, Edu	-0.0479	0.0588
$G_1(2,3)$, Black	-0.0481	0.7715
$G_1(2,4)$, Hispanic	-0.3362	0.7206
$G_2(1,1)$	1.0056	0.3820
$G_2(2,2)$	3.6136	1.0506
$G_4(1,1)$	-11.5363	0.5403
$G_4(2,2)$	-0.4015	0.0612
$A_3(1,1)$	0.3188	0.0066
$A_3(2,2)$	0.0009	0.0305
$\bar{x}_{1,0}(1)$	-0.2144	0.0225
$X(1,1)$, AFQT	0.1327	0.0180
$X(1,2)$, Edu	0.0399	0.0020
$X(1,3)$, Black	-0.0248	0.0144
$X(1,4)$, Hispanic	0.0519	0.0156
$\bar{x}_{1,0}(2)$	0.4954	0.0826
$X(2,1)$, AFQT	0.0892	0.0650
$X(2,2)$, Edu	-0.0049	0.0075
$X(2,3)$, Black	0.1024	0.0510
$X(2,4)$, Hispanic	0.0938	0.0537

Note: The utility from job characteristics is given by $v(x_t, \bar{x}_t, s_t, \tilde{v}_t) = (g_0 + G_1 d + G_2 s_t + \tilde{v}_t)' x_t + x_t' G_3 x_t + (x_t - \bar{x}_t)' G_4 (x_t - \bar{x}_t)$. Parameter G_3 is normalized as the negative of an identity matrix. The transition equation of work habit is $\bar{x}_{t+1} = A_3 \bar{x}_t + (I - A_3) x_t$ where I is a (2×2) identity matrix. The initial work habit is given by $\bar{x}_1 = \bar{x}_{1,0} + Xd$ where d is a vector of AFQT score, education, and dummy variables for race (whites are the reference group.) The first element of vectors and (1,1) element of matrices are for cognitive skills and the second element of vectors and (2,2) element of matrices are motor skills.

Table 6: Policy Function

Notation	Estimates	Std. Error
$c_{0,t=1}(1)$	-0.0655	0.0121
$c_{0,t=1}(2)$	-0.1103	0.1020
$C_{1,t=1}(1,1)$, AFQT	0.0373	0.0201
$C_{1,t=1}(1,2)$, Edu	0.0088	0.0015
$C_{1,t=1}(1,3)$, Black	-0.0142	0.0188
$C_{1,t=1}(1,4)$, Hispanic	0.0069	0.0176
$C_{1,t=1}(2,1)$, AFQT	0.1156	0.3244
$C_{1,t=1}(2,2)$, Edu	-0.0202	0.0232
$C_{1,t=1}(2,3)$, Black	-0.0192	0.3043
$C_{1,t=1}(3,4)$, Hispanic	-0.1327	0.2841
$C_{2,t=1}(1,1)$	0.1132	0.0374
$C_{2,t=1}(2,2)$	1.4943	0.4112
$C_{3,t=1}(1,1)$	0.7628	0.0057
$C_{3,t=1}(2,2)$	0.2383	0.0199
$\Sigma_\nu(1,1)$	0.0258	0.0001
$\Sigma_\nu(2,1)$	-0.0006	0.0001
$\Sigma_\nu(2,2)$	0.0178	0.0004

Note: The policy function is $x_t^* = c_{0,t} + C_{1,t}d + C_{2,t}s_t + C_{3,t}\bar{x}_t + \nu_t$ where $\nu_t \sim N(0, \Sigma_\nu)$ The first element of vectors and (1,1) element of matrices are for cognitive skills and the second element of vectors and (2,2) element of matrices are motor skills.

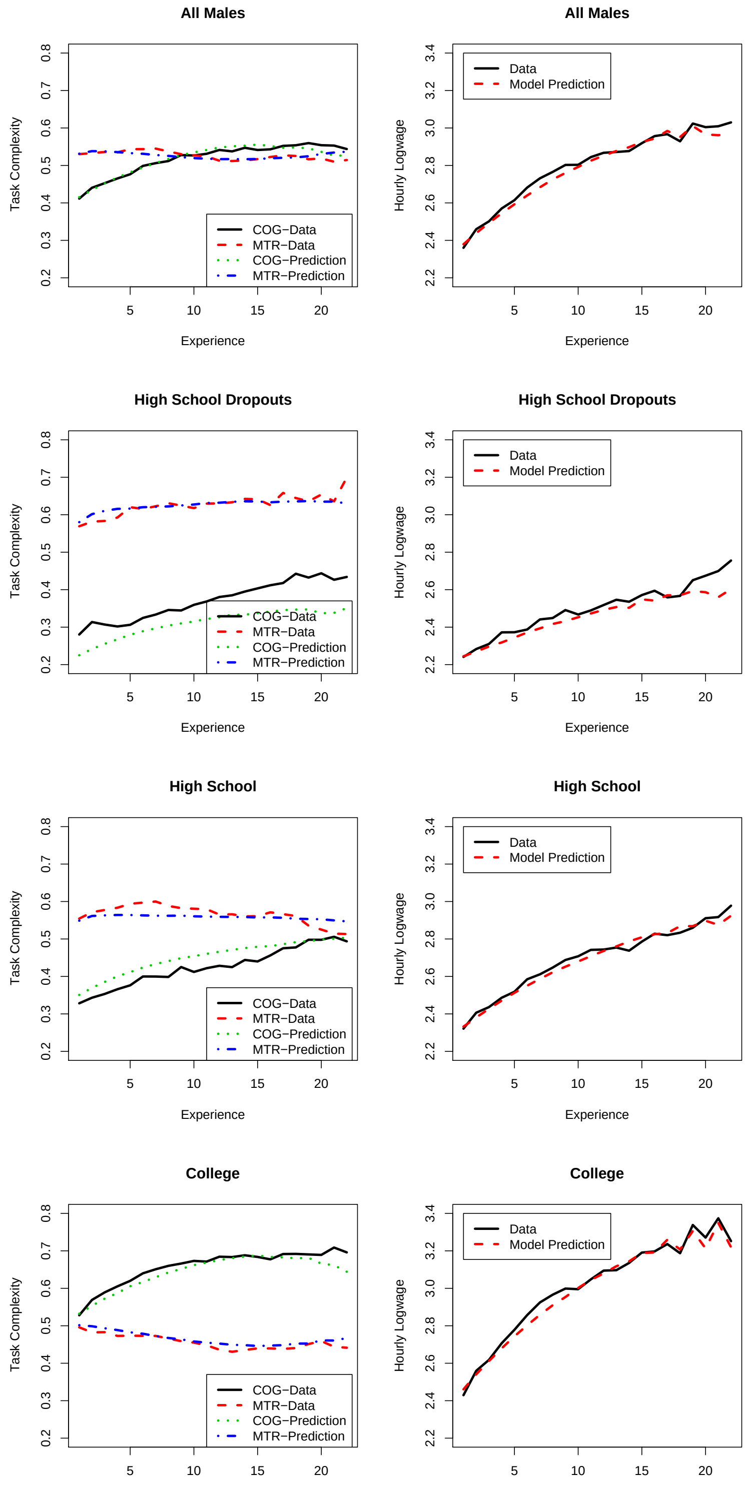
Table 7: Logwage Variance Decomposition

	Cognitive	Motor	Covariance	Error	Total
Year 1					
All	0.318	0.226	−0.401	0.012	0.154
Dropouts	0.293	0.234	−0.398	0.012	0.140
High School	0.303	0.224	−0.392	0.012	0.147
College	0.310	0.224	−0.397	0.012	0.150
Year 10					
All	0.569	0.299	−0.644	0.012	0.236
Dropouts	0.456	0.297	−0.581	0.012	0.184
High School	0.459	0.289	−0.577	0.012	0.182
College	0.506	0.290	−0.598	0.012	0.210
Year 20					
All	0.726	0.332	−0.773	0.012	0.297
Dropouts	0.521	0.318	−0.647	0.012	0.203
High School	0.519	0.313	−0.647	0.012	0.196
College	0.607	0.314	−0.687	0.012	0.246

Table 8: Task Complexity Variance Decomposition

	Cognitive			Motor		
	Model	Error	Total	Model	Error	Total
Year 1						
All	0.018	0.026	0.044	0.022	0.018	0.039
Dropouts	0.003	0.026	0.029	0.021	0.018	0.039
High School	0.002	0.026	0.028	0.020	0.018	0.038
College	0.010	0.026	0.035	0.021	0.018	0.039
Year 10						
All	0.048	0.026	0.074	0.048	0.018	0.066
Dropouts	0.028	0.026	0.054	0.045	0.018	0.063
High School	0.028	0.026	0.054	0.043	0.018	0.061
College	0.037	0.026	0.063	0.046	0.018	0.064
Year 20						
All	0.055	0.026	0.080	0.054	0.018	0.072
Dropouts	0.031	0.026	0.057	0.047	0.018	0.065
High School	0.031	0.026	0.057	0.046	0.018	0.064
College	0.041	0.026	0.067	0.050	0.018	0.068

B Figures



Note: COG is for profiles of cognitive task complexity and MTR is for profiles of motor task complexity.

Figure 1: Model Fit (Task Complexity Profile)

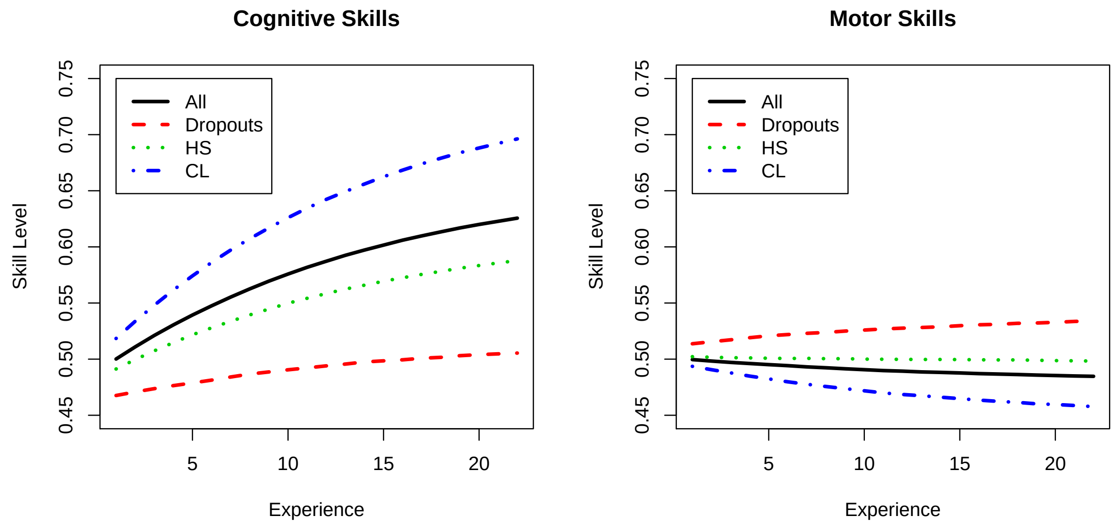


Figure 2: Skill Growth Profiles

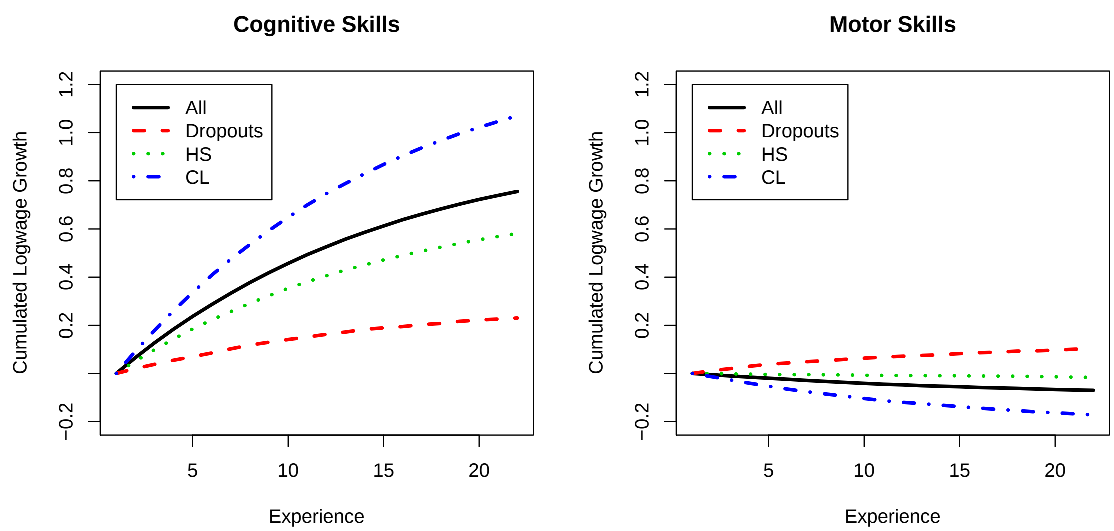


Figure 3: Contribution of Skills to Wage Growth

C Model Solution

The goal of this section is to prove that the optimal policy function is a linear function of time-invariant demographic variables d, skills s_t, work habit (the weighted average of the task complexity of the past jobs) $\bar{x}_t$, and preference shock ν_t

$$x_t^* \quad = \quad c_{0,t} + C_{1,t} d + C_{2,t} s_t + C_{3,t} \bar{x}_t + \nu_t. \tag{37}$$

To simplify the problem, rewrite the original Bellman equation (9) in the following form

$$
\begin{aligned}
V_t(z_t, d, \tilde{\nu}_t) &= \max_{x_t} r_0 + r_1^{'} x_t + r_2^{'} z_t + x_t^{'} R_3 x_t + 2x_t^{'} R_4 z_t + z_t^{'} R_5 z_t + \beta E V_{t+1}(z_{t+1}, d, \tilde{\nu}_{t+1}) && (38) \\
\text{s.t.} & \\
z_{t+1} &= l_0 + L_1 z_t + L_2 x_t + L_3 d + \xi_{t+1} && (39) \\
V_{T+1} &= 0 && (40)
\end{aligned}
$$

where

$$
\begin{aligned}
z_t^{'} &= (s_t^{'}\ \bar{x}_t^{'}) && (41) \\
\xi_t^{'} &= (\varepsilon_t^{'}\ 0^{'}) && (42) \\
r_0 &= p_0 + \eta_t && (43) \\
r_1 &= g_0 + G_1 d + \tilde{\nu}_t && (44) \\
r_2^{'} &= (p_1^{'}\ 0^{'}) && (45) \\
R_3 &= (G_3 + G_4) && (46) \\
R_4 &= \begin{pmatrix} 0.5 \cdot (P_2 + G_2) & -G_4 \end{pmatrix} && (47) \\
R_5 &= \begin{pmatrix} 0 & 0 \\ 0 & G_4 \end{pmatrix} && (48) \\
l_0^{'} &= (a_0^{'}\ 0^{'}) && (49) \\
L_1 &= \begin{pmatrix} D & 0 \\ 0 & A_3 \end{pmatrix} && (50) \\
L_2 &= \begin{pmatrix} A_1 \\ I - A_3 \end{pmatrix} && (51) \\
L_3 &= \begin{pmatrix} A_2 \\ 0 \end{pmatrix}. && (52)
\end{aligned}
$$

First, I show that expectation of the value function for period $t = 2$ to $T+1$ can be written as a quadratic function of state variables. This claim is true for $t = T+1$ because $EV_{T+1} = 0$ by the assumption that the problem has a finite horizon. I am going to show that this claim is also true in period t when it is true in period $t+1$. Specifically, assume that the expected value of the value function in period $t+1$ can be written as a quadratic function of state variables,

$$
\begin{aligned}
EV_{t+1} &= q_{0,t+1} + (q_{1,t+1} + Q_{2,t+1} d)^{'} z_{t+1} + z_{t+1}^{'} Q_{3,t+1} z_{t+1} && (53) \\
&\equiv q_{0,t+1} + q_{A,t+1}^{'} z_{t+1} + z_{t+1}^{'} Q_{3,t+1} z_{t+1}. && (54)
\end{aligned}
$$

Substituting the transition equation (3) for the next period state variables, the value function in period t can be written as

$$\begin{aligned}
V_t(z_t, \varepsilon_t, \tilde{v}_t) &= \max_{x_t} r_0 + r_1^{'} x_t + r_2^{'} z_t + x_t^{'} R_3 x_t + 2x_t^{'} R_4 z_t + z_t^{'} R_5 z_t + \\
&\quad \beta [q_{0,t+1} + q_{A,t+1}^{'} (l_0 + L_1 z_t + L_2 x_t + L_3 d) + \\
&\quad (l_0 + L_1 z_t + L_2 x_t + L_3 d)^{'} Q_{3,t+1} (l_0 + L_1 z_t + L_2 x_t + L_3 d)]. \qquad (55)
\end{aligned}$$

The first order condition for optimality is characterized by

$$0 = r_1 + 2R_3 x_t^* + 2R_4 z_t + \beta [L_2^{'} q_{A,t+1} + 2L_2^{'} Q_{3,t+1}^{'} (l_0 + L_1 z_t + L_2 x_t^* + L_3 d)]. \qquad (56)$$

Solving this equation for x_t^* to find

$$\begin{aligned}
x_t^* &= -\frac{1}{2} (R_3 + \beta L_2^{'} Q_{3,t+1} L_2)^{-1} [(g_0 + G_1 d + \tilde{v}_t) + \\
&\quad \beta L_2^{'} \{(q_{1,t+1} + Q_{2,t+1} d) + 2Q_{3,t+1}^{'} (l_0 + L_3 d)\} + \{2R_4 + 2\beta L_2^{'} Q_{3,t+1}^{'} L_1\} z_t] \qquad (57) \\
&= -\frac{1}{2} (R_3 + \beta L_2^{'} Q_{3,t+1} L_2)^{-1} [\{g_0 + \beta L_2^{'} (q_{1,t+1} + 2Q_{3,t+1}^{'} l_0\} + \\
&\quad \{G_1 + \beta L_2^{'} (Q_{2,t+1} + 2Q_{3,t+1}^{'} L_3)\} d + \{2R_4 + 2\beta L_2^{'} Q_{3,t+1}^{'} L_1\} z_t + \tilde{v}_t] \qquad (58) \\
&\equiv b_{0,t} + B_{1,t} d + B_{2,t} z_t + v_t \qquad (59)
\end{aligned}$$

where I substitute $g_0 + G_1 d + \tilde{v}_t$ for r_1 (see Equation 44) and

$$\begin{aligned}
b_{0,t} &= -\frac{1}{2} (R_3 + \beta L_2^{'} Q_{3,t+1} L_2)^{-1} [g_0 + \beta L_2^{'} (q_{1,t+1} + 2Q_{3,t+1}^{'} l_0)] \qquad (60) \\
B_{1,t} &= -\frac{1}{2} (R_3 + \beta L_2^{'} Q_{3,t+1} L_2)^{-1} [G_1 + \beta L_2^{'} (Q_{2,t+1} + 2Q_{3,t+1}^{'} L_3)] \qquad (61) \\
B_{2,t} &= -\frac{1}{2} (R_3 + \beta L_2^{'} Q_{3,t+1} L_2)^{-1} 2[R_4 + \beta L_2^{'} Q_{3,t+1}^{'} L_1] \qquad (62) \\
v_t &= -\frac{1}{2} (R_3 + \beta L_2^{'} Q_{3,t+1} L_2)^{-1} \tilde{v}_t. \qquad (63)
\end{aligned}$$

I will show that the expected value function in period t is also a quadratic function of state variables like Equation (53). To do so, I first write the expected value function in period $t+1$ in terms of state variables in period t using the transition equation (39). To simplify notation, define

$$b_{A,t} \equiv b_{0,t} + B_{1,t} d + v_t. \qquad (64)$$

The expected value function in period $t+1$ is

$$
\begin{aligned}
EV_{t+1} &= q_{0,t+1} + q'_{A,t+1} z_{t+1} + z'_{t+1} Q_{3,t+1} z_{t+1} && (65)\\
&= q_{0,t+1} + q'_{A,t+1}(l_0 + L_1 z_t + L_2(b_{A,t} + B_{2,t} z_t) + L_3 d) + \\
&\quad (l_0 + L_1 z_t + L_2(b_{A,t} + B_{2,t} z_t) + L_3 d)' Q_{3,t+1}(l_0 + L_1 z_t + L_2(b_{A,t} + B_{2,t} z_t) + L_3 d) && (66)\\
&= \{q_{0,t+1} + q'_{A,t+1}(l_0 + L_2 b_{A,t} + L_3 d) + (l_0 + L_2 b_{A,t} + L_3 d)' Q_{3,t+1}(l_0 + L_2 b_{A,t} + L_3 d)\} + \\
&\quad \{(L_1 + L_2 B_{2,t})'(q_{A,t+1} + 2Q_{3,t+1}(l_0 + L_2 b_{A,t} + L_3 d))\}' z_t + \\
&\quad z'_t (L_1 + L_2 B_{2,t})' Q_{3,t+1}(L_1 + L_2 B_{2,t}) z_t. && (67)
\end{aligned}
$$

Next I write the current utility in terms of the current state variables,

$$
\begin{aligned}
V_t - \beta EV_{t+1} &= r_0 + r'_1 x^*_t + r'_2 z_t + x^{*'}_t R_3 x^*_t + 2x^{*'}_t R_4 z_t + z'_t R_5 z_t && (68)\\
&= r_0 + r'_1(b_{A,t} + B_{2,t} z_t) + r'_2 z_t + (b_{A,t} + B_{2,t} z_t)' R_3 (b_{A,t} + B_{2,t} z_t) + \\
&\quad 2(b_{A,t} + B_{2,t} z_t)' R_4 z_t + z'_t R_5 z_t && (69)\\
&= \{r_0 + r'_1 b_{A,t} + b'_{A,t} R_3 b_{A,t}\} + \{B'_{2,t} r_1 + r_2 + (2R_3 B_{2,t} + 2R_4)' b_{A,t}\}' z_t + \\
&\quad z'_t \{B'_{2,t} R_3 B_{2,t} + B'_{2,t} R_4 + R'_4 B_{2,t} + R_5\} z_t. && (70)
\end{aligned}
$$

So, the expectation of the value function in period t is

$$
\begin{aligned}
EV_t &= E[\{r_0 + r'_1 b_{A,t} + b'_{A,t} R_3 b_{A,t}\} + \\
&\quad \beta\{q_{0,t+1} + q'_{A,t+1}(l_0 + L_2 b_{A,t} + L_3 d) + (l_0 + L_2 b_{A,t} + L_3 d)' Q_{3,t+1}(l_0 + L_2 b_{A,t} + L_3 d)\}] + \\
&\quad E[\{B'_{2,t} r_1 + r_2 + 2(R_3 B_{2,t} + R_4)' b_{A,t}\} + \\
&\quad \beta\{(L_1 + L_2 B_{2,t})'(q_{A,t+1} + 2Q_{3,t+1}(l_0 + L_2 b_{A,t} + L_3 d))\}]' z_t + \\
&\quad E z'_t [\{B'_{2,t} R_3 B_{2,t} + B'_{2,t} R_4 + R'_4 B_{2,t} + R_5\} + \beta\{(L_1 + L_2 B_{2,t})' Q_{3,t+1}(L_1 + L_2 B_{2,t})\}] z_t && (71)\\
&\equiv q_{0,t} + q'_{A,t} z_t + z'_t Q_{3,t} z_t. && (72)
\end{aligned}
$$

Notice that $B'_{2,t} R_4 + R'_4 B_{2,t}$ is symmetric and so $Q_{3,t}$ is. Remember that $b_{A,t}$ and r_1 are random variables and that $Eb_{A,t} = b_{0,t} + B_{1,t} d$ and $Er_1 = g_0 + G_1 d$. Given these facts, $q_{A,t}$ can be written as

$$
\begin{aligned}
q_{A,t} &= B'_{2,t}(g_0 + G_1 d) + r_2 + 2(R_3 B_{2,t} + R_4)' Eb_{A,t} + \\
&\quad \beta\{(L_1 + L_2 B_{2,t})'(q_{A,t+1} + 2Q_{3,t+1}(l_0 + L_2 Eb_{A,t} + L_3 d))\} && (73)\\
&= B'_{2,t}(g_0 + G_1 d) + r_2 + 2\beta(L_1 + L_2 B_{2,t})' Q_{3,t+1}(l_0 + L_3 d) + \\
&\quad \beta(L_1 + L_2 B_{2,t})'(q_{1,t+1} + Q_{2,t+1} d) +
\end{aligned}
$$

$$2\{(R_3B_{2,t}+R_4)^{'}+\beta(L_1+L_2B_{2,t})^{'}Q_{3,t+1}L_2\}(b_{0,t}+B_{1,t}d). \tag{74}$$

To simplify notation, define variables $F_{1,t}$ and $F_{2,t}$ such that

$$\begin{aligned} F_{1,t} &= L_1+L_2B_{2,t} && (75)\\ F_{2,t} &= (R_3B_{2,t}+R_4)^{'}+\beta F_{1,t}^{'}Q_{3,t+1}L_2. && (76)\end{aligned}$$

Then the Equation (74) can be written as

$$\begin{aligned} q_{A,t} &= B_{2,t}^{'}(g_0+G_1d)+r_2+2\beta F_{1,t}^{'}Q_{3,t+1}(l_0+L_3d)+\\ &\quad \beta F_{1,t}^{'}(q_{1,t+1}+Q_{2,t+1}d)+2F_{2,t}(b_{0,t}+B_{1,t}d) && (77)\\ &= [r_2+\beta F_{1,t}^{'}(2Q_{3,t+1}l_0+q_{1,t+1})+2F_{2,t}b_{0,t}]+\\ &\quad [B_{2,t}^{'}G_1+\beta F_{1,t}^{'}(Q_{2,t+1}+2Q_{3,t+1}L_3)+2F_{2,t}B_{1,t}]d && (78)\\ &\equiv q_{1,t}+Q_{2,t}d && (79)\end{aligned}$$

where

$$\begin{aligned} q_{1,t} &= r_2+\beta F_{1,t}^{'}(2Q_{3,t+1}l_0+q_{1,t+1})+2F_{2,t}b_{0,t} && (80)\\ Q_{2,t} &= B_{2,t}^{'}G_1+\beta F_{1,t}^{'}(Q_{2,t+1}+2Q_{3,t+1}L_3)+2F_{2,t}B_{1,t}. && (81)\end{aligned}$$

This shows that the expected value function in period t has the same form as Equation (53). Thus, I prove the claim that expectation of the value function for period $t=2$ to $T+1$ can be written as a quadratic function of state variables. I have also shown by Equation (59) that the optimal policy function in period t is a linear function of state variables, worker characteristics, and preference shocks, when the expected value function is given by Equation (53). Because z_t is a vector of skills and work habit (see Equation 41) and I can write

$$B_{2,t} = \begin{pmatrix} C_{2,t} & 0 \\ 0 & C_{3,t} \end{pmatrix}. \tag{82}$$

Equation (37) indeed holds for $t=1$ to T.