Tasks and Heterogeneous Human Capital

Shintaro Yamaguchi. (2012).
Journal of Labor Economics, 30(1): pp. 1–53.

Econ 350, Winter 2021
This draft, January 12, 2021 5:55pm

THE UNIVERSITY OF CHICAGO

I. Introduction

II. Model

- Task complexity: x_t
- Skills: s_t

Skills vs. Task Complexity

- 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.
- The **task complexity** indexes take nonnegative values, and sufficiently many occupations exist so that an individual can choose any occupation in the task complexity space.
- **Skills** in year t are denoted by a K-dimensional vector s_t 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.
- Each firm offers jobs of a single type of complexity, which implies that the products of each firm can be characterized by a task complexity vector.
- 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

- The marginal value product of a worker with skill s_t in an occupation with task complexity x_t is

$$w_t = \pi(x_t) q(x_t, s_t) \exp(\eta_t), \tag{1}$$

- $\eta_t \sim N(0, \sigma_\eta^2)$ 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.
- (1) is assumed.

- Labor productivity:

$$\ln q(x_t, s_t) = \theta'(x_t) s_t, \tag{2}$$

- $\theta(x_t)$ 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.

- I parameterize the output price and productivity:

$$\ln \pi(x_t) = p_0 + p_1' x_t \tag{3}$$

and

$$\ln q(x_t, s_t) = \theta'(x_t) s_t = [p_2 + P_3' x_t]' s_t, \tag{4}$$

- 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

$$\ln w_t = p_0 + p_1' x_t + [p_2 + P_3' x_t]' s_t + \eta_t. \tag{5}$$

B. Skill Formation

- 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 $\Sigma_\varepsilon : \varepsilon_t \sim N(0, \Sigma_\varepsilon)$.
- Skills grow from year t to year $t+1$ according to the following skill transition equation:

$$s_{t+1} = Ds_t + a_0 + A_1 x_t + A_2 d + \varepsilon_{t+1}, \tag{6}$$

- 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.

THE UNIVERSITY OF CHICAGO

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

$$s_1 = h + Hd + \varepsilon_1, \tag{7}$$

- h 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 $\varepsilon_1 \sim N(0, \Sigma_{\varepsilon_1})$.

THE UNIVERSITY OF CHICAGO

C. Job Preferences

- Utility derived from work:

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

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

- g_0 is a K-dimensional vector of preference parameters;
- G_1 is a $K \times L$ matrix of preference parameters;
- $\tilde{\nu}_t$ 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 UNIVERSITY OF CHICAGO

- The utility from job tasks varies across individuals according to their individual characteristics d, skill levels s_t, a preference shock $\tilde{\nu}_t$, and work habits $\bar{x}_t$.
- Skilled workers prefer complex tasks if the parameter matrix G_2 is positive definite.
- G_3: negative definite.
- 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.

- Individuals form their work habits $\bar{x}_{t+1}$ through the following transition equation:

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

- 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.

- The initial condition for $\bar{x}_t$ varies across individuals according to initial observed characteristics d such that

$$\bar{x}_1 = \bar{x}_{1,0} + Xd, \tag{11}$$

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

D. Bellman Equation

The Bellman equation for an individual is given by

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

$$\ln w_t = p_0 + p_1' x_t + [p_2 + P_3' x_t]' s_t + \eta_t, \quad (13)$$

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

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

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

$$s_1 = h + Hd + \varepsilon_1, \quad (17)$$

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

$$x_t^* = c_{0,t} + C_{1,t}d + C_{2,t}s_t + C_{3,t}\bar{x}_t + \nu_t, \tag{19}$$

- $c_{0,t}$ is a K-dimensional vector,
- $C_{1,t}$ is a $K \times L$ matrix,
- $C_{2,t}$ and $C_{3,t}$ are K-dimensional diagonal matrices, and
- ν_t is a K-dimensional vector of rescaled preference shocks (i.e., I can write $\nu_t = M_t\tilde{\nu}_t$, where M_t is a K-dimensional diagonal matrix).

THE UNIVERSITY OF CHICAGO

III. Estimation Strategy

A. Identification

- 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.
- Normalize skills by assuming that the unconditional mean and variance of initial skills are 0 and 1, respectively.

- Specifically,
  $E(s_1) = h + HE(d) = 0$,
- $\text{diag}[\text{Var}(s_1)] = \text{diag}[HE(dd')H' + \Sigma_{\varepsilon_1}] = \text{diag}[I]$, 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.
- 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.
- 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.

THE UNIVERSITY OF CHICAGO

- The parameters for job preferences are, in principle, 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 parameters in the policy function (19).
- To see this unidentifiability, consider the optimal choice of occupation at the terminal period T.
- The optimal task complexity x_T^* is given by

$$
\begin{aligned}
x_T^* &= -(1/2)(G_3 + G_4)^{-1}[g_0 + G_1 d + (P_3 + G_2)s_T - 2G_4\bar{x}_T + \tilde{\nu}_T] \\
&\equiv c_{0,T} + C_{1,T} d + C_{2,T} s_T + C_{3,T}\bar{x}_T + \nu_T.
\end{aligned}
$$

THE UNIVERSITY OF CHICAGO

B. Kalman Filter

- Use 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

$$E(s_t|x_1, w_1, \ldots, x_{t-1}, w_{t-1}; \bar{x}_1, d) \equiv E(s_t|Y_{t-1}) \quad (20)$$
$$\equiv \hat{s}_{t|t-1}, \quad (21)$$

$$\text{Var}(s_t|x_1, w_1, \ldots, x_{t-1}, w_{t-1}; \bar{x}_1, d) \equiv \text{Var}(s_t|Y_{t-1}) \quad (22)$$
$$\equiv \Sigma^s_{t|t-1}, \quad (23)$$

- where T_{t-1} summarizes all the information up to year $t-1$.

- The conditional mean and variance of x_t given Y_t are

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

$$\text{Var}(x_t|Y_{t-1}) = C_{2,t}\Sigma^s_{t|t-1}C'_{2,t} + \Sigma_\nu. \tag{25}$$

THE UNIVERSITY OF CHICAGO

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

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

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

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

THE UNIVERSITY OF CHICAGO

- 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 and Y_{t-1} and x_t.
- The conditional mean and variance of the log wage are

$$E(\ln w_t | Y_{t-1}, x_t) = p_0 + p_1' x_t + [p_2 + P_3' x_t]' E(s_t | Y_{t-1}, x_t), \quad (28)$$

$$\text{Var}(\ln w_t | Y_{t-1}, x_t) = [p_2 + P_3' x_t]' \text{Var}(s_t | Y_{t-1}, x_t)[p_2 + P_3' x_t] + \sigma_\eta^2. \quad (29)$$

THE UNIVERSITY OF CHICAGO

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

$$
\begin{aligned}
& E(s_t | Y_{t-1}, x_t, w_t) \\
= & E(s_t | Y_{t-1}, x_t) + \text{Var}(s_t | Y_{t-1}, x_t)[p_2 + P_3' x_t][\text{Var}(\ln w_t | Y_{t-1} x_t)]^{-1} \hat{\eta}_t,
\end{aligned} \tag{30}
$$

$$\text{Var}(s_t | Y_{t-1}, x_t, w_t)$$
$$= \text{Var}(s_t | Y_{t-1}, x_t)$$
$$- \text{Var}(s_t | Y_{t-1}, x_t)[p_2 + P_3' x_t][\text{Var}(\ln w_t | Y_{t-1} x_t)]^{-1}[p_2 + P_3' x_t]' \text{Var}(s_t | Y_{t-1}, x_t), \tag{31}$$

- where $\hat{\eta}_t$ is the log wage residual, and
- $\hat{n}_t = \ln w_t - E(\ln w_t | Y_{t-1}, x_t)$.

- 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:

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

$$\Sigma^s_{t+1|t} = D\,\mathrm{Var}(s_t | Y_{t-1}, x_t, w_t) D + \Sigma_\varepsilon. \tag{33}$$

- 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

$$E(s_{t+1}|Y_{t-1}, x_t) = Ds_t + a_0 + A_1 x_t + A_2 d$$

$$= D\hat{s}_{t|t-1} + a_0 + A_1 x_t + A_2 d, \tag{34}$$

$$\mathrm{Var}(s_{t+1}|Y_{t-1}, x_t) = D\Sigma^s_{t|t-1} D + \Sigma_\varepsilon, \tag{35}$$

- which replace equations (32) and (33) in this case.

- The likelihood contribution of individual i is

$$
\begin{aligned}
& l(w_{i1}, x_{i1}, \ldots, w_{iT_i}, x_{iT_i} | \bar{x}_{i1}, d_i) \\
= \; & l(x_{i1} | \bar{x}_{i1}, d_i) l(w_{i1} | x_{i1}; \bar{x}_{i1}, d_i) \times \ldots \times l(x_{iT_i} | Y_{iT_i - 1}) l(w_{i1} | Y_{iT_i - 1}, x_{iT_i}).
\end{aligned}
\tag{36}
$$

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

$$
l = \mathbf{\Pi}_{i=1}^{N} l(w_{i1}, x_{i1}, \ldots, w_{iT_i}, x_{iT_i} | \bar{x}_{i1}, d_i). \tag{37}
$$

IV. Data

A. Dictionary of Occupational Titles

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

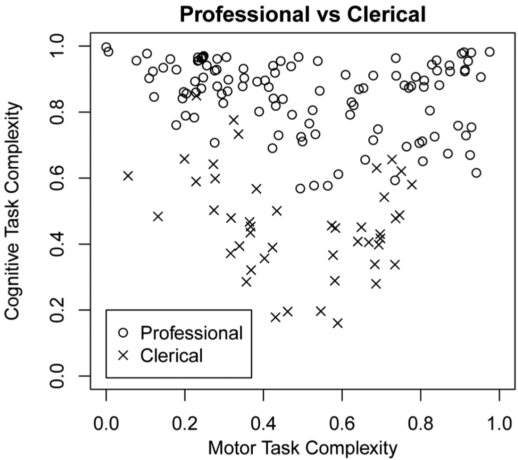
	Cognitive Task		Motor Task		
	Mean	SD	Mean	SD	No. Observations
Professional	.85	.14	.45	.33	7,522
Manager	.79	.15	.21	.21	5,538
Sales	.57	.17	.23	.15	3,748
Clerical	.49	.16	.56	.22	9,270
Craftsmen	.52	.20	.82	.20	6,557
Operatives	.20	.18	.58	.20	5,824
Transport	.28	.15	.63	.10	1,774
Laborer	.15	.16	.46	.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.

B. National Longitudinal Survey of Youth 1979

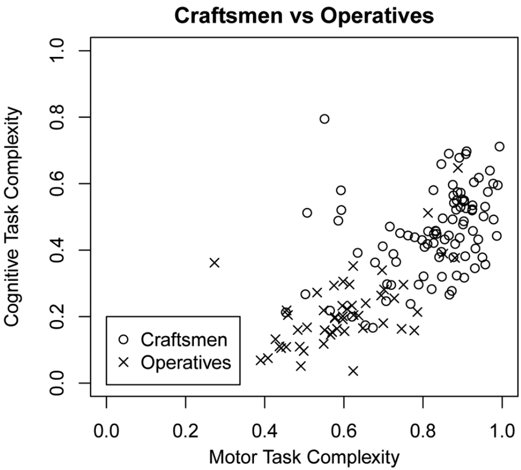
THE UNIVERSITY OF CHICAGO

Figure 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).

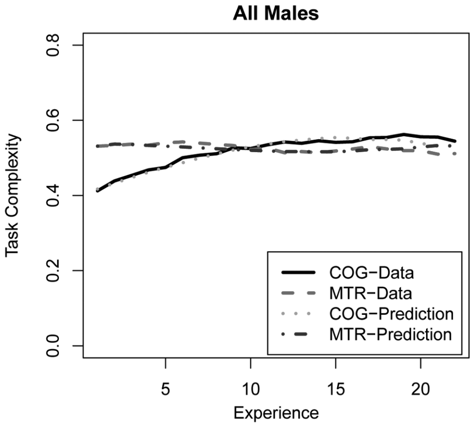
Figure 1: Task complexity comparison



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).

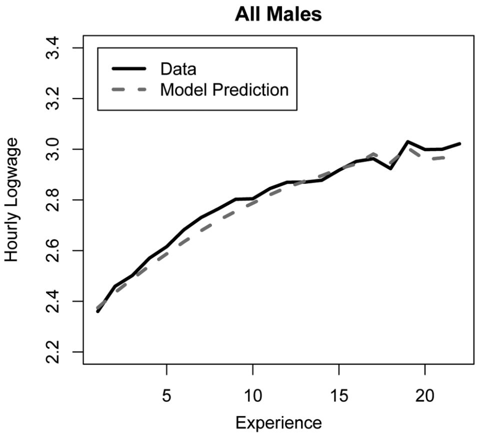
C. Career Progression Patterns

Figure 2: Model fit (all men)



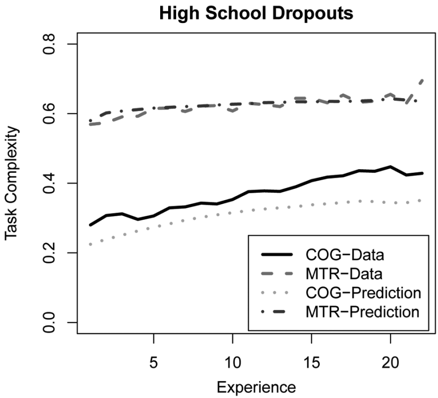
COG = profiles of cognitive task complexity; MTR = profiles of motor task complexity. *Source*: Author's estimates from the National Longitudinal Survey of Youth 1979–2000. Sample consists of 325 high school dropouts, 1,009 high school graduates, and 1,083 college workers. Color version available as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

Figure 2: Model fit (all men)



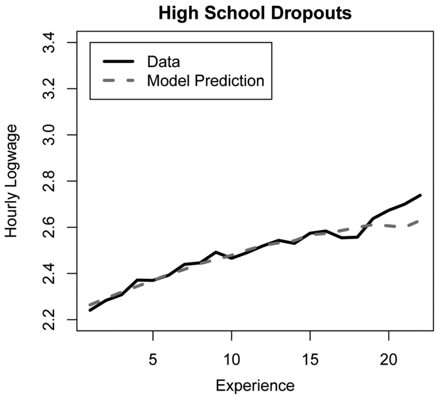
COG = profiles of cognitive task complexity; MTR = profiles of motor task complexity. *Source*: Author's estimates from the National Longitudinal Survey of Youth 1979–2000. Sample consists of 325 high school dropouts, 1,009 high school graduates, and 1,083 college workers. Color version available as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

Figure 3: Model fit (high school dropouts)



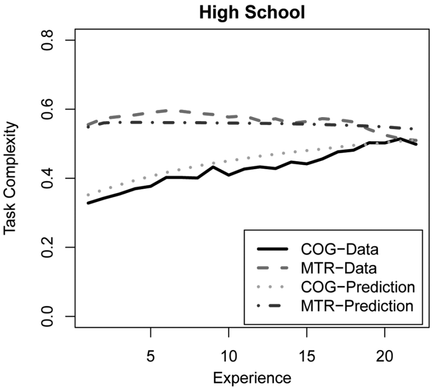
COG = profiles of cognitive task complexity; MTR = profiles of motor task complexity.
Source: Author's estimates from the National Longitudinal Survey of Youth 1979–2000.
Sample includes 325 high school dropouts. Color version available as an online enhancement.

Figure 3: Model fit (high school dropouts)



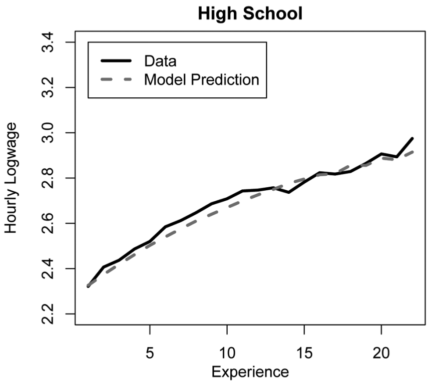
COG = profiles of cognitive task complexity; MTR = profiles of motor task complexity.
Source: Author's estimates from the National Longitudinal Survey of Youth 1979–2000.
Sample includes 325 high school dropouts. Color version available as an online enhancement.

Figure 4: Model fit (high school workers)



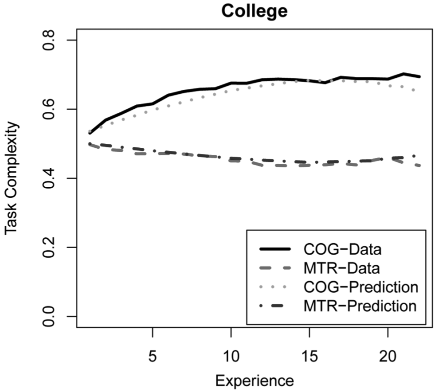
COG = profiles of cognitive task complexity; MTR = profiles of motor task complexity. *Source*: Author's estimates from the National Longitudinal Survey of Youth 1979–2000. Sample includes 1,009 high school graduates. Color version available as an online enhancement.

Figure 4: Model fit (high school workers)



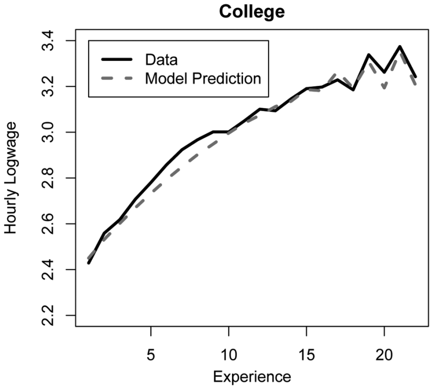
COG = profiles of cognitive task complexity; MTR = profiles of motor task complexity. *Source*: Author's estimates from the National Longitudinal Survey of Youth 1979–2000. Sample includes 1,009 high school graduates. Color version available as an online enhancement.

Figure 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.

Figure 5: Model fit (college workers)

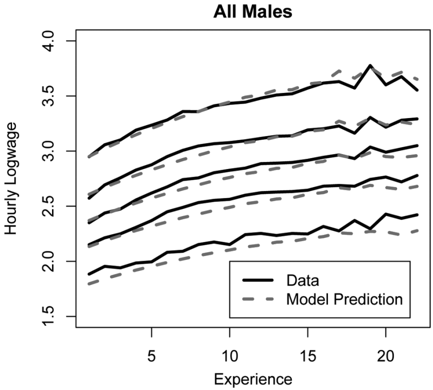


COG = profiles of cognitive task complexity; MTR = profiles of motor task complexity. *Source*: Author's estimates from the National Longitudinal Survey of Youth 1979–2000. Sample includes 1,083 college workers. Color version available as an online enhancement.

V. Estimation Results

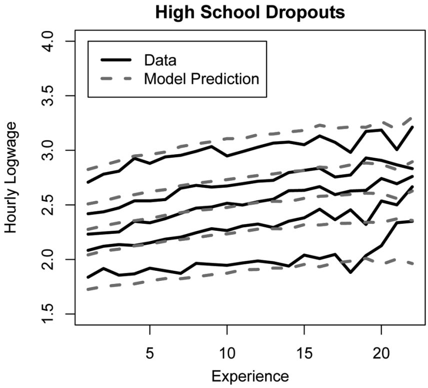
A. Model Fit

Figure 6: Model fit: wage growth by percentile (all men and high school dropouts)



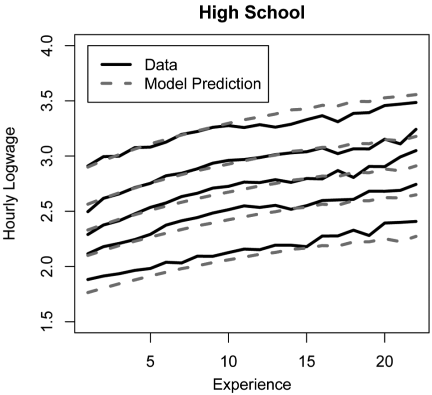
Tenth, 30th, 50th, 70th, and 90th percentiles of hourly log wages are plotted over time. *Source*: Author's estimates from the National Longitudinal Survey of Youth 1979–2000. Sample consists of 325 high school dropouts, 1,009 high school graduates, and 1,083 college workers. Color version available as an online enhancement.

Figure 6: Model fit: wage growth by percentile (all men and high school dropouts)



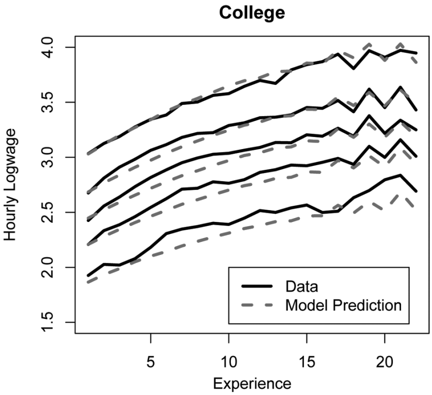
Tenth, 30th, 50th, 70th, and 90th percentiles of hourly log wages are plotted over time. *Source*: Author's estimates from the National Longitudinal Survey of Youth 1979–2000. Sample consists of 325 high school dropouts, 1,009 high school graduates, and 1,083 college workers. Color version available as an online enhancement.

Figure 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.

Figure 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 $\ln w_t = p_0 + p_1' x_t + (p_2 + P_3' x_t)' s_t + \eta_t$, where $\eta_t \sim N(0, \sigma_\eta^2)$. The first element of vectors and (1, 1) element of matrices are for cognitive skills, and the second element of vectors and (2, 2) element of matrices are motor skills.

B. Parameter Estimates

1. Wage Equation

2. Skill Transition

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
$\Sigma_\varepsilon(1, 1)$	.156	.019
$\Sigma_\varepsilon(2, 1)$	−.125	.019
$\Sigma_\varepsilon(2, 2)$	.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
$\Sigma_{\varepsilon_1}(1, 1)$	.654	.075
$\Sigma_{\varepsilon_1}(2, 1)$	−.519	.080
$\Sigma_{\varepsilon_1}(2, 2)$	.685	.089

3. Job Preference

VI. Discussion

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

THE UNIVERSITY OF CHICAGO

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

- The wage can be written as

$$\ln w_t = p_0 + [p_1^C x_t^C + p(x_t)^C s_t^C] + [p_1^M x_t^M + p(x_t)^M s_t^M] + \eta_t,$$

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

THE UNIVERSITY OF CHICAGO

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.

Table 6: Log Wage Variance When Initial Conditions Are Homogeneous

		Homogeneous			
Year	Benchmark	Preference	Initial Skills	Learning Ability	All
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.

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

		Homogeneous		
Year	Benchmark	Preference	Initial Skills	Leaning 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.

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

Table 9: Accumulated Wage Growth by Skill Type and Education

	Benchmark				All Men	
Years since Entry	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.

THE UNIVERSITY OF CHICAGO

VII. Conclusion

Appendix A
Identification

- Denote the kth 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 kth element affects the kth 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.

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

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

- Consider the conditional mean of the log wage:

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

- 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

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

- 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.

THE UNIVERSITY OF CHICAGO

Appendix B
Robustness: Normally Distributed Task Indexes

Table 10: Estimates of Correlation Coefficients for Unobserved Variables

	Percentile Score		Normally Distributed Index	
Notation	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 11: 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 12: 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.