Skill Formation and Utilization in the Labor Market

Shintaro Yamaguchi*

May 2008

VERY PRELIMINARY AND INCOMPLETE

COMMENTS WELCOME

Abstract

This paper attempts to understand occupational mobility and wage growth of young workers in terms of basic labor market skills; cognitive, interpersonal, and motor skills. I emphasize that occupations are heterogeneous in formation and utilization of these skills. The model is structurally estimated using occupational task measures from the Dictionary of Occupational Titles and career history from the NLSY79. I find that observed occupational choice is largely rationalized by skills, not by preference. The results also indicate that wage growth of high school graduates are driven by the growth of motor skills at the very early career. However, interpersonal skills take over the position as the main contributor to the wage growth subsequently.

1 Introduction

Labor market skills are inherently multidimensional and usefulness of heterogeneous skills are substantially different across jobs. However, for simplicity, most empirical models ignore heterogeneity of skills and jobs or analyze them in a quite limited way. This paper attempts to explain wage growth and occupational mobility in terms of low-dimensional basic skills; cognitive, interpersonal, and motor skills. I emphasize that skills are differently produced and rewarded across occupations. The model is structurally estimated using occupational task complexity measures

*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.

from the Dictionary of Occupational Titles (DOT) and career histories of white males from the National Longitudinal Survey of Youth 1979 (NLSY).

In the model, occupations are characterized by a task mix and substantive complexity of tasks. For example, some occupations require more interpersonal communication, while others require more manual dexterity. In addition, occupations that use a similar set of tasks can be ranked by substantive complexity. Individuals have multidimensional labor market skills, which are developed through learning-by-doing. The amount of learning depends on how intensely skills are used to perform tasks in the current occupation. Occupations also differ in how they reward worker skills. Individuals have to enter an occupation where their comparative advantage is appreciated to be rewarded. For example, mathematicians are not rewarded for their physical strength and laborers are not rewarded for their analytical skills even if they have those skills plentifully. In addition, skill levels are important to be better rewarded. In occupations where a certain type of skill is intensely used, the labor productivity is sensitive to the worker's skill level. A low-skill worker would choose to work in a less skill-intensive occupation to avoid being unproductive, while a high-skill worker would enter a skill-intensive occupation to fully utilize her skill. Hence, workers are sorted across occupations in terms of their skill level as well as their comparative advantages. As individuals develop their skills over the career through learning-by-doing, they gradually move up to more skill-intensive occupations.

This paper contributes to the literature of human capital formation. The earliest studies such as Ben-Porath (1967) develop human capital investment models to interpret life-cycle earnings profiles. Most papers in the literature assume human capital is single-dimensional for simplicity. Recently, Heckman, Stixrud, and Urzua (2006) propose a two-skill model and find evidence that not only cognitive, but also noncognitive skills strongly influence labor market outcomes and social behavior. Cunha and Heckman (2007) model the technology of human capital production where multidimensional skills are formed over life cycle. These papers primarily focus on skill formation before individuals enter the labor market and multidimensional skills are equally utilized in all jobs. The present paper is complement to these papers, because it provides an extensive analysis of post-schooling skill formation and utilization.

Another strand of literature considers how workers are rewarded through occupational self-selection, which is pioneered by Roy (1951). In Willis (1986), skills are occupation specific and workers self-select into occupations on their skill quantity and prices. Keane and Wolpin (1997) combine features of Roy model and human capital production. Rubinstein and Weiss (2006) also presents a theoretical model in which multidimensional skills are differently formed and utilized across occupations. In these models, occupations are horizontally differentiated and worker allocation is driven by comparative advantage. In contrast, Gibbons, Katz, Lemieux, and Parent (2005) constructs a model where workers are sorted across vertically differentiated occupations. Empirical evidence strongly suggests that occupations are heterogeneous both in substantive task complexity and in task combination. This paper departs from the previous contributions by providing a unified analysis of upward occupational mobility and comparative advantage.

The model is structurally estimated by the Maximum Likelihood. The parameter estimates are intuitive and the model fit to some interesting features of the data is reasonably well. Main empirical findings are as follows. First, observed occupational choice is largely rationalized by worker skills. A variance decomposition of task complexity measures indicates that more than a half of the variance is explained by skill heterogeneity. Preference heterogeneity plays a relatively small role in explaining occupational difference across individuals. Second, the important contributor of wage growth for high school graduate workers switches from motor skill to interpersonal skill early in their careers, although cognitive skill is the biggest contributor throughout the sample period.

The rest of the paper is organized as follows: Section 2 describes the data set including the occupational characteristics in the DOT and occupational histories from the NLSY. The main patterns of the data are also reported in this section. Section 3 lays out the model. Section 4 discusses the estimation strategy. The estimation results are presented and discussed in Section 5. Section 6 concludes.

2 Data

2.1 Dictionary of Occupational Titles

The occupational characteristics are drawn from the 1994 revised fourth edition of the Dictionary of Occupational Titles (DOT). The occupational definitions are provided for 12,099 occupations in terms of 44 characteristics, which are based on information collected between 1978 and 1990. The survey period of the DOT largely matches that 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 generate occupational definitions through their on-site observations of jobs and information obtained from 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.

The occupational characteristics in the DOT are aggregated to occupations defined by the 1970 Census 3-digit classification system, because the DOT contains more occupations than the Census classification. To construct occupational characteristics for the Census classification, I use the April 1971 Current Population Survey augmented by the fourth edition of the DOT which was compiled by the Committee on Occupational Classification and Analysis at the National Academy of Sciences. Notice that this augmented CPS file contains occupation code for the fourth edition of the DOT, not the revised fourth edition. Some occupations are deleted, or integrated into other occupations, while some are newly added in the revised fourth edition. I update the occupation code in the augmented CPS file using the conversion table in the revised fourth edition. Following Autor, Levy, and Murnane (2003), the DOT variables are converted into percentile scores, because most DOT variables are ordinal or even binary. Occupational characteristics for each occupation in the 1970 Census classification are constructed by averaging over individuals in the augmented CPS file.

As I examine the textual definitions of the DOT variables, I assume that occupational characteristics can be categorized into three broad types of tasks or skill requirements. These categories are originally considered by Bacolod and Blum (2005) and similar ones are also found in Ingram and Neumann (2006). The first type is cognitive requirement. The DOT variables that measure cognitive requirement include the following seven variables: worker functions related to data, reasoning development, mathematical development, language development, verbal aptitude, numerical aptitude, and a temperament of making judgments and decisions. The second type of task is interpersonal skill requirement. This is captured by the following four DOT variables: worker functions related to people, a temperament of influencing people in their opinions, attitudes, and judgments, a temperament of directing, controlling, or planning activities of others, and a temperament of dealing with people. The third type of task is motor-skill requirement, which is measured by the following 9 variables: worker functions related to things, spatial aptitude, form perception, motor coordination, finger dexterity, manual dexterity, eye-hand-foot coordination, color discrimination, and a temperament of attaining precise set limits, tolerances, and standards.

It is theoretically possible to include all of these 20 variables in the structural model presented in the following section to extract occupational information completely. However, I summarize the information by a small number of variables, which is particularly useful for descriptive analysis. To reduce the dimension of task measures, I simply take an average over tasks within each task category, which I call standardized scores of task complexity. An alternative conventional method is to construct factor scores by a principal component analysis. But, all results in this paper are essentially unchanged.

To see if the constructed standardized scores reasonably characterize occupations, mean and standard deviation of task complexity are reported for each 1-digit occupation in Table 1. Tasks of professionals require the highest level of cognitive skills, which is followed by managers. Laborers and household service workers are required the lowest level of cognitive skills. This cognitive task complexity measure largely matches the conventional notion of skill in the empirical literature where skill is single dimensional. However, this index alone is not rich enough to describe het-

Table 1: Task Complexity by Occupation at 1 Digit Classification

	Mean			Std. Dev.			Nobs.
	Cognitive	Interpersonal	Motor	Cognitive	Interpersonal	Motor	
Professional	0.675	0.453	0.299	0.144	0.255	0.265	7522
Manager	0.596	0.456	0.115	0.148	0.154	0.159	5538
Sales	0.353	0.455	0.131	0.161	0.161	0.106	3748
Clerical	0.347	0.164	0.271	0.149	0.141	0.129	9270
Craftsmen	0.387	0.106	0.530	0.150	0.175	0.156	6557
Operatives	0.126	0.032	0.343	0.142	0.084	0.167	5824
Transport	0.132	0.146	0.376	0.104	0.171	0.090	1774
Laborer	0.077	0.037	0.225	0.115	0.087	0.158	2818
Farmer	0.491	0.277	0.351	0.139	0.096	0.061	1117
Farm Laborer	0.134	0.040	0.295	0.140	0.097	0.138	882
Service	0.182	0.161	0.255	0.160	0.130	0.181	6834
HH Service	0.072	0.140	0.099	0.060	0.096	0.106	1469
ALL	0.350	0.226	0.285	0.246	0.228	0.207	53353

Source: The Revised Fourth Edition of the DOT (1991) and the augmented 1971 April CPS.
Note: HH service is household service occupation.

erogeneous tasks across occupations. For example, cognitive task complexity is similar between sales occupations and craft occupations. But, the nature of tasks are very different; sales workers communicate with their customers and craft workers use tools and labor to make things.

Interpersonal skill requirement and motor skill requirement more clearly characterize the nature of occupation than a single skill index. Interpersonal skills are useful in professional, managerial, and sales occupations. In these occupations, workers have to direct their subordinates and persuade their clients. Laborers and household service workers use little of their interpersonal skills, because their tasks do not involve interactions with people. Motor skills are most required by tasks of craftsmen such as automobile mechanics and carpenters. Tasks of household service workers, managers, and sales workers require little motor skills. These patterns are quite intuitive, and thus, the statistics provide evidence for usefulness of the DOT task measures.

2.2 National Longitudinal Survey of Youth 1979

The data for career history are taken from the NLSY which includes information on the weekly work history of individuals from 1978. The survey subjects comprise individuals who were between 14 and 21 years old as of January 1, 1979. The NLSY is particularly suitable for this study because it contains a detailed career history of individuals and because it focuses on young workers who change occupations more frequently than older workers. The DOT variables are added to the NLSY using the 1970 Census three-digit occupation code. Observations from 1979 through 1994 are used in the analysis, because occupation change is not reported on an annual basis in later surveys. [1]

2.2.1 Sampling Criteria

A sample of white male high school graduates is taken from the NLSY, because this is a relatively homogeneous demographic group and their labor force attachment is strong. The present paper focuses on high school graduates because they make career progression through occupational changes, while college graduates do so within the same occupation such as professionals and managers. Yamaguchi (2007, 2008) provide evidence for these career patterns across education groups.

I concentrate on individuals who make a long-term transition to the labor market during the sample period. I define a long-term transition to occur when an individual spends 3 consecutive years primarily working. An individual is classified as “primarily working” if he worked (1) at least 26 weeks and (2) 30 hours per week in the job where he spends the longest hours in a given year.

There are 2,236 white males in the NLSY. I dropped 305 individuals who did not make a long-term transition to the labor market during the sample period. At the time of the long-term transition, 855 individuals graduated from high school. When a 3-digit occupation code is missing after the long-term transition, observations of the individual from that year on are dropped from the sample. After removing individuals whose occupation code at the time of the long-term transition is missing, 768 individuals still remain in the sample. I have 6,671 person-year observations of occupational choice.

Hourly wages are deflated by the 2002 CPI. Some recorded hourly wages are extremely high or low. If the recorded hourly wage is greater than \$100 or less than one dollar, they are regarded as missing. I finally have 6,420 person-year observations of logwage.

Summary statistics of the sample are reported in Table 2. Labor market experience is measured

[1] In surveys later than 1994, an occupation change can be identified only when an individual also changes employers.

Table 2: Summary Statistics

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Nobs.
Logwage	0.005	2.117	2.430	2.421	2.733	4.428	6420
Experience	0.000	2.000	4.000	4.974	8.000	15.000	6671
Cognitive	0.013	0.122	0.286	0.304	0.489	0.881	6671
Interpersonal	0.000	0.032	0.064	0.153	0.242	0.765	6671
Motor	0.005	0.172	0.333	0.341	0.465	0.732	6671

Source: NLSY and DOT

Note: Nobs. is the number of observations in the pooled sample (person-year.) There are 768 white male high school graduates in the sample.

by the number of years after the long-term transition to the labor market. Thus, years of experience plus one equals the number of periods in the sample. On average, observations are available for 6 years for an individual. Skill requirement indexes in Table 2 are averages of percentile scores of task complexity within each skill category (summary statistics for all 20 tasks are reported in Table 7.) The mean skill requirement indexes suggest that high school graduates have comparative advantage in motor skills. This is consistent with the fact that the proportion of craftsmen and operatives are higher than other education groups.

2.3 Dynamic Occupational Mobility

Figure 1 shows the evolution of task complexity. Both cognitive skill requirement and interpersonal skill requirement increase with labor market experience. At the time of long-term transition to the labor market, average task complexity indexes of cognitive tasks and interpersonal tasks are 0.24 and 0.11, respectively. They are comparable to average cognitive requirement of service workers and average interpersonal skill requirement of craftsmen. Individuals take on more and more complex cognitive and interpersonal tasks over time; cognitive task complexity index reaches 0.34 in 10 years and 0.39 in 15 years. The cognitive requirement at 15 years of experience is comparable to that of sales, clerical and crafts occupations. Similarly, interpersonal task complexity index reaches 0.19 in 10 years and 0.26 in 15 years. The upward trends of cognitive and interpersonal task complexity indexes are statistically significant, when a task complexity index is regressed on experience. The reported upward occupational mobility is consistent with the fact that high school graduate workers are gradually promoted to managers and craftsmen. Increasing cognitive requirement reflects increasing proportions of these two occupations. But, increasing interpersonal skill requirement is mainly driven by promotions to manager, because tasks of craftsmen do not include many interactions with people.

In contrast to cognitive and interpersonal task complexity, growth of motor skill requirement is not monotonic. At the time of the long-term labor market transition, the average motor skill requirement is 0.32, which is close to the motor skill requirement for operatives. The index peaks

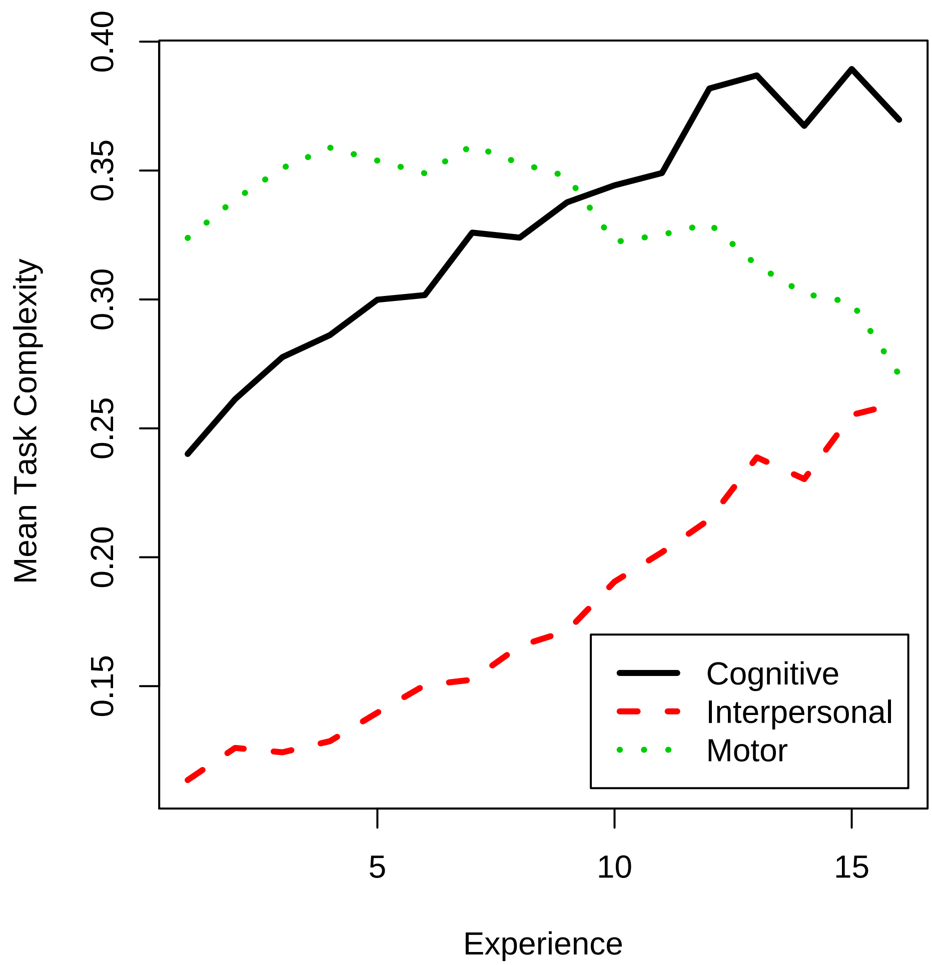

Figure 1: Dynamic Occupational Mobility

around 0.36 in 4-7 years of labor market experience. Then, it gradually decreases and reaches about 0.30 in 15 years, which is below the initial motor-skill requirement. This hump-shaped motor skill requirement profile is explained by the following occupational mobility patterns. 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-skill requirement index. 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. This switch from motor-skill intensive tasks into interpersonal-skill intensive tasks is captured by a declining motor-skill requirement.

2.4 Workers Move To A Similar Occupation.

How do workers move across occupations? The previous subsection describes the average time profiles of task complexity of high school graduate workers. I provide evidence that the observed patterns are achieved by workers moving to an occupation similar to the current one, not by workers randomly moving without directions.

To see if workers move to a similar occupation, I measure similarity of a pair of occupations by the Mahalanobis distance,

$$d(x_A, x_B) = (x_A - x_B)' \Sigma^{-1} (x_A - x_B) \tag{1}$$

where x_A and x_B are task vectors of occupation A and B, and Σ is the variance-covariance matrix of task complexity vectors. The Mahalanobis distance is scale-invariant and accounts for the correlation of observations between task dimensions. The Euclidean distance is incorporated as a special case in which Σ is an identity matrix.

Table 3 presents summary statistics of observed moves using a sample of individuals who change occupations between two survey years. To see if individuals move to a similar occupation, I compare observed occupational moves with predicted occupational moves when individuals move at random without direction. Formally, I say that an occupational move is non-directional when

$$F_B(x_B|x_A) = F_B(x_B) \tag{2}$$

where x_A and x_B are task complexity vectors of the origin and the destination, respectively, and F_B is a cumulative distribution function of x_B. Assuming that equation (2) holds, the conditional mean occupational move given the current occupation is $\bar{d}(x_A) = \int d(x_A, B) dF_B$. Thus, the unconditional mean occupational move $\bar{d}$ is given by integrating out x_A. I estimated F_B and the distribution of x_A nonparametrically using a sample of occupational changers.

Table 4 reports the point estimate of $\bar{d}$ and its standard deviation. If an occupational move is non-directional, the mean move would be 14.384, which is much greater than the sample mean, 4.330. The null hypothesis that the predicted mean move equals the average of observed moves is strongly rejected. Moreover, the predicted mean move is greater than the observed move at the 95th percentile.

This result is consistent with predictions of models where skills are multidimensional. In Lazear (2004), skills are fully transferred across occupations, but they are used by different weights. A worker would suffer wage loss if she moves to an occupation where her comparative advantage is underutilized. Gibbons and Waldman (2006) also propose a similar concept, i.e., task specific skills. In their model, workers accumulate task specific skills through learning-by-doing. A worker would lose the task specific skills if she moves to a job characterized by a different task from the current one. In both models, a worker typically moves to an occupation with similar tasks to avoid wage loss due to under-utilization or loss of skills. An alternative interpretation of the result is based on task preference, which is missing in both papers. A worker may choose similar jobs, because she likes the task common to those jobs. The structural model presented in the subsequent section accounts for both multidimensional skills and task preference, because the skill explanation and the preference explanation are complement for understanding occupational mobility.

Table 3: Observed Move of Occupational Change

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Nobs.
Obs. Distance	0.006	1.000	2.997	4.330	6.747	43.960	3697

Note: Computed from a sample of occupational changers. Distance between the current and the new occupations is defined by the Mahalanobis distance (see equation 1 for definition.)

Table 4: Predicted Mean Move of Occupational Change When Move is Non-Directional

	Estimates	Std. Dev.
$\bar{d}$	14.384	0.180

3 Model

I construct a model of skill formation and occupational mobility where skills are multidimensional and occupations are heterogeneous in learning opportunity and utilization of skills. The model is capable of explaining the key features of the data presented in the preceding section and the empirical findings in the literature.

The model considers occupational choice of high school graduate workers after they make a long-term transition to the labor market. Decisions of skill investment and utilization are made through choice of occupations. Individuals have a J-dimensional vector of skills s. Examples of skills include cognitive skill, interpersonal skill, motor skill and so on. Occupations are characterized by a K-dimensional vector of task complexity x. All occupations involve all of K tasks, but task complexity is different across occupations. To perform a task, one of J skills is used and the same skill can be used for similar, but different tasks. For example, cognitive skills can be used for numerical analysis and reading comprehension. This implies that task dimension is equal or greater than skill dimension ($K \geq J$). Let x^j be a sub-vector of tasks that use skill j, i.e., $x = \{x^1, \cdots, x^J\}$. The number of elements in the vector x^j is K_j and $K = \sum_{j=1}^{J} K_j$.

Wage Function Skills are differently utilized across occupations according to task complexity. Let $p^j(x_t^j)$ be a measure of utilization of skill j when complexity of the relevant tasks in period t

is x_t^j. Wages are given by the following equation

$$\ln w_t \quad = \quad c(x_t) + \sum_{j=1}^{J} p^j(x_t^j) s_t^j \tag{3}$$

where subscript t is for period and $c(x_t)$ captures a pure occupation effect, because it is constant across individuals within the same occupation regardless of worker skills. I assume that skills are more utilized with task complexity ($\partial p^j(x_j)/\partial x^j > 0$) at a non-increasing rate ($\partial^2 p^j(x_j)/\partial^2 x^j \leq 0$), and that skills are productive ($p^j(x_j) > 0$) for all level of task complexity x^j. I also assume that a pure occupation effect $c(x)$ is decreasing and convex in task complexity $\partial c/\partial x < 0$ and $\partial^2 c/\partial^2 x < 0$.

Under this wage function, it is straightforward to show that the wage-maximizing task complexity is increasing in the relevant skill. When tasks are simple, worker skills little affect productivity. Simple tasks, such as house keeping, can be satisfactory performed by a low-skill worker. In addition, she is unlikely to be far outperformed by a high-skill worker in such a task. In contrast, productivity of complex tasks is sensitive to worker skills. Managerial task is an example of skill sensitive tasks. Because quality of a manager affects productivity of her subordinates, a small difference in managerial skills can be translated into a large productivity difference. Thus, a low-skill worker is likely to perform managerial tasks quite poorly and may not produce anything. If this is the case, a low-skill worker is paid better in a simple task. Similarly, a high-skill worker is paid better in a complex task, because her skills are better utilized. This relationship between skills, wages, and tasks are illustrated in Figure 2.

The proposed wage function combines an efficient worker allocation model by Gibbons, Katz, Lemieux, and Parent (2005) and the skill-weight approach by Lazear (2004). In Gibbons, Katz, Lemieux, and Parent (2005), occupations are vertically differentiated in returns to a single dimensional unobserved skill. They show that high-skill workers are gradually sorted into occupations that offer higher returns to the skill, as worker skills are gradually revealed. Their model provides a plausible interpretation on growths of cognitive task complexity and wage of young workers. However, heterogeneous nature of occupations can be captured even better when this is combined with multidimensional tasks and skills. In particular, comparative advantage of workers is easily incorporated in the model. In Lazear's skill-weight approach, jobs are horizontally differentiated in skill weight. A worker optimally chooses a job so that her comparative advantage is heavily used; some workers prefer interpersonal skill demanding jobs while others prefer motor skill demanding jobs. The proposed wage function combines these two previous models to allow for a unified analysis of upward occupational mobility and comparative advantage.

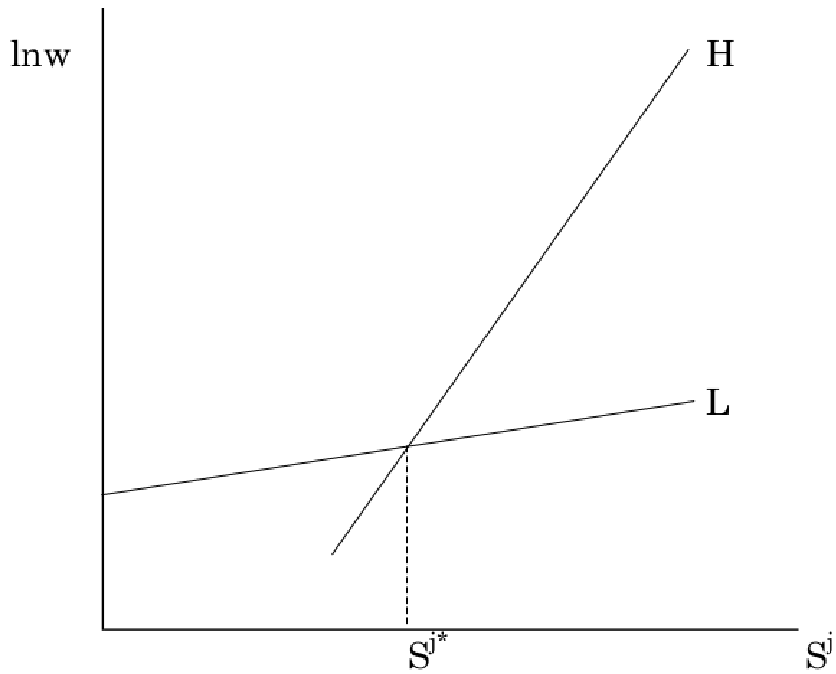

Figure 2: Skill Utilization and Wages Across Occupations

The wage equation is parametrized as follows in a matrix notation

$$\ln w_t = [c_0 + c_1^{'} x_t + x^{'} C x_t] + [p_0 + P^{'} x_t]^{'} s_t \tag{4}$$

where $c_1^{'} = (c_1, \cdots, c_K)$ is a K dimensional vector, C is a negative definite $(K \times K)$ matrix, $p_0^{'} = (p_0^1, \cdots, p_0^J)$ is a J dimensional vector of the minimum skill utilization, and P is a $(K \times J)$ matrix of skill utilization parameters such that

$$P = \begin{pmatrix} p_1^1 & 0 & & 0 \\ \vdots & \vdots & & \vdots \\ p_{K_1^1} & 0 & & \vdots \\ 0 & p_1^2 & & \vdots \\ \vdots & \vdots & \cdots & \vdots \\ \vdots & p_{K_2}^2 & & 0 \\ \vdots & 0 & & p_1^J \\ \vdots & \vdots & & \vdots \\ 0 & 0 & & p_{K_J}^J \end{pmatrix}$$

where p_k^j is the marginal utilization of skill j with respect to task k. It is assumed that all elements in p_0 and P are positive, while all elements in c_1 is negative. I allow for permanent heterogeneity in wage level. Specifically, the wage intercept c_0 is normally distributed with mean $\bar{c}_0$ and variance σ_c^2. Because this wage component is unaffected by skills or tasks, it does not affect a worker's behavior.

Skill Learning Skills are acquired through learning-by-doing. Individuals learn more skills by using them intensely. Let $a^j(x_t^j)$ be a gross increase of skill j when complexity of the relevant tasks in the current occupation is x_t^j. Evolution of skill is given by

$$s_{t+1}^j = \delta^j s_t^j + a^j(x_t^j) + \varepsilon_{t+1}^j$$

where

$$a^j(x_t^j) = a_0^j + \sum_{k=1}^{K_j} a_k^j x_{k,t}^j,$$

$1 - \delta^j$ is a depreciation rate of skill j, and ε_t^j is a stochastic component of a change of skill j. A vector of skill shock $\varepsilon_t = (\varepsilon_t^1, \cdots, \varepsilon_t^J)$ is normal, independent and identically distributed with mean zero and variance Σ_ε. The parameter a_k^j captures the marginal learning of skill j with respect to task k, $x_{k,t}^j$. Ability of learning is allowed to vary across individuals. Specifically, a vector of parameters $a_0 = (a_0^1, \ldots, a_0^J)$ is normally distributed with mean $\bar{a}_0$ and variance Σ_a. Because a learning ability is assumed constant over time, I do not explicitly take into account this heterogeneity in solving the model. It is explicitly treated when constructing the likelihood.

A possible alternative to this learning-by-doing assumption is an on-the-job training (or skill investment) model by Ben-Porath (1967). Although these two models of skill formation have different economic implications, they are hard to be distinguished without training data (see Heckman, Lochner, and Cossa (2002) for an extensive discussion.) However, using the NLSY, Kuruscu (2006) finds that an increase in lifetime income from training is less than 1 percent. This suggests that on-the-job training plays a limited role in understanding a long-term wage growth.

The skill evolution can be written in a matrix notation as follows

$$s_{t+1} = Ds_t + a_0 + A' x_t + \varepsilon_{t+1} \tag{5}$$

where D is a $(J \times J)$ diagonal matrix of which (j, j) element is δ^j, $a_0 = (a_0^1, \cdots, a_0^J)$ is a J dimensional vector, and A is a $(K \times J)$ matrix of parameters of marginal skill learning with respect to tasks such that

$$A = \begin{pmatrix} a_1^1 & 0 & & 0 \\ \vdots & \vdots & & \vdots \\ a_{K_1^1} & 0 & & \vdots \\ 0 & a_1^2 & & \vdots \\ \vdots & \vdots & \dots & \vdots \\ \vdots & a_{K_2}^2 & & 0 \\ \vdots & 0 & & a_1^J \\ \vdots & \vdots & & \vdots \\ 0 & 0 & & a_{K_J}^J \end{pmatrix}.$$

Job Preference I allow for occupational choice based on preference for tasks, which has two features. First, job preference is heterogeneous across individuals. Some individuals may prefer cognitive-skill demanding tasks, while other may prefer motor-skill demanding tasks, for example. Not only difference in skill endowment, but also difference in job preference can also rationalize observed occupational choices. Second, disutility of work increases with task complexity. Physically demanding tasks certainly make an individual fatigued. Similarly, highly cognitive demanding tasks can exhaust a worker mentally. This preference story also explains why some individuals do not enter a high skill occupation, which complements the explanations based on qualification and skills.

The disutility of work is given by the following quadratic function of tasks,

$$v_t = g' x_t + x_t G x_t$$

where G is a $(K \times K)$ negative semi-definite matrix and a vector of parameters g is normally distributed with mean $\bar{g}$ and variance Σ_g. I assume that g is constant over time. Thus, this heterogeneity is not explicitly treated when solving the model. It is considered when constructing the likelihood.

Bellman Equation The objective of an individual is to maximize net utility over infinite horizon. When individuals enter the labor market, they possess the initial labor market skill s_1 which is normal, independent, and identically distributed with mean μ_1^s and covariance matrix Σ_1^s. The

Bellman equation for an individual is given by

$$\begin{aligned} V(s_t) &= \max_{x_t} \ln w(x_t, s_t) + v(x_t, s_t) + \beta EV(s_{t+1}) \qquad (6) \\ \text{s.t.} & \\ s_{t+1} &= Ds_t + a_0 + A^{'} x_t + \varepsilon_{t+1} \\ \ln w_t &= c_0 + c_1^{'} x_t + x_t^{'} C x_t + [p_0 + P^{'} x_t]^{'} s_t \\ v_t &= g^{'} x_t + x_t G x_t. \end{aligned}$$

Because this is a stochastic optimal linear regulator problem, as proved in Appendix B, the value function is a quadratic function of skills and the optimal policy function is given by a linear function of skills

$$\begin{aligned} v(s_t) &= q_0 + q_1^{'} s_t + s_t^{'} Q s_t \qquad (7) \\ x_t^* &= b + Bs_t \qquad (8) \end{aligned}$$

where

$$\begin{aligned} q_0 &= c_0 + [(c_1 + g)^{'} + b^{'}(C + G)]b + \beta \left[q_0 + \left[q_1^{'} + (a_0 + A^{'} b)^{'} Q \right] (a_0 + A^{'} b) + trQ\Sigma_\varepsilon \right] \qquad (9) \\ q_1 &= p_0 + B^{'}(c_1 + g) + (P^{'} + 2B^{'}(C + G)^{'})b + \beta \left[(D + A^{'} B)^{'} [q_1 + 2Q^{'}(a_0 + A^{'} b)] \right] \qquad (10) \\ Q &= B^{'} [P + (C + G)B] + \beta (D + A^{'} B)^{'} Q (D + A^{'} B) \qquad (11) \\ b &= -\frac{1}{2} [C + G + \beta AQA^{'}]^{-1} [c_1 + g + \beta A(q_1 + 2Qa_0)] \qquad (12) \\ B &= -\frac{1}{2} [C + G + \beta AQA^{'}]^{-1} [P + 2\beta AQD]. \qquad (13) \end{aligned}$$

Numerical solutions to the value function and the policy function are obtained by a policy function iteration, although many other methods are available, as surveyed by Anderson, Hansen, McGrattan, and Sargent (1996). More specifically, first set $q_0 = q_1 = Q = 0$. Substitute them into the right hand side of equations (12) and (13). The resulting b and B are then substituted into equations (9), (10), and (11) to update the value function. This operation is repeated until convergence is achieved.

The optimal choice of task complexity is better understood by seeing the Euler equation. Substituting the value function (7) and skill transition equation (5) into the right hand side of equation (6) and differentiating with respect to task complexity x_t, I find

$$Ps + \beta A^{'} \frac{\partial EV}{\partial s} = -(c_1 + 2Cx^* + g + 2Gx^*).$$

The left hand side collects marginal returns from task complexity, while the right hand side collects marginal costs. The first term on the left hand side is the marginal logwage return due to skill utilization and the second term captures the marginal return to skill learning. The first two terms on the right hand side are the marginal logwage loss because the pure occupation effect on wages decreases with task complexity. The last two terms are the marginal disutility of work. At the optimum, the marginal return and costs are equal.

I show that skills converge to the steady-state level under certain conditions. Substituting the policy function (8) into the skill transition function (5), I have a system of difference equations for skills

$$s_{t+1} = a_0 + A^{'}b + \varepsilon_{t+1} + (D + A^{'}B)s_t. \tag{14}$$

If the absolute value of all eigenvalues of the matrix $D + A^{'}B$ is less than one, the system converges so that

$$\lim_{t\to\infty} E_{t=0}s_t = (I - D - A^{'}B)^{-1}(a_0 + A^{'}b)$$

for any given initial condition s_1 where I is an identity matrix. Generally speaking, it is analytically complex to characterize the conditions for stability of equation (14) in terms of model primitives (see Anderson, Hansen, McGrattan, and Sargent (1996) for a detailed discussion.) However, the system is confirmed stable at the estimated parameter values given below.

4 Estimation Strategy

4.1 Measurement Error

Many empirical papers including Neal (1999) point out that occupation codes in the NLSY79 are riddled with coding errors. Measurement error in occupation codes will bias parameter estimates for the coefficients of task complexity toward zero, which is known as an attenuation bias. One approach to correcting measurement error is to use other job information such as employer changes. Neal (1999) assumes all occupation changes within the same employer are false. Kambourov and Manovskii (2008) inspects a change of tasks as well as a change of employers in the PSID. Another approach is to model the measurement error process. Keane and Wolpin (2001) and Keane and Sauer (2006) develop an estimation procedure under classification error.

I examine my sample from the NLSY and find many occupational changes within the same employer. However, the majority of them is promotions to a managerial position, and thus, they seem to be a real occupation change.[2] Assuming that all within-employer occupation changes are false would understate the increase of task complexity over a worker's career. Because any obviously useful variables to correct measurement error are not available in the NLSY, I model the process of measurement error. Previous papers deal with an occupation code as a categorical variable and estimate the probability that an occupation is miscategorized. But, this paper considers an occupational classification error as a mismeasured task complexity. An observed task complexity measure $\tilde{x}_t$ is related to the true task complexity x_t by

$$\tilde{x}_t = x_t + \nu_t \tag{15}$$

where ν_t is a vector of measurement errors such that $\nu_t \sim N(0, \Sigma_\nu)$ and serially uncorrelated. This approach assumes that an occupation tends to be misclassified to another occupation that is similar to the truth. Although this assumption is not testable with variables at hand, I find evidence in the NLSY: many of observed occupation changes are moves between two occupations. For example, an individual is classified as a sales worker in year 1 and is classified as a manager in year 2. In year 3, she is again classified as a sales worker. A possible explanation for this cycle of occupation codes is that her true task is close to the boarder between a sales worker and a manager. This story is particularly plausible for a cycle of occupation codes within the same employer. Indeed, I find about 40% of the observed occupation changes within the same employer are cycles of occupation codes. The proposed approach to correcting misclassification seems reasonable.

I also allow for the classical measurement error on wage. An observed wage $\tilde{w}_t$ is related to the true wage w_t by

$$\ln \tilde{w}_t = \ln w_t + \eta_t \tag{16}$$

where η_t is an i.i.d. measurement error such that $\eta_t \sim N(0, \sigma_\eta^2)$.

4.2 Likelihood Function

The likelihood is calculated using the Kalman filter. Let μ_t^s and Σ_t^s be the conditional mean and the conditional covariance matrix of skills at period t given observed tasks up to period $t-1$, $\{\tilde{x}_\tau\}_{\tau=1}^{t-1} = (\tilde{x}_1, \ldots, \tilde{x}_{t-1})$. Because s_1 and ε_t are normally distributed and the skill transition function (5) is linear, the distribution of skills in period $t \geq 1$ is normally distributed given the past tasks up to $t-1$. Using the skill transition equation (5) and the definition of measurement error (15), the

[2] Neal (1999) considers a broadly defined occupation. Thus, he does not classify a promotion as a change of occupation. However, in this paper, an occupation is quite narrowly defined. I view a promotion as a real change of tasks, because the main task of a worker changes to supervising her subordinates.

conditional mean and variance can be recursively calculated by

$$
\begin{aligned}
E(s_t|\{\tilde{x}_\tau\}_{\tau=1}^{t-1}) &= \mu_t^s \\
&= D\mu_{t-1}^s + a_0 + A'\tilde{x}_{t-1} \quad (t>1) \qquad (17)\\
V(s_t|\{\tilde{x}_\tau\}_{\tau=1}^{t-1}) &= \Sigma_t^s \\
&= D\Sigma_{t-1}^s D + \Sigma_\varepsilon - A'\Sigma_\nu A \quad (t>1) \qquad (18)
\end{aligned}
$$

and μ_1^s and Σ_1^s are given. Clearly from the policy function (8) and the wage equation (4), tasks and wages are linear in skills, and thus, they are also normally distributed. The conditional distributions of observed wage and tasks are given by

$$
\begin{aligned}
\tilde{x}_t &\sim N\left(b + B\mu_t^s,\ B\Sigma_t^s B' + \Sigma_\nu \mid \{\tilde{x}_\tau\}_{\tau=1}^{t-1}\right)\\
\ln \tilde{w}_t &\sim N\left((c_0 + c_1' x_t + [p_0 + P' x_t]' \mu_t^s,\ [p_0 + P' x_t]' \Sigma_t^s [p_0 + P' x_t] - c_1' \Sigma_\nu c_1 + \sigma_\eta^2 \mid \{\tilde{x}_\tau\}_{\tau=1}^{t}\right).
\end{aligned}
$$

I have observations of a wage and task trajectory $(\tilde{w}_1, \tilde{x}_1, \cdots, \tilde{w}_T, \tilde{x}_T)$ where T is the last sample period. The likelihood contribution of an individual wage and task history is decomposed into a product of the conditional likelihood of each observation of wage $\tilde{w}_t$ and task $\tilde{x}_t$

$$
\begin{aligned}
& l(\tilde{w}_1, \tilde{x}_1, \cdots, \tilde{w}_T, \tilde{x}_T; a_0, c_0, g)\\
= \ & l(\tilde{x}_1; a_0, g) l(\tilde{w}_1|\tilde{x}_1; a_0, c_0, g) \cdots l(\tilde{x}_T|\tilde{x}_1, \ldots, \tilde{x}_{T-1}; a_0, c_0, g) l(\tilde{w}_T|\tilde{x}_1, \ldots, \tilde{x}_T; a_0, c_0, g) \qquad (19)
\end{aligned}
$$

where a_0, c_0, and g are parameters that vary permanently across individuals. The marginal likelihood contribution is computed by integrating them out. Let $(\tilde{w}^i, \tilde{x}^i) = (\tilde{w}_1, \tilde{x}_1, \cdots, \tilde{w}_T, \tilde{x}_T)$ be a vector of all observations for individual i. The log-likelihood for the whole sample is

$$
\ln L(\{\tilde{w}^i, \tilde{x}^i\}_{i=1}^N|\Theta) = \sum_{i=1}^N \ln\left(\int\int\int l(\tilde{w}^i, \tilde{x}^i|a_0, c_0, g) dF_a dF_c dF_g\right)
$$

where N is the number of individuals in the sample, F_a, F_c, and F_g are cumulative distribution functions of a_0, c_0, and g, respectively. I set the discount factor $\beta = 0.95$. Integration is approximated by a quasi-Monte Carlo method. More specifically, evaluating points are drawn from the Halton sequence.

4.3 Identification

As discussed in the literature of factor models, skills in this model are identified up to affine transformation. The location of skills can take any positive value, but a negative skill does not make an economic sense. It is normalized by setting the mean initial skill to one, $\mu_1^s = 1$. Skills and their utilization are not separately identified, because only the inner product affects wages. A high wage can be rationalized by a high skill utilization or by a large amount of skill. The scale of skills is normalized by setting the minimum skill utilization to 0.1, $p_0 = 0.1$.

4.4 Implementation

Restrictions such as positivity and negative semi-definiteness are imposed on some selected parameters during parameter search in the Maximum Likelihood. Skill utilization parameters p_0, P and skill learning parameters A are restricted to be non-negative. A matrix $C+G$ is restricted to be negative semi-definite to obtain an interior solution, but I do not impose negative semi-definiteness on C and G. Of course, all covariance matrices must be a positive definite matrix. Lastly, skill depreciation parameters δ^j are imposed to lie between 0 and 1.

Positivity restriction can be easily imposed by an exponential transformation. A restriction that a parameter lie between 0 and 1 is imposed by a logit transformation. Negative (or positive) semi-definiteness is imposed using the Cholesky decomposition. Specifically, a negative semi-definite matrix X can be decomposed such that $X = -L_X L_X^{'}$ where L_X is a lower triangular matrix that is unrestricted otherwise.

To decrease computational burden, the distribution of intercepts b in the policy function (12) is estimated as free parameters, rather than functions of structural parameters that are solved by iteration. Notice that a vector of parameters g appears only in equation (12). The distribution of structural parameter g is easily recovered by solving the equation (12) for g and by the delta method for variance estimation.

5 Estimation Results

5.1 Parameter Estimates

Parameter estimates for policy function are reported in Table 8. The diagonal elements of matrix B are the effects of skills on the optimal complexity of the relevant tasks. They are positive in all skill and task dimensions, but the effects of interpersonal skill are imprecisely estimated. The results indicate that individuals move to occupations with more complex task, as workers accumulate their skills.

Table 9 present parameter estimates for skill transition equation (5). They indicate that skills grow more rapidly when individuals are engaged in complex tasks. Cognitive skill accumulation is particularly strongly affected by task complexity, while interpersonal skill accumulation is less so. Depreciation of skills are precisely estimated and they are between 0.38 and 0.44. Motor skills depreciate more quickly than the other skills.

Parameter estimates for wage equation (4) are reported in Table 10. Notice that coefficients c_0, c_1, and C_2 capture pure occupation effects and are estimated without restriction. Pure occupation effects decrease with interpersonal-skill requirement, as expected, but they increase with cognitive-skill and motor-skill requirement. Skill utilization parameters P are estimated with positivity constraints. Utilization of interpersonal and motor-skills increases with task complexity, however, cognitive skill utilization does not change with task complexity. The results suggest that upward occupational mobility is driven by interpersonal skills.

Table 11 presents parameters for measurement error. They are jointly significant, but some of them are imprecisely estimated. In the presence of stochastic component in skill growth, distinguishing them from measurement error turns out to be difficult. The results indicate that measurement error does not strongly affect estimates if the model is flexible.

5.2 Model Fit

To assess the empirical performance of the model, I compare the model prediction with the observed counterpart for trajectories of mean task complexity and logwage. Figure 3 presents the model fit to these features of the data. The model predictions for task complexity are reasonably close to the data. For most of the sample periods, they lie in 95% confidence interval of the sample mean. However, the predicted logwage trajectory is significantly different from the observed trajectory. The predicted logwage profile is too concave; wage grows more quickly than the data during the first five years or so, but the growth slows down afterward.

5.3 Skill or Preference?

The model provides two possible explanations for occupational choice. One is based on skill endowment of an individual. Skilled workers move to occupations with complex tasks to pursuit higher returns to skills, while unskilled workers remain in occupations with less complex tasks. Due to the wage structure presented in Figure 2, workers are sorted into different occupations according to their skills. Another explanation is based on preference for tasks. When individuals are heterogeneous in task preference, those who have identical skills choose different occupations.

I assess the relative importance of the two stories by decomposing the variance of task complexity measures. The variance of observed tasks can be decomposed in the following way,

$$\begin{aligned} Var(\tilde{x}_t) &= Var(b + Bs_t + \nu_t) \\ &= Var(b) + B\Sigma_t^s B' + \Sigma_\nu \end{aligned}$$

Table 5: Sources of Task Complexity Variance

Year = 5

	Cognitive	Interpersonal	Motor
Skill	52.7%	97.4%	71.5%
Preference	17.7%	0.2%	1.2%
Measurement Error	29.6%	2.4%	27.3%

Year = 10

	Cognitive	Interpersonal	Motor
Skill	58.9%	98.3%	81.5%
Preference	15.3%	0.1%	0.8%
Measurement Error	25.7%	1.6%	17.7%

Year = 15

	Cognitive	Interpersonal	Motor
Skill	61.4%	98.6%	85.4%
Preference	14.4%	0.1%	0.6%
Measurement Error	24.2%	1.3%	14.0%

where the first term captures the effects of permanent preference heterogeneity, the second term captures the effects of skill heterogeneity, and the third term captures the variance from measurement error. Skills affect not only choice of tasks, but also wages. Variations of observed tasks unexplained by skills are categorized into preference effects or measurement error. Preference effects are identified by a persistent component of residuals of the observed task complexity, while measurement error is identified by a transient component.

This variance decomposition exercise is applied to period $t = 5$, 10, and 15 and results are summarized in Table 5. For all dimensions of tasks, skill heterogeneity explains the largest part of task complexity. It accounts for about 50-60% of the variance of the cognitive requirement index. For interpersonal skill requirement, most of the variance is explained by skill heterogeneity. Motor skill heterogeneity accounts for about 70-85% of the variance. Preference heterogeneity plays only a limited role in explaining the task complexity variance. Although it accounts for about 15% of cognitive skill requirement, interpersonal skill requirement and motor skill requirement are explained very little. This exercise demonstrates that the proposed skill model is able to explain a substantially large part of heterogeneous choice of occupations.

5.4 Sources of Wage Growth

Individuals develop their skills through occupational experience. I examine the contributions of each skill type to wage growth early in a worker's career. One might argue that the relationship between skills and wage growth can be learned from the observed trajectories of task complexity

Table 6: Sources of Wage Growth

Year	Occ. Effect	Cognitive	Interpersonal	Motor	Total
5	0.011	0.174	0.036	0.147	0.367
10	−0.028	0.226	0.088	0.167	0.453
15	−0.072	0.243	0.136	0.153	0.460

and wage in Figure 3. However, this is not necessarily true. First, an observed task complexity is not fully explained by skills. As shown in the previous subsection, non-negligible amount of task complexity variance remains unexplained by skill heterogeneity. Second, learning may be small, or equivalently utilization may be low, for some skill type. If this is the case, even if a growing task complexity is observed, it may not be translated into wage growth. To understand wage growth in terms of basic labor market skills in a reliable way, a formal econometric model is required.

A mean accumulated wage growth is given by a mean logwage difference between the initial period ($t = 1$) and period t. It is decomposed into the pure occupation effect and the effect of each skill type,

$$
\begin{aligned}
E(\ln w_t - \ln w_1) &= Ec(x_t) - Ec(x_1) + \sum_{j=1}^{J} E[p^j(x_t^j)s_t^j] - E[p^j(x_1^j)s_1^j] \\
&= E\Delta c(x_t) + \sum_{j=1}^{J} E\Delta[p^j(x_t^j)s_t^j]
\end{aligned}
$$

where the first term is the contribution of pure occupation effect, because it is independent of worker skills. Each element of the second term $\Delta[p^j(x_t^j)s_t^j]$ is the contribution of the j-th skill.

The estimated contributions to wage growth are summarized in Table 6. During the first five years in the career, wage growth is largely driven by cognitive skills, which is followed by motor skills. Interpersonal skill does not much contribute to wage growth very early in the career. However, in the next ten years, interpersonal skill is the biggest contributor to the wage growth. The contribution of cognitive skills is a little smaller than interpersonal skill during this period. Interestingly, motor skill contributes very little after the first five years. During this period, many individuals switch from motor skill intensive jobs to interpersonal skill intensive jobs. Therefore, the main contributor to the wage growth is switched from motor skill to interpersonal skill.

6 Conclusion

This paper constructs and estimates a model of human capital formation where skills are multi-dimensional. An important feature of the model is heterogeneous occupations in terms of skill learning opportunity and skill utilization by using the task complexity measures from the DOT. This task approach allows the model for a unified analysis of upward occupational mobility and comparative advantage. The parameter estimates are intuitive, and the model fits the data reasonably well.

Main empirical findings are as follows. First, observed occupational choices are largely rationalized by worker skills. The variance decomposition of task complexity measures indicates that at least more than a half of the variance is explained by skill heterogeneity. Preference heterogeneity plays a relatively small role in accounting for occupational differences across individuals. Second, the important contributor of wage growth for high school graduate workers switches from motor skills to interpersonal skills, although cognitive skill is the biggest contributor throughout the sample period.

The proposed model can be applied to answer a number of interesting questions. For example, an analysis of gender occupational segregation is a relevant application. Males and females might differ in skills and preference for jobs. The model can quantify which factor is important for explaining the gender difference. Another application is an analysis of wage inequality. If the model is extended to allow for time-varying skill utilization (or skill prices), it could explain the widening general wage gap and the narrowing gender wage gap (see Welch (2000)) during the last three decades. These applications are studied in a future work.

References

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 (2005): “Two Sides of the Same Coin: U.S. “Residual" Inequality and the Gender Gap,” University of California-Irvine.

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

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

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.

GIBBONS, R., AND M. WALDMAN (2006): "Enriching a Theory of Wage and Promotion Dynamics inside Firms," *Journal of Labor Economics*, 24(1), 59–107.

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.

KEANE, M. P., AND R. M. SAUER (2006): "Classification Error in Dynamic Discrete Choice Models: Implications for Female Labor Supply Behavior," IZA Discussion Papers No. 2332.

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

——— (2001): "The Effects of Parental Transfers and Borrowing Constraints on Educational Attainment," *International Economic Review*, 42(4), 1051–1103.

KURUSCU, B. (2006): "Training and Lifetime Income," *American Economic Review*, 96(3), 832–846.

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

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

ROY, A. D. (1951): "Some Thoughts on the Distribution of Earnings," *Oxford Economic Papers*, 3(2), 135–146.

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.

WELCH, F. (2000): "Growth in Women's Relative Wages and in Inequality among Men: One Phenomenon or Two?," *American Economic Review*, 90(2), 444–49.

WILLIS, R. J. (1986): "Wage Determinants: A Survey and Reinterpretation of Human Capital Earnings Functions," in *Handbook of Labor Economics*, ed. by O. Ashenfelter, and R. Layard. North-Holland, vol. 1, pp. 525–602.

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 and Figures

Table 7: Summary Statistics for Task Complexity

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
Cognitive Requirement						
Data	0.000	0.113	0.356	0.360	0.613	0.970
GED: Reasoning	0.021	0.138	0.220	0.312	0.477	0.969
GED: Mathematical	0.000	0.149	0.337	0.347	0.488	0.975
GED: Language	0.000	0.135	0.285	0.315	0.472	0.962
Verbal Aptitude	0.000	0.044	0.254	0.247	0.321	0.964
Numerical Aptitude	0.001	0.098	0.265	0.279	0.393	0.941
TMPR: Judgement	0.000	0.075	0.367	0.266	0.422	0.454
Interpersonal Skill Requirement						
Poeple	0.000	0.069	0.193	0.269	0.423	0.923
TMPR: Influencing People	0.000	0.005	0.025	0.130	0.076	0.774
TMPR: Directing	0.000	0.000	0.003	0.068	0.014	0.867
TMPR: Dealing w/ People	0.000	0.009	0.028	0.146	0.293	0.513
Motor Skill Requirement						
Things	0.000	0.084	0.440	0.416	0.663	0.984
Spatial Aptitude	0.001	0.204	0.391	0.413	0.617	0.929
Form Perception	0.001	0.142	0.319	0.339	0.482	0.910
Motor Coordination	0.000	0.088	0.329	0.304	0.485	0.889
Finger Dexterity	0.007	0.031	0.199	0.260	0.501	0.936
Manual Dexterity	0.000	0.194	0.376	0.349	0.412	0.933
Eye-Hand-Foot Coordination	0.000	0.084	0.240	0.346	0.605	0.914
Color Discrimination	0.000	0.187	0.354	0.330	0.461	0.987
TMPR: Precise	0.000	0.053	0.333	0.309	0.564	0.621

Source: NLSY and DOT

Note: Sample size is 6,671.

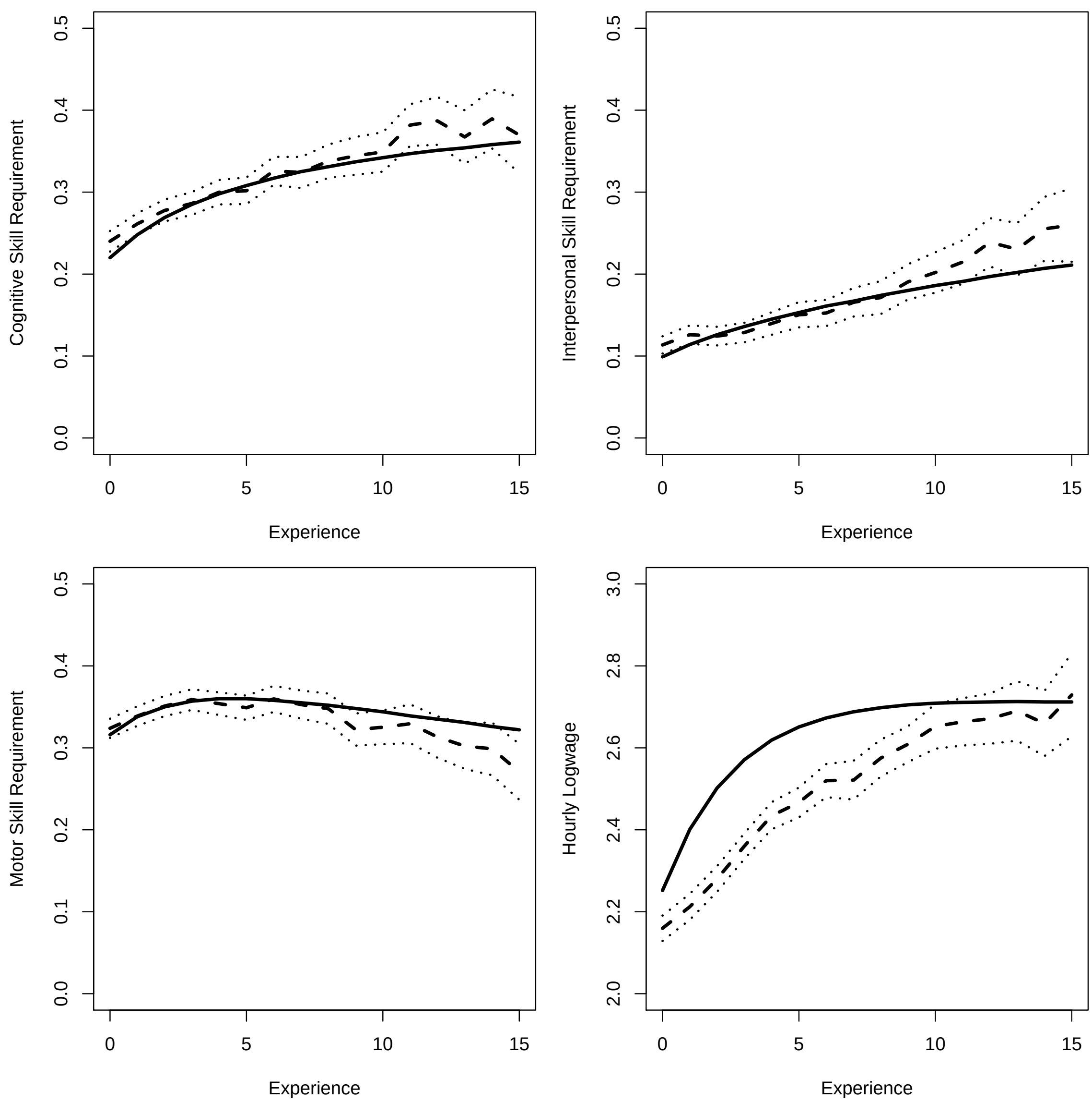

Figure 3: Model Fit

Note: Solid lines are predicted unconditional means of the model. Dashed lines are means of the sample with 95% confidence interval given by dots.

Table 8: Parameter Estimates (Policy Function)

Notation	Estimates	Std. Dev.
$b(1,1)$	0.70608	0.44932
$b(2,1)$	-0.41683	0.59316
$b(3,1)$	0.49601	0.35225
$\Sigma_b(1,1)$	0.06058	0.00235
$\Sigma_b(2,1)$	0.03502	0.00331
$\Sigma_b(2,2)$	-0.00296	0.00308
$\Sigma_b(3,1)$	0.04518	0.00308
$\Sigma_b(3,2)$	0.00975	0.00380
$\Sigma_b(3,3)$	0.02109	0.00262
$B(1,1)$	0.15715	0.02805
$B(1,2)$	-0.42960	0.39775
$B(1,3)$	-0.21357	0.07182
$B(2,1)$	-0.00006	0.00005
$B(2,2)$	0.67381	0.61614
$B(2,3)$	-0.15791	0.04865
$B(3,1)$	-0.00007	0.00006
$B(3,2)$	-0.41438	0.38317
$B(3,3)$	0.23452	0.06717

Table 9: Parameter Estimates (Skill Transition Equation)

Notation	Estimates	Std. Dev.
$a_0(1,1)$	0.64593	0.06461
$a_0(2,1)$	0.40168	0.03923
$a_0(3,1)$	0.36418	0.03930
$\Sigma_a(1,1)$	0.21348	0.03189
$\Sigma_a(2,1)$	0.02196	0.01973
$\Sigma_a(2,2)$	0.00044	0.00136
$\Sigma_a(3,1)$	0.08622	0.02338
$\Sigma_a(3,2)$	0.00777	0.00578
$\Sigma_a(3,3)$	0.02595	0.00884
$A(1,1)$	1.90544	0.23924
$A(2,2)$	0.32612	0.29637
$A(3,3)$	0.93271	0.26629
$D(1,1)$	0.61934	0.00928
$D(2,2)$	0.64048	0.01065
$D(3,3)$	0.56529	0.01329
$\Sigma_1^s(1,1)$	2.86062	0.75149
$\Sigma_1^s(2,1)$	0.34973	0.69750
$\Sigma_1^s(2,2)$	0.05066	0.16375
$\Sigma_1^s(3,1)$	0.64256	1.83431
$\Sigma_1^s(3,2)$	0.05447	0.23886
$\Sigma_1^s(3,3)$	0.26172	0.82120
$\Sigma_\varepsilon(1,1)$	0.84206	0.85605
$\Sigma_\varepsilon(2,1)$	0.13862	0.39942
$\Sigma_\varepsilon(2,2)$	0.02282	0.60227
$\Sigma_\varepsilon(3,1)$	0.01446	1.18141
$\Sigma_\varepsilon(3,2)$	0.00238	0.15189
$\Sigma_\varepsilon(3,3)$	0.00025	5.46089

Table 10: Parameter Estimates (Wage Equation)

Notation	Estimates	Std. Dev.
c_0	1.63226	0.03478
σ_c^2	0.42937	0.00608
$c_1(1,1)$	0.17770	0.15496
$c_1(2,1)$	-1.10223	0.58057
$c_1(3,1)$	0.75074	0.19765
$C2(1,1)$	0.89477	0.23172
$C2(2,1)$	-0.37429	0.24638
$C2(2,2)$	0.63982	0.38486
$C2(3,1)$	-0.77863	0.17669
$C2(3,2)$	0.33770	0.25435
$C2(3,3)$	-0.69364	0.29659
$P(1,1)$	0.00008	0.00008
$P(2,2)$	0.58729	0.58062
$P(3,3)$	0.23864	0.10537

Table 11: Parameter Estimates (Measurement Error)

Notation	Estimates	Std. Dev.
$\Sigma_v(1,1)$	0.01163	0.83206
$\Sigma_v(2,1)$	-0.00019	0.09607
$\Sigma_v(2,2)$	0.00117	0.04921
$\Sigma_v(3,1)$	0.00879	0.09230
$\Sigma_v(3,2)$	0.00185	0.03717
$\Sigma_v(3,3)$	0.01005	0.96028
σ_η^2	0.05461	3.18237

B Model Solution

This section proves by a guess-and-verify that the value function is given by a quadratic function of skills and the optimal policy function is given by a linear function of skills. The Bellman equation for a worker is given by

$$\begin{aligned}
V(s_t) &= \max_{x_t} \ln w(x_t, s_t) + v(x_t, s_t) + \beta EV(s_{t+1}) \\
\text{s.t.} & \\
s_{t+1} &= Ds_t + a_0 + A^{'} x_t + \varepsilon_t \\
\ln w_t &= c_0 + c_1^{'} x_t + x_t^{'} C x_t + [p_0 + P^{'} x_t]^{'} s_t \\
v_t &= g^{'} x_t + x_t G x_t.
\end{aligned}$$

Suppose the value function is given by $V(s_t) = q_0 + q_1^{'} s_t + s_t^{'} Q s_t$. Using the transition law to eliminate skills in the next period, the Bellman equation becomes

$$\begin{aligned}
V(s_t) &= \max_{x_t} \left[c_0 + c_1^{'} x_t + x_t^{'} C x_t + [p_0 + P^{'} x_t]^{'} s_t \right] + \left[g^{'} x_t + x_t G x_t \right] + \\
&\beta E \left[q_0 + q_1^{'} [Ds_t + a_0 + A^{'} x_t + \varepsilon_t] + [Ds_t + a_0 + A^{'} x_t + \varepsilon_t]^{'} Q [Ds_t + a_0 + A^{'} x_t + \varepsilon_t] \right]
\end{aligned} \tag{20}$$

The first order necessary condition for the maximum problem on the right hand side of equation is

$$-2[C + G + \beta AQA^{'}] x_t^* = c_1 + Ps_t - g + \beta E[A\{q_1 + 2Q(Ds_t + a_0 + \varepsilon_t)\}].$$

Because $E\varepsilon_t = 0$, the optimal task x_t^* is given by

$$\begin{aligned}
x_t^* &= -\frac{1}{2}[C + G + \beta AQA^{'}]^{-1} [c_1 - g + \beta A(q_1 + 2Qa_0) + (P + 2\beta AQD)s_t] && (21) \\
&\equiv b + Bs_t \\
b &= -\frac{1}{2}[C + G + \beta AQA^{'}]^{-1} [c_1 - g + \beta A(q_1 + 2Qa_0)] && (22) \\
B &= -\frac{1}{2}[C + G + \beta AQA^{'}]^{-1} [P + 2\beta AQD]. && (23)
\end{aligned}$$

Substituting this optimizer (21) in the Bellman equation (20), I have a closed form solution to the Bellman equation. The current period utility is given by

$$\begin{aligned}
\ln w(x_t^*, s_t) - v(x_t^*, s_t) &= c_0 + (c_1 + g)^{'} x_t^* + x_t^{*'} (C + G) x_t^* + x^{*'} P s_t + p_0^{'} s_t \\
&= c_0 + (c_1 + g)^{'} (b + Bs_t) + (b + Bs_t)^{'} (C + G)(b + Bs_t) + (b + Bs_t)^{'} P s_t + p_0^{'} s_t \\
&= c_0 + [(c_1 + g)^{'} + b^{'} (C + G)] b + \left[p_0 + B^{'} (c_1 + g) + (P^{'} + 2B^{'} (C + G)^{'}) b \right]^{'} s_t +
\end{aligned}$$

$$s_t^{'}B^{'}\left[P+(C+G)B\right]s_t. \tag{24}$$

The expected value function is given by

$$\begin{aligned}
EV(s_{t+1}) &= q_0+Eq_1^{'}[Ds_t+a_0+A^{'}(b+Bs_t)+\varepsilon_t]+ \\
&\quad E[Ds_t+a_0+A^{'}(b+Bs_t)+\varepsilon_t]^{'}Q[Ds_t+a_0+A^{'}(b+Bs_t)+\varepsilon_t] \\
&= q_0+\left[q_1^{'}+(a_0+A^{'}b)^{'}Q\right](a_0+A^{'}b)+E\varepsilon_t^{'}Q\varepsilon_t+ \\
&\quad [q_1^{'}+2(a_0+A^{'}b)^{'}Q](D+A^{'}B)s_t+s_t^{'}(D+A^{'}B)^{'}Q(D+A^{'}B)s_t. \qquad (25)
\end{aligned}$$

Notice that $E\varepsilon_t^{'}Q\varepsilon_t = tr[E(\varepsilon_t^{'}Q\varepsilon_t)] = tr[QE\varepsilon_t\varepsilon_t^{\prime}] = tr(Q\Sigma_\varepsilon)$. Both equations (24) and (25) are quadratic in s_t. Thus, it is confirmed that the value function is quadratic and the policy function is linear in skills s_t. Specifically, coefficients are given by

$$\begin{aligned}
q_0 &= c_0+[(c_1+g)^{'}+b^{'}(C+G)]b+\beta\left[q_0+\left[q_1^{'}+(a_0+A^{'}b)^{'}Q\right](a_0+A^{'}b)+trQ\Sigma_\varepsilon\right] \\
q_1 &= p_0+B^{'}(c_1+g)+(P^{'}+2B^{'}(C+G)^{'})b+\beta\left[(D+A^{'}B)^{'}[q_1+2Q^{'}(a_0+A^{'}b)]\right] \\
Q &= B^{'}\left[P+(C+G)B\right]+\beta(D+A^{'}B)^{'}Q(D+A^{'}B)
\end{aligned}$$

where b and B are defined by equations (22) and (23).