



The Society of Labor Economists

NORC at the University of Chicago

Tasks and Heterogeneous Human Capital

Author(s): Shintaro Yamaguchi

Source: *Journal of Labor Economics*, Vol. 30, No. 1 (January 2012), pp. 1-53

Published by: The University of Chicago Press on behalf of the Society of Labor Economists and the NORC at the University of Chicago

Stable URL: <https://www.jstor.org/stable/10.1086/662066>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <https://about.jstor.org/terms>



JSTOR

The University of Chicago Press, Society of Labor Economists and NORC at the University of Chicago are collaborating with JSTOR to digitize, preserve and extend access to *Journal of Labor Economics*

Tasks and Heterogeneous Human Capital

Shintaro Yamaguchi, *McMaster University*

This article proposes a new approach to modeling heterogeneous human capital using task data from the Dictionary of Occupational Titles. The key feature of the model is that it departs from the Roy model, which treats occupations as distinct categories and conceives of occupations as bundles of tasks. The advantages of this approach are that it can accommodate many occupations without computational burden and provide a clear interpretation as to how and why skills are differently rewarded across occupations. The model is structurally estimated by the Kalman filter using the National Longitudinal Survey of Youth 1979.

I. Introduction

Heterogeneity in worker skills is a central feature of labor economics. One approach to understanding the heterogeneity of human capital is to assume that workers classified by such criteria as race, sex, and education possess homogeneous skills. Another approach pays close attention to the sector affiliation of workers. In the Roy model, workers are sorted

I am grateful for comments from Chris Ferrall, Hiro Kasahara, John Kennan, Salvador Navarro, Ronni Pavan, Chris Robinson, Yuya Takahashi, Chris Worrick, and other participants of the fourteenth meeting of the Society of Labor Economists, the forty-third annual conference of the Canadian Economics Association, and the Society for Economic Dynamics annual meeting and seminar participants at Alberta, Bank of Japan, Bureau of Labor Statistics, Concordia, Duke, Hitotsubashi, Institute of Social and Economic Research at Osaka, International University of Japan, Keio, Ryerson, Tokyo, University of Pennsylvania, Western Ontario, Wisconsin, and Yokohama. The use of the Shared Hierarchical Academic Research Computing Network computational facilities is gratefully acknowledged. Contact the author at yamtaro@mcmaster.ca.

[*Journal of Labor Economics*, 2012, vol. 30, no. 1]
© 2012 by The University of Chicago. All rights reserved.
0734-306X/2012/3001-0001\$10.00

across occupations in which they have a comparative advantage. This approach is more economically meaningful because it clearly relates the unobserved skills to the observed sector affiliation and wages. Heckman and Sedlacek (1985), Keane and Wolpin (1997), and many subsequent papers including Lee (2005), Lee and Wolpin (2006), and Sullivan (2010) are based on the Roy model and study issues in which heterogeneous human capital plays a central role. While this article is in the same spirit as the Roy model, it departs from the previous Roy-type models by taking a more direct look at what describes the tasks workers perform on the job using the occupational task information provided by the Dictionary of Occupational Titles (DOT).

This article conceives of an occupation as a bundle of tasks. Different occupations involve different tasks at different levels of complexity. For example, high levels of complicated quantitative analysis would be required to perform the tasks of an actuary, while the tasks carried out by a cashier require relatively low levels. Additionally, the tasks performed by a tailor require high levels of finger and manual dexterity, while much lower levels are necessary for the tasks performed by a sales worker. The data from the DOT allow me to construct a low-dimensional vector of continuous measures of task complexity. Occupations are therefore characterized in a continuous and multidimensional space of task complexity in the model. This approach is very different from the Roy-type models mentioned above because these models deal with each occupation as a distinct category. When occupations are treated as categories, one cannot quantify how they are similar to each other. However, in the proposed approach, similarity of occupations is readily measured by the distance between occupations in the task complexity space.

The proposed approach has two advantages over other Roy-type models. First, the model accommodates many occupations without a large number of parameters and state variables because the preference and wage functions are not defined over occupations but over a low-dimensional vector of occupational tasks. My approach can be interpreted as an application of the characteristic approach proposed by Lancaster (1966) to the Roy model and, thus, has a strength similar to that of the characteristic approach. In Roy-type models, returns to skills differ across occupations in an arbitrary way. This implies that the number of parameters increases with the number of occupations in the model because the returns to skills are estimated for each occupation in the model. When skills are measured by occupation-specific experiences as in Keane and Wolpin (1997), the number of state variables also increases with the number of occupations. Because the computational burden of solving a dynamic programming problem exponentially increases with the number of state variables, this model structure has prevented previous papers from including several occupations. This is not satisfactory because the apparent heterogeneity

within broadly defined occupations is ignored. Indeed, empirical evidence from the DOT indicates that tasks are significantly different across three-digit census occupations that are grouped in the same one-digit census occupation. This article overcomes this limitation by assuming that the returns to skills change with task complexity. In an occupation that involves relatively complex tasks, worker skills are more intensely used and thus better rewarded. This assumption makes the number of parameters and state variables independent of the number of occupations in the model, and thus the model is able to deal with about 500 occupations at the census three-digit level. This modeling strategy is useful not only for complicated structural estimations but also for reduced-form wage regressions. Economists may want to estimate wage equations across occupations to understand the extent of skill transferability or differences in the wage structure across occupations. However, there is not enough occupational mobility in the data to do so with any reasonable precision. A common solution to this problem is either to aggregate occupations or to impose a strong restriction on the transferability of skills across occupations. The task-based approach provides a way of dealing with this data limitation.

The second and more important advantage is that the proposed approach provides a clear explanation as to why and how skills are differently rewarded and transferable across occupations. In Keane and Wolpin (1997) and subsequent papers that adopt their framework, it is not quite clear how and why experience accumulated in one occupation is rewarded in others. While this is less problematic if the model includes only a couple of occupations, interpretation of the differences between occupations is not necessarily straightforward when the model includes more than a few occupations, as in Johnson and Keane (2007) and Sullivan (2010). In this article, I interpret differences in returns to skills across occupations from the viewpoint of tasks. Skills are similarly rewarded between occupations that involve similar tasks, while they are less so between occupations that involve very different tasks. Skill transferability changes with the task complexity-measured similarity of occupations.

In the model, workers are endowed with heterogeneous (or multidimensional) skills. They synthesize all types of skills to perform the tasks involved in any occupation, but their skills are differently rewarded across occupations. Occupations also offer different learning opportunities according to their task complexity. For example, if skills are acquired through learning by doing, performing more complicated tasks helps workers develop their skills faster. Workers also possess preferences over task complexity itself. For example, some workers like complicated manual tasks, while other workers may find them unpleasant. Due to the presence of differing job preferences, workers with identical skills may choose different occupations. In each period, a worker chooses an oc-

cupation that maximizes the expected present value of lifetime utility. Unlike the previous career decision models, occupational choice is modeled as a continuous choice problem rather than a discrete choice problem because choosing an occupation is equivalent to choosing a set of tasks in a continuous space. Individual heterogeneity in initial skill endowment, learning ability, and job preferences is included in the model.

I show that, under my functional-form assumptions, the optimal policy function for occupational choice can be represented by a linear function of observed worker characteristics, unobserved skills, and unobserved preference shocks. This result implies that the observed task complexity can be interpreted as a noisy signal of unobserved skills. Hence, I can estimate the dynamics of unobserved skills from the observed task complexity. This closed-form solution also allows me to estimate the model with little computational burden. The previous models based on discrete choice dynamic programming, such as those in Keane and Wolpin (1997), require a numerical solution for each grid point in the discretized state space, which could take several hours or even days. It is also worth noting that the current model can include many observed worker characteristics, including continuous variables. This is often intractable for dynamic structural models because the value function and the policy function are solved for all types of individuals as defined by observed characteristics.

I draw data for the measurement of task complexity from the DOT, which defines occupations in terms of several tasks. For example, the complexity of tasks in terms of mathematical development is evaluated on a six-point scale. Tasks at the lowest level of complexity involve adding and subtracting two-digit numbers, while tasks at the highest level of complexity involve the use of advanced calculus and statistical inference. The DOT measures the complexity of tasks with respect to 62 characteristics, but many of them are highly correlated and redundant. I therefore summarize the information of the DOT by constructing a complexity measurement for two broadly defined tasks: cognitive tasks and motor tasks. In order to construct wage and task complexity histories for young male workers, the task complexity measures are combined with the National Longitudinal Survey of Youth (NLSY) 1979 by using the three-digit census occupation code.

I estimate the model by the Kalman filter. The results indicate that returns to skills and the rate of learning differ across occupations according to task complexity. The parameter estimates are used to evaluate the roles of cognitive and motor skills in determining cross-sectional wage variance and wage growth. I find that both cognitive and motor skills account for a considerable fraction of the wage variance. The differences existing between workers before labor market entry account for about 35% of the log wage variance 10 and 20 years after entry. For wage growth, cognitive skills are the main driver, accounting for all of the wage growth for those

who completed high school or higher. The growth of motor skills is important for high school dropouts; it accounts for half of their wage growth.

This article is also related to the growing literature that attempts to look into the skills associated with job tasks. My contribution to this literature is a novel use of task information that makes a clear distinction between worker skills and job tasks. A common approach in the literature (see, e.g., Autor, Levy, and Murnane 2003; Ingram and Neumann 2006; Bacolod and Blum 2010) is to treat tasks as proxies for unobserved worker skills. Under such a framework, tasks do not play any substantial role in the model. These models also assume skill prices to be constant across occupations. For example, Ingram and Neumann (2006) and Bacolod and Blum (2010) include task complexity measures in their wage regressions and interpret the coefficients as returns to skills.

However, this approach has limitations from theoretical and empirical viewpoints. On the theoretical side, Rosen (1983) criticizes the model presented by Welch (1969) in which skill prices are uniform across occupations. He shows that when workers are sorted across occupations, uniform skill prices arise in a very special circumstance, and skill prices are likely to differ across occupations. On the empirical side, Heckman and Scheinkman (1987) and other papers based on the Roy models provide evidence that returns to skills are systematically different across sectors.¹ Building on the Roy model, my model distinguishes between tasks and skills and characterizes the relationship between the two. The derived policy function for occupational choice suggests that observed tasks can be interpreted as a noisy signal of unobserved skills. It is also shown that the structural model can be estimated by the Kalman filter under parametric assumptions.

Poletaev and Robinson (2008) and Gathmann and Schönberg (2010) examine wage losses after occupation changes, using U.S. and German data, respectively. They find that losses increase with the task distance measure between occupations. This article provides an economic model that helps interpret their empirical findings. Autor and Handel (2009) consider the relationship among skills, tasks, and occupations. They have task information at the individual level, unlike task-based approach papers that use the occupation level data contained in the DOT. Autor and Handel (2009) assume that skills are measured by the observed tasks and are rewarded differently across occupations. However, in contrast to this article, they do not conceive of occupations as bundles of tasks; therefore, the differences between occupations remain ambiguous. Consequently,

¹ Firpo, Fortin, and Lemieux (2011) also argue in favor of the Roy model approach, over uniform skill price models, in order to study the changes in the U.S. wage structure using occupational task data.

like in the Roy model, it is not quite clear how and why skills are rewarded differently across occupations in their model. My article also differs in that I empirically analyze the dynamics of skills and wages, while Autor and Handel (2009) test a few predictions in a static setup.

Lazear (2009) presents an idea similar to this article. In his skill-weights approach, returns to skills change with the weights that a job places on skills. Workers are attracted to jobs that place more weight on skills in which the workers have a comparative advantage. Lazear (2009) assumes that a high weight on one skill necessarily means a low weight on the other skills. This trade-off is necessary to avoid a situation in which all workers enter the same occupation. In contrast, in this article some occupations may reward all types of skills better than other occupations. This is consistent with empirical evidence from the DOT that demonstrates that some occupations involve more complex tasks in all dimensions than other occupations. Lazear (2009) develops a theoretical analysis of the model and provides a useful insight into the relationship between skills and job tasks, but it is not quite clear how the model could be estimated or tested due to the lack of skill weights data. The main contribution of this article is to propose a new model that can be empirically implemented with task data, such as the DOT, and provide an extensive empirical analysis of the dynamics of skills and wages.

The rest of the article is organized as follows: Section II lays out the structural dynamic model of occupational choice and shows that a closed-form solution to the policy function exists. Section III explains the estimation strategy. Section IV describes how task complexity measures are constructed using the DOT and sampling criteria for the NLSY. The estimation results including the model fit are presented in Section V. Section VI discusses the effects of each type of skill on the distribution and growth of wages. Section VII concludes. Appendix E, available in the online version of *Journal of Labor Economics*, extensively discusses the robustness of the results.

II. Model

In this section, I describe a dynamic model of occupational choice and skill formation. In the model, each individual who has made a long-term transition to the full-time labor market has a finite decision horizon ending in an exogenously fixed retirement age. Jobs in the same occupation are homogeneous in terms of task complexity; that is, jobs and occupations are not distinguished. In each year t , an individual chooses an occupation that lies in a K -dimensional continuous space of task complexity x_t that is observable and takes nonnegative values.² The task complexity indexes take nonnegative values, and sufficiently many occupations exist so that

² See Sec. IV.A for a detailed discussion of task complexity measurement.

an individual can choose any occupation in the task complexity space.³ Skills in year t are denoted by a K -dimensional vector s_t that is unobserved by the econometrician. The skills index s_t can take any real number, including negative values.

Labor is the only factor of production in the economy. Each firm offers jobs of a single type, which implies that the products of each firm can be characterized by a task complexity vector.⁴ The products are heterogeneous and consumed by households. The price of the product characterized by task x_t is denoted by $\pi(x_t)$. The productivity of a worker with skill s_t in a job with task complexity x_t is $q(x_t, s_t)$.

A. Wage Function

Workers are paid in accordance with their marginal value product. The marginal value product of a worker with skill s_t in an occupation with task complexity x_t is

$$w_t = \pi(x_t)q(x_t, s_t) \exp(\eta_t), \quad (1)$$

where $\eta_t \sim N(0, \sigma_\eta^2)$ can be interpreted as an independent and identically distributed productivity shock or as measurement error. As in the Roy model, skills are rewarded differently across occupations. I assume that labor productivity may be represented by the inner product of skills and their implicit prices:

$$\ln q(x_t, s_t) = \theta'(x_t)s_t, \quad (2)$$

where $\theta(x_t)$ is a K -dimensional vector of implicit skill prices and represents the contribution of skills s_t to an occupation with task x_t . Skills are more intensely used and contribute to productivity more, when the corresponding tasks are complex $\partial\theta_k(x)/\partial x_k > 0$, where subscript k is an index for the task dimension.⁵ For example, quantitative skills are essential for actuaries but do not contribute as much to a truck driver's productivity. Similarly, finger and manual dexterity are required for a tailor's tasks but are not very useful for an accountant's tasks. I parametrize the output price and productivity as

³ This assumption is not restrictive for two reasons. First, the model contains about 500 occupations at the census three-digit level. Second and more important, the observed distribution of occupations in task complexity space does not necessarily reflect the constraints faced by workers. It may reflect the equilibrium outcomes of workers and firms without constraints in their supply of and demand for occupations. In other words, certain types of occupations may not be seen in the data, simply because no workers find them attractive.

⁴ This assumption is for expositional convenience. Alternatively, the products of each job can be viewed as intermediate outputs for homogeneous final products. Both interpretations are observationally equivalent.

⁵ Teulings (1995) and Gibbons et al. (2005) also assume that returns to skills increase with task complexity. This article extends their wage equations to the case in which skills and tasks are multidimensional.

$$\ln \pi(x_i) = p_0 + p_1'x_i \quad (3)$$

and

$$\ln q(x_i, s_i) = \theta'(x_i)s_i = [p_2 + P_3'x_i]'s_i, \quad (4)$$

where p_0 is a scalar, p_1 and p_2 are K -dimensional vectors, and P_3 is a K -dimensional diagonal matrix. Therefore, the log wage is given by

$$\ln w_i = p_0 + p_1'x_i + [p_2 + P_3'x_i]'s_i + \eta_i. \quad (5)$$

In a Roy-type model, such as in Keane and Wolpin (1997), the output price π and implicit skill prices θ are allowed to differ across occupations in an arbitrary way. Since this approach does not impose a restriction on how they vary, the parameters π and θ must be estimated for each occupation. Many Roy-type models do not include more than a few occupations because the number of parameters increases with the number of occupations. My article overcomes this limitation by imposing structure from the viewpoint of occupational tasks on how the parameters π and θ vary across occupations. If two occupations are similar in terms of tasks x —say a cognitive-skilled worker switches occupations from bank officer to office manager—then the wage change after a transition from one occupation to the other would be small. This can be interpreted as the existence of high skill transferability between the two occupations. In contrast, if two jobs involve very different tasks—say a motor-skilled worker changes occupations from mechanic to clerk—then a large wage loss would occur. This implies low skill transferability exists between the two occupations. In a Roy-type model, it is not necessarily clear why experience in one occupation is rewarded in others. The current model can answer this question because similarity between occupations is clearly characterized in terms of tasks. The proposed approach facilitates an interpretation of the wage structure and decreases computational burden by reducing the number of parameters.

This wage function also provides an interpretation of job match quality introduced by Jovanovic (1979). While many empirical papers find that wage gains from job search are substantial for young workers, their models do not explicitly explain why match quality varies across jobs.⁶ In my model, a worker receives a wage gain when he moves to an occupation in which the returns to skills are high. The size of this wage gain depends on the skill endowment of the worker. Consider a worker with high motor skills and low cognitive skills. This worker can expect a large wage gain by taking a job that involves complex motor tasks but cannot expect a significant wage gain by taking a job that involves complex cognitive tasks. From the viewpoint of wage gains, desirable occupations vary across individuals according to their skill endowments. The job search process

⁶ The contributions in this area include Topel and Ward (1992), Schönberg (2007), Sullivan (2010), Yamaguchi (2010b), and Pavan (2011).

could be interpreted as a worker attempting to reach an occupation that offers a higher return to skills that the worker has an abundance of in an environment in which the worker does not know the location of the job he prefers, while the current model assumes he does.

B. Skill Formation

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

$$s_{t+1} = Ds_t + a_0 + A_1x_t + A_2d + \varepsilon_{t+1}, \quad (6)$$

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

Two interpretations of the model are possible, depending on the sign of A_1 . When A_1 is positive definite, this specification can be interpreted as an extension of a learning-by-doing model to the case in which jobs are heterogeneous. For example, if a worker spends many years in occupations that involve complex motor tasks, this person should have developed significant motor skills. When A_1 is negative definite, this skill formation technology explains workers' occupational mobility over their careers, which is shown by Rubinstein and Weiss (2006). Notice that a trade-off between returns to skills and skill-learning opportunities exists in this case: occupations with complex tasks provide higher returns to skills, but skill growth is slow in those occupations. In this environment, a typical young worker starts his career in an occupation with simple tasks in order to develop more skills. Later in his career, he moves to an occupation characterized by complex tasks in order to receive higher returns to the skills he has developed in the early stages of his career. Both interpretations are plausible and are empirically examined below.

Individuals start their careers with initial skills s_1 that differ across individuals in both observable and unobservable ways. The initial skill endowment is given by

$$s_1 = b + Hd + \varepsilon_1, \quad (7)$$

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

C. Job Preferences

I examine the role that job preferences play in occupational choices for two reasons. First, it is plausible that workers with identical skills may choose different occupations according to their differing tastes for tasks. Accounting for job preferences is particularly important to understanding differences in skills between demographic groups. For example, significant occupational gaps exist between men and women, and blacks and whites. The gaps may reflect differences in their skills to some extent, but they may also be due to differences in job preferences. Ignoring differences in job preferences will result in biased estimates of skill endowments across demographic groups. The second reason to consider job preferences is to rationalize the observed occupational choices. The proposed wage function and skill formation technology have the potential to sort workers into different occupations on the basis of their skills. However, the sign conditions discussed above may not be satisfied empirically. Even if they are, the model without job preferences may not fit the data very well. The empirical results below strongly support the inclusion of job preferences in the model.

The following quadratic function of task complexity determines the utility derived from work:

$$v_t = v(x_t, \bar{x}_t, s_t, \tilde{v}_t; d) \quad (8)$$

$$= (g_0 + G_1 d + G_2 s_t + \tilde{v}_t)' x_t + x_t' G_3 x_t + (x_t - \bar{x}_t)' G_4 (x_t - \bar{x}_t), \quad (9)$$

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

The utility from job tasks varies across individuals according to their individual characteristics d , skill levels s_t , a preference shock \tilde{v}_t , and work habits \bar{x}_t . Skilled workers prefer complex tasks if the parameter matrix G_2 is positive definite. I assume that the matrix G_3 is negative definite. This restriction implies that, for a very high value of x_t , the marginal utility from task complexity is negative; this is the cost of entering an occupation with complex tasks. The parameters g_0 , G_1 , and G_2 are unrestricted. The last term in the above equation captures the effect of work habits on utility.

The work habits of an individual are measured by the weighted average of the task complexities of previous occupations held by the individual. Individuals may have difficulty in adjusting themselves to a new work environment. The mental and physical costs of this adjustment are high when entering into an occupation that is very different from the past occupations held. This effect can also be interpreted as a sort of search friction because individuals are unlikely to receive a job offer for a position

that involves very different tasks compared to the recent jobs held. This work habit effect is introduced to approximate the fact that workers do not change occupations every year.⁷ It is true that the model can predict more realistic worker mobility patterns by introducing a fixed entry cost to a new occupation or by assuming that workers do not receive job offers every period, but the proposed specification is necessary to derive the linear policy function shown below. I argue that the benefit from this approximation exceeds its loss of realism.

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

Individuals form their work habits through the following transition equation:

$$\bar{x}_{t+1} = A_3 \bar{x}_t + (I - A_3)x_t, \quad (10)$$

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

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

$$\bar{x}_1 = \bar{x}_{1,0} + Xd, \quad (11)$$

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

⁷ It is well known in the literature that an individual's choices are serially correlated. One common approach to accounting for persistence in a worker's choices is to include a lagged choice variable. For example, in Keane and Wolpin (1997), a worker pays a job-finding cost if he did not work in the occupation in the previous period. Altug and Miller (1998) allow for work disutility to depend on hours worked in previous years. The work habits in this article are in the same line as these papers.

D. Bellman Equation

The Bellman equation for an individual is given by

$$V_t(s_t, \bar{x}_t, \tilde{v}_t, \eta_t; d) = \max_{x_t} \ln w(x_t, s_t, \eta_t) + v(x_t, \bar{x}_t, s_t, \tilde{v}_t; d) \\ + \beta EV_{t+1}(s_{t+1}, \bar{x}_{t+1}, \tilde{v}_{t+1}, \eta_{t+1}; d), \quad (12)$$

such that

$$\ln w_t = p_0 + p'_1 x_t + [p_2 + P'_3 x_t]' s_t + \eta_t, \quad (13)$$

$$v_t = (g_0 + G_1 d + G_2 s_t + \tilde{v}_t)' x_t + x_t' G_3 x_t + (x_t - \bar{x}_t)' G_4 (x_t - \bar{x}_t), \quad (14)$$

$$s_{t+1} = D s_t + a_0 + A_1 x_t + A_2 d + \varepsilon_{t+1}, \quad (15)$$

$$\bar{x}_{t+1} = A_3 \bar{x}_t + (I - A_3) x_t, \quad (16)$$

$$s_1 = b + H d + \varepsilon_1, \quad (17)$$

$$\bar{x}_1 = \bar{x}_{1,0} + X d. \quad (18)$$

The objective of workers is to maximize the present value of lifetime utility by choosing occupation x_t . Workers take into consideration that different occupations offer different wage returns, skill-learning opportunities, and disutilities according to the tasks that comprise the occupations. Note that the differing wage returns represented by equation (5) can account for worker sorting across occupations based on skills on its own, although all three factors affect workers' decisions. To see this point, differentiate the wage equation by x_t : $\partial \ln w_t / \partial x_t = p_1 + P'_3 s_t$. This indicates that the more skills workers have, the more benefit they receive. Notice that s_t can take a negative value. Thus, unskilled workers (i.e., workers with large negative values of s_t) do not want to enter an occupation with high x_t because doing so would reduce their wages.

The optimal policy function is a linear function of skills, work habits, individual characteristics, and preference shocks because the flow utility $\ln w + v$ is a quadratic function of the state variables, and the transition equations of the state variables are linear in the state variables.⁸ Notice that normality of the random variables is not necessary to derive the linear policy function. It can be expressed as

$$x_t^* = c_{0,t} + C_{1,t} d + C_{2,t} s_t + C_{3,t} \bar{x}_t + v_t, \quad (19)$$

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 v_t is a K -dimensional vector of

⁸ This class of problems is known as a stochastic optimal linear regulator problem.

rescaled preference shocks (i.e., I can write $v_t = M_t \tilde{v}_t$, where M_t is a K -dimensional diagonal matrix). The proof is shown in appendix C. The rescaled preference shocks v_t are normal, independent, and identically distributed random variables with a zero mean and a nondiagonal variance matrix Σ_v . The parameters $c_{0,t}$, $C_{1,t}$, $C_{2,t}$, and $C_{3,t}$ are functions of structural parameters and are not estimated as free parameters. Because the problem has a finite horizon, I solve the value function and the policy function through backward recursion.⁹

The derived policy function provides a useful interpretation of the observed occupational choice: it is a noisy signal of underlying skills. Hence, I can estimate the dynamics of unobserved skills from the observed task complexity measures. This is an important departure from previous papers that implicitly assume that workers in the same occupation have identical occupational skills and use the observed occupation as a proxy for skills. My model distinguishes occupation and worker skills and characterizes the relationship between the two through the policy function.

This analytical solution also dramatically reduces the computational time for estimation. In many papers that estimate a structural dynamic model of occupational choice, such as Keane and Wolpin (1997), no analytical solution is available, and the model is solved for at each grid point in the discretized state space for each parameter value throughout the estimation.¹⁰ With the analytical solution, I can avoid calculating the numerical solutions that account for most of the computational burden in those papers. Another advantage of having an analytical solution is that it allows me to accommodate many state variables, including continuous variables for observed worker characteristics, which is intractable in previous structural models. This article accounts for the heterogeneity of occupations and workers in a rich way with little computation cost.

Although modeling occupational choice as a continuous choice problem instead of a discrete choice problem has advantages in interpretation and computation, there is a limitation in the predicted pattern of worker mobility. The model predicts that x_t^* changes every period due to skill and preference shocks ε_t and v_t , which means workers change occupations every year. This is different from what is observed in the data; most workers do not change occupations every year, and thus the time profile of task complexity x_t at the individual level is a step function. While the model cannot match this feature of the data, it can approximate the profile of task complexity over time. A fixed cost of an occupation change would

⁹ Many other methods are available for infinite horizon problems. See Anderson et al. (1996) for a survey.

¹⁰ Hotz and Miller (1993) and many subsequent papers develop a method to avoid calculating numerical solutions when estimating a structural dynamic discrete choice model. Imai, Jain, and Ching (2009) also develop a Bayesian approach to reduce computational burden when estimating a model of the same class.

allow the model to overcome this limitation. But the linear policy function would also be lost. Because the purpose of the article is not to match the frequency of occupational change, I do not include a fixed cost of an occupation change in the model.

III. Estimation Strategy

A. Identification

Notice that this model is an application of a state-space model. Like other state-space models, there is an identification issue related to unobserved skills. The scale parameters of skills are not identified because observed variables (i.e., wage and task complexity) are the product of unobserved skills and unknown parameters such as the returns to skills. Observed high wages can be rationalized by either a large amount of skills or high returns to skills. The location parameters of skills are also not identified because no natural measures of skills exist. Hence, I normalize skills by assuming that the unconditional mean and variance of initial skills are 0 and 1, respectively.¹¹ This normalization of the location parameter affects the intercept p_0 and the coefficient vector p_1 in the wage equation. However, the marginal effect of x on $\ln w$ (i.e., $\partial \ln w / \partial x = p_1 + P_3 s$) can be consistently estimated regardless of how I normalize the location parameter.¹² The covariances of the initial unobserved skills and skill shocks (i.e., off-diagonal elements of Σ_ε and Σ_{ε_1}) are identified by the conditional wage variance given tasks. For example, the estimated covariance of the unobserved components of the skills is negative, if the conditional log wage variance decreases with the product of cognitive and motor task complexity indexes. A more detailed argument can be found in appendix A. The parameters for job preferences are identified by the occupational choices characterized by the policy function. However, they are not separately identified because the number of these parameters in the job preference equation (9) is greater than the number of the param-

¹¹ Specifically, it is imposed that $E(s_1) = b + HE(d) = 0$, $\text{diag}[\text{Var}(s_1)] = \text{diag}[HE(dd')H' + \Sigma_{\varepsilon_1}] = \text{diag}[I]$, where 0 is a vector of zeros, I is an identity matrix, and diag is an operator that converts a matrix into a vector that consists of diagonal elements of the matrix. In the estimation, b and $\text{diag}[\Sigma_{\varepsilon_1}]$ are not estimated as free parameters but recovered from these restrictions. The off-diagonal elements of Σ_{ε_1} are unrestricted and estimated as free parameters.

¹² To understand this point, let s_1 be $s_1 = \mu_s + \tilde{s}_1$, where μ_s is a vector of constants and \tilde{s}_1 is a vector of normal random variables with zero mean. In this article, I assume $\mu_s = 0$ for normalization, but let us suppose not. The wage equation can be rewritten as $\ln w_1 = (p_0 + p'_2 \mu_s) + (p_1 + P_3 \mu_s + P_3 \tilde{s}_1)x_1 + (p'_2 \tilde{s}_1 + \eta_1)$. With parametric assumptions, $(p_0 + p'_2 \mu_s)$ and $(p_1 + P_3 \mu_s)$ are identified. This implies that the parameter values of p_0 and p_1 depend on μ_s .

eters in the policy function (19).¹³ To avoid this unidentifiability, it is imposed that G_3 be the negative of an identity matrix. This normalization is innocuous as long as G_3 is negative definite because the policy function parameters are consistently estimated regardless of how I normalize G_3 . The structural parameters G_1 , G_2 , and G_4 are identified by the variations of worker characteristics d , skills s_t , and work habits \bar{x}_t , respectively. The parameter g_0 is identified by the constant term of the policy function.

B. Kalman Filter

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

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

$$E(s_t | x_1, w_1, \dots, x_{t-1}, w_{t-1}; \bar{x}_1, d) \equiv E(s_t | Y_{t-1}) \quad (20)$$

$$\equiv \hat{s}_{t|t-1}, \quad (21)$$

$$\text{Var}(s_t | x_1, w_1, \dots, x_{t-1}, w_{t-1}; \bar{x}_1, d) \equiv \text{Var}(s_t | Y_{t-1}) \quad (22)$$

$$\equiv \Sigma_{t|t-1}^s, \quad (23)$$

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 eq. [19]) is linear in skills, work habits (i.e., the weighted average of the task complexities of the past occupations), and preference shocks \tilde{v}_t . The conditional mean and variance of x_t given Y_t are

$$E(x_t | Y_{t-1}) = c_{0,t} + C_{1,t}d + C_{2,t}\hat{s}_{t|t-1} + C_{3,t}\bar{x}_t, \quad (24)$$

$$\text{Var}(x_t | Y_{t-1}) = C_{2,t}\Sigma_{t|t-1}^s C_{2,t}' + \Sigma_v. \quad (25)$$

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

¹³ To see this unidentifiability, consider the optimal choice of occupation at the terminal period T . The optimal task complexity x_T^* is given by

$$x_T^* = -(1/2)(G_3 + G_4)^{-1}[g_0 + G_1d + (P_3 + G_2)s_T - 2G_4\bar{x}_T + \tilde{v}_T]$$

$$\equiv c_{0,T} + C_{1,T}d + C_{2,T}s_T + C_{3,T}\bar{x}_T + \nu_T.$$

Here, the number of structural parameters to be estimated is $(4 + L)K$, while the number of parameters in the policy function is $(3 + L)K$. To proceed, K parameters have to be fixed. The exact relationships between the structural parameters and the policy function parameters are outlined in app. C.

$$E(s_t|Y_{t-1}, x_t) = \hat{s}_{t|t-1} + \Sigma_{t|t-1}^s C'_{2,t} (C_{2,t} \Sigma_{t|t-1}^s C'_{2,t} + \Sigma_\nu)^{-1} \hat{\nu}_t, \quad (26)$$

$$\text{Var}(s_t|Y_{t-1}, x_t) = \Sigma_{t|t-1}^s - \Sigma_{t|t-1}^s C'_{2,t} (C_{2,t} \Sigma_{t|t-1}^s C'_{2,t} + \Sigma_\nu)^{-1} C_{2,t} \Sigma_{t|t-1}^s, \quad (27)$$

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

$$E(\ln w_t|Y_{t-1}, x_t) = p_0 + p'_1 x_t + [p_2 + P'_3 x_t] E(s_t|Y_{t-1}, x_t), \quad (28)$$

$$\text{Var}(\ln w_t|Y_{t-1}, x_t) = [p_2 + P'_3 x_t]' \text{Var}(s_t|Y_{t-1}, x_t) [p_2 + P'_3 x_t] + \sigma_\eta^2. \quad (29)$$

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

$$\begin{aligned} & E(s_t|Y_{t-1}, x_t, w_t) \\ &= E(s_t|Y_{t-1}, x_t) + \text{Var}(s_t|Y_{t-1}, x_t) [p_2 + P'_3 x_t] [\text{Var}(\ln w_t|Y_{t-1}, x_t)]^{-1} \hat{\eta}_t, \end{aligned} \quad (30)$$

$$\begin{aligned} & \text{Var}(s_t|Y_{t-1}, x_t, w_t) \\ &= \text{Var}(s_t|Y_{t-1}, x_t) \end{aligned} \quad (31)$$

$$- \text{Var}(s_t|Y_{t-1}, x_t) [p_2 + P'_3 x_t] [\text{Var}(\ln w_t|Y_{t-1}, x_t)]^{-1} [p_2 + P'_3 x_t]' \text{Var}(s_t|Y_{t-1}, x_t),$$

where $\hat{\eta}_t$ is the log wage 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 eq. [6]). Because skills in year $t+1$ are linear in current skills and task complexity, they are also normally distributed with mean and variance:

$$\hat{s}_{t+1|t} = DE(s_t|Y_{t-1}, x_t, w_t) + a_0 + A_1 x_t + A_2 d, \quad (32)$$

$$\Sigma_{t+1|t}^s = D \text{Var}(s_t|Y_{t-1}, x_t, w_t) D + \Sigma_\epsilon. \quad (33)$$

When the wage is missing in the data, it is integrated out to construct the likelihood. Given the linear skill transition equation (6), the conditional distribution of s_{t+1} given Y_{t-1} and x_t is normal with mean and variance

$$\begin{aligned} E(s_{t+1}|Y_{t-1}, x_t) &= Ds_t + a_0 + A_1 x_t + A_2 d \\ &= D\hat{s}_{t|t-1} + a_0 + A_1 x_t + A_2 d, \end{aligned} \quad (34)$$

$$\text{Var}(s_{t+1}|Y_{t-1}, x_t) = D\Sigma_{t|t-1}^s D + \Sigma_\epsilon, \quad (35)$$

which replace equations (32) and (33) in this case.¹⁴ This algorithm allows me to calculate the conditional distribution of skills, wages, and occupational tasks 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 $(w_{i1}, x_{i1}, \dots, w_{iT_i}, x_{iT_i})$, 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

$$l(w_{i1}, x_{i1}, \dots, w_{iT_i}, x_{iT_i} | \bar{x}_{i1}, d_i) \quad (36)$$

$$= l(x_{i1} | \bar{x}_{i1}, d_i) l(w_{i1} | x_{i1}; \bar{x}_{i1}, d_i) \times \dots \times l(x_{iT_i} | Y_{iT_i-1}) l(w_{iT_i} | Y_{iT_i-1}, x_{iT_i}).$$

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

$$l = \prod_{i=1}^N l(w_{i1}, x_{i1}, \dots, w_{iT_i}, x_{iT_i} | \bar{x}_{i1}, d_i). \quad (37)$$

IV. Data

A. Dictionary of Occupational Titles

The DOT contains information on 12,099 occupations defined by the tasks performed by workers in those individual occupations. It was constructed by the U.S. Department of Labor to provide standardized occupational information for an employment service matching job applicants with job openings. The information included in the DOT is based on the on-site observation of jobs as they are performed in diverse business establishments and, for jobs that are difficult to observe, on information obtained from professional and trade associations.¹⁵ On this basis, in the revised fourth edition of the DOT, analysts rate each occupation with respect to 62 characteristics that include the aptitudes, temperaments, and interests necessary for adequate performance; the training time necessary

¹⁴ When x_t is missing, all the observations from that period on are dropped. This is because the Kalman filter algorithm here relies on the observation of \bar{x} . If x is missing, \bar{x} is an additional unobserved state variable that is serially correlated. If \bar{x} is dropped from the model, i.e., $G_4 = 0$, the method here can be easily extended to the case in which x_t is missing.

¹⁵ One might be concerned that task complexity cannot be correctly measured by observing jobs performed because what analysts observe is a realized combination of job tasks and worker skills in equilibrium: even when the task is simple, an analyst might consider it complex if the worker is skilled. This confusion should be at least partially avoided because the job information is obtained from other sources as well (e.g., interviewing incumbents and supervisors). It is also worth noting that the DOT explicitly states that each occupation is defined on the basis of the tasks performed. See Miller et al. (1980) for a critical review of the DOT.

to prepare for an occupation; the physical demands of the occupation; and the working conditions under which work in the occupation typically occurs.

Many characteristics are measured by a multipoint scale and have detailed definitions. For example, the variable DATA measures the complexity of tasks in relation to information, knowledge, and conceptions by integers from 0 to 6. Tasks at the lowest level of complexity involve judging the readily observable characteristics of data. Examples include sorting hats according to color and size as specified, comparing invoices of incoming articles with the actual number and weights of articles, and so on. Tasks at the intermediate level of task complexity involve compiling information. Examples include summarizing details of transactions; collecting, classifying, and recording data; and receiving customer complaints to record and file them for future processing. Tasks at the highest level of complexity involve integrating analysis of data to discover facts or developing knowledge concepts of interpretations. Examples include formulating hypotheses and experimental designs, writing critical reviews of art for publication, and conducting research. Other tasks such as interpersonal communication and operating machines or equipment are also evaluated in a similar manner. Some tasks, particularly those measuring physical demand, are measured by a binary variable that takes the value of one if the occupation involves the task and zero otherwise.

To facilitate interpretation of the data, I summarize the detailed information in the DOT by constructing a low-dimensional vector of occupational tasks through principal component analysis (PCA). Previous studies take two different approaches. The first approach, which is employed by Bacolod and Blum (2010) and Yamaguchi (2010a), assumes that a subset of DOT variables measures the complexity of a single task. For example, it is assumed that three general educational development (reasoning, mathematics, and language) variables measure cognitive task complexity but do not measure other types of tasks such as motor tasks. This approach requires a priori knowledge about which variables measure which types of tasks. The second approach assumes that a DOT variable contains information about several underlying task complexities that are orthogonally distributed. For example, unlike the first approach, it is assumed that the general educational development variables above contain information about both cognitive and motor task complexity. The second approach does not require the a priori knowledge mentioned above but does impose that task complexity be orthogonally distributed. Ingram and Neumann (2006) employ a method similar to this.

In general, it seems impossible to determine which approach is better than the other because these two approaches impose different restrictions. Nevertheless, given the data, I find the first approach to be more suitable for this article because the constructed task complexity vector has a clear

interpretation. This is not necessarily the case under the orthogonality assumption in the second approach, particularly when two seemingly unrelated DOT variables have high loadings on the same factor. For example, it is hard to interpret a factor when a variable related to worker intelligence and a variable related to physical strength have high loadings for that factor. In the first approach, this possibility is excluded by construction: a variable related to worker intelligence is not used to construct a physical strength measure.

In light of the job analysis literature, as well as previous economics papers that use the DOT, this article assumes that tasks are broadly categorized into either cognitive tasks or motor tasks. While I assume these two task categories a priori, PCA conducted with the orthogonality assumption also yields a similar set of factors.¹⁶

Autor et al. (2003) consider different task groups, including routine and nonroutine manual tasks, in order to understand the role of technological change in the labor market, but these tasks capture different aspects of the motor tasks needed to perform a job. More narrowly defined task categories could also be considered instead of cognitive and motor tasks. For example, cognitive tasks could be separated into intelligent and communication tasks. However, these two are highly correlated with each other (the correlation coefficient is .73) and do not seem to provide any additional insights.¹⁷ Breaking down tasks increases the number of free parameters and, thus, complicates the analysis.¹⁸

By examining the textual definitions of the DOT variables, I assume which variables measure which type of task complexity. The following choices seem reasonable, and I confirm that the constructed task complexity index is robust to the choice of the DOT variables.¹⁹ The DOT variables that measure cognitive task complexity consist of two worker function variables (data and people), three general educational development variables (reasoning, mathematical, and language), three aptitude variables (intelligence, verbal, and numerical), and three adaptability var-

¹⁶ See app. E for a robustness check on this issue.

¹⁷ I construct communication task complexity indexes and estimate the structural model with them. With three tasks, the communication tasks as well as the cognitive and motor tasks, the estimation results are not robust to the choice of the sampling criteria. As can be seen in app. E, the results for the presented model with two tasks are robust.

¹⁸ Although cognitive and motor tasks seem most suitable for my sample of men, different task categorization may be more useful for a different demographic group. For example, if one is interested in the highly educated, considering numerical tasks and verbal tasks may make more sense than cognitive and motor tasks because the highly educated are unlikely to be in occupations in which they are rewarded for their physical abilities. Task categories should be chosen depending on the purpose of analysis.

¹⁹ See app. E for a robustness check on this issue.

iables (influencing people, accepting responsibility for direction, and dealing with people). Motor task complexity is measured by one worker function variable (things), motor coordination, finger dexterity, manual dexterity, eye-hand-foot coordination, spatial perception, form perception, color discrimination, setting limits, tolerance or standards, and 20 physical demand variables.

In the PCA, factor loadings are calculated so that the variation of the data explained by the constructed variables is maximized (or to minimize the information loss, equivalently).²⁰ For this purpose, I take a sample of occupational characteristics from the April 1971 Current Population Survey (CPS) augmented by the fourth edition of the DOT.²¹ This augmented CPS file contains the 1970 census occupation code, the DOT occupation code, and the DOT variables. I update occupational characteristics using the revised fourth edition by matching the DOT occupation code.²² The result of the PCA is available in appendix E and shows that the factor loadings are intuitive. Following Autor et al. (2003), I further convert these first principal components into percentile scores using the weights from the augmented CPS file.²³ The resulting indexes at the individual level are aggregated into the level of the 1970 census occupation by taking the sample mean for each of the three-digit occupations so that they can be merged with the NLSY.

To see whether the constructed variables characterize occupations reasonably, I report the mean and standard deviation of the task complexity measures for each census one-digit occupation in table 1. The cognitive tasks of professionals are the most complex, followed by those of managers. Cognitive tasks of laborers and household service workers are at the lowest level of complexity. This cognitive task complexity measure largely matches the conventional one-dimensional notion of skill levels seen in the empirical literature (e.g., Gibbons et al. 2005). However, this index alone is not rich enough to describe heterogeneous tasks across occupations. For example, cognitive task complexity is similar between

²⁰ See app. E for a technical discussion of PCA.

²¹ Most previous studies, including Ingram and Neumann (2006), Poletaev and Robinson (2008), and Bacolod and Blum (2010), use the 1995 DOT that does not contain the number of workers holding each occupation. Those authors are forced to assume that the weights of each DOT occupation are equal, despite the fact that, in the real economy, some occupations may have few workers and some occupations may have thousands. This limitation may bias the task complexity index in an unpredictable direction. I avoid this problem because the augmented CPS contains the number of workers in each DOT occupation.

²² The DOT occupational code crosswalk between the fourth and the revised editions is used for matching.

²³ I also estimate the model using the first principal components without converting them into percentile scores. The results are almost unchanged. See app. E for details.

Table 1
Task Complexity by Occupation at One-Digit Classification

	Cognitive Task		Motor Task		No. Observations
	Mean	SD	Mean	SD	
Professional	.85	.14	.45	.33	7,522
Manager	.79	.15	.21	.21	5,538
Sales	.57	.17	.23	.15	3,748
Clerical	.49	.16	.56	.22	9,270
Craftsmen	.52	.20	.82	.20	6,557
Operatives	.20	.18	.58	.20	5,824
Transport	.28	.15	.63	.10	1,774
Laborer	.15	.16	.46	.13	2,818
Farmer	.68	.19	.78	.14	1,117
Farm laborer	.18	.19	.53	.16	882
Service	.32	.22	.44	.24	6,834
Household service	.20	.11	.24	.23	1,469
All occupations	.49	.29	.50	.29	53,353

NOTE.—Sample consists of all working individuals in the 1971 April Current Population Survey augmented with occupational characteristics variables from the revised fourth edition of the Dictionary of Occupational Titles (1991). Sample size is 53,353. Task complexity measures are percentile scores divided by 100.

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 the most complex, while those of sales workers, household service workers, and managers are the least complex. These features are quite intuitive, and the proposed measurement is a useful description of the heterogeneity of occupations. Another important finding here is that task complexity varies within one-digit occupation. The reported standard deviation is large in all one-digit occupations. To assess the extent of heterogeneity within occupations more formally, I decompose the total task complexity variance into the within-occupation variance and the between-occupation variance.²⁴ For cognitive task complexity, about 59% of the total variance is explained by the within-occupation variance. For motor task complexity, the within-occupation variance explains a larger fraction of the total variance at 75%. This variance decomposition indicates that tasks are greatly heterogeneous within one-digit occupations and that quite a large part of task complexity variation would be lost if one relied on the one-digit occupation code. Dealing with occupations at the three-digit level is essential to accounting for the heterogeneity of jobs thoroughly.

²⁴ The law of total variance states that

$$\text{Var}(X) = E(\text{Var}(X|O)) + \text{Var}(E(X|O)),$$

where X is the task complexity index, and O is the occupational affiliation at one-digit level. The first term is the within-occupation variance, and the second term is the between-occupation variance.

Dispersion of the task complexity index within one-digit occupation is so large that two different one-digit occupations overlap in terms of task complexity. Figure 1 is a scatter plot of the task complexity indexes for three-digit occupations. In the top panel, I show a plot of task complexity indexes of three-digit occupations that are broadly categorized into professional and clerical occupations at the one-digit level. By and large, professionals' cognitive tasks are more complex than those of clerical workers. However, for some professional occupations, cognitive tasks are less complex than those of some clerical workers. The bottom panel also shows a similar pattern for craftsmen and operatives. Broadly speaking, craftsmen work on more complex cognitive and motor tasks than operatives. But, there are some operatives whose cognitive and motor tasks are more complex than those of some craftsmen. The results indicate that the idea that a certain one-digit occupation is uniformly more skilled than the other is questionable.

B. National Longitudinal Survey of Youth 1979

The NLSY 1979 is particularly suitable for this study because it contains detailed individual career histories and focuses on youths; young individuals change occupations more frequently than older individuals. I concentrate on male workers who make a long-term transition to the full-time labor market during 1979–2000.²⁵ I define a long-term transition to occur when an individual spends 3 consecutive years working 30 hours per week or more. After removing those who do not make a long-term transition and those with missing values, I have the career histories of 2,417 men, with 32,849 person-year observations of occupational choices and 27,063 person-year observations of wages. The full details of data set construction and the robustness of the estimates to sampling criteria are discussed in appendix D and appendix E, respectively. The sample mean AFQT (Armed Forces Qualification Test) score divided by 100 is 0.50, and the sample standard deviation is 0.29. The sample mean years of education is 13.22, and the sample standard deviation is 2.45. In the sample, 10% of individuals are black, and 7% are Hispanic.

Previous empirical papers, including Neal (1999) and Sullivan (2009),

²⁵ Observations after 2000 are not included in the sample because in surveys later than 2002 the occupation code is not compatible with that used until 2000. Since the 2002 survey, the NLSY adopted the 2000 census three-digit occupation code, and it is quite different from the 1970 census code. Although Integrated Public Use Microdata Series-CPS provides an occupational crosswalk, some occupations in the 2000 census code do not exist in the 1970 census code, and thus the DOT occupational characteristics are missing. O*NET, which is the successor of the DOT, contains occupational characteristics for these recent occupations, but the O*NET variables are not compatible with those of the DOT. Due to these limitations, I drop the post-2000 observations in the NLSY.

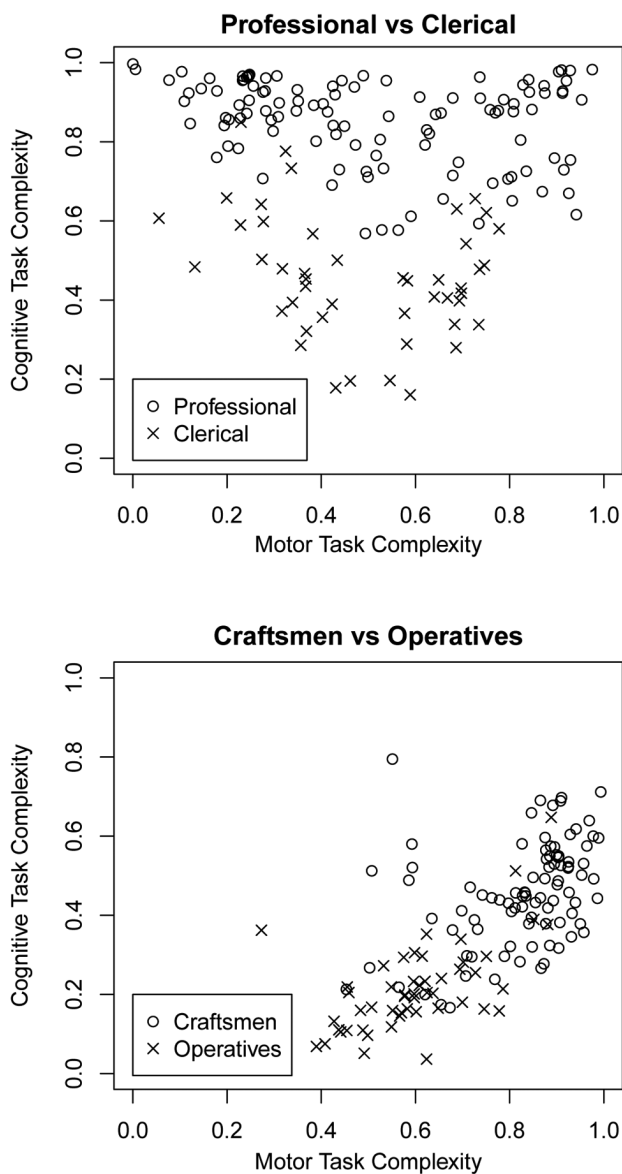


FIG. 1.—Task complexity comparison. Task complexity measures are percentile scores divided by 100. Source: 1971 April Current Population Survey augmented with occupational characteristics variables from the revised fourth edition of the Dictionary of Occupational Titles (1991).

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), Yamaguchi (2010b), and Pavan (2011) 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 3 years regardless of whether he changes an employer or not. I edit the 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. I estimate the model using both the unedited data and the data edited by the method outlined here. The results available in appendix E indicate that this editing method affects the parameter estimates little. A possible explanation for this is that an occupation is often miscoded to other occupations that are similar in terms of job characteristics. If this is the case, when the NLSY falsely classifies the same occupation as different, it may result in only a small measurement error in task complexity. Because there is no evidence that this editing method makes the data close to the truth and editing makes little difference in the parameter estimates, I take the data as they are.

C. Career Progression Patterns

The time profiles of the average wage and occupational task complexity are presented in figures 2–5. At the point of long-term transition to the labor market, the average cognitive task complexity index of men is 0.41. Individuals take on more and more complex cognitive tasks over time; the cognitive task complexity index reaches .52 in 10 years and .55 in 20 years. The mean motor task complexity slightly decreases over time: at the entry to the labor market, it is .53 and decreases to .52 in 10 years and .51 in 20 years. Mean wages grow by 48% in 10 years and by 64% in 20 years as measured by log wage differences.

These profiles are also presented for three different schooling levels: high school dropouts (education less than 12 years), high school graduates (12 years of education), and college workers (education more than 12 years). The cognitive task complexity indexes at labor market entry are substantially different across education groups, with more educated workers taking jobs that involve more complex cognitive tasks. The average indexes at entry are .28 for dropouts, .33 for high school graduates, and .53 for college workers. But, for all education groups, cognitive task com-

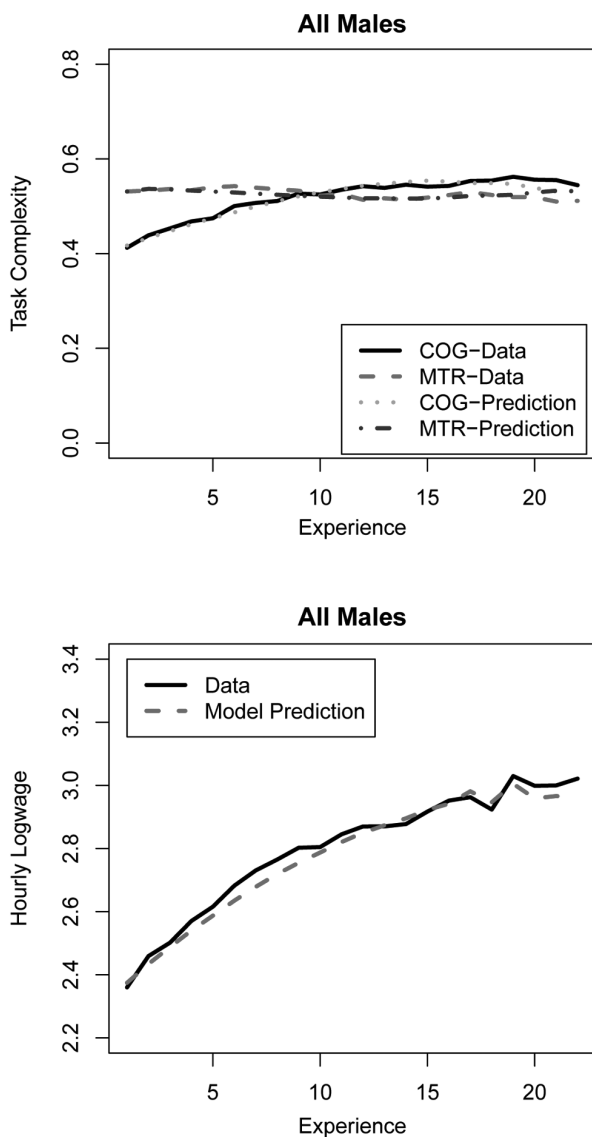


FIG. 2.—Model fit (all men). COG = profiles of cognitive task complexity; MTR = profiles of motor task complexity. Source: Author's estimates from the National Longitudinal Survey of Youth 1979–2000. Sample consists of 325 high school dropouts, 1,009 high school graduates, and 1,083 college workers. Color version available as an online enhancement.

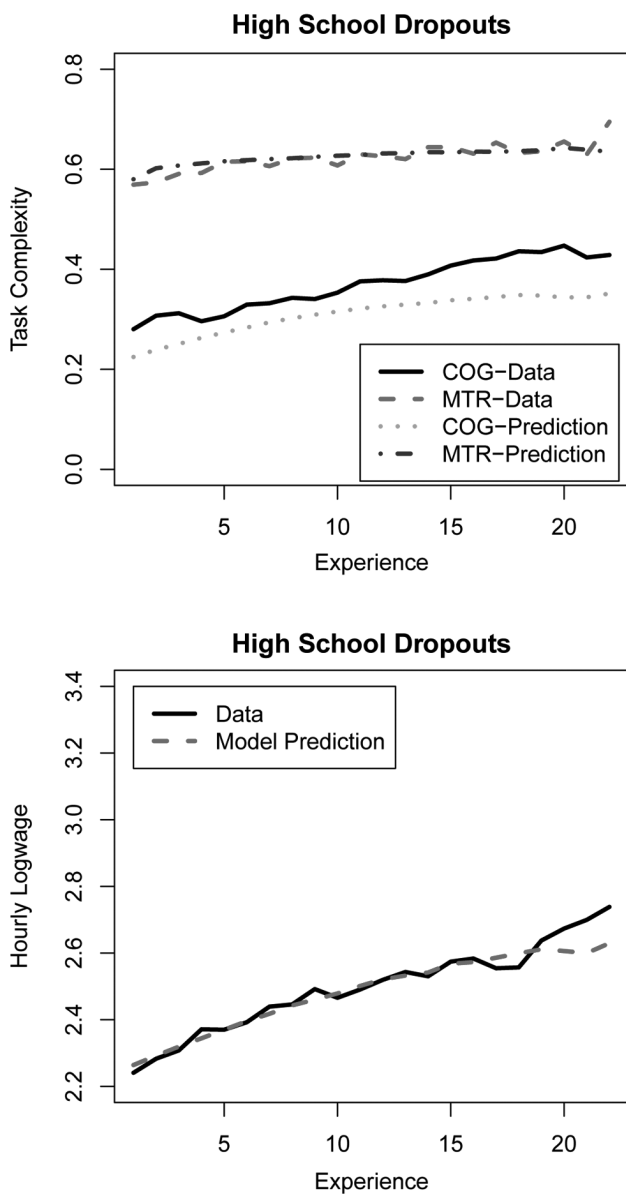


FIG. 3.—Model fit (high school dropouts). COG = profiles of cognitive task complexity; MTR = profiles of motor task complexity. Source: Author's estimates from the National Longitudinal Survey of Youth 1979–2000. Sample includes 325 high school dropouts. Color version available as an online enhancement.

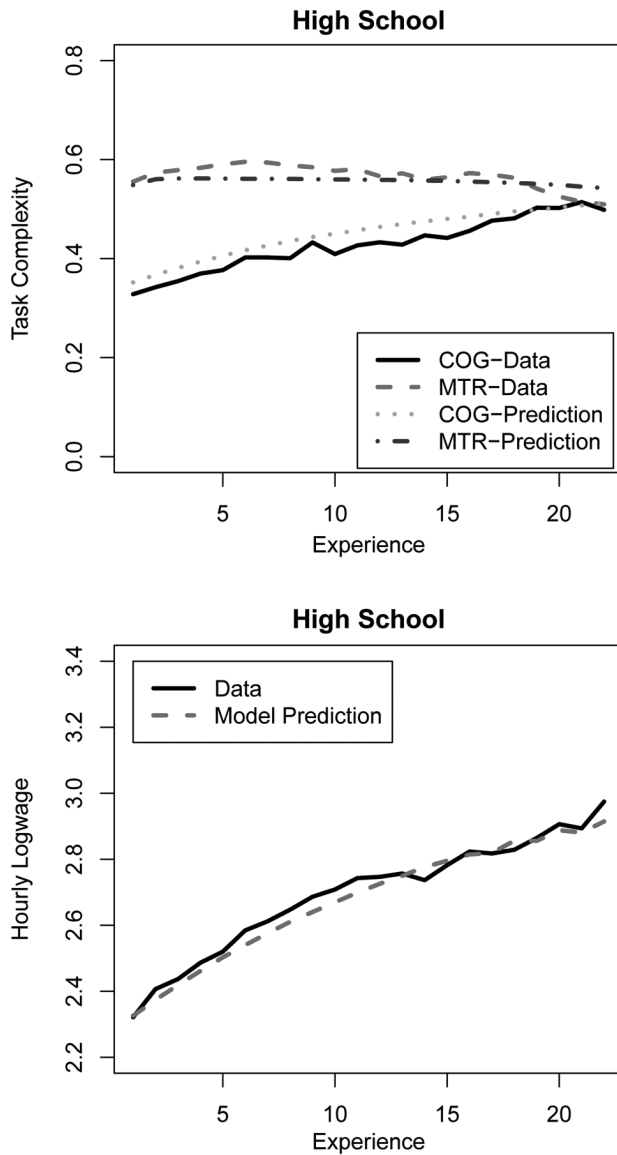


FIG. 4.—Model fit (high school workers). COG = profiles of cognitive task complexity; MTR = profiles of motor task complexity. Source: Author's estimates from the National Longitudinal Survey of Youth 1979–2000. Sample includes 1,009 high school graduates. Color version available as an online enhancement.

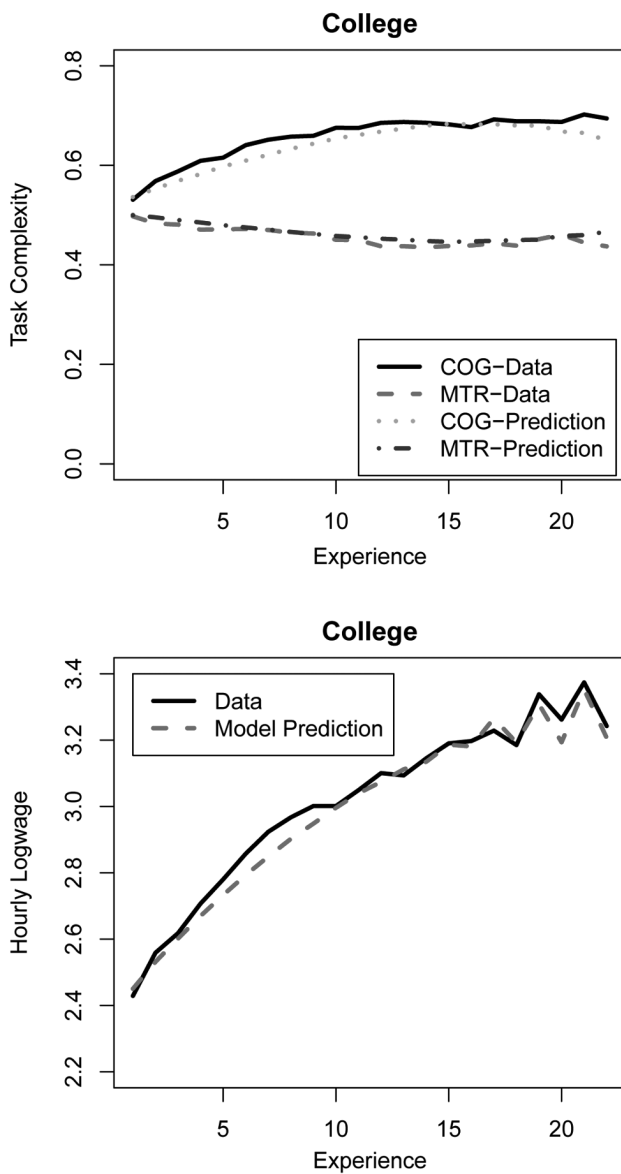


FIG. 5.—Model fit (college workers). COG = profiles of cognitive task complexity; MTR = profiles of motor task complexity. Source: Author's estimates from the National Longitudinal Survey of Youth 1979–2000. Sample includes 1,083 college workers. Color version available as an online enhancement.

plexity increases over time. It grows to .36 in 10 years and .44 in 20 years for dropouts and .41 in 10 years and .50 in 20 years for high school graduates. The profile for college workers is concave. The cognitive task complexity for college workers grows to .67 in 10 years, but the speed of the growth slows down, and it reaches .69 in 20 years.

The profiles of motor task complexity are also very different across education groups. Less educated workers tend to take jobs that involve more complex motor tasks. The motor task complexity at labor market entry is .57 for dropouts, .55 for high school graduates, and .50 for college workers. Unlike cognitive task complexity, motor task complexity does not increase over time for all education levels. For high school dropouts, it monotonically grows to .63 in 10 years and .64 in 20 years. The profile for high school graduates is hump shaped. It peaks at .60 in 6 years and decreases to .53 in 20 years. The motor task complexity for college workers monotonically decreases to .45 in 10 years and .44 in 20 years.

V. Estimation Results

A. Model Fit

Using the estimated parameters, I calculate the predicted paths of the mean wage and task complexity of occupations through simulation. I simulate each individual in the sample 1,000 times, from his entrance to the full-time labor market until the last year when he is seen in the data. To avoid potential attrition problems, if information about wage or occupational choice is missing in a certain year, I treat the corresponding simulation outcomes as missing.

Figures 2–5 compare the actual and predicted profiles of mean task complexity and hourly log wage over time. The predicted profiles for all men are very close to the actual profiles from the data. For high school dropouts, the model fit to the profiles of motor task complexity and log wage is good, but the predicted cognitive task complexity is lower than the actual profile. For high school graduates, cognitive task complexity is slightly overpredicted, but the level of the predicted motor task complexity and log wage is close to that of the data. Finally, for college workers, the model fits very well for all three dimensions. All in all, the model shows an ability to fit these interesting features of the data.

To see how well the model is able to match the variation in wages over time and across people within the broadly defined educational categories, figures 6 and 7 show wage profiles at the 10th, 30th, 50th, 70th, and 90th percentiles of the wage distribution in the actual and simulated data. At the 10th percentile, the model underpredicts wages, but the differences are small. The model shows a good fit to the data in this dimension as well.

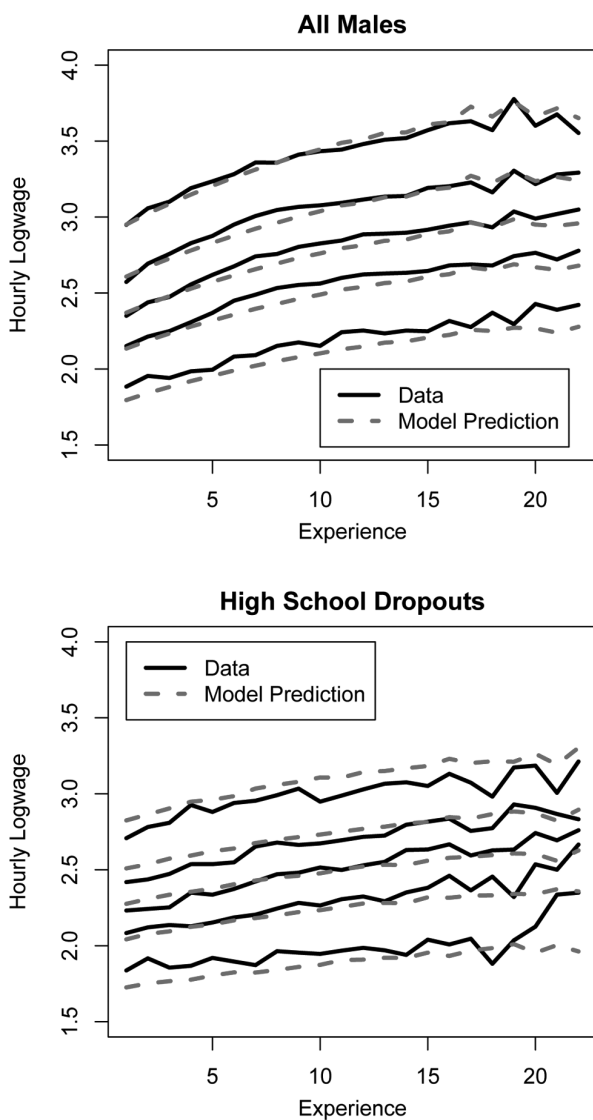


FIG. 6.—Model fit: wage growth by percentile (all men and high school dropouts). Tenth, 30th, 50th, 70th, and 90th percentiles of hourly log wages are plotted over time. Source: Author's estimates from the National Longitudinal Survey of Youth 1979–2000. Sample consists of 325 high school dropouts, 1,009 high school graduates, and 1,083 college workers. Color version available as an online enhancement.

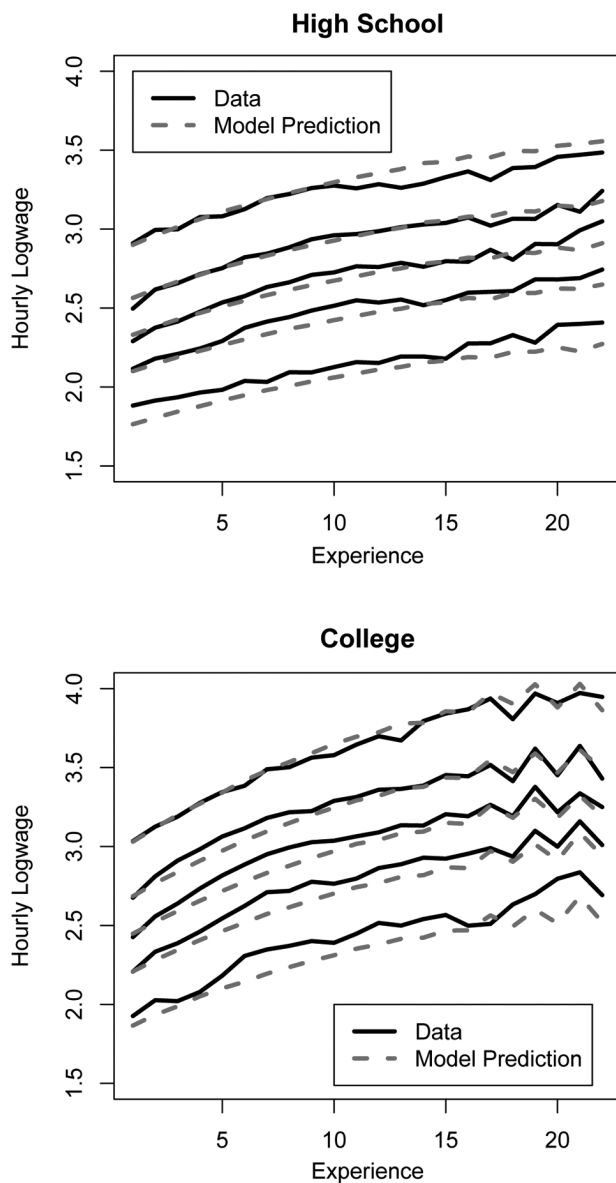


FIG. 7.—Model fit: wage growth by percentile (high school and college workers). Tenth, 30th, 50th, 70th, and 90th percentiles of hourly log wages are plotted over time. Source: Author's estimates from the National Longitudinal Survey of Youth 1979–2000. Sample includes 1,009 high school graduates and 1,083 college workers. Color version available as an online enhancement.

Table 2
Wage Equation

Notation	Estimate	SE
p_0	2.294	.014
$p_1(1)$.015	.012
$p_1(2)$.106	.019
$p_2(1)$.646	.065
$p_2(2)$.575	.078
$P_3(1, 1)$.108	.010
$P_3(2, 2)$.047	.012
σ_η^2	.061	.001

SOURCE.—National Longitudinal Survey of Youth 1979–2000. Sample consists of 2,417 men.

NOTE.—Parameter estimates are for the wage equation $\ln w_i = p_0 + p_1'x_i + (p_2 + P_3x_i)s_i + \eta_i$, where $\eta_i \sim N(0, \sigma_\eta^2)$. 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.

B. Parameter Estimates

1. Wage Equation

Table 2 presents the parameter estimates of the wage equation (see eq. [5]) and their standard errors. The implicit skill prices are given by $p_2 + P_3'x_i$ and significantly increase with task complexity x_i . The estimated implicit skill prices imply that an increase in cognitive skills by one unit (i.e., 1 SD of the initial skills) raises the log wage by 0.74 for a job at the 90th percentile of cognitive task complexity and by 0.65 for a job at the 10th percentile. An increase of motor skills by one unit raises the log wage by 0.62 for a job at the 90th percentile of motor task complexity and by 0.58 for a job at the 10th percentile. These estimates indicate that differences in returns to skills across occupations are sizable.

Using the estimated wage equation, I calculate the potential wage losses after job displacement. This exercise can also be interpreted as measuring the extent of skill transferability across jobs. To see the cognitive skill transferability, consider an average college worker with 10 years of experience. He has 1.405 units of cognitive skills. Suppose that this worker occupies a job at the 90th percentile of cognitive task complexity. If he moves to a job with the same motor tasks but with cognitive tasks at the 10th percentile of cognitive task complexity, then he suffers a 13% wage loss. Next, to demonstrate motor skill transferability, consider an average high school dropout worker with 10 years of experience. He has 0.871 units of motor skills. If this worker moves from a job at the 90th percentile of motor task complexity to a job at the 10th percentile while cognitive task complexity remains the same, his wage loss would be 12%. These estimates show that the displacement of a worker to a very different occupation results in a significant wage loss.

How could these estimates be interpreted in the context of the related literature? Kambourov and Manovskii (2009) estimate the returns to oc-

cupation-specific experience through the instrumental variables approach proposed by Altonji and Shakotko (1987) and find that 5 years of occupation-specific experience increases wages by 12%–20%.²⁶ Note that they do not estimate the contribution of occupation match quality to the wage growth experienced during the first 10 years. Pavan (2011) estimates a structural model of two-stage job search in the style of Neal (1999). His wage growth decomposition results show that a worker with 10 years of experience would suffer a 5%–12% wage loss after an exogenous job displacement. Yamaguchi (2010b) also estimates a model similar to Pavan (2011) and finds that the wage loss would be 20%–23%. These results must be carefully interpreted because all of these papers and this article use differing samples and models.

Nevertheless, the smaller wage losses after job displacement (or more skill transferability, equivalently) found in this article than in Kambourov and Manovskii (2009) and Yamaguchi (2010b) may suggest a limitation of the current model: a lack of truly occupation-specific skills. Two occupations could have very similar cognitive and motor tasks but could also require very different occupation-specific knowledge. For example, using the definition of occupations presented in the model, the three-digit occupations of economist and actuary are very similar since both require high cognitive skills but very low motor skills. However, there are certainly large amounts of economist- and actuary-specific knowledge that distinguish these occupations. In addition, a firm-specific component of wages is missing. The job search literature shows that a substantial amount of wage growth is derived from job search. This effect may be absorbed into the skill shocks in the current model. These issues should be addressed in future research to improve the understanding of skill transferability.

2. Skill Transition

Parameter estimates for the skill transition equation (see eq. [6]) and the initial skill endowment (see eq. [7]) are reported in table 3. The parameters for all diagonal elements of A_1 are positive and significant, indicating that skills grow faster when tasks are more complex. This result is consistent with skill acquisition through learning by doing and does not support the occupational choice mechanism postulated by Rubinstein and Weiss (2006).

Education and AFQT scores are also positively associated with faster cognitive skill growth, while they are negatively associated with motor skill growth. No significant differences in learning ability are found across race. Skill shocks are about 15% of one standard deviation for the initial

²⁶ Although occupation is differently defined, Yamaguchi (2010b) reports that average occupation-specific experience is about 5 years for workers with 10 years of experience.

Table 3
Transition Equation and Initial Conditions
for Skills

Notation	Estimate	SE
$D(1, 1)$.925	.002
$D(2, 2)$.911	.008
$a_0(1)$	−.309	.034
$a_0(2)$.227	.048
$A_1(1, 1)$.057	.009
$A_1(2, 2)$.101	.036
$A_2(1, 1)$, AFQT	.071	.032
$A_2(1, 2)$, Edu	.026	.003
$A_2(1, 3)$, Black	.015	.025
$A_2(1, 4)$, Hispanic	−.040	.026
$A_2(2, 1)$, AFQT	−.044	.044
$A_2(2, 2)$, Edu	−.020	.005
$A_2(2, 3)$, Black	−.036	.033
$A_2(2, 4)$, Hispanic	.050	.035
$\Sigma_e(1, 1)$.156	.019
$\Sigma_e(2, 1)$	−.125	.019
$\Sigma_e(2, 2)$.145	.021
$H(1, 1)$, AFQT	.514	.410
$H(1, 2)$, Edu	.184	.040
$H(1, 3)$, Black	.500	.327
$H(1, 4)$, Hispanic	−.237	.344
$H(2, 1)$, AFQT	−.051	.469
$H(2, 2)$, Edu	−.202	.046
$H(2, 3)$, Black	−.638	.366
$H(2, 4)$, Hispanic	.371	.388
$\Sigma_{e_1}(1, 1)$.654	.075
$\Sigma_{e_1}(2, 1)$	−.519	.080
$\Sigma_{e_1}(2, 2)$.685	.089

SOURCE.—National Longitudinal Survey of Youth 1979–2000. Sample consists of 2,417 men.

NOTE.—Parameter estimates are for the skill transition equation $s_{t+1} = Ds_t + a_0 + A_1x_t + A_2d + \varepsilon_{t+1}$, where $\varepsilon \sim N(0, \Sigma_e)$. The initial skills are given by $s_1 = b + Hd + \varepsilon_1$, where $\varepsilon_1 \sim N(0, \Sigma_{e_1})$ and d is a vector of the Armed Forces Qualification Test percentile score divided by 100, years of education, and dummy variables for race with whites being the reference group. The unconditional mean and variance of initial skills are normalized to 0 and 1, respectively. 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.

skill distribution. Cognitive and motor skill shocks are highly and negatively correlated with a correlation coefficient of $-.83$, which is calculated from the estimates for the variance-covariance matrix Σ_e . The annual skill depreciation rates for cognitive and motor skills are slightly less than 10%, which implies that skills are highly persistent over time.²⁷

The initial skill endowment differs across individuals in an unobserved

²⁷ These estimates seem reasonable, although no other papers estimate parameters that can be directly compared with mine. Mincer and Ofek (1982) and many others estimate the depreciation of human capital by extracting the variation of time out of work. This article cannot apply this identification strategy because all individuals in the model work full time. Nevertheless, skill depreciation parameters are still identified by the extent to which the task complexity of previous jobs affects the current wage and occupational choice.

way. For both cognitive and motor skills, about two-thirds of the skill variance present at labor market entry is due to unobserved skills. The unobserved component of the initial skill endowment is negatively correlated with a correlation coefficient of $-.77$, which is calculated from the estimates for the variance-covariance matrix Σ_{ε_1} . The AFQT score is positively associated with initial cognitive skills, while it is negatively associated with initial motor skills, although neither association is statistically significant. A year of education is significantly associated with .18 more units of initial cognitive skills and .20 fewer units of initial motor skills. I do not find that initial skill endowments are significantly different across race after controlling for AFQT scores and education.

3. Job Preference

The parameter estimates for job preferences (see eq. [9]) are reported in table 4. The AFQT score and education are positively and significantly correlated with preferences for complex cognitive tasks. They are also positively correlated with preferences for complex motor tasks, although not significantly. Blacks significantly dislike complex cognitive tasks. Otherwise, no significant differences in job preferences are found across race. The results indicate that substantial heterogeneity in job preferences exists across individuals. The parameter G_2 is positive for both tasks, although it is only significant for motor tasks. This implies that workers endowed with plenty of motor skills enjoy complex motor tasks.

The estimate of the parameter G_4 suggests that workers suffer a large adjustment cost when entering into an occupation that involves tasks that differ greatly from what they have experienced. The transition parameter for work habits is significantly positive for cognitive tasks at .475 but almost zero for motor tasks, which implies that the tasks of jobs held more than a year ago do not influence occupational choice through job preferences. Initial work habits also vary across individuals. The AFQT score and education are positively and significantly correlated with the initial conditions for work habits for cognitive tasks, while they are positively but insignificantly correlated with the initial conditions for work habits for motor tasks. Hispanics have slightly but significantly higher initial conditions for work habits for cognitive tasks. Otherwise, no significant differences in initial work habits are found across race.

VI. Discussion

A. Do Cognitive and Motor Skills Account for Wage Variance?

To assess the importance of each type of skill in explaining the variation of wages, I decompose the log wage variance. Because the wage consists of two skill components and a white noise term, the log wage variance is the sum of the variance of the cognitive skill component, the variance

Table 4
Job Preference and Work Habits

Notation	Estimate	SE
$g_0(1)$.075	.119
$g_0(2)$	−.171	.278
$G_1(1, 1)$, AFQT	.295	.074
$G_1(1, 2)$, Edu	.038	.009
$G_1(1, 3)$, Black	−.190	.056
$G_1(1, 4)$, Hispanic	.055	.063
$G_1(2, 1)$, AFQT	.047	.172
$G_1(2, 2)$, Edu	.043	.020
$G_1(2, 3)$, Black	.142	.134
$G_1(2, 4)$, Hispanic	−.208	.141
$G_2(1, 1)$.037	.020
$G_2(2, 2)$.326	.030
$G_4(1, 1)$	−13.048	.780
$G_4(2, 2)$	−.146	.045
$A_3(1, 1)$.475	.007
$A_3(2, 2)$.000	.074
$\bar{x}_{1,0}(1)$	−.218	.024
$X(1, 1)$, AFQT	.152	.019
$X(1, 2)$, Edu	.040	.002
$X(1, 3)$, Black	−.017	.015
$X(1, 4)$, Hispanic	.050	.017
$\bar{x}_{1,0}(2)$.255	.198
$X(2, 1)$, AFQT	.158	.155
$X(2, 2)$, Edu	.007	.017
$X(2, 3)$, Black	.219	.125
$X(2, 4)$, Hispanic	.259	.131

SOURCE.—National Longitudinal Survey of Youth 1979–2000. Sample consists of 2,417 men.

NOTE.—Utility from occupational characteristics is given by $v(x_i, \bar{x}_i, s_i, \bar{s}_i) = (g_0 + G_1d + G_2s_i + \bar{v}_i)x_i + x_i'G_3\bar{x}_i + (x_i - \bar{x}_i)'G_4(x_i - \bar{x}_i)$, where parameter G_3 is normalized as the negative of an identity matrix. The transition equation of work habit is $\bar{x}_{i+1} = A_3\bar{x}_i + (I - A_3)x_i$, 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 (Armed Forces Qualification Test) percentile score divided by 100, years of education, and dummy variables for race with whites being 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.

of the motor skill component, the covariance of the two, and the variance of the white noise term.²⁸

Table 5 presents the results for different education levels for years 1, 10, and 20 after labor market entry. In year 1 (i.e., at labor market entry), for all education groups, both cognitive and motor skill wage components have large variances. As the level of education increases, the ratio of the variance of the cognitive skill component to that of the motor skill component increases. The covariance term is large and negative, which reflects

²⁸ The wage can be written as

$$\ln w_i = p_0 + [p_1^C x_i^C + p(x_i)^C s_i^C] + [p_1^M x_i^M + p(x_i)^M s_i^M] + \eta_i,$$

where superscript C is for cognitive skills, and M is for motor skills. I call the term $p_1^C x_i^C + p(x_i)^C s_i^C$ the cognitive skill component and the term $p_1^M x_i^M + p(x_i)^M s_i^M$ the motor skill component.

Table 5
Log Wage Variance Decomposition

	Cognitive	Motor	Covariance	Error	Total
Year 1:					
All men	.482	.378	-.715	.061	.205
High school dropouts	.353	.339	-.565	.061	.188
High school	.321	.287	-.471	.061	.198
College	.421	.323	-.600	.061	.204
Year 10:					
All men	.913	.546	-1.226	.061	.293
High school dropouts	.565	.470	-.871	.061	.224
High school	.502	.399	-.729	.061	.233
College	.755	.458	-.997	.061	.277
Year 20:					
All men	1.167	.617	-1.484	.061	.360
High school dropouts	.652	.520	-.991	.061	.242
High school	.571	.431	-.813	.061	.249
College	.934	.506	-1.173	.061	.328

NOTE.—Author's estimates from the National Longitudinal Survey of Youth 1979–2000. Sample consists of 325 high school dropouts, 1,009 high school graduates, and 1,083 college workers.

the estimation result that the initial unobserved cognitive and motor skills as well as the cognitive and motor skills shocks are highly and negatively correlated.²⁹ As time in the labor market increases, the log wage variance grows to .293 in year 10 and to .360 in year 20 for all men. This increase is achieved through the growth of both cognitive and motor skill wage components. However, the ratio of the variance of the cognitive skill component to that of the motor skill component grows over time for all levels of education.

What are the sources of the skill and wage differences across individuals? Are they due to the differences established at labor market entry, or are they the result of shocks that individuals receive over their careers? Individuals enter the labor market with different initial skill endowments, learning abilities, job preferences, and work habits. While the effect of the differences in initial skill endowments and work habits on inequality later in life decreases over time due to depreciation, differences in learning abilities and job preferences may have persistent effects on inequality throughout the lifecycle. Job preferences affect wages through the choice of occupation because implicit skill prices and learning opportunities are different across occupations. To evaluate the effects of initial conditions on lifecycle skill and wage inequality, the model is simulated under the restriction that individuals are homogeneous at labor market entry. When eliminating het-

²⁹ This result is partly driven by the use of percentile scores as task complexity indexes. When task indexes are normally distributed, the estimated correlation coefficient is greater (or smaller in absolute value). See app. B for a detailed discussion on this issue.

Table 6
Log Wage Variance When Initial Conditions Are Homogeneous

Year	Benchmark	Homogeneous			All
		Preference	Initial Skills	Learning Ability	
1	.206	.204	.061	.206	.061
10	.292	.260	.234	.241	.190
20	.359	.297	.335	.257	.232

NOTE.—Author’s estimates from the National Longitudinal Survey of Youth 1979–2000. Sample consists of 2,417 men.

erogeneity, I assume that the distributions are degenerated at the mean.³⁰ This assumption does not change the profiles of mean skills. Five simulations are conducted under five different assumptions: (1) a benchmark simulation in which initial conditions are heterogeneous, (2) task preferences are homogeneous, (3) initial skill endowments are homogeneous, (4) learning abilities are homogeneous, and (5) all initial conditions are homogeneous.

Table 6 presents the results for the variance decomposition of the log wage over time. In year 1, about 70% ($1 - 0.061/0.206$) of the log wage variance is explained by differences in initial skill endowments, and the majority of the remaining variance is due to the white noise term. Differences in preferences and learning abilities do not affect the log wage variance in year 1. In year 10, the log wage variance in the benchmark simulation is 0.292. It decreases to 0.260 when preferences are homogeneous, to 0.234 when initial skill endowments are homogeneous, to 0.241 when learning abilities are homogeneous, and to 0.190 when all the initial conditions are homogeneous. These results imply that about 35% of the log wage variance in year 10 is explained by the initial conditions and that these three different factors have similar effects. In year 20, the log wage variance in the benchmark simulation is 0.359. It decreases to 0.297 when preferences are homogeneous, to 0.335 when initial skill endowments are homogeneous, to 0.257 when learning abilities are homogeneous, and to 0.232 when all the initial conditions are homogeneous. All together, the initial conditions account for about 35% of the log wage variance in year 20. This result is similar to that found in year 10, but the three factors affect the log wage variance differently in each year. Notice that the effect of the initial skill endowments on the log wage variance almost disappears in year 20, while differences in preferences substantially affect the log wage variance, and differences in learning ability have the largest effect on the log wage variance in the same year.

The college/high school log wage gaps are also examined using the same

³⁰ To simulate homogeneous agents, replace the vector of individual characteristics d_i by the sample mean \bar{d} and set $\Sigma_{\epsilon_1} = 0$. All other parameters remain the same.

Table 7
College/High School Log Wage Gaps When Initial Conditions Are Homogeneous

Year	Benchmark	Homogeneous		
		Preference	Initial Skills	Learning Ability
1	.126	.107	.011	.126
10	.340	.231	.283	.155
20	.460	.308	.432	.167

NOTE.—Author’s estimates from the National Longitudinal Survey of Youth 1979–2000. Sample consists of 2,417 men.

set of simulations. Table 7 presents the results with the exception of simulation 5, in which all initial conditions are homogeneous because no differences exist between the two groups by construction. Remember that college workers in this article are those who attended college for at least 1 year and may not have earned a degree. In year 1, the log wage gap in the benchmark simulation is 0.126 and is explained almost entirely by differences in initial skill endowments. The effect of the differences in initial skill endowments on the log wage gap diminishes over time: they account for 0.028 ($0.460 - 0.432$) out of a 0.460 log wage gap in year 20. The differences in preferences significantly account for 0.109 ($0.340 - 0.231$) of the log wage gap in year 10 and 0.152 ($0.460 - 0.308$) in year 20, which implies that heterogeneity in preferences composes about a third of the college/high school wage gap in year 20. The differences in learning ability have the largest (and increasing) influence on the log wage gap. When high school and college workers have the same learning ability, the log wage gap decreases by 0.185 ($0.340 - 0.155$) in year 10 and by 0.293 ($0.460 - 0.167$) in year 20, which implies that heterogeneity in learning ability accounts for about two-thirds of the wage gap in year 20.

The variance decomposition exercise reveals that differences in both cognitive and motor skills substantially account for the total variance of log wages, but cognitive skills are more important than motor skills for educated and experienced individuals. Differences established at labor market entry account for about 35% of the log wage variance 10 and 20 years after labor market entry. The remaining fraction of inequality is explained by shocks received over the individuals’ careers.³¹ The college/high school wage gap is largely explained by differences in learning ability. Differences in task preferences also account for a substantial fraction of the gap, about 30%. The effect of the initial skill endowments gradually decreases and almost disappears by year 20.

³¹ Differences in task complexity contribute to skill differences. However, for individuals with homogeneous initial conditions, the task complexity differences are due to independent and identically distributed shocks ν_t and ε_t . This is why all differences are either due to initial conditions or due to shocks.

Table 8
Mean Skill Profiles by Education

Year	All Men	High School Dropouts	High School	College
Cognitive skills:				
1	.000	-.813	-.269	.498
10	.631	-.650	.206	1.405
20	.996	-.539	.489	1.923
Motor skills:				
1	.000	.731	.240	-.448
10	-.066	.871	.240	-.637
20	-.108	.950	.238	-.750

NOTE.—Author’s estimates from the National Longitudinal Survey of Youth 1979–2000. Sample consists of 325 high school dropouts, 1,009 high school graduates, and 1,083 college workers.

B. The Growth of Skills and Wages

How do unobserved skills grow over the careers of workers? Are the skill growth patterns different across groups? Remember that the observed task complexity does not necessarily mirror underlying skills because not only skills, but also job preferences, affect the choice of occupations. It is possible for one worker to possess more skills than the other even though the former’s tasks are not as complex than the latter’s. Time profiles of the underlying skills cannot be uncovered without the model.

Using the parameter estimates, I present the calculated time profiles of mean unobserved skills by education in table 8. In year 1, initial skill endowments differ according to education. Remember that the mean of the initial skill distribution is set to zero and that the skill scale is normalized by making the standard deviation of the initial skill distribution equal to one. The initial cognitive skills are -0.813 for high school dropouts, -0.269 for high school graduates, and 0.498 for college workers. In contrast, initial motor skills decrease with education: they are 0.731 for high school dropouts, 0.240 for high school graduates, and -0.448 for college workers. Cognitive skills grow over time for all levels of education but grow faster for the educated for two reasons. First, educated workers have a higher learning ability. Second, educated workers tend to enter occupations with high cognitive task complexity and, thus, enjoy more on the job-skill-learning opportunities. Consequently, the cognitive skill gap across education levels grows over time because of self-selection into occupations and skill accumulation through learning by doing. Motor skills grow for high school dropouts, but they are constant for high school graduates and decreasing for college workers, which results in a gap in motor skills across education levels that grows over time. This result is also driven by different learning abilities, self-selection, and learning by doing.

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. Notice that the profile of skills

Table 9
Accumulated Wage Growth by Skill Type and Education

Years since Entry	Benchmark				All Men	
	Dropouts (1)	High School (2)	College (3)	All Men (4)	CF 1 (5)	CF 2 (6)
Cognitive skills:						
5	.068	.209	.416	.282	.276	.271
10	.124	.362	.716	.487	.472	.454
15	.166	.472	.927	.634	.612	.580
20	.197	.549	1.074	.737	.710	.665
Motor skills:						
5	.061	.003	-.069	-.021	-.026	-.027
10	.098	.003	-.120	-.038	-.045	-.044
15	.124	.002	-.156	-.052	-.059	-.055
20	.141	.000	-.183	-.063	-.071	-.062
Total:						
5	.129	.212	.347	.261	.249	.244
10	.222	.365	.596	.448	.427	.410
15	.291	.474	.771	.582	.552	.525
20	.337	.549	.892	.674	.639	.603

NOTE.—Author’s estimates from the National Longitudinal Survey of Youth 1979–2000. Sample consists of 325 high school dropouts, 1,009 high school graduates, and 1,083 college workers. CF = counterfactual.

itself, even if observed, does not answer this question because a rapid growth of skills does not translate into a sizable wage growth if skill prices are low. Table 9 presents the wage growth accumulated during the first 5, 10, 15, and 20 years after labor market entry. Cognitive skills are the main source of wage growth for all education groups, with their contribution to wage growth increasing as higher levels of education are attained. Cognitive skills increase wages by about 12% for high school dropouts, 36% for high school graduates, and 72% for college workers during the first 10 years. The wage growth due to cognitive skill growth slows down in the next 10 years for all levels of education. Motor skills contribute to wage growth differently across education. For college workers, wages decrease due to the deterioration of motor skills by 7% in 5 years, 12% in 10 years, and 18% by 20 years. Motor skill growth contributes little to the wage growth of high school graduates. However, for high school dropouts, increasing motor skills are an important source of wage growth. Their wages grow by 13% in 5 years, 22% in 10 years, and 34% in 20 years, which implies that motor skills account for about half of high school dropouts’ wage growth. The results indicate that cognitive skills are generally the main source of wage growth and that they are more important for more highly educated workers. Motor skills contribute only to the wage growth of high school dropouts and account for about half of their wage growth in the first 20 years after their full-time transition to the labor market.

Wage growth is driven by changes in tasks as well as changes in skills. To see the contribution of changes in tasks to wage growth, I calculate

counterfactual wages under two assumptions. In the first simulation, to evaluate the overall contribution of changes in task complexity over time, I calculate wages over time when workers are stuck in their first jobs. The results in column 6 of table 9 show that wages in this simulation are about 4 and 7 percentage points lower than those in the benchmark simulation in years 10 and 20, respectively. The wage differences are the result of lower task complexity leading to lower skill prices and fewer opportunities for skill accumulation. To isolate the effects of the lower skill prices resulting from lower task complexity, I calculate counterfactual wages when task complexity is held fixed at the first period level but with skills identical to those in the benchmark simulations. The results in column 5 of table 9 show that the wage growth rate of men 20 years since labor market entry drops by 3 percentage points. These modest effects stemming from changes in task complexity may again indicate that the model does not fully account for truly occupation-specific skills.

VII. Conclusion

In this article I propose a new approach to modeling heterogeneous human capital through a novel application of the DOT. While task information is usually used as a skill proxy, this article clearly distinguishes between tasks and skills. A key feature of the model is that an occupation is defined as a bundle of tasks, and thus it is characterized in a continuous space of task complexity. This approach has two advantages over the previously employed Roy-type models. First, it can account for many heterogeneous occupations without computational burden and a loss of precision in the parameter estimates because the number of parameters and state variables do not increase with the number occupations in the model. This feature is appealing both to structural estimation and to reduced-form estimation. Second, the model facilitates an interpretation as to how and why skills are differently rewarded across occupations. In addition, I show that a closed-form solution for optimal occupational choice exists under some functional-form assumptions. This analytical simplicity dramatically reduces computational burden and allows me to include many observed worker characteristics, which are often intractable in structural dynamic models. The solution also allows for an interpretation of the observed task complexity as a noisy signal of underlying skills, which helps me identify the dynamics of unobserved skills from the observed task information.

The model is estimated by the Kalman filter, using the task complexity measures from the DOT and career and wage histories from the NLSY. The parameter estimates indicate that returns to skills increase with task complexity and that skills grow faster when the worker is employed in an occupation characterized by more complex tasks. The log wage variance

decomposition analysis shows that both cognitive and motor skills account for a significant fraction of cross-sectional log wage variances. In terms of wage growth, cognitive skills play a central role. They account for all of the wage growth for high school and college workers. However, for high school dropouts, the growth of motor skills accounts for about half of their wage growth.

While this article demonstrates that the proposed model is useful in empirical human capital research, it still leaves some important areas for extensions and future research. First, truly occupation-specific and firm-specific components of wages are missing in the current model. Second, it would be useful to make schooling decisions endogenous. The empirical results indicate that large differences in skills and wages exist across education, but the sources of the differences remain unknown. Third, it would be useful to allow for endogenous labor force participation and hours of work. Although assuming that all individuals work full time does not seem a serious concern for male workers, this extension is essential for the analysis of female workers. The extended model would be useful for understanding gender gaps in labor market outcomes from the viewpoint of tasks and skills.

Appendix A

Identification

This appendix discusses identification of the wage equation, skill equation, and the covariance skill distribution in a simplified setting. Denote the k th element of a vector z by $z(k)$. Similarly, denote the (k, l) element of a matrix Z by $Z(k, l)$. The wage equation can be rewritten as

$$\ln w = p_0 + \sum_{k=1}^K p_1(k)x(k) + \sum_{k=1}^K [p_2(k) + P_3(k, k)x(k)]s(k) + \eta.$$

To simplify the discussion, assume that skills are given by

$$s(k) = \alpha_1(k)z_1(k) + \alpha_2(k)z_2 + \tilde{s}(k),$$

where z_1 is a vector of variables of which the k th element affects the k th skills only (i.e., task complexity of the past job), z_2 is a scalar variable that can affect all types of skills (i.e., education, AFQT score, race), and \tilde{s} is a vector of unobserved components of skills. Because skills have no natural unit, they must be normalized. Here, I normalize skills by setting $\alpha_1(k) = 1$ and $E(\tilde{s}(k)) = 0$ because it simplifies the analysis below. In the main text, I normalized skills by setting the initial skill variance to be one and the mean of the initial skills to be zero because it facilitates interpretation of the empirical results. Both ways of normalization are valid. Substituting the skill equation, I rewrite the wage equation as

$$\ln w = p_0 + \sum_{k=1}^K p_1(k)x(k) + \sum_{k=1}^K [p_2(k) + P_3(k, k)x(k)][z_1(k) + \alpha_2(k)z_2 + \tilde{s}(k)] + \eta.$$

Rearranging terms, I have

$$\begin{aligned} \ln w = p_0 &+ \sum_{k=1}^K [p_1(k) + P_3(k, k)\tilde{s}(k)]x(k) + \sum_{k=1}^K p_2(k)z_1(k) + \sum_{k=1}^K p_2(k)\alpha_2(k)z_2 \\ &+ \sum_{k=1}^K P_3(k, k)x(k)z_1(k) + \sum_{k=1}^K P_3(k, k)\alpha_2(k)x(k)z_2 + \left(\sum_{k=1}^K p_2(k)\tilde{s}(k) + \eta \right). \end{aligned}$$

Consider the conditional mean of the log wage:

$$\begin{aligned} E(\ln w | x, z_1, z_2) &= p_0 + \sum_{k=1}^K p_1(k)x(k) + \sum_{k=1}^K p_2(k)z_1(k) + \sum_{k=1}^K p_2(k)\alpha_2(k)z_2 \\ &+ \sum_{k=1}^K P_3(k, k)x(k)z_1(k) + \sum_{k=1}^K P_3(k, k)\alpha_2(k)x(k)z_2. \end{aligned}$$

The variation of x , z_1 , z_2 , xz_1 , xz_2 , and the conditional mean of the log wage identifies the parameters p_0 , p_1 , p_2 , P_3 , and α_2 . The variance and covariance of unobserved skills are identified by the conditional variance of log wage, which is given by

$$\begin{aligned} V(\ln w | x, z_1, z_2) &= \sum_{k=1}^K \sum_{l=1}^K [p_2(k) + P_3(k, k)x(k)] \\ &\quad \times [p_2(l) + P_3(l, l)x(l)] \text{Cov}[\tilde{s}(k), \tilde{s}(l)] + \sigma_\eta^2. \end{aligned}$$

Given that p_2 and P_3 are identified by the conditional mean log wage, the variance and covariance of unobserved skills are identified by the variation of $x(k)x(l)$ and the conditional variance.

Appendix B

Robustness: Normally Distributed Task Indexes

This appendix examines whether the main results are robust to the choice of method used to construct the task complexity indexes. In the main text of the article, I use the percentile scores because they are easily interpreted. However, any monotonic transformation of the percentile scores is also valid as a measurement of task complexity because the original DOT variables are ordinal. To examine the robustness of the main results, I estimate and simulate the model using task complexity indexes that are normally distributed. While this is not the only possible alternative to the percentile score, this is a sensible choice for at least two reasons. First, when the model is true, task indexes are normally distributed given

the observed characteristics. Because the skills and other stochastic variables are assumed to be normally distributed and tasks are linear in these variables, the task indexes are normally distributed as well. Of course, these stochastic variables may not be normally distributed, but assuming that tasks are normally distributed is consistent with the model's assumptions. Second, the shape of a normal distribution is very different from that of the uniform distribution implied by the percentile scores. Hence, this exercise is suitable for an examination of the possible effects of a distributional assumption on the main results.

The percentile scores are converted into normally distributed indexes in the following way. They are truncated at 0.001 and 0.999 to avoid outliers and are then mapped onto a normal distribution with mean 0.5 and standard deviation 0.16. The standard deviation is chosen so that the new task indexes lie between 0 and 1 like the percentile scores.

For most parameters, the estimates are similar to those in the main text. However, an important difference is found for the estimates of the correlation coefficients of unobserved skills. Table B1 presents the correlation coefficient estimates for the unobserved variables for two different task measures. The correlation coefficients for the unobserved initial skills are negative for both indexes. However, the magnitude is significantly different: it is -0.776 when the percentile scores are used, while it is lower at -0.575 when the normally distributed task indexes are used. A similar pattern can be found for the correlation coefficient for the skill shocks: it is -0.833 for the percentile scores and -0.687 for the normal task indexes. The correlation coefficient for the preference shock is little affected by the choice of task indexes. These changes in the correlation coefficient estimates are reflected in the simulation results. Table B2 presents the log wage variance decomposition results for the normally distributed task indexes. As expected, the covariance term is significantly smaller than that in the main results presented in table 5. However, the qualitative result that both cognitive and motor skills are important when accounting for the log wage variance remains the same. Table B3 shows the results for the decomposition of wage growth for the normally distributed task indexes. Because the correlation coefficient for skills is smaller, the contributions of both cognitive and motor skills to wage growth shift toward zero. However, again, the qualitative results remain the same.

The exercise in this appendix demonstrates that most parameter estimates and simulation results are robust to the choice of the task complexity indexes with one exception: the correlation coefficient of skills. Unfortunately, it is not possible to unanimously decide which index is better than the other because both are valid. This problem comes from the restriction that the task is a linear function of unobserved skills. When the distribution of skills itself is of interest, one should allow for a more flexible functional form for the policy function for the task choice, although it means that the analytical solution and Kalman filter algorithm are no longer available.

Table B1
Estimates of Correlation Coefficients for Unobserved Variables

Notation	Percentile Score		Normally Distributed Index	
	Estimate	SE	Estimate	SE
Corr ($\varepsilon_1(1)$, $\varepsilon_1(2)$)	-.7757	.0498	-.5751	.0779
Corr ($\varepsilon(1)$, $\varepsilon(2)$)	-.8332	.0351	-.6865	.0587
Corr ($\nu(1)$, $\nu(2)$)	-.0295	.0038	.0102	.0038

SOURCE.—National Longitudinal Survey of Youth 1979–2000. Sample consists of 2,417 men.
 NOTE.—Variables ε_1 , ε , and ν are initial unobserved skills, skill shocks, and preference shocks, respectively. For each variable, the first element is for cognitive skills/tasks, and the second element is for motor skills/tasks.

Table B2
Log Wage Variance Decomposition for Normally Distributed Task Indexes

	Cognitive	Motor	Covariance	Error	Total
Year 1:					
All men	.466	.301	-.622	.061	.206
Dropouts	.287	.219	-.374	.061	.193
High school	.221	.136	-.219	.061	.198
College	.347	.209	-.412	.061	.205
Year 10:					
All men	.818	.393	-.978	.061	.293
Dropouts	.434	.281	-.547	.061	.228
High school	.329	.186	-.341	.061	.235
College	.589	.276	-.654	.061	.272
Year 20:					
All men	1.025	.435	-1.165	.061	.355
Dropouts	.499	.304	-.617	.061	.247
High school	.370	.202	-.382	.061	.251
College	.720	.304	-.771	.061	.313

NOTE.—Author’s estimates from the National Longitudinal Survey of Youth 1979–2000. Sample consists of 325 high school dropouts, 1,009 high school graduates, and 1,083 college workers.

Table B3
Accumulated Wage Growth by Skill Type and Education for Normally Distributed Task Indexes

Years since Entry	All Men	High School Dropouts	High School	College
Cognitive skills:				
5	.272	.086	.210	.387
10	.466	.153	.361	.662
20	.699	.237	.542	.983
Motor skills:				
5	-.008	.041	.006	-.036
10	-.015	.066	.009	-.062
20	-.030	.093	.008	-.100
Total:				
5	.264	.128	.216	.352
10	.451	.219	.370	.599
20	.669	.328	.550	.884

NOTE.—Author’s estimates from the National Longitudinal Survey of Youth 1979–2000. Sample consists of 325 high school dropouts, 1,009 high school graduates, and 1,083 college workers.

Appendix C

Model Solution

The goal of this appendix is to prove that (1) the value function is a quadratic function of the state variables $z'_t = \{d', s'_t, \bar{x}'_t, \tilde{v}'_t\}$ and (2) the optimal policy function is a linear function of the state variables z_t . To simplify the problem, rewrite the original Bellman equation (12) in the following form:

$$V_t(z_t) = \max_{x_t} r_0 + r'_1 x_t + r'_2 z_t + x'_t R_3 x_t + x'_t R_4 z_t + z'_t R_5 z_t + \beta E V_{t+1}(z_{t+1}), \quad (C1)$$

such that

$$z_{t+1} = l_0 + L_1 z_t + L_2 x_t + \xi_{t+1}, \quad (C2)$$

$$V_{T+1} = 0, \quad (C3)$$

where

$$z'_t = (d', s'_t, \bar{x}'_t, \tilde{v}'_t), \quad (C4)$$

$$\xi'_t = (0', \varepsilon'_t, 0', 0'), \quad (C5)$$

$$r_0 = p_0 + \eta_t, \quad (C6)$$

$$r_1 = p_1 + g_0, \quad (C7)$$

$$r'_2 = (0', p'_2, 0', 0'), \quad (C8)$$

$$R_3 = G_3 + G_4, \quad (C9)$$

$$R_4 = (G_1, P_3 + G_2, -2G_4, I), \quad (C10)$$

$$R_5 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & G_4 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad (C11)$$

$$l'_0 = (0', a'_0, 0', 0'), \quad (C12)$$

$$L_1 = \begin{pmatrix} I & 0 & 0 & 0 \\ A_2 & D & 0 & 0 \\ 0 & 0 & A_3 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad (\text{C13})$$

$$L_2 = \begin{pmatrix} 0 \\ A_1 \\ I - A_3 \\ 0 \end{pmatrix}. \quad (\text{C14})$$

Claim 1 is trivially true in $t = T + 1$ because $V_{T+1}(z_{T+1}) = 0$. I show below that the value function in period t can be represented by a quadratic function of the state variables z_t , if the value function in period $t + 1$ is given by a quadratic function of the state variables:

$$V_{t+1}(z_{t+1}) = q_{0,t+1} + q'_{1,t+1}z_{t+1} + z'_{t+1}Q_{2,t+1}z_{t+1}, \quad (\text{C15})$$

where $Q_{2,t+1}$ is a symmetric matrix. Consider the value function in period t :

$$\begin{aligned} V_t(z_t) = & \max_{x_t} r_0 + r'_1x_t + r'_2z_t + x'_tR_3x_t + x'_tR_4z_t + z'_tR_5z_t \\ & + \beta E[q_{0,t+1} + q'_{1,t+1}(l_0 + L_1z_t + L_2x_t + \xi_{t+1}) \\ & + (l_0 + L_1z_t + L_2x_t + \xi_{t+1})'Q_{2,t+1}(l_0 + L_1z_t + L_2x_t + \xi_{t+1})]. \end{aligned} \quad (\text{C16})$$

The first-order condition for optimality is

$$\begin{aligned} 0 = & r_1 + 2R_3x_t + R_4z_t \\ & + \beta E[L'_2\{q_{1,t+1} + 2Q_{2,t+1}(l_0 + L_1z_t + \xi_{t+1})\} + 2L'_2Q_{2,t+1}L_2x_t]. \end{aligned} \quad (\text{C17})$$

Solving this equation for x_t , I find the optimal policy function:

$$\begin{aligned} x_t^* = & -\frac{1}{2}(R_3 + L'_2Q_{2,t+1}L_2)^{-1}[r_1 + \beta L'_2(q_{1,t+1} + 2Q_{2,t+1}l_0) \\ & + (R_4 + 2\beta L'_2Q_{2,t+1}L_1)z_t] \end{aligned} \quad (\text{C18})$$

$$\equiv b_{0,t} + B_{1,t}z_t, \quad (\text{C19})$$

where

$$b_{0,t} = -\frac{1}{2}(R_3 + \beta L'_2Q_{2,t+1}L_2)^{-1}[r_1 + \beta L'_2(q_{1,t+1} + 2Q_{2,t+1}l_0)], \quad (\text{C20})$$

$$B_{1,t} = -\frac{1}{2}(R_3 + \beta L'_2Q_{2,t+1}L_2)^{-1}[R_4 + 2\beta L'_2Q_{2,t+1}L_1]. \quad (\text{C21})$$

Substituting the optimal policy function (C19) into the instantaneous utility function, I find

$$\begin{aligned}
& r_0 + r'_1 x_t^* + r'_2 z_t + x_t^{*'} R_3 x_t^* + x_t^{*'} R_4 z_t + z_t' R_5 z_t \\
&= r_0 + r'_1 (b_{0,t} + B_{1,t} z_t) + r'_2 z_t + (b_{0,t} + B_{1,t} z_t)' R_3 (b_{0,t} + B_{1,t} z_t) \\
&\quad + (b_{0,t} + B_{1,t} z_t)' R_4 z_t + z_t' R_5 z_t \quad (C22) \\
&= [r_0 + r'_1 b_{0,t} + b_{0,t}' R_3 b_{0,t}] + [B_{1,t}' r_1 + r_2 + 2B_{1,t}' R_3 b_{0,t} + R_4' b_{0,t}]' z_t \\
&\quad + z_t' [B_{1,t}' R_3 B_{1,t} + B_{1,t}' R_4 + R_5] z_t.
\end{aligned}$$

Notice that, using this expression for the optimal task complexity x_t^* , I can rewrite the transition equation for the state variables as

$$z_{t+1} = l_0 + L_1 z_t + L_2 (b_{0,t} + B_{1,t} z_t) + \xi_{t+1} \quad (C23)$$

$$= (l_0 + L_2 b_{0,t}) + (L_1 + L_2 B_{1,t}) z_t + \xi_{t+1} \quad (C24)$$

$$\equiv f_{0,t} + F_{1,t} z_t + \xi_{t+1}, \quad (C25)$$

where $f_{0,t} = l_0 + L_2 b_{0,t}$, and $F_{1,t} = L_1 + L_2 B_{1,t}$. Substituting this expression into the expected value function, I find

$$EV_{t+1}(z_{t+1}) = E[q_{0,t+1} + q'_{1,t+1} z_{t+1} + z'_{t+1} Q_{2,t+1} z_{t+1}] \quad (C26)$$

$$\begin{aligned}
&= q_{0,t+1} + q'_{1,t+1} (f_{0,t} + F_{1,t} z_t) \\
&\quad + (f_{0,t} + F_{1,t} z_t)' Q_{2,t+1} (f_{0,t} + F_{1,t} z_t) + E[\xi'_{t+1} Q_{2,t+1} \xi_{t+1}] \quad (C27)
\end{aligned}$$

$$\begin{aligned}
&= [q_{0,t+1} + q'_{1,t+1} f_{0,t} + f_{0,t}' Q_{2,t+1} f_{0,t} + E[\xi'_{t+1} Q_{2,t+1} \xi_{t+1}]] \\
&\quad + [F'_{1,t} (q_{1,t+1} + 2Q_{2,t+1} f_{0,t})]' z_t + z_t' [F'_{1,t} Q_{2,t+1} F_{1,t}] z_t. \quad (C28)
\end{aligned}$$

The value function for period t is therefore given by

$$V_t(z_t) = q_{0,t} + q'_{1,t} z_t + z_t' Q_{2,t} z_t, \quad (C29)$$

where

$$\begin{aligned}
q_{0,t} &= [r_0 + r'_1 b_{0,t} + b_{0,t}' R_3 b_{0,t}] \\
&\quad + \beta [q_{0,t+1} + q'_{1,t+1} f_{0,t} + f_{0,t}' Q_{2,t+1} f_{0,t} + E[\xi'_{t+1} Q_{2,t+1} \xi_{t+1}]], \quad (C30)
\end{aligned}$$

$$q_{1,t} = [B'_{1,t} r_1 + r_2 + 2B'_{1,t} R_3 b_{0,t} + R_4' b_{0,t}] + \beta [F'_{1,t} (q_{1,t+1} + 2Q_{2,t+1} f_{0,t})], \quad (C31)$$

$$Q_{2,t} = [B'_{1,t} R_3 B_{1,t} + B'_{1,t} R_4 + R_5] + \beta [F'_{1,t} Q_{2,t+1} F_{1,t}]. \quad (C32)$$

It remains to be shown that $Q_{2,t}$ is symmetric. Rewrite the matrix $Q_{2,t}$ such that

$$\begin{aligned}
Q_{2,t} &= [B'_{1,t}R_3B_{1,t} + B'_{1,t}R_4 + R_5] \\
&\quad + \beta[L'_{1,t}Q_{2,t+1}L_{1,t} + 2B'_{1,t}L'_{2,t}Q_{2,t+1}L_{1,t} + B'_{1,t}L'_{2,t}Q_{2,t+1}L_{2,t}B_{1,t}] \quad (C33) \\
&= R_5 + \beta L'_{1,t}Q_{2,t+1}L_{1,t} \\
&\quad + B'_{1,t}[(R_3 + \beta L'_{2,t}Q_{2,t+1}L_{2,t})B_{1,t} + (R_4 + 2\beta L'_{2,t}Q_{2,t+1}L_{1,t})]. \quad (C34)
\end{aligned}$$

Notice that

$$\begin{aligned}
&B'_{1,t}[(R_3 + \beta L'_{2,t}Q_{2,t+1}L_{2,t})B_{1,t} + (R_4 + 2\beta L'_{2,t}Q_{2,t+1}L_{1,t})] \\
&= B_{1,t}[-\frac{1}{2}(R_4 + 2\beta L'_{2,t}Q_{2,t+1}L_{1,t}) + (R_4 + 2\beta L'_{2,t}Q_{2,t+1}L_{1,t})] \quad (C35)
\end{aligned}$$

$$= -\frac{1}{4}[R_4 + 2\beta L'_{2,t}Q_{2,t+1}L_{1,t}](R_3 + \beta L'_{2,t}Q_{2,t+1}L_{2,t})^{-1}[R_4 + 2\beta L'_{2,t}Q_{2,t+1}L_{1,t}]. \quad (C36)$$

Substituting this expression into equation (C34), I find

$$\begin{aligned}
Q_{2,t} &= R_5 + \beta L'_{1,t}Q_{2,t+1}L_{1,t} \\
&\quad - \frac{1}{4}[R_4 + 2\beta L'_{2,t}Q_{2,t+1}L_{1,t}](R_3 + \beta L'_{2,t}Q_{2,t+1}L_{2,t})^{-1} \\
&\quad \times [R_4 + 2\beta L'_{2,t}Q_{2,t+1}L_{1,t}]. \quad (C37)
\end{aligned}$$

Because $Q_{2,t+1}$ and R_5 are symmetric by assumption, this equation implies that $Q_{2,t}$ is symmetric. Claim 1 has been proved through mathematical induction.

Claim 2 is also proved in equation (C19). Notice that the parameter $B_{1,t}$ can be written as $B_{1,t} = [C_{1,t}, C_{2,t}, C_{3,t}, C_{4,t}]$. Hence, the policy function can be written as equation (19) because

$$x_t^* = b_{0,t} + B_{1,t}z_t \quad (C38)$$

$$= c_{0,t} + C_{1,t}d + C_{2,t}s_t + C_{3,t}\tilde{x}_t + C_{4,t}\tilde{\nu}_t \quad (C39)$$

$$= c_{0,t} + C_{1,t}d + C_{2,t}s_t + C_{3,t}\tilde{x}_t + \nu_t, \quad (C40)$$

where $c_{0,t} = b_{0,t}$, and $\nu_t = C_{4,t}\tilde{\nu}_t$.

Appendix D

Details of Sampling Criteria for NLSY79

In the NLSY cross-section sample, 3,003 men are included. I dropped 294 individuals who served actively in the armed forces during the sample period. Out of 2,709 individuals, 149 did not make a long-term transition to the full-time labor market until 2000. I also excluded 11 individuals who made the long-term transition at age 16 or younger because they are likely to be mismeasured. I also dropped 131 out of 2,549 remaining individuals whose AFQT scores are missing. Finally, 132 individuals are excluded because there are fewer than 3 years of observations for these individuals. Hourly wages are deflated by the 2005 constant dollar using

the consumer price index. If the recorded hourly wage is greater than \$100 or less than \$1, it is regarded as missing because it is likely to be mismeasured. After imposing all sample restrictions, the sample contains the career histories of 2,417 men and contains 32,849 person-year observations of occupational choices and 27,063 person-year observations of wages. As a robustness check, I relax the restriction of full-time work and include part-time jobs in the sample. This change does not affect the main results greatly because male workers take part-time jobs in the beginning of their careers only. The detailed results for this robustness check are available in appendix E.

References

- Altonji, Joseph G., and Robert A. Shakotko. 1987. Do wages rise with job seniority? *Review of Economic Studies* 54, no. 3:437–59.
- Altug, Sumru, and Robert A. Miller. 1998. The effect of work experience on female wages and labor supply. *Review of Economic Studies* 65:45–85.
- Anderson, Evan W., Lars P. Hansen, Ellen R. McGrattan, and Thomas J. Sargent. 1996. Mechanics of forming and estimating dynamic linear economies. In *Handbook of computational economics*, vol. 1, ed. Hans M. Amman, David A. Kendrick, and John Rust, 171–252. Amsterdam: North-Holland Elsevier.
- Autor, David, and Michael Handel. 2009. Putting tasks to the test: Human capital, job tasks, and wages. Working Paper no. 15116, National Bureau of Economic Research, Cambridge, MA.
- Autor, David H., Frank Levy, and Richard J. Murnane. 2003. The skill content of recent technological change: An empirical exploration. *Quarterly Journal of Economics* 118, no. 4:1279–1333.
- Bacolod, Marigee, and Bernardo S. Blum. 2010. Two sides of the same coin: U.S. “residual” inequality and the gender gap. *Journal of Human Resources* 45, no. 1:197–242.
- Firpo, Sergio, Nicole Fortin, and Thomas Lemieux. 2011. *Occupational tasks and changes in the wage structure*. Vancouver: University of British Columbia.
- Gathmann, Christina, and Uta Schönberg. 2010. How general is specific human capital? A task-based approach. *Journal of Labor Economics* 28, no. 1:1–50.
- Gibbons, Robert, Lawrence F. Katz, Thomas Lemieux, and Daniel Parent. 2005. Comparative advantage, learning and sectoral wage determination. *Journal of Labor Economics* 23, no. 4:681–724.
- Heckman, James, and Jose Scheinkman. 1987. The importance of bundling

- in a Gorman-Lancaster model of earnings. *Review of Economic Studies* 54, no. 2:243–55.
- Heckman, James J., and Guilherme Sedlacek. 1985. Heterogeneity, aggregation, and market wage functions: An empirical model of self-selection in the labor market. *Journal of Political Economy* 93, no. 6: 1077–1125.
- Hotz, Joseph, and Robert Miller. 1993. Conditional choice probabilities and the estimation of dynamic models. *Review of Economic Studies* 60, no. 3:497–529.
- Imai, Susumu, Neelam Jain, and Andrew Ching. 2009. Bayesian estimation of dynamic discrete choice models. *Econometrica* 77, no. 6:1865–99.
- Ingram, Beth F., and George R. Neumann. 2006. The returns to skill. *Labour Economics* 13:35–59.
- Johnson, Matthew, and Michael Keane. 2007. A dynamic equilibrium model of the U.S. wage structure. Working paper, University of Technology, Sydney.
- Jovanovic, Boyan. 1979. Job matching and the theory of turnover. *Journal of Political Economy* 87, no. 5:972–90.
- Kambourov, Gueorgui, and Iouri Manovskii. 2009. Occupational specificity of human capital. *International Economic Review* 50, no. 1:63–115.
- Keane, Michael P., and Kenneth I. Wolpin. 1997. The career decisions of young men. *Journal of Political Economy* 105, no. 3:473–522.
- Lancaster, Kelvin J. 1966. A new approach to consumer theory. *Journal of Political Economy* 74, no. 2:132–55.
- Lazear, Edward P. 2009. Firm-specific human capital: A skill-weights approach. *Journal of Political Economy* 117, no. 5:914–40.
- Lee, Donghoon. 2005. An estimable dynamic general equilibrium model of work, schooling, and occupational choice. *International Economic Review* 46, no. 1:1–34.
- Lee, Donghoon, and Kenneth I. Wolpin. 2006. Intersectoral labor mobility and the growth of the service sector. *Econometrica* 74, no. 1:1–46.
- Miller, Ann R., Donald J. Treiman, Pamela S. Cain, and Patricia A. Roos, eds. 1980. *Work, jobs, and occupations: A critical review of the Dictionary of Occupational Titles*. Washington, DC: National Academy.
- Mincer, Jacob, and Haim Ofek. 1982. Interrupted work careers: Depreciation and restoration of human capital. *Journal of Human Resources* 27, no. 1:3–24.
- Neal, Derek. 1999. The complexity of job mobility among young men. *Journal of Labor Economics* 17:237–61.
- Pavan, Ronni. 2011. Career choice and wage growth. *Journal of Labor Economics* 29, no. 3:549–87.
- Poletaev, Maxim, and Chris Robinson. 2008. Human capital specificity: Evidence from the Dictionary of Occupational Titles and Displaced

- Worker Surveys, 1984–2000. *Journal of Labor Economics* 26, no. 4:387–420.
- Rosen, Sherwin. 1983. A note on aggregation of skills and labor quality. *Journal of Human Resources* 18, no. 3:425–31.
- Rubinstein, Yona, and Yoram Weiss. 2006. Post schooling wage growth: Investment, search and learning. In *Handbook of the economics of education*, vol. 1, ed. Eric Alan Hanushek and Finis Welch, 1–67. Amsterdam: North-Holland Elsevier.
- Schönberg, Uta. 2007. Wage growth due to human capital accumulation and job search: A comparison between the United States and Germany. *Industrial and Labor Relations Review* 60:562–86.
- Sullivan, Paul J. 2009. Estimation of an occupational choice model when occupations are misclassified. *Journal of Human Resources* 44, no. 2: 495–535.
- . 2010. A dynamic analysis of educational attainment, occupational choices, and job search. *International Economic Review* 51, no. 1:289–317.
- Teulings, Coen. 1995. The wage distribution in a model of the assignment of skills to jobs. *Journal of Political Economy* 103, no. 2:180–215.
- Topel, Robert, and Michael Ward. 1992. Job mobility and the careers of young men. *Quarterly Journal of Economics* 107, no. 2:439–79.
- Welch, Finis. 1969. Linear synthesis of skill distribution. *Journal of Human Resources* 4, no. 3:311–27.
- Yamaguchi, Shintaro. 2010a. Career progression and comparative advantage. *Labour Economics* 17, no. 4:679–89.
- . 2010b. The effect of match quality and specific experience on career decisions and wage growth. *Labour Economics* 17, no. 2:407–23.