Formation of Heterogeneous Skills and Wage Growth

Shintaro Yamaguchi*

April 2009

Abstract

This paper examines how primitive skills associated with occupations are formed and rewarded in the labor market over the careers of young male high school graduates. The objective task complexity measurement from the Dictionary of Occupational Titles enables a more direct look into primitive skills of workers. Through characterizing a worker's optimal occupational choice, I show that the task complexity of a worker's occupation can be interpreted as a noisy signal of his unobserved 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 raise wages by 10% during the first three years. However, motor skills contribute little to subsequent wage growth. In contrast, cognitive skill growth continues to drive wage growth and raises wages by 30% during the first 10 years.

1 Introduction

The recent empirical evidence[1] suggests that human capital accumulation is the main source of wage growth for young individuals. However , human capital is most often assumed homogeneous for simplicity and the content of human capital is not well understood. Using objective task complexity measures from the Dictionary of Occupational Titles, this paper takes a closer look at the human capital accumulation process in order to understand skill formation and wage growth. To uncover the growth of unobserved cognitive and motor skills, I show that, by characterizing the optimal occupational choice of an individual, the task complexity of an individual's occupation can be interpreted as a noisy signal of his unobserved skills. Understanding the growth of these

*Address: Department of Economics, McMaster University, 1280 Main St. West, Hamilton, ON., Canada L8S 4M4, URL: `http://socserv.mcmaster.ca/yamtaro`, Email: `yamtaro@mcmaster.ca`. The author acknowledges the use of SHARCNET computational facilities.

[1]Topel and Ward (1992) is the best known example. Schönberg (2007), Barlevy (2008), Liu (2009) also find that human capital contributes more to wage growth than job search.

two different skills and their contributions to wages is useful for designing training programs and other labor market policies.

In the model, individuals are characterized in a skill space and occupations are characterized in a task complexity space. 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 job. Similarly, occupations 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 occupations. This skill and task complexity space approach allows for a clearer interpretation of the data by directly analyzing primitive skills rather than by proxying them with occupational specific experience. 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.

Heterogeneous occupations 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 an occupation with a complex cognitive task. Second, 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 occupation will have accumulated a large amount of motor skills. Third, individuals suffer more disutility when their tasks are complex. This is the cost of entering a skill demanding occupation where skills are better rewarded and developed. Entering an occupation with complex tasks is not beneficial for low skill individuals due to the high disutility. Heterogeneous individuals sort themselves into different occupations according to their skill endowment.

The assumption of a quadratic instantaneous reward function and a linear skill growth equation allows the analytical solution of the optimal policy function for occupational choice to be expressed as a linear function of skills, demographic variables, and preference shocks. This policy function implies that the task complexity of the optimal occupation can be interpreted as a noisy signal of skills. Using a sample of male high school graduates' wage and occupational choice histories from the NLSY97, the growth of unobserved cognitive and motor skills as well as the structural parameters are estimated using the Kalman filter.

The estimation results indicate that cognitive and motor skills have considerably different growth paths. During the first three years, cognitive and motor skills contribute equally to wage growth; each raising wages by 10%. However, the growth of motor skills quickly slows down and contributes little to later wage growth. In contrast, cognitive skills steadily increase (but at a decreasing pace) and drive the subsequent wage growth. During the first 10 years, cognitive skills raise wages by 30%, while the contribution of motor skills remains at 10%. I also find that the depreciation rate of motor skills is very high, implying that training programs which enhance motor skills are unlikely to have a significant long-term effect.

2 Related Literature

The extensive empirical literature on human capital formation recognizes that skills are intrinsically heterogeneous. In their pioneering work, Keane and Wolpin (1997) construct and estimate a structural model of human capital formation in which heterogeneous skills are represented by years of experience in blue-collar, white-collar, and military jobs. Neal (1995) and Parent (2000) find evidence for industry specific human capital. Kambourov and Manovskii (2009), Pavan (2006) and Yamaguchi (2007) estimate returns to occupational specific experience. All of these papers find substantial returns to sectoral experience, which implies that skills are heterogeneous across sectors.

This paper departs from these previous contributions by taking a more direct look at the primitive skills of workers, rather than using occupational specific experience as a proxy for skills. This approach allows for a clearer interpretation of the estimation results. For example, Keane and Wolpin (1997) find one additional year of experience in an occupation raises an individual's wages, but it is not quite clear what the underlying abilities that make an individual earn a higher wage in that occupation are. Moreover, the results indicate that experience in one occupation also positively contributes to wages in other occupations, but, again, the exact reason for this is not well understood. This paper identifies the growth of primitive skills and shows how skills learned in one occupation are also rewarded in other occupations.

Another advantage of the present model is that it can account for the heterogeneity of occupations in a richer way. Keane and Wolpin (1997) include only three occupations in their model because of computational tractability. Other papers mentioned above can account for many more occupations, but for tractability reasons they impose that returns to occupation specific experience are the same across all occupations and that occupational skills are completely nontransferable. These limitations are quite restrictive because some occupations may provide more skill learning opportunities and individuals may transfer substantial occupational skills to occupations similar to the current one. This paper overcomes these limitations by characterizing all occupations in terms of a low dimension task complexity vector. This skill and task complexity space approach enables a more extensive analysis of skills and occupations than that found in previous papers.

This paper is also related to a research program that investigates primitive skills. Heckman, Stixrud, and Urzua (2006) find evidence that not only cognitive, but also noncognitive skills, strongly influence labor market outcomes and social behavior. Cunha, Heckman, Lochner, and Masterov (2006) and Cunha and Heckman (2007) utilize models of human capital production in which individuals form heterogeneous skills over the life cycle. These papers focus primarily on skill formation before individuals enter the labor market, and assume that heterogeneous skills are equally rewarded across different jobs. This paper complements these contributions because it analyzes how skills are rewarded and formed in the labor market.

Some recent papers attempt to look into the skills associated with jobs. Ingram and Neumann (2006) and Bacolod and Blum (2008) construct skill measures by conducting a principal component analysis using the DOT. They include these skill indices as measures of worker skills in their OLS regressions. These reduced-form analyses provide suggestive evidence about the relationship between primitive skills and wages. However, they do not distinguish between the skills that are required for a job and the skills that individuals actually have. In general, they are different. This shortcoming makes it difficult to interpret the estimated coefficients as returns to skills. Indeed, Poletaev and Robinson (2008) find evidence in favor of a distinction between the two. Poletaev and Robinson (2008) use the Displaced Worker Surveys, as well as the DOT, to examine wage changes following displacement. They find a greater wage loss for those who have moved to jobs that require a different skill portfolio than their pre-displacement jobs. Using German data, Gathmann and Schönberg (2007) find a similar pattern. This paper builds on these previous contributions by distinguishing between skills and task complexity to illustrate how skills are rewarded. This paper uses a structural estimation approach that allows for this important distinction to be made and acts as a complement to earlier reduced form based research.

3 Data

3.1 Dictionary of Occupational Titles

I draw measures of occupational task complexity from the Dictionary of Occupational Titles and wage and occupational choice histories from the National Longitudinal Survey of Youth (NLSY) 1979. Among the many editions of the DOT, I use the 1994 revised fourth edition, of which data are collected between 1978 and 1990. This edition is used because the survey period largely matches the sample period of the NLSY. The main purpose of the DOT is to provide standardized occupational information for an employment service matching workers and jobs. Expert occupational analysts defined 12,099 occupations with respect to 44 characteristics using information obtained through their on-site observations of jobs and information provided by professional associations. The occupational definitions describe necessary or desirable worker characteristics as well as occupational tasks, which can be broadly grouped into seven categories: worker functions; required General Educational Development; aptitudes; temperaments; interests; physical demands; and work-environment conditions.

To merge the task complexity measures of the DOT with the occupations in the NLSY, I use the task complexity measures of 12,099 occupations present in the DOT to construct task complexity measures for about 500 occupations contained in the 1970 Census 3-digit classification system. I map the DOT occupation codes onto the 1970 Census occupation codes using the April 1971 Current Population Survey augmented by the fourth edition of the DOT and compiled by the Committee on Occupational Classification and Analysis at the National Academy of Sciences. Note that this augmented CPS file contains occupation codes for the fourth edition of the DOT, not the revised fourth edition. In the revised fourth edition, some occupations are deleted or integrated into other occupations, while others are newly added. I update the task complexity measures and the DOT occupation code in the augmented CPS file using the conversion table in the revised fourth edition. I calculate the task complexity measures of each 1970 Census occupation by averaging over individuals in that occupation in the augmented CPS file.

This paper considers two broadly defined tasks: cognitive tasks and motor tasks, which are similar to the task groups analyzed by Ingram and Neumann (2006) and Bacolod and Blum (2008). Autor, Levy, and Murnane (2003) consider different task groups including routine manual task and non-routine manual task to understand the role of technological change in the labor market. But, these tasks capture different aspects of the motor tasks to perform a job. For simplicity, I use the broadly defined task categories, which allows me to focus on the analysis of dynamics of unobserved skills. Ingram and Neumann (2006) and Bacolod and Blum (2008) find evidence that these tasks are particularly strongly correlated with wages.

One possibly important task category that is missing in this paper is inter-personal task. Borghans, ter Weel, and Weinberg (2006) argue that people skills are important determinants of labor market outcomes. But, they do not show direct empirical evidence about how these tasks and skills affect wages. Bacolod and Blum (2008) include an inter-personal task complexity index in their wage regression and find it insignificant. I construct an inter-personal task complexity index in the way similar to Bacolod and Blum (2008) and find it is highly correlated with the cognitive task complexity index (the correlation coefficient is 0.8), which implies two things. First, this high correlation prevents me from estimating inter-personal skills precisely. Second, if there is any independent effect of inter-personal skills on wages, they are mostly absorbed in cognitive skills.

The DOT includes 7 variables that measure cognitive task complexity: worker functions related to data, reasoning development, mathematical development, language development, verbal aptitude, numerical aptitude, and a temperament for making judgments and decisions. Another set of 9 DOT variables measures motor task complexity: worker functions related to things, spatial aptitude, form perception, motor coordination, finger dexterity, manual dexterity, eye-hand-foot coordination, color discrimination, and a temperament for attaining precise set limits, tolerances, and standards.

In order to allow for a simple interpretation and to reduce the number of parameters, I summarize task complexity by using one variable for each task category, although it is theoretically possible to extract occupational information completely by including all of these 16 variables in the model. Following previous papers that use the DOT, such as Autor, Levy, and Murnane (2003), Ingram and Neumann (2006), and Bacolod and Blum (2008), I summarize the DOT variables in each task category by the first principal component. The first column of Table 1 shows that the first principal component of cognitive task complexity is positively correlated with all seven measures and explains 85% of the variation. Similarly, the second column shows that the first principal component of motor task complexity is positively correlated with all nine measures and explains 74% of the variation. These results imply that both cognitive and motor task complexity measures can be summarized by a single variable without losing much information. Following Autor, Levy, and Murnane (2003), I further convert these first principal components into percentile scores using the weights from the augmented CPS file.

Of course, this variable construction method is not the only possible way to construct a single summary variable. However, I find that the estimation results are robust to alternative ways to construct the indexes. I examine other DOT variables to construct the summary task complexity indexes. The constructed indexes are found to be robust to addition and deletion of the DOT variables, because relevant variables are highly correlated with each other in a given task category. I also examine if converting the first principal components into percentile scores affects the results. To do so, I estimate the model using the first principal component without converting into percentile scores. I find that the main results of the paper are also robust in this respect (see Appendix D for details.)

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.

3.2 National Longitudinal Survey of Youth 1979

3.2.1 Sampling Criteria

The NLSY is particularly suitable for this study because it contains detailed individual career histories and focuses on young individuals who change occupations more frequently than older individuals. I take a sample of male high school graduates because this is a relatively homogeneous demographic group and their labor force attachment is strong. More importantly, high school graduates experience career progression through occupational changes, while college graduates do so within the same occupation. Because unobserved skill growth is identified by the observed occupational choice of individuals, the high occupational mobility of high school graduates is suitable for the model. Yamaguchi (2007, 2008) provides evidence for these types of career patterns across education groups.

I concentrate on high school graduates who make a long-term transition to the full-time labor market during the period between 1979 and 1994. Observations after 1994 are not in the sample, because in surveys later than 1994 occupational changes are not reported on an annual basis. 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, 1,289 individuals graduated from high school and did not pursue any post-secondary education by 1994. I then dropped 158 individuals who served actively in the armed forces during the sample period. Out of 1,131 individuals, 81 individuals did not make a long-term transition to the full-time labor market. I also excluded 214 individuals who made the long-term transition at age 17 or younger, or at age 23 or older, either because they are likely to be mismeasured or because their full-time labor force attachment is weak. Finally, out of the remaining 836 individuals, I dropped 41 whose AFQT scores are missing. Hourly wages are deflated by the 2002 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. The final sample contains the career histories of 795 high school graduates, and contains 8,971 person-year observations of occupational choices and 8,695 person-year observations of wages. I change these sampling criteria and estimate the model in order to check the robustness of the parameter estimates. The results indicate that the sampling criteria are not crucial for the main results of the paper (see Appendix D for a detailed discussion.)

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 (2007) 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. 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 1,896 cases out of 8,504 apparent occupation changes and reduces the annual occupation change rate at the 3-digit level from 75% to 58%. This rate is still high, but consistent with the rate reported by Moscarini and Vella (2003) who use the NLSY.[2]

Table 3 reports summary statistics. The mean hourly wage and the mean cognitive task complexity in the 10th year following the labor market transition are not significantly different between whites and hispanics, while those of blacks are significantly lower than the other two. The mean motor task complexity in the 10th year is highest for whites, second highest for blacks, and that of hispanics is the lowest. The percentile AFQT scores vary substantially across race. Whites record the highest mean score, which is followed by hispanics, and blacks record the lowest. The average number of observations per individual is more than 10 for all races. This relatively large number of observations across the time dimension helps identify the persistent component of unobserved skills.

3.3 Career Progression Patterns

The time profiles of the average task complexity of occupations of male high school graduates are presented in Figure 3. At the point of long-term transition to the labor market, the average cognitive task complexity index is 0.32. This is comparable to the average cognitive task complexity of service workers. Individuals take on more and more cognitive skill demanding tasks over time; the cognitive task complexity index reaches 0.44 in 10 years and 0.47 in 15 years. I find that the upward trend in cognitive task complexity is statistically significant. This career progression pattern is consistent with the fact that high school graduates are gradually promoted to managers and craftsmen. The increase in cognitive task complexity reflects the increasing proportions of these two occupations in the population.

Unlike cognitive task complexity, the growth of motor task complexity is not monotonic. At the point of long-term labor market transition, the average motor task complexity index is 0.53, which is close to the motor task complexity for operatives. The index peaks around 0.58 in 5 years; then it gradually decreases to about 0.50 in 15 years, which is below the initial motor

[2]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.

task complexity index. The following occupational mobility patterns explain this hump-shaped motor task complexity profile. During the first 5 years, the share of craftsmen rises from 21% to 30%, while the share of operatives and laborer decreases from 37% to 26%. This shift to craft occupations raises the average motor task complexity. However, after the peak, many craftsmen, particularly foremen, move into managerial occupations. During the next 10 years, the share of craftsmen decreases from 30% to 25%, while the share of managers increases from 8% to 25%. The flow into managerial occupations is not only from white-collar jobs such as sales and clerical occupations, but also from blue-collar jobs including craftsmen. I find a substantial flow of foremen into managerial occupations, which is fairly intuitive, because one of their primary tasks is to supervise other crews. The decline in motor task complexity captures this switch from motor skill intensive tasks to cognitive skill intensive tasks.

I examine whether or not year effects, instead of the effects of age or experience, drive the task complexity profiles. Given the widely reported rising (cognitive) skill price, the upward trend in cognitive task complexity is of particular concern. With a sample of male high school graduates taken from the 1968-2002 CPS, I construct age profiles of cognitive and motor task complexity using cohorts defined by birth year. The year effect and the age (or experience) effect can be distinguished in this sample because it includes several cohorts. Figure 1 shows that an upward trend in cognitive task complexity is observed in all cohorts. Figure 2 shows the profiles of motor task complexity by cohort. Except for the youngest cohort that born between 1970 and 1979, hump-shaped profiles are observed. These two figures provide evidence in favor of that year effects do not drive the above results.

4 Model

In this section I describe a model of skill formation and occupational choice. In the model, each individual who made a long term transition to the full-time labor market has a finite decision horizon beginning at year 1 and ending in year 45. 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.

4.1 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 task complexity of the job in year t is x_t. Wages depend on skill quantity and its returns;

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

where p_0 is a constant term. 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 satisfactory. 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 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

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

where p_1 is a K-dimensional vector and P_2 is a $K \times K$ diagonal matrix of skill price parameters.

4.2 Skill Formation

An individual develops his skills through learning-by-doing. He learns more skills by using them more intensely. For example, individuals accumulate more motor skills by taking on motor skill intensive tasks. Let A_1 denote a $K \times K$ diagonal matrix of the marginal effects of task complexity on learning. Let a_0 be a K-dimensional vector of skill learning parameters. Skills grow from year t to year $t+1$ according to the following skill transition equation

$$s_{t+1} = Ds_t + a_0 + A_1^{'} x_t + \varepsilon_{t+1}, \quad (3)$$

where D is a K-dimensional diagonal matrix, and a_0 and ε_t are K-dimensional vectors. Skills depreciate every year at the rate $I - D$ where I is a K-dimensional identity matrix. A vector of skill shocks ε_t is normal, independent and identically distributed with mean zero and variance Σ_ε: $\varepsilon_t \sim N(0, \Sigma_\varepsilon)$. Notice that this shock has a persistent effect on skills unless $D = 0$.

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 the features observed in the data. Moreover, Altonji and Spletzer (1991) find that a skill demanding job offers more skill training even after controlling for workers' educations. The key assumption 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 mean initial skills vary across individuals according to a L-dimensional vector of time-invariant demographic variables d. Given observed individual characteristics d, the initial skills are normally distributed with mean and variance

$$E(s_1|d) = h_0 + Hd \tag{4}$$

$$Var(s_1|d) = \Sigma_{s1}, \tag{5}$$

where h_0 is a scalar and H is a $K \times L$ matrix of parameters.

Work Disutility Work disutility, as well as skill endowment, can rationalize the observed occupational choices. The following quadratic function of task complexity determines work disutility,

$$v_t = v(x_t, \bar{x}_t, s_t, \tilde{v}_t) \tag{6}$$

$$(g_0 + G_1 d + \tilde{v}_t)' x_t + x_t' G_2 x_t + (x_t - \bar{x}_t)' G_{3,t} (x_t - \bar{x}_t), \tag{7}$$

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 and $G_{3,t}$ are $K \times K$ symmetric negative definite matrices, and $\bar{x}_t$ is a K-dimensional vector of work habits.

The first two terms capture the effect of the task complexity of the current job on work disutility. Motor skill intensive tasks make an individual physically fatigued. Similarly, cognitive skill demanding tasks can exhaust an individual mentally. This disutility component varies across individuals according to observed permanent individual characteristics d as well as a preference shock $\tilde{v}_t$.

The third term captures the effect of work habits on disutility. The work habits of individuals are measured by the weighted average of the task complexity of previous occupations. 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. Kambourov and Manovskii (2008) shows that occupational mobility decreases with age. To account for this empirical finding, I allow for this disutility component to vary with age. Specifically, the parameter matrix $G_{3,t}$ is given by

$$G_{3,t} = F_0 \odot \exp(t \cdot F_1) \tag{8}$$

where $\odot$ is an element-by-element multiplication operator, F_0 is a K-dimensional negative definite diagonal matrix, and F_1 is a K-dimensional diagonal matrix.

Individuals form their work habits according to the following transition equation

$$\bar{x}_{t+1} = A_2\bar{x}_t + (I - A_2)x_t \tag{9}$$

where A_2 is a K-dimensional diagonal matrix of which elements take values between zero and one and I is a K-dimensional identity matrix. Hence, the work habit $\bar{x}$ is a weighted average of the task complexity of the past jobs. When $A_2 = 0$, only the tasks in the last occupation affect work disutility. In contrast, when $A_2 = I$, work habits remain constant at the initial value $\bar{x}_1$. For all other cases where the elements of A_2 are greater than 0 and less than 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. The initial condition for $\bar{x}_t$ varies across individuals according to observed characteristics d such that

$$\bar{x}_1 = \bar{x}_{1,0} + Xd, \tag{10}$$

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

Bellman Equation The Bellman equation for an individual is given by

$$\begin{aligned}
V_t(s_t, \bar{x}_t, \tilde{v}_t) &= \max_{x_t} \ln w(x_t, s_t) - v(x_t, \bar{x}_t, s_t, \tilde{v}_t) + \beta E V_{t+1}(s_{t+1}, \bar{x}_{t+1}, \tilde{v}_{t+1}) \qquad (11)\\
\text{s.t.} & \\
\ln w_t &= p_0 + [p_1 + P_2' x_t]' s_t \\
v_t &= (g_0 + G_1 d + \tilde{v}_t)' x_t + x_t' G_2 x_t + (x_t - \bar{x}_t)' G_{3,t}(x_t - \bar{x}_t) \\
s_{t+1} &= D s_t + a_0 + A_1' x_t + \varepsilon_{t+1} \\
\bar{x}_{t+1} &= A_2 \bar{x}_t + (I - A_2) x_t \\
s_1 &\sim N(h_0 + Hd,\ \Sigma_{s1}) \\
\bar{x}_1 &= \bar{x}_{1,0} + Xd. \qquad (12)
\end{aligned}$$

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

$$x_t^* = c_{0,t} + C_{1,t} d + C_{2,t} s_t + C_{3,t} \bar{x}_t + \nu_t, \tag{13}$$

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 available 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. Due to the finite horizon nature of the problem, I numerically solve the value function and the policy function by backward recursion.[3]

4.3 Discussion

The following first order condition characterizes an individual's optimal occupational choice

$$\frac{\partial v}{\partial x_t} = \frac{\partial \ln w}{\partial x_t} + \beta \frac{\partial s_{t+1}}{\partial x_t} \frac{\partial EV_{t+1}}{\partial s_{t+1}} + \beta \frac{\partial \bar{x}_{t+1}}{\partial x_t} \frac{\partial EV_{t+1}}{\partial \bar{x}_{t+1}}. \tag{14}$$

The left hand side is the marginal cost of taking on a complex task. The right hand side is the marginal return from choosing such a task and it can be divided into three components. The first term is the returns to skills. The sign of this term is positive, suggesting that returns to skills are higher in a job with complex tasks. The second term is the skill investment value. The sign of this term is also positive, implying that individuals learn more when undertaking relatively more complex tasks. The third term is the value of habit formation in tasks. The sign of this term is ambiguous and depends on parameter values and the state variables. At the optimum, the marginal disutility equals the sum of these three separate marginal returns.

Gibbons, Katz, Lemieux, and Parent (2005) proposes a similar model of occupational sorting. In their static model, uni-dimensional skills are differently rewarded across 1-digit occupations and individuals sort themselves into different occupations according to their skill endowment. The present model differs from their model in two major ways. First, it incorporates skill investment. The model of Gibbons, Katz, Lemieux, and Parent (2005) is static and they do not analyze dynamics of skills. Second, it allows for multidimensional skills, which enables me to analyze worker

[3] Many other methods are available for infinite horizon problems, as surveyed by Anderson, Hansen, McGrattan, and Sargent (1996).

skills and heterogeneity of occupations in greater detail. This approach makes it possible to deal with hundreds of occupations at three digit level, while Gibbons, Katz, Lemieux, and Parent (2005) include occupations only at one digit level.

Lazear (2004) also shows that his skill-weight approach explains worker mobility and wage changes associated with tasks. In his model, two-dimensional skills are utilized with different weights across occupations. Because the weight must add up to one by definition, utilizing one skill intensely necessarily means that the other skills are less utilized. Individuals choose a job that puts more weight on skills in which they have a comparative advantage. A difference between Lazear (2004) and the model that I present is that my model does not impose a trade off between skill weights. Some jobs utilize all types of skills more intensely than other jobs. Another difference is that Lazear (2004) assumes that individuals develop their skills through on-the-job training (skill investment) and the skill investment decision is a separate problem from job choice. In my model, the skill investment decision is integrated into the occupational choice problem. Returns to skills, skill investment, and work disutility affect the optimal choice of occupation, which differs from Lazear's model in that only returns to skills matter in choosing a job.

5 Estimation Strategy

5.1 Identification Restrictions

I estimate Equations (2) and (13), given the transition equations (3) and (9), under the parameter constraints imposed on C by the structural model. Notice that this model is an application of a state-space model with time-varying coefficients. The work disutility parameters g_0, G_1, G_2 and $G_{3,t}$ are associated with the policy function parameters $c_{0,t}$, $C_{1,t}$, $C_{2,t}$ and $C_{3,t}$, respectively. For identification, each of the work disutility parameter vector and matrices must contain equal or fewer parameters than the corresponding policy function parameters. These conditions are satisfied in my specification, because the number of unknown parameters in each of the work disutility parameter vector and matrices is equal to the number of parameters in the corresponding parameter vector and matrices in the policy function. The relationships between the structural parameters and the policy function parameters are outlined in Appendix C.

Skills are identified up to an affine transformation because they do not have a natural scale. The conditional mean and variance of skills given time-invariant individual characteristics d are Hd and one, respectively:

$$E(s_1|d) = Hd \tag{15}$$

$$Var(s_1|d) = I, \tag{16}$$

where I is a K-dimensional identity matrix. In other words, I impose restrictions on Equations (4) and (5) such that $h_0 = 0$ and $\Sigma_{s1} = I$. The intuition for the unidentifiability of skill scale is that observed variables (i.e. wage and task complexity) are the product of unobserved skills and unknown parameters such as the rate of return to skills. Either a large amount of skills or a high rate of return are needed to rationalize an observed high wage.

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.

5.2 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 condition for the tasks of past occupations $\bar{x}_1$. The conditional mean and variance of skills are

$$
\begin{aligned}
E(s_t|x_1, w_1, \cdots, x_{t-1}, w_{t-1}; \bar{x}_1) &\equiv E(s_t|Y_{t-1}) && (17)\\
&\equiv \hat{s}_{t|t-1} && (18)\\
Var(s_t|x_1, w_1, \cdots, x_{t-1}, w_{t-1}; \bar{x}_1) &\equiv Var(s_t|Y_{t-1}) && (19)\\
&\equiv \Sigma^s_{t|t-1} && (20)
\end{aligned}
$$

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 13) is linear in skills, 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

$$
\begin{aligned}
E(x_t|Y_{t-1}) &= c_{0,t} + C_{1,t}\hat{s}_{t|t-1} + C_{2,t}\bar{x}_t && (21)\\
Var(x_t|Y_{t-1}) &= C_{1,t}\Sigma^s_{t|t-1}C^{'}_{1,t} + \Sigma_V. && (22)
\end{aligned}
$$

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

$$
E(s_t|Y_{t-1}, x_t) = \hat{s}_{t|t-1} + \Sigma^s_{t|t-1}C^{'}_{1,t}(C_{1,t}\Sigma^s_{t|t-1}C^{'}_{1,t} + \Sigma_V)^{-1}\hat{v}_t \quad (23)
$$

$$Var(s_t|Y_{t-1},x_t) = \Sigma^s_{t|t-1} - \Sigma^s_{t|t-1}C'_{1,t}(C_{1,t}\Sigma^s_{t|t-1}C'_{1,t}+\Sigma_\nu)^{-1}C_{1,t}\Sigma^s_{t|t-1}, \quad (24)$$

where $\hat{\nu}_t$ is a vector of residuals and $\hat{\nu}_t = x_t - E(x_t|Y_{t-1})$. 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

$$E(\ln w_t|Y_{t-1},x_t) = p_0 + [p_1 + P'_2 x_t]' E(s_t|Y_{t-1},x_t) \quad (25)$$

$$Var(\ln w_t|Y_{t-1},x_t) = [p_1 + P'_2 x_t]' \Sigma^s_{t|t}[p_1 + P'_2 x_t] + \sigma^2_\eta. \quad (26)$$

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

$$E(s_t|Y_{t-1},x_t,w_t) = E(s_t|Y_{t-1},x_t) + Var(s_t|Y_{t-1},x_t)[p_1 + P'_2 x_t][Var(\ln w_t|Y_{t-1}x_t)]^{-1}\hat{\eta}_t \quad (27)$$

$$Var(s_t|Y_{t-1},x_t,w_t) = Var(s_t|Y_{t-1},x_t) - Var(s_t|Y_{t-1},x_t)[p_1 + P'_2 x_t][Var(\ln w_t|Y_{t-1}x_t)]^{-1}[p_1 + P'_2 x_t]' Var(s_t|Y_{t-1},x_t) \quad (28)$$

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,

$$\hat{s}_{t+1|t} = D\hat{s}_{t|t} + a_0 + A_1 x_t \quad (29)$$

$$\Sigma^s_{t+1|t} = D\Sigma^s_{t|t}D + \Sigma_\varepsilon. \quad (30)$$

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

$$\begin{aligned} & l(w_{i1}, x_{i1}, \cdots, w_{iT_i}, x_{iT_i}|\bar{x}_{i1}) \\ = \; & l(x_{i1}|\bar{x}_{i1}) l(w_{i1}|x_{i1};\bar{x}_{i1}) \times l(x_{i2}|Y_{i1}) l(w_{i1}|Y_{i1},x_{i2}) \times \cdots \times l(x_{iT_i}|Y_{iT_i-1}) l(w_{i1}|Y_{iT_i-1},x_{iT_i}). \end{aligned} \quad (31)$$

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

$$l = \Pi_{i=1}^{N} l(w_{i1}, x_{i1}, \cdots, w_{iT_i}, x_{iT_i} | \bar{x}_{i1}). \tag{32}$$

6 Estimation Results

6.1 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 survey. To account for potential attrition problems, if an observation is missing in the data, I treat the corresponding simulation outcomes as missing.

Figure 3 shows the profiles of the mean task complexity of individuals' occupations over time. The model replicates the observed time profiles of both cognitive and motor task complexity very well. Figure 4 presents the hourly logwage profiles. The model's prediction largely matches the observed mean wage profile, but the predicted wage profile is less concave. Consequently, predicted wages between year 5 and 12 are slightly lower than those found in the data. The model has a small limitation in this respect because the curvature of the wage profile is generated by a single skill depreciation parameter. Introducing a more flexible skill depreciation function might help the model fit the data better, but that would make the model too complicated to derive an analytical solution to the policy function, which is certainly the strength of the model. All in all, the model shows an ability to fit these interesting features of the data.

6.2 Parameter Estimates

6.2.1 Wage Equation

Table 4 presents the parameter estimates of the wage equation (see equation 2) and their standard errors. Rates of return to skills are positive and significant. The rates of return to cognitive skills across occupations vary from 0.31 to 0.33 and those to motor skills vary from 0.050 to 0.056. Although task complexity of the current job statistically significantly increases the rate of return, its effect on wages seems modest. Consider an average individual in his 10th year since labor market entry. He has 0.85 units of cognitive skills and 1.83 units of motor skills. His wages can change by about 2% due to the difference in returns to cognitive skills and by 1% due to the difference in returns to motor skills.

These modest effects may be accounted for by occupational classification errors producing a potential attenuation bias in the estimates of P_2. Many papers including Sullivan (2009) report evidence of mis-classification of occupation. When occupations are misclassified in the sample, the task complexity of the jobs is necessarily measured with error. This measurement error is particularly problematic for estimating P_2, because it is identified by variation in the product of the task complexity of the current job and skills that are a function of the task complexity of past jobs. If the measurement error accounts for a part of this variation, the parameter estimates of P_2 suffer from attenuation bias.

6.2.2 Skill Transition

Gross skill learning significantly increases with the task complexity of the job. Table 5 indicates that, within a year, cognitive skill learning ranges from 0.11 to 0.22, while motor skill learning ranges from 0.43 to 1.46. Unlike the returns to skills, the amount of skill learning differs considerably across occupations. Although these parameters also suffer from attenuation bias due to the measurement error, the bias is probably smaller because the sources of measurement error are fewer.

Although the gross amount of learning is greater for motor skills than for cognitive skills, motor skills depreciate much faster. The annual skill depreciation rates for cognitive skill and motor skill are 6% and 55%, respectively. These estimates imply that motor skills are more strongly influenced by recent job experience, while cognitive skills persistently remain productive.

The variance of skill shocks Σ_ε is smaller than that of the initial skill distribution Σ_{s1}, which seems reasonable. The initial skills are significantly different according to AFQT scores and across race. AFQT scores and initial cognitive skills have a significant and strongly positive relationship, while AFQT scores and motor skills have a negative (but insignificant) relationship. Blacks have significantly less initial cognitive skills than whites and hispanics. Their initial motor skills are greater than those of whites and less than those of hispanics, although the differences are insignificant. Hispanics have almost the same amount of initial cognitive skills as whites and significantly more initial motor skills than whites. These results for initial skills are consistent with the large literature of Mincerian wage regression[4] , which finds that the coefficient for AFQT score is strongly positive and significant, blacks earn substantially lower than whites, and hispanics earn almost the same as whites.

6.2.3 Work Disutility

The parameter estimates reported in Table 6 show that the disutility of work is increasing and convex in both cognitive and motor task complexity. Blacks receive more disutility from both

[4]See Rubinstein and Weiss (2006), for example.

cognitive and motor tasks than whites and hispanics, while hispanics receive less disutility from cognitive tasks and more disutility from motor tasks than whites. But, the differences across race are insignificant. High AFQT score holders receive less work disutility for both cognitive and motor skill demanding tasks.

The parameter estimates of $G_{3,t}$ indicate that individuals suffer more work disutility when they work a job that is different from previously held jobs. In addition, this disutility component significantly increases with age. This result reflects the fact, widely reported in the literature [5], that occupational mobility is decreasing over the careers of individuals. The transition parameters A_2 are significantly greater than zero for both cognitive and motor task complexity, which means that work disutility is indeed influenced by task complexity more than a year ago. The estimates for the initial values of the task complexity of past jobs $\bar{x}_1$ indicate that blacks and hispanics have lower initial values than whites, although the differences are insignificant. Those who have high AFQT scores start their careers with higher values of $\bar{x}_1$ for both cognitive and motor skills.

7 Discussion

7.1 Skills and Wage Growth

Using the parameter estimates, I calculate, through simulation, the time profiles of unobserved skills and their contributions to the wage growth. I simulate each individual in the sample for 300 times from year 1 to year 16 and, unlike the simulations used to evaluate the model's fit, I treat no simulation outcomes as missing.[6] I normalize the mean initial skills to zero and the conditional variance of the initial skills given AFQT scores and race to one.

Figure 5 illustrates the growth of average skills. Cognitive skills steadily increase over time, though at a decreasing pace. Motor skills grow rapidly during the first three years, but this growth quickly slows in the years following. Average motor skills increase to 1.83 in year 9, and then gradually decline to 1.77 in year 16. The profile of motor skills is hump-shaped because the skill investment value of work decreases over time and the motor skill depreciation rate is high. The intercepts and coefficients of the policy function are decreasing for the same reason.

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. During the first three years, the contributions of cognitive skills and motor skills are roughly the same and each type of skill increases wages by about 10% (see Figure 6). However, as motor skill growth slows down, the contribution of motor skills remains around 10%. In contrast, cognitive skills

[5] See Kambourov and Manovskii (2008), for example

[6] The results change very little when the same simulation rule as the model fit evaluation is applied.

continue to contribute to wage growth over the careers of individuals. Cognitive skills raise wages by 29% during the first 10 years and by 39% during the first 15 years, while motor skills raise wages by 9.8% during the first 10 years and by 9.5% during the first 15 years.[7] The results indicate that both cognitive and motor skills substantially contribute to wage growth. However, motor skills contribute to wage growth only early in the career, cognitive skills are the driving factor of wage growth afterwards.

7.2 How Wage Return, Skill Formation, and Habit Formation Affect Occupational Choice

To assess the importance of wage return, skill formation, and habit formation for the occupational choice decision, I report in Table 7 the age profiles of the marginal cost, the marginal wage returns to skills, the marginal value of skill formation, and the marginal value of habit formation. Equation (14) shows that, at the optimal occupational choice, the marginal cost equals the sum of the three marginal returns .

The marginal wage return to cognitive skills steadily increases over time, while the wage return to motor skills quickly increases during the first three years, but remains unchanged afterwards. These patterns reflect the profiles of cognitive and motor skill growth. The marginal value of skill formation of both cognitive and motor skills decreases gradually. The value function is concave with respect to skills and the approaching retirement period lowers the marginal return to skill formation. Both of these factors reduce the incentive to work at a job with a complex task. The value of habit formation in cognitive task complexity is considerably large at the beginning of the career, but quickly declines to zero in year 16. Being accustomed to a complex cognitive task is valuable, because it allows an individual to enter a cognitive skill demanding job in which he can expect a better wage return and skill learning opportunity. The value of habit formation in motor task is very small and does not have a significant effect on the occupational choice decision. These results indicate that skill formation is the primary factor that affects occupational choice decisions.

8 Conclusion

This paper constructs and estimates a structural dynamic model of occupational choice. The model departs from previous contributions by distinguishing between the skills possessed by individuals

[7]The total wage growth during the first 10 years in the labor market is 39%, which is slightly lower than the wage growth rate reported by Rubinstein and Weiss (2006), not only because of the downward bias in the predicted wage level discussed in subsection 6.1, but also because the sample criteria of this paper is more restrictive. The sample used in this paper includes individuals who have made a long-term transition to the full-time labor market, which excludes some of initial experiences in the labor market.

and the skills that are required for an occupation, and by characterizing all occupations in a continuous multidimensional task complexity space. This task space approach enables me to derive a simple policy function and analyze the heterogeneity of occupations in greater depth.

I estimate the model using the Kalman filter and data from the DOT and NLSY. The empirical results indicate that both cognitive and motor skills raise wages by about 10% during the first three years. Cognitive skills drive subsequent wage growth and raise wages by 30% during the first ten years, while motor skills do not increase wages after the first three years. I also find that the depreciation of motor skills is very high, which suggests that, in order to generate a persistent effect, training programs should aim to enhance cognitive skills instead of motor skills if the costs are similar.

The model has some limitations. First, it does not allow for part-time work and non participation, which are important aspects of the labor market, particularly for female labor supply. Second, the model treats pre-labor market skill investment as exogenous. This paper is unable to answer questions regarding the the effect that schooling may have on individual cognitive and motor skills. Third, job search is missing in the model. Previous papers find that it has a substantial effect on wage growth. The proposed model could be integrated with a search model in order to understand wage growth and career mobility better. Lastly, an estimation method that accounts for occupation classification error needs to be developed. These important and interesting issues are to be addressed in future research.

References

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 (2008): "Two Sides of the Same Coin: U.S. "Residual" Inequality and the Gender Gap," University of California-Irvine.

BARLEVY, G. (2008): "Identification of Search Models using Record Statistics," *Review of Economic Studies*, 75(1), 29–64.

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

BORGHANS, L., B. TER WEEL, AND B. A. WEINBERG (2006): "People People: Social Capital and The Labor-Market Outcomes of Underrepresented Groups," NBER Working Paper 11985.

CUNHA, F., AND J. J. HECKMAN (2007): "The Technology of Skill Formation," *American Economic Review*, 97(2), 31–47.

CUNHA, F., J. J. HECKMAN, L. LOCHNER, AND D. V. MASTEROV (2006): "Interpreting the Evidence on Life Cycle Skill Formation," in *Handbook of the Economics of Education*, ed. by E. A. Hanushek, and F. W. (eds.). Elsevier Science Publisher (North-Holland), vol. 1.

GATHMANN, C., AND U. SCHÖNBERG (2007): "How General is Specific Human Capital," Stanford University.

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.

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., 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.

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.

LIU, H. (2009): "Life Cycle Human Capital Formation, Search Intensity, and Wage Dynamics," University of Western Ontario.

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.

RUBINSTEIN, Y., AND Y. WEISS (2006): "Post Schooling Wage Growth: Investment, Search and Learning," in *Handbook of the Economics of Education*, ed. by E. A. Hanushek, and F. W. (eds.). Elsevier Science Publisher (North-Holland), vol. 1.

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. (2009): "Estimation of an Occupational Choice Model when Occupations are Misclassified," *Journal of Human Resources*.

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

YAMAGUCHI, S. (2007): "The Effect of Match Quality and Specific Experience on Career Decisions and Wage Growth," McMaster University.

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

A Tables

Table 1: Factor Loadings

Skill Requirement Measures	Cognitive Skill	Motor Skill
Data	0.626	
GED: Reasoning	0.345	
GED: Mathematical	0.377	
GED: Language	0.409	
Intelligence	0.242	
Verbal	0.258	
Numerical	0.234	
Things		0.930
Motor Coordination		0.161
Finger Dexterity		0.165
Manual Dexterity		0.134
Eye-Hand-Foot Coordination		0.056
Color Discrimination		0.081
Form Perception		0.146
Spatial		0.152
Torelance		0.103
Proportion of Variance	0.849	0.745

Source: The 1971 April CPS augmented with job characteristics variables from the Revised Fourth Edition of the DOT (1991).
Note: Factor loadings of the first principal components and the proportions of the variance explained by the first principal components are reported.

Table 2: Task Complexity by Occupation at 1-Digit Classification

	Cognitive	Motor	Nobs.
Professional	0.902	0.490	7716
Manager	0.863	0.253	5538
Sales	0.530	0.166	3949
Clerical	0.524	0.611	9821
Craftsmen	0.558	0.856	6841
Operatives	0.183	0.549	6099
Transport	0.235	0.652	1774
Laborer	0.100	0.353	2818
Farmer	0.744	0.853	1117
Farm Laborer	0.156	0.485	882
Service	0.298	0.420	7249
HH Service	0.153	0.234	1469
ALL	0.506	0.506	55274

Source: The 1971 April CPS augmented with job characteristics variables from the Revised Fourth Edition of the DOT (1991).
Note: HH service is household service occupation.

Table 3: Summary Statistics

	White			Black			Hispanic		
	Mean	S.D.	Nobs.	Mean	S.D.	Nobs.	Mean	S.D.	Nobs.
Hourly Wage (t=10)	14.36	5.81	451	11.30	5.11	63	14.30	6.52	30
Cognitive (t=10)	0.45	0.27	463	0.33	0.25	64	0.45	0.29	31
Motor (t=10)	0.56	0.28	463	0.53	0.25	64	0.52	0.28	31
AFQT	43.96	22.52	647	14.79	14.84	100	32.44	24.83	48
Nobs. Per Individual	11.50	4.01	647	10.39	3.87	100	10.25	4.23	48

Source: NLSY and The Revised Fourth Edition of the DOT
Note: S.D. is standard deviation and Nobs. is number of observations. Hourly wages are deflated by 2000 CPI. Statistics of hourly wage, cognitive and motor task complexity measurement are evaluated in the 10th year since the long term transition to the labor market.

Table 4: Wage Equation: $\ln w_t = p_0 + (p_1^{'} + P_2 x_t)^{'} s_t + \eta_t$.

Notation	Estimates	Std. Error
p_0	1.9990	0.0381
$p_1(1)$	0.3060	0.0069
$p_1(2)$	0.0502	0.0084
$P_2(1,1)$	0.0203	0.0054
$P_2(2,2)$	0.0056	0.0024
σ_η^2	0.0107	0.0004

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 5: Skill Transition: $s_{t+1} = Ds_t + a_0 + A_1 x_t + \varepsilon_t$.

Notation	Estimates	Std. Error
$a_0(1)$	0.1062	0.0180
$a_0(2)$	0.4318	0.4856
$A_1(1,1)$	0.1092	0.0279
$A_1(2,2)$	1.0283	0.2494
$D(1,1)$	0.9386	0.0068
$D(2,2)$	0.4495	0.0332
$\Sigma_\varepsilon(1,1)$	0.1936	0.0119
$\Sigma_\varepsilon(2,2)$	0.3503	0.0712
$H(1,1)$, AFQT	1.2686	0.2398
$H(1,2)$, Black	−0.4579	0.1792
$H(1,3)$, Hispanic	0.0352	0.2105
$H(2,1)$, AFQT	−1.0495	1.4768
$H(2,2)$, Black	1.6200	1.0338
$H(2,3)$, Hispanic	2.5384	1.1303

Note: The skill transition equation is $s_{t+1} = Ds_t + a_0 + A_1 x_t$. The conditional initial mean skills are given by $E(s_1|d) = Hd$ where d is a vector of AFQT scores and dummy variables for race (whites are the reference group.) The conditional initial skill variance matrix Σ_{s1} is an identity matrix given d, for normalization. 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 6: Work Disutility

$$g(x_t,\bar{x}_t,\tilde{v}_t) = (g_0 + G_1 d + \tilde{v}_t)' x_t + x_t' G_2 x_t + (x_t - \bar{x}_t)' G_{3,t}(x_t - \bar{x}_t)$$

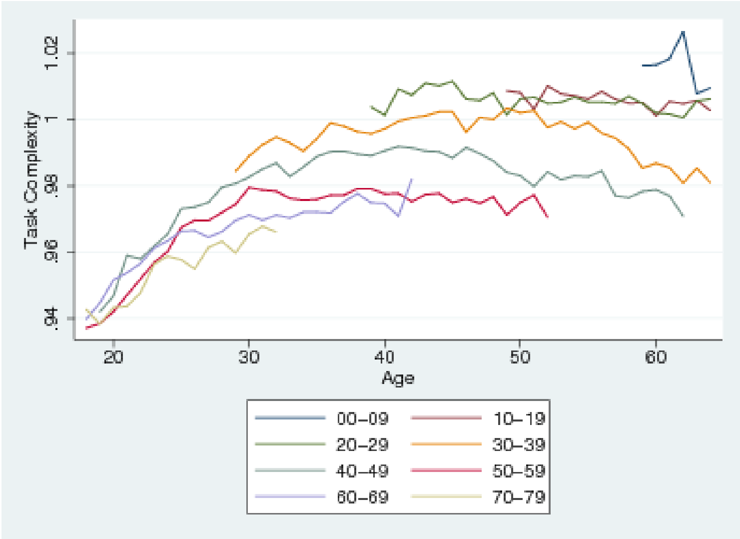
Notation	Estimates	Std. Error
$g_0(1)$	−0.1727	0.0732
$G_1(1,1)$, AFQT	0.1059	0.0430
$G_1(1,2)$, Black	−0.0145	0.0188
$G_1(1,3)$, Hispanic	0.0044	0.0221
$g_0(2)$	−0.0866	0.0195
$G_1(2,1)$, AFQT	0.0001	0.0002
$G_1(2,2)$, Black	−0.0004	0.0002
$G_1(2,3)$, Hispanic	−0.0003	0.0003
$G_2(1,1)$	−0.1981	0.0695
$G_2(2,2)$	−0.0124	0.0069
$F_0(1,1)$	−1.8244	0.6804
$F_0(2,2)$	−0.0073	0.0024
$F_1(1,1)$	0.0623	0.0236
$F_1(2,2)$	0.1887	0.0226
$\bar{x}_{1,0}(1)$	0.2246	0.0263
$X(1,1)$, AFQT	0.1512	0.0487
$X(1,2)$, Black	−0.0396	0.0420
$X(1,3)$, Hispanic	−0.0137	0.0433
$\bar{x}_{1,0}(2)$	1.2233	0.4345
$X(2,1)$, AFQT	0.5694	0.7224
$X(2,2)$, Black	−0.8491	0.5204
$X(2,3)$, Hispanic	−1.2545	0.6324
$A_2(1,1)$	0.4146	0.0152
$A_2(2,3)$	0.3690	0.0160

Note: The work disutility function is $g(x_t,\bar{x}_t,v_t) = (g_0 + G_1 d + v_t)' x_t + x_t' G_2 x_t + (x_t - \bar{x}_t)' G_{3,t}(x_t - \bar{x}_t)$ and $G_{3,t} = F_0 \odot \exp(t \cdot F_1)$. The transition equation of work habit is $\bar{x}_{t+1} = A_2 \bar{x}_t + (I - A_2) x_t$ 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 scores 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 7: Marginal Cost and Marginal Returns Over The Career

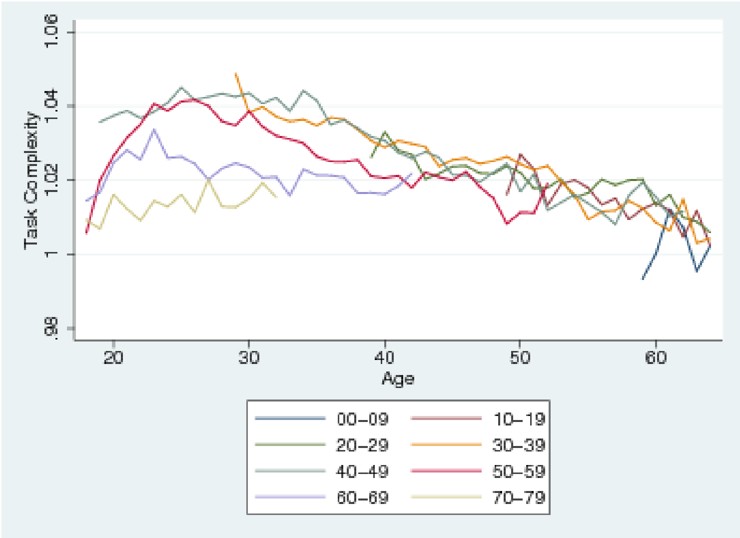
Year	Cognitive Skills				Motor Skills			
	MC	Wage	Skill	Pref.	MC	Wage	Skill	Pref.
1	4.56	0.00	2.94	1.62	0.87	0.00	0.92	−0.05
2	4.41	0.03	2.94	1.44	0.95	0.06	0.91	−0.02
3	4.28	0.05	2.94	1.29	0.99	0.09	0.91	−0.01
4	4.17	0.08	2.94	1.15	1.00	0.10	0.91	0.00
5	4.08	0.10	2.94	1.03	1.01	0.10	0.91	0.00
6	3.99	0.13	2.94	0.92	1.01	0.11	0.91	0.00
7	3.90	0.15	2.94	0.82	1.01	0.11	0.91	−0.01
8	3.83	0.17	2.93	0.73	1.01	0.11	0.91	−0.01
9	3.75	0.19	2.93	0.63	1.00	0.11	0.91	−0.01
10	3.68	0.21	2.93	0.54	1.00	0.11	0.91	−0.02
11	3.60	0.22	2.93	0.46	1.00	0.11	0.91	−0.02
12	3.53	0.24	2.92	0.37	0.99	0.11	0.91	−0.02
13	3.46	0.26	2.92	0.29	0.98	0.11	0.90	−0.03
14	3.38	0.27	2.91	0.20	0.98	0.11	0.90	−0.03
15	3.29	0.29	2.91	0.10	0.97	0.11	0.90	−0.04
16	3.20	0.30	2.90	0.00	0.97	0.11	0.90	−0.04

B Figures



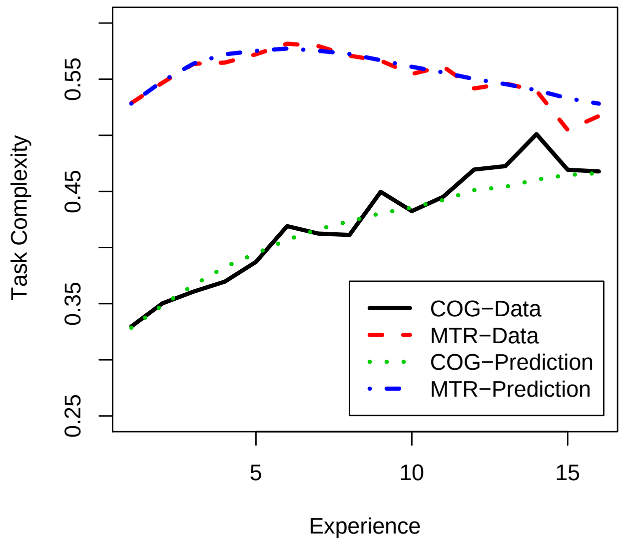
Source: CPS 1968-2002
Note: Cohort is defined by birth year.

Figure 1: Cognitive Task Complexity Profile by Cohort



Source: CPS 1968-2002
Note: Cohort is defined by birth year.

Figure 2: Motor Task Complexity Profile by Cohort



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

Figure 3: Model Fit (Task Complexity Profile)

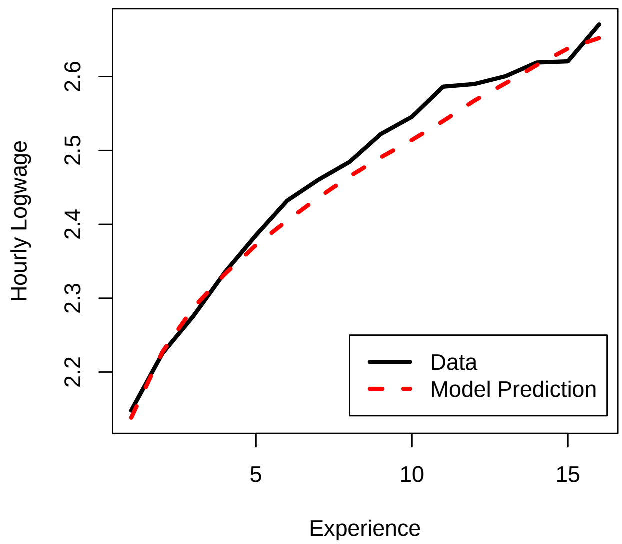


Figure 4: Model Fit (Hourly Logwage Profile)

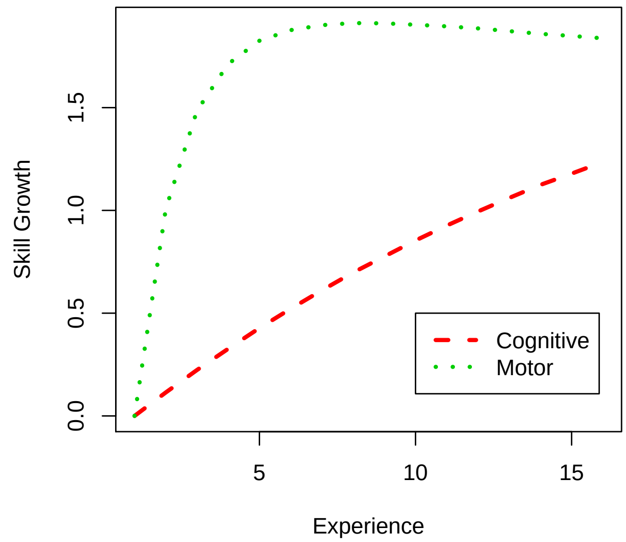


Figure 5: Skill Growth Profiles

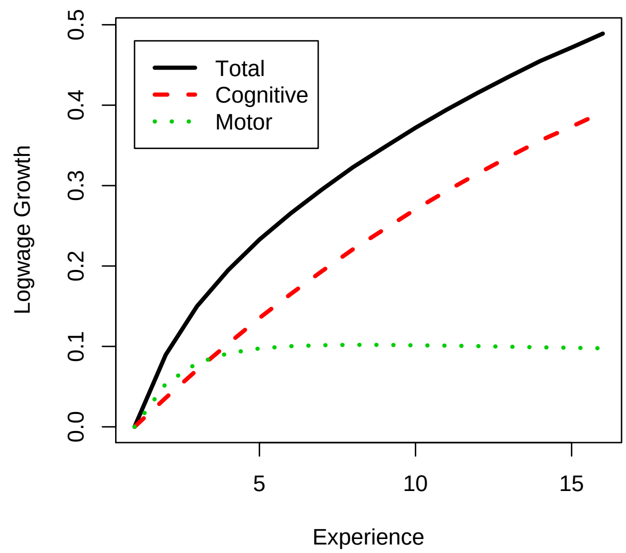


Figure 6: Contribution of Skills to Wage Growth

C Model Solution

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

$$x_t^* = c_{0,t} + C_{1,t}d + C_{2,t}s_t + C_{3,t}\bar{x}_t + \nu_t. \tag{33}$$

I first rewrite the original Bellman equation in the following form

$$\begin{aligned}
V_t(z_t, \varepsilon_t, \tilde{\nu}_t) &= \max_{x_t} r_0 + r_1^{'}x_t + r_2^{'}z_t + x_t^{'}R_3x_t + 2x_t^{'}R_4z_t + z_t^{'}R_5z_t + \beta EV_{t+1}(z_{t+1}, \varepsilon_{t+1}, \tilde{\nu}_{t+1}) && (34)\\
&\text{s.t.} \\
z_{t+1} &= l_0 + L_1z_t + L_2x_t && (35)\\
V_{T+1} &= 0 && (36)
\end{aligned}$$

where

$$\begin{aligned}
z_t^{'} &= (s_t^{'}\ \bar{x}_t^{'}) && (37)\\
r_0 &= p_0 && (38)\\
r_1 &= g_0 + G_1d + \tilde{\nu}_t && (39)\\
r_2^{'} &= (p_1^{'}\ 0^{'}) && (40)\\
R_3 &= (G_2 + G_3) && (41)\\
R_4 &= \begin{pmatrix} 0.5 \cdot P_2 & -G_3 \end{pmatrix} && (42)\\
R_5 &= \begin{pmatrix} 0 & 0 \\ 0 & G_3 \end{pmatrix} && (43)\\
l_0^{'} &= (a_0^{'}\ 0^{'}) && (44)\\
L_1 &= \begin{pmatrix} D & 0 \\ 0 & I - A_2 \end{pmatrix} && (45)\\
L_2 &= \begin{pmatrix} A_1 \\ A_2 \end{pmatrix}. && (46)
\end{aligned}$$

Suppose the expected value of the value function in period $t+1$ can be written as a quadratic function of state variables,

$$EV_{t+1} = q_{0,t+1} + (q_{1,t+1} + Q_{2,t+1}d)^{'}z_{t+1} + z_{t+1}^{'}Q_{3,t}z_{t+1} \tag{47}$$

$$\equiv \quad q_{0,t+1} + q^{'}_{A,t+1} z_{t+1} + z^{'}_{t+1} Q_{3,t} z_{t+1} \tag{48}$$

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

$$\begin{aligned} V_t(z_t, \varepsilon_t, \tilde{v}_t) \quad = \quad & \max_{x_t} r_0 + r^{'}_1 x_t + r^{'}_2 z_t + x^{'}_t R_3 x_t + 2x^{'}_t R_4 z_t + z^{'}_t R_5 z_t + \\ & \beta[q_{0,t+1} + q^{'}_{A,t+1}(l_0 + L_1 z_t + L_2 x_t) + (l_0 + L_1 z_t + L_2 x_t)^{'} Q_{3,t+1}(l_0 + L_1 z_t + L_2 x_t)]. \end{aligned}$$

The first order condition for optimality is characterized by

$$0 \quad = \quad r_1 + 2R_3 x^*_t + 2R_4 z_t + \beta[L^{'}_2 q_{A,t+1} + 2L^{'}_2 Q^{'}_{3,t+1}(l_0 + L_1 z_t) + 2L^{'}_2 Q_{3,t+1} L_2 x^*_t].$$

Solving this equation for x^*_t to find

$$\begin{aligned} x^*_t \quad = \quad & -\frac{1}{2}(R_3 + \beta L^{'}_2 Q_{3,t+1} L_2)^{-1}[g_0 + G_1 d + \tilde{v}_t + \\ & \beta L^{'}_2((q_{1,t+1} + Q_{2,t+1} d) + 2Q^{'}_{3,t+1} l_0) + \{2R_4 + 2\beta L^{'}_2 Q^{'}_{3,t+1} L_1\} z_t] \qquad (49) \\ = \quad & -\frac{1}{2}(R_3 + \beta L^{'}_2 Q_{3,t+1} L_2)^{-1}[\{g_0 + \beta L^{'}_2(q_{1,t+1} + 2Q^{'}_{3,t+1} l_0\} + \\ & \{G_1 + \beta L^{'}_2 Q_{2,t+1}\} d + \{2R_4 + 2\beta L^{'}_2 Q^{'}_{3,t+1} L_1\} z_t + \tilde{v}_t] \qquad (50) \\ \equiv \quad & b_{0,t} + B_{1,t} d + B_{2,t} z_t + v_t \qquad (51) \end{aligned}$$

where I substitute $g_0 + G_1 d + \tilde{v}_t$ for r_1 (see Equation 39) and

$$\begin{aligned} b_{0,t} \quad &= \quad -\frac{1}{2}(R_3 + \beta L^{'}_2 Q_{3,t+1} L_2)^{-1}[g_0 + \beta L^{'}_2(q_{1,t+1} + 2Q^{'}_{3,t+1} l_0)] \qquad (52) \\ B_{1,t} \quad &= \quad -\frac{1}{2}(R_3 + \beta L^{'}_2 Q_{3,t+1} L_2)^{-1}[G_1 + \beta L^{'}_2 Q_{2,t+1}] \qquad (53) \\ B_{2,t} \quad &= \quad -\frac{1}{2}(R_3 + \beta L^{'}_2 Q_{3,t+1} L_2)^{-1} 2[R_4 + \beta L^{'}_2 Q^{'}_{3,t+1} L_1] \qquad (54) \\ v_t \quad &= \quad -\frac{1}{2}(R_3 + \beta L^{'}_2 Q_{3,t+1} L_2)^{-1} \tilde{v}_t. \qquad (55) \end{aligned}$$

Because z_t is a vector of acquired skills and task complexity of the past jobs (see Equation 37), Equation (33) indeed holds when the expected value function is quadratic like Equation (47). Notice that I can write

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

I will show that the expected value function in period t is also a quadratic function of state variables like Equation (47). 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 (35). To simplify notation, define

$$b_{A,t} \equiv b_{0,t} + B_{1,t}d + v_t. \tag{56}$$

The expected value function is

$$\begin{aligned}
EV_{t+1} &= q_{0,t+1} + q'_{A,t+1}z_{t+1} + z'_{t+1}Q_{3,t}z_{t+1} & (57)\\
&= q_{0,t+1} + q'_{A,t+1}(l_0 + L_1z_t + L_2(b_{A,t} + B_{2,t}z_t)) + \\
&\quad (l_0 + L_1z_t + L_2(b_{A,t} + B_{2,t}z_t))'Q_{3,t+1}(l_0 + L_1z_t + L_2(b_{A,t} + B_{2,t}z_t)) & (58)\\
&= \{q_{0,t+1} + q'_{A,t+1}(l_0 + L_2b_{A,t}) + (l_0 + L_2b_{A,t})'Q_{3,t+1}(l_0 + L_2b_{0,t})\} + \\
&\quad \{(L_1 + L_2B_{2,t})'(q_{A,t+1} + 2Q_{3,t+1}(l_0 + L_2b_{A,t}))\}'z_t + \\
&\quad z'_t(L_1 + L_2B_{2,t})'Q_{3,t+1}(L_1 + L_2B_{2,t})z_t. & (59)
\end{aligned}$$

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

$$\begin{aligned}
V_t - \beta EV_{t+1} &= r_0 + r'_1x_t + r'_2z_t + x'_tR_3x_t + x'_tR_4z_t + z'_tR_5z_t & (60)\\
&= r_0 + r'_1(b_{A,t} + B_{2,t}z_t) + r'_2z_t + (b_{A,t} + B_{2,t}z_t)'R_3(b_{A,t} + B_{2,t}z_t) + \\
&\quad 2(b_{A,t} + B_{2,t}z_t)'R_4z_t + z'_tR_5z_t & (61)\\
&= \{r_0 + r'_1b_{A,t} + b'_{A,t}R_2b_{A,t} + \{B'_{2,t}r_1 + r_2 + (2R_3B_{2,t} + 2R_4)'b_{A,t}\}'z_t + \\
&\quad z'_t\{B'_{2,t}R_3B_{2,t} + B'_{2,t}R_4 + R'_4B_{2,t} + R_5\}z_T. & (62)
\end{aligned}$$

So, the value function in period t is

$$\begin{aligned}
EV_t &= E[\{r_0 + r'_1b_{A,t} + b'_{A,t}R_3b_{A,t}\} + \\
&\quad \beta\{q_{0,t+1} + q'_{A,t+1}(l_0 + L_2b_{A,t}) + (l_0 + L_2b_{A,t})'Q_{3,t+1}(l_0 + L_2b_{0,t})\}] + \\
&\quad E[\{B'_{2,t}r_1 + r_2 + 2(R_3B_{2,t} + R_4)'b_{A,t}\} + \\
&\quad \beta\{(L_1 + L_2B_{2,t})'(q_{A,t+1} + 2Q_{3,t+1}(l_0 + L_2b_{A,t}))\}]'z_t + \\
&\quad Ez'_t[\{B'_{2,t}R_3B_{2,t} + B'_{2,t}R_4 + R'_4B_{2,t} + R_5\} + \beta\{(L_1 + L_2B_{2,t})'Q_{3,t+1}(L_1 + L_2B_{2,t})\}]z_t & (63)\\
&\equiv q_{0,t} + q'_{A,t}z_t + z'_tQ_{3,t}z_t. & (64)
\end{aligned}$$

Notice that $B'_{2,t}R_4 + R'_4R_{2,t}$ 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 $Er_1 = G_1d$. Given these facts, $q_{A,t}$ can be written as

$$q_{A,t} = B'_{2,t}G_1d + r_2 + 2(R_3B_{2,t} + R_4)'Eb_{A,t} +$$

$$
\begin{aligned}
& \beta\{(L_1+L_2B_{2,t})^{'}(q_{A,t+1}+2Q_{3,t+1}(l_0+L_2Eb_{A,t}))\} && (65)\\
= \quad & B_{2,t}^{'}G_1d+r_2+2\beta(L_1+L_2B_{2,t})^{'}Q_{3,t+1}l_0+ \\
& \beta(L_1+L_2B_{2,t})^{'}(q_{1,t+1}+Q_{2,t+1}d)+ \\
& 2\{(R_3B_{2,t}+R_4)^{'}+\beta(L_1+L_2B_{2,t})^{'}Q_{3,t+1}L_2\}(b_{0,t}+B_{1,t}d). && (66)
\end{aligned}
$$

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

$$
\begin{aligned}
F_{1,t} &= L_1+L_2B_{2,t} && (67)\\
F_{2,t} &= (R_3B_{2,t}+R_4)^{'}+\beta F_{1,t}^{'}Q_{3,t+1}L_2. && (68)
\end{aligned}
$$

Then the Equation (66) can be written as

$$
\begin{aligned}
q_{A,t} &= B_{2,t}^{'}G_1d+r_2+2\beta F_{1,t}^{'}Q_{3,t+1}l_0+\beta F_{1,t}^{'}(q_{1,t+1}+Q_{2,t+1}d)+2F_{2,t}(b_{0,t}+B_{1,t}d) && (69)\\
&= [r_2+\beta F_{1,t}^{'}(2Q_{3,t+1}l_0+q_{1,t+1})+2F_{2,t}b_{0,t}]+[B_{2,t}^{'}G_1+\beta F_{1,t}^{'}Q_{2,t+1}+2F_{2,t}B_{1,t}]d && (70)\\
&\equiv q_{1,t}+Q_{2,t}d && (71)
\end{aligned}
$$

where

$$
\begin{aligned}
q_{1,t} &= r_2+\beta F_{1,t}^{'}(2Q_{3,t+1}l_0+q_{1,t+1})+2F_{2,t}b_{0,t} && (72)\\
Q_{2,t} &= B_{2,t}^{'}G_1+\beta F_{1,t}^{'}Q_{2,t+1}+2F_{2,t}B_{1,t} && (73)
\end{aligned}
$$

This shows that the expected value function in period t has the same form as Equation (47). The expected value function in period $T+1$ is a special case of this because $EV_{T+1}=0$. Thus, the optimal policy function in period t can be written as Equation (33).

D Robustness

This section examines, from two major viewpoints, the robustness of the parameter estimates presented above. In the first set of exercises, the sample restrictions outlined in Section 3.2.1 are modified. In the second set of exercises, an alternative measurement of task complexity is used.

D.1 Including part-time jobs

Because the model is intended to describe the behavior of individuals whose main activity is work, jobs in which individuals work less than 30 hours per week are excluded from the sample. This subsection examines if the main results are sensitive to this restriction by imposing a less restrictive sample selection criterion. Specifically, jobs in which individuals work 10 hours per week or more are included in the sample. Due to this change, part-time jobs held while individuals are in high school are counted as a part of their careers. Accordingly, the age restriction at the transition to the labor market is relaxed so that individuals must enter the labor market between age 16 and 22, instead of 18 and 22. This alternative sample contains 896 high school graduates, which comprises 10,955 person-year observations of occupational choices and 10,565 person-year observations of wages.

Tables 8 through 10 report the parameter estimates when using the alternative sample. A noticeable difference is that scale of motor skill transition parameters in the alternative sample is greater than that in the preferred sample. Otherwise, the parameter estimates are very similar to those of the preferred sample. Figures 7 and 8 show the average skill growth paths and the average contribution of skills to wage growth, respectively. The biggest difference is that motor skills grow as high as 2.10 in year 10 in the alternative sample, while they grow to 1.82 in the same year in the preferred sample. This results in a greater contribution of motor skills to wage growth. In the alternative sample, wages grow by 12% during the first 10 years, which is 2 percentage points higher than in the preferred sample. No substantial difference is found for cognitive skills.

Including part-time jobs results leads to more growth in motor skills. This is quite intuitive because the part-time jobs of young male individuals are relatively motor skill intensive. Nevertheless, the skill growth profiles and the contributions of skills to wage growth are very similar to those of the preferred sample. The main results of the paper are robust to the choice of hours of work restrictions.

D.2 No restriction on age at the transition to the full-time labor market.

Individuals who transit to the full-time labor market before age 18 are excluded, because it is unlikely that they work full-time while they are in high school. In addition, high school graduates who transit to the labor market at age 23 or later are also excluded, either because this is unlikely or because their labor force attachment is particularly weak. The model is meant to explain the behavior of individuals whose main activity is work. However, this age restriction is essentially arbitrary. I estimate the model using a sample that does not impose this age restriction. This alternative sample contains 988 high school graduates, which comprises 10,671 person-year observations of occupational choices and 10,323 person-year observations of wages.

Tables 11 through 13 report the parameter estimates of the alternative sample, and Figures 9 and 10 show the average skill growth paths and the average contribution of skills to wage growth, respectively. I do not find any substantial differences in the results between the alternative sample and the preferred sample. I conclude that the age restriction imposed in the preferred sample does not drive the main results of the paper.

D.3 Self-employed workers are excluded.

Excluding self-employed workers is a sensible restriction, because their wage determination mechanism and nature of work may be significantly different from those of employed workers. However, a large drawback is its small sample size: if I exclude individuals who have been self-employed at least for a year, the sample size decreases from 795 to 582. Although I decide to include self-employed workers in the preferred sample because of its larger sample size, it is worth estimating the model using a sample that excludes the self-employed. This alternative sample contains 6,649 person-year observations of occupational choices and 6,507 person-year observations of wages.

Tables 14 through 16 report the parameter estimates of the alternative sample. The only difference worth noticing is that the parameter estimates of the motor skill transition equation are larger than those in the preferred sample. Consequently, motor skills grow as high as 2.36 in year 10, as shown in Figure 11. However, this is offset by a lower motor skill price. As a result, the contribution of motor skills to wage growth is 10% during the first 10 years (see Figure 12), which is almost the same as the estimates in the preferred sample. This exercise demonstrates that excluding self-employed workers does not affect the main results of the paper.

D.4 Standardized DOT scores (instead of percentile scores)

The task complexity measures in the DOT are ordinal, although the model requires a cardinal task complexity index. To partially address this issue, following Autor, Levy, and Murnane (2003), I converted the composite DOT score constructed by the Principal Component Analysis into percentile scores. Nevertheless, this approach is not completely satisfactory, because other monotonic transformations of the DOT score are also possible. Because the model is not compatible with an ordinal task complexity index, examining the robustness of the results in this respect is important.

In doing so, I use the composite DOT score as a task complexity index, instead of the percentile scores. The composite DOT scores are standardized so that the minimum is zero and the maximum is one. Of course, sample size remains the same as the preferred sample.

Tables 17 through 19 report the parameter estimates of the alternative sample, and Figures 13 and 14 show the average skill growth paths and the average contribution of skills to wage growth, respectively. The parameter estimates are very similar to those of the preferred specification (i.e. percentile scores) and none of the substantive results are different. Although it is impossible to test all possible task complexity index constructions (because there are infinitely many), this exercise provides limited evidence that the main results are insensitive to the choice of the task complexity index construction method.

D.5 Tables

D.5.1 Including part-time jobs.

Table 8: Wage Equation

Notation	Estimates	Std. Error
p_0	2.0572	0.0148
$p_1(1)$	0.3134	0.0067
$p_1(2)$	0.0547	0.0067
$P_2(1,1)$	0.0222	0.0055
$P_2(2,2)$	0.0052	0.0018
σ_η^2	0.0119	0.0004

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 9: Skill Transition

Notation	Estimates	Std. Error
$a_0(1)$	0.0899	0.0156
$a_0(2)$	0.4232	0.2282
$A_1(1,1)$	0.1101	0.0254
$A_1(2,2)$	1.0599	0.2250
$D(1,1)$	0.9323	0.0062
$D(2,2)$	0.5137	0.0365
$\Sigma_\varepsilon(1,1)$	0.1901	0.0111
$\Sigma_\varepsilon(2,2)$	0.5955	0.0984
$H(1,1)$, AFQT	1.2169	0.2320
$H(1,2)$, Black	−0.3856	0.1889
$H(1,3)$, Hispanic	0.2171	0.2132
$H(2,1)$, AFQT	−1.5740	1.3256
$H(2,2)$, Black	1.6165	1.0781
$H(2,3)$, Hispanic	−0.4503	1.0684

Note: The skill transition equation is $s_{t+1} = Ds_t + a_0 + A_1 x_t$. The conditional initial mean skills are given by $E(s_1|d) = Hd$ where d is a vector of AFQT scores and dummy variables for race (whites are the reference group.) The conditional initial skill variance matrix Σ_{s1} is an identity matrix given d, for normalization. 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 10: Work Disutility

Notation	Estimates	Std. Error
$g_0(1)$	-0.1560	0.0670
$G_1(1,1)$, AFQT	0.1163	0.0394
$G_1(1,2)$, Black	-0.0176	0.0166
$G_1(1,3)$, Hispanic	0.0021	0.0164
$g_0(2)$	-0.1072	0.0190
$G_1(2,1)$, AFQT	0.0000	0.0003
$G_1(2,2)$, Black	-0.0005	0.0003
$G_1(2,3)$, Hispanic	-0.0004	0.0003
$G_2(1,1)$	-0.2121	0.0660
$G_2(2,2)$	-0.0139	0.0060
$F_0(1,1)$	-1.4711	0.4830
$F_0(2,2)$	-0.0054	0.0016
$F_1(1,1)$	0.0807	0.0194
$F_1(2,2)$	0.2358	0.0194
$\bar{x}_{1,0}(1)$	0.2249	0.0269
$X(1,1)$, AFQT	0.0398	0.0500
$X(1,2)$, Black	-0.0230	0.0424
$X(1,3)$, Hispanic	-0.0372	0.0466
$\bar{x}_{1,0}(2)$	0.9860	0.4899
$X(2,1)$, AFQT	1.2164	0.8844
$X(2,2)$, Black	-0.9994	0.6838
$X(2,3)$, Hispanic	0.3568	0.6612
$A_2(1,1)$	0.4094	0.0132
$A_2(2,3)$	0.3702	0.0146

Note: The work disutility function is $g(x_t, \bar{x}_t, \nu_t) = (g_0 + G_1 d + \nu_t)' x_t + x_t' G_2 x_t + (x_t - \bar{x}_t)' G_{3,t} (x_t - \bar{x}_t)$ and $G_{3,t} = F_0 \odot \exp(t \cdot F_1)$. The transition equation of work habit is $\bar{x}_{t+1} = A_2 \bar{x}_t + (I - A_2) x_t$ 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 scores 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 12: Skill Transition

Notation	Estimates	Std. Error
$a_0(1)$	0.0769	0.0150
$a_0(2)$	0.3142	0.2088
$A_1(1,1)$	0.1096	0.0249
$A_1(2,2)$	0.9976	0.2153
$D(1,1)$	0.9411	0.0061
$D(2,2)$	0.4668	0.0350
$\Sigma_\varepsilon(1,1)$	0.1895	0.0109
$\Sigma_\varepsilon(2,2)$	0.3495	0.0623
$H(1,1)$, AFQT	1.3085	0.2114
$H(1,2)$, Black	−0.5120	0.1552
$H(1,3)$, Hispanic	−0.0076	0.1963
$H(2,1)$, AFQT	−1.0565	1.0666
$H(2,2)$, Black	0.9686	0.7228
$H(2,3)$, Hispanic	1.4281	0.8149

Note: The skill transition equation is $s_{t+1} = Ds_t + a_0 + A_1 x_t$. The conditional initial mean skills are given by $E(s_1|d) = Hd$ where d is a vector of AFQT scores and dummy variables for race (whites are the reference group.) The conditional initial skill variance matrix Σ_{s1} is an identity matrix given d, for normalization. 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.

D.5.2 No restriction on age at the transition to the full-time labor market.

Table 11: Wage Equation

Notation	Estimates	Std. Error
p_0	2.1182	0.0140
$p_1(1)$	0.3121	0.0067
$p_1(2)$	0.0646	0.0103
$P_2(1,1)$	0.0182	0.0052
$P_2(2,2)$	0.0076	0.0029
σ_η^2	0.0116	0.0004

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 13: Work Disutility

Notation	Estimates	Std. Error
$g_0(1)$	−0.1739	0.0718
$G_1(1,1)$, AFQT	0.1172	0.0458
$G_1(1,2)$, Black	−0.0110	0.0166
$G_1(1,3)$, Hispanic	0.0010	0.0204
$g_0(2)$	−0.1097	0.0189
$G_1(2,1)$, AFQT	0.0000	0.0003
$G_1(2,2)$, Black	−0.0005	0.0003
$G_1(2,3)$, Hispanic	−0.0004	0.0003
$G_2(1,1)$	−0.2037	0.0721
$G_2(2,2)$	−0.0168	0.0081
$F_0(1,1)$	−1.8968	0.6992
$F_0(2,2)$	−0.0101	0.0033
$F_1(1,1)$	0.0559	0.0212
$F_1(2,2)$	0.1974	0.0205
$\bar{x}_{1,0}(1)$	0.2170	0.0233
$X(1,1)$, AFQT	0.1509	0.0432
$X(1,2)$, Black	−0.0304	0.0354
$X(1,3)$, Hispanic	0.0003	0.0385
$\bar{x}_{1,0}(2)$	0.9756	0.3061
$X(2,1)$, AFQT	0.6258	0.5272
$X(2,2)$, Black	−0.5017	0.3521
$X(2,3)$, Hispanic	−0.6730	0.4115
$A_2(1,1)$	0.4041	0.0136
$A_2(2,3)$	0.3898	0.0147

Note: The work disutility function is $g(x_t, \bar{x}_t, \nu_t) = (g_0 + G_1 d + \nu_t)' x_t + x_t' G_2 x_t + (x_t - \bar{x}_t)' G_{3,t}(x_t - \bar{x}_t)$ and $G_{3,t} = F_0 \odot \exp(t \cdot F_1)$. The transition equation of work habit is $\bar{x}_{t+1} = A_2 \bar{x}_t + (I - A_2) x_t$ 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 scores 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.

D.5.3 Self-employed workers are excluded.

Table 14: Wage Equation

Notation	Estimates	Std. Error
p_0	2.1570	0.0171
$p_1(1)$	0.3051	0.0091
$p_1(2)$	0.0361	0.0116
$P_2(1,1)$	0.0114	0.0051
$P_2(2,2)$	0.0076	0.0038
σ_η^2	0.0067	0.0004

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 15: Skill Transition

Notation	Estimates	Std. Error
$a_0(1)$	0.0670	0.0173
$a_0(2)$	0.6338	0.4416
$A_1(1,1)$	0.1604	0.0330
$A_1(2,2)$	1.2306	0.3921
$D(1,1)$	0.9316	0.0075
$D(2,2)$	0.4403	0.0342
$\Sigma_\varepsilon(1,1)$	0.1790	0.0126
$\Sigma_\varepsilon(2,2)$	0.4544	0.0984
$H(1,1)$, AFQT	1.5768	0.2768
$H(1,2)$, Black	−0.4016	0.1990
$H(1,3)$, Hispanic	0.2536	0.2460
$H(2,1)$, AFQT	−3.0234	2.1861
$H(2,2)$, Black	0.7447	1.3125
$H(2,3)$, Hispanic	0.3684	2.0560

Note: The skill transition equation is $s_{t+1} = Ds_t + a_0 + A_1 x_t$. The conditional initial mean skills are given by $E(s_1|d) = Hd$ where d is a vector of AFQT scores and dummy variables for race (whites are the reference group.) The conditional initial skill variance matrix Σ_{s1} is an identity matrix given d, for normalization. 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 16: Work Disutility

Notation	Estimates	Std. Error
$g_0(1)$	-0.3264	0.0779
$G_1(1,1)$, AFQT	0.0641	0.0374
$G_1(1,2)$, Black	-0.0011	0.0136
$G_1(1,3)$, Hispanic	0.0066	0.0153
$g_0(2)$	-0.0773	0.0182
$G_1(2,1)$, AFQT	0.0003	0.0004
$G_1(2,2)$, Black	-0.0004	0.0003
$G_1(2,3)$, Hispanic	-0.0003	0.0004
$G_2(1,1)$	-0.1121	0.0578
$G_2(2,2)$	-0.0189	0.0132
$F_0(1,1)$	-1.4359	0.7267
$F_0(2,2)$	-0.0083	0.0034
$F_1(1,1)$	0.0391	0.0282
$F_1(2,2)$	0.2085	0.0328
$\bar{x}_{1,0}(1)$	0.2238	0.0280
$X(1,1)$, AFQT	0.1581	0.0518
$X(1,2)$, Black	-0.0492	0.0433
$X(1,3)$, Hispanic	-0.0446	0.0429
$\bar{x}_{1,0}(2)$	1.0978	0.7154
$X(2,1)$, AFQT	1.6575	1.4532
$X(2,2)$, Black	-0.5280	0.7559
$X(2,3)$, Hispanic	-0.2027	1.1415
$A_2(1,1)$	0.4357	0.0163
$A_2(2,3)$	0.3900	0.0199

Note: The work disutility function is $g(x_t, \bar{x}_t, \nu_t) = (g_0 + G_1 d + \nu_t)' x_t + x_t' G_2 x_t + (x_t - \bar{x}_t)' G_{3,t} (x_t - \bar{x}_t)$ and $G_{3,t} = F_0 \odot \exp(t \cdot F_1)$. The transition equation of work habit is $\bar{x}_{t+1} = A_2 \bar{x}_t + (I - A_2) x_t$ 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 scores 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.

D.5.4 Standardized DOT scores (instead of percentile scores)

Table 17: Wage Equation

Notation	Estimates	Std. Error
p_0	2.1366	0.0155
$p_1(1)$	0.3036	0.0070
$p_1(2)$	0.0529	0.0090
$P_2(1,1)$	0.0322	0.0077
$P_2(2,2)$	0.0057	0.0024
σ_η^2	0.0107	0.0004

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 18: Skill Transition

Notation	Estimates	Std. Error
$a_0(1)$	0.0737	0.0175
$a_0(2)$	0.4688	0.2949
$A_1(1,1)$	0.1499	0.0384
$A_1(2,2)$	1.0549	0.2415
$D(1,1)$	0.9376	0.0068
$D(2,2)$	0.4212	0.0365
$\Sigma_\varepsilon(1,1)$	0.1943	0.0119
$\Sigma_\varepsilon(2,2)$	0.2908	0.0724
$H(1,1)$, AFQT	1.3880	0.2392
$H(1,2)$, Black	−0.4130	0.1759
$H(1,3)$, Hispanic	0.0893	0.2055
$H(2,1)$, AFQT	−1.8461	1.4499
$H(2,2)$, Black	1.2602	0.9600
$H(2,3)$, Hispanic	2.0457	1.1006

Note: The skill transition equation is $s_{t+1} = Ds_t + a_0 + A_1 x_t$. The conditional initial mean skills are given by $E(s_1|d) = Hd$ where d is a vector of AFQT scores and dummy variables for race (whites are the reference group.) The conditional initial skill variance matrix Σ_{s1} is an identity matrix given d, for normalization. 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 19: Work Disutility

Notation	Estimates	Std. Error
$g_0(1)$	−0.1558	0.1098
$G_1(1,1)$, AFQT	0.1621	0.0614
$G_1(1,2)$, Black	−0.0211	0.0289
$G_1(1,3)$, Hispanic	0.0088	0.0340
$g_0(2)$	−0.0897	0.0215
$G_1(2,1)$, AFQT	0.0000	0.0003
$G_1(2,2)$, Black	−0.0004	0.0002
$G_1(2,3)$, Hispanic	−0.0002	0.0003
$G_2(1,1)$	−0.4180	0.1318
$G_2(2,2)$	−0.0126	0.0069
$F_0(1,1)$	−3.9082	1.3336
$F_0(2,2)$	−0.0097	0.0034
$F_1(1,1)$	0.0624	0.0235
$F_1(2,2)$	0.1648	0.0227
$\bar{x}_{1,0}(1)$	0.2262	0.0192
$X(1,1)$, AFQT	0.1088	0.0354
$X(1,2)$, Black	−0.0305	0.0301
$X(1,3)$, Hispanic	−0.0159	0.0322
$\bar{x}_{1,0}(2)$	0.7321	0.3010
$X(2,1)$, AFQT	0.7590	0.5516
$X(2,2)$, Black	−0.4816	0.3602
$X(2,3)$, Hispanic	−0.7846	0.4450
$A_2(1,1)$	0.4147	0.0150
$A_2(2,3)$	0.3773	0.0159

Note: The work disutility function is $g(x_t, \bar{x}_t, \nu_t) = (g_0 + G_1 d + \nu_t)' x_t + x_t' G_2 x_t + (x_t - \bar{x}_t)' G_{3,t} (x_t - \bar{x}_t)$ and $G_{3,t} = F_0 + t \cdot F_1$. The transition equation of the past job characteristics is $\bar{x}_{t+1} = A_2 \bar{x}_t + (I - A_2) x_t$ where I is a (2×2) identity matrix. The initial condition of the past job characteristics is given by $\bar{x}_1 = \bar{x}_{1,0} + Xd$ where d is a vector of AFQT scores 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.

D.6 Figures

D.6.1 Including part-time jobs

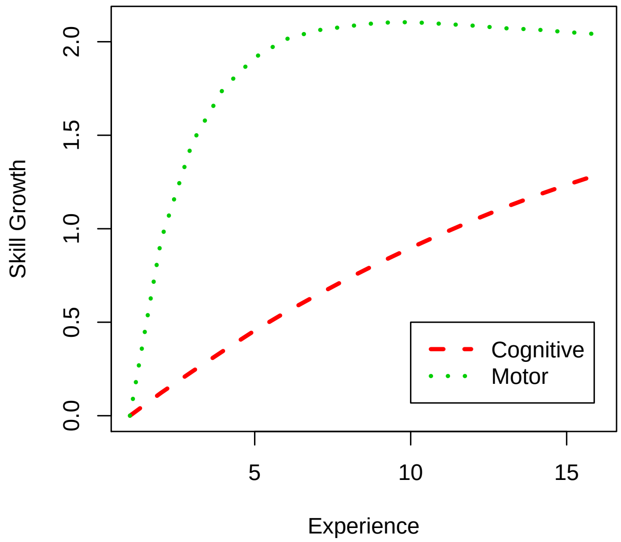


Figure 7: Skill Growth Profiles

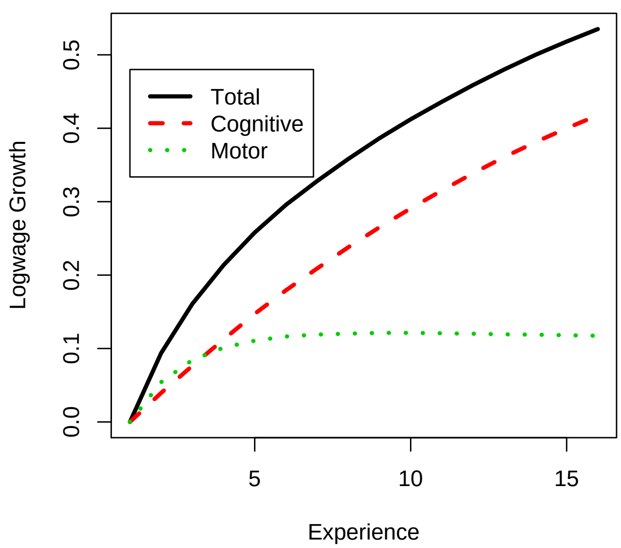


Figure 8: Contribution of Skills to Wage Growth

D.6.2 No restriction on age at the transition to the full-time labor market.

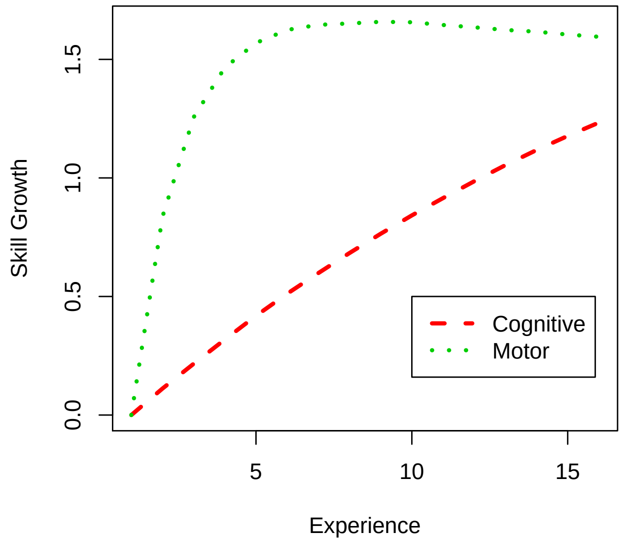


Figure 9: Skill Growth Profiles

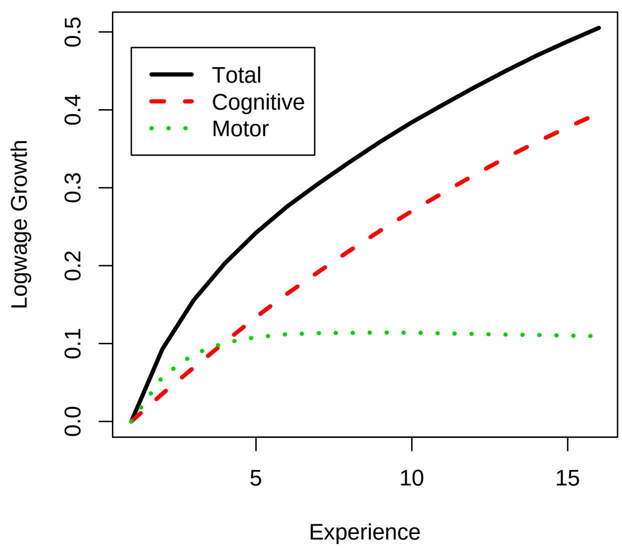


Figure 10: Contribution of Skills to Wage Growth

D.6.3 Self-employed workers are excluded.

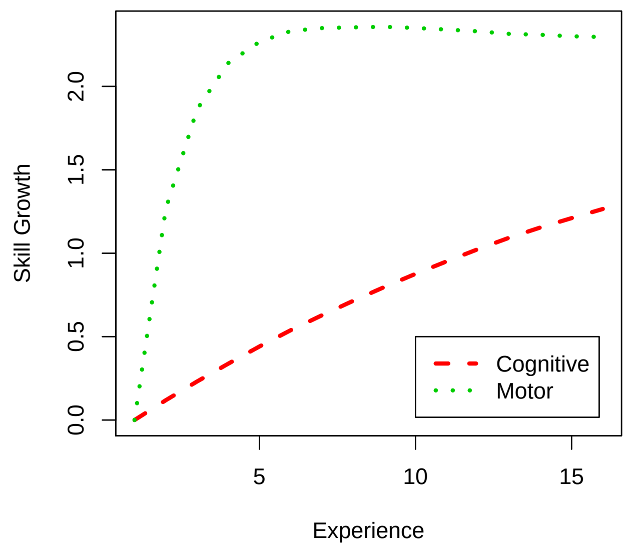


Figure 11: Skill Growth Profiles

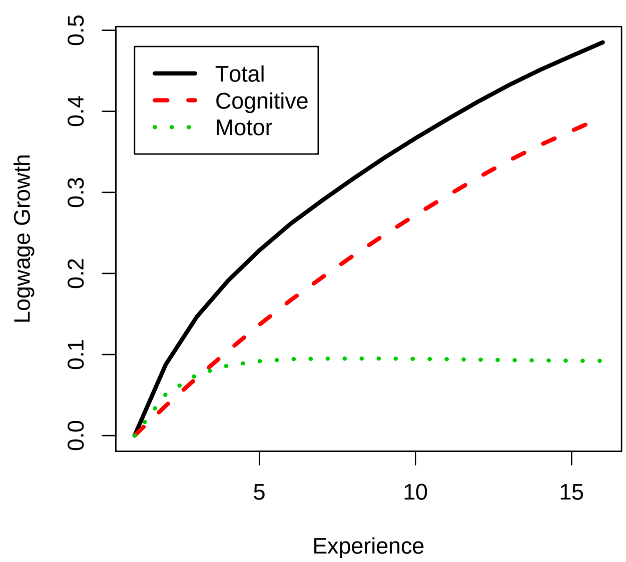


Figure 12: Contribution of Skills to Wage Growth

D.6.4 Standardized DOT scores (instead of percentile scores)

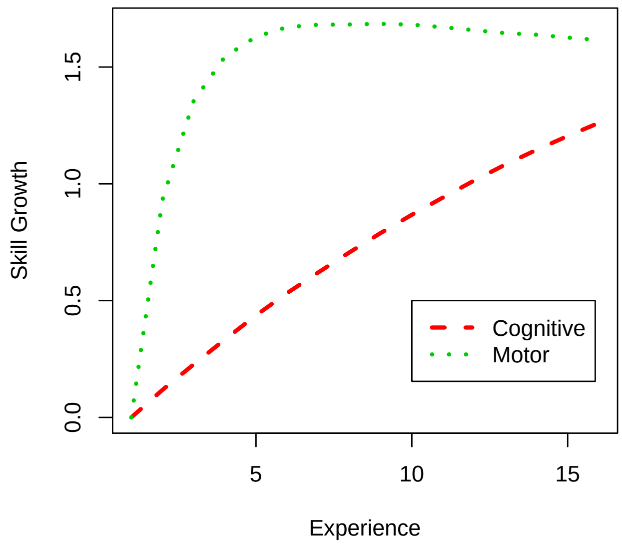


Figure 13: Skill Growth Profiles

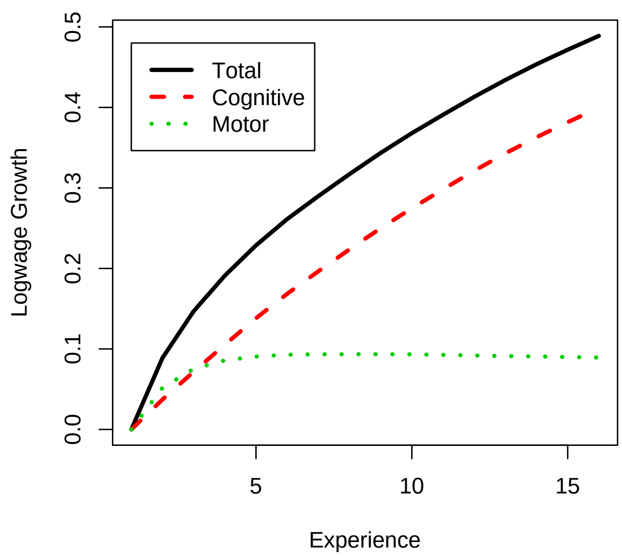


Figure 14: Contribution of Skills to Wage Growth