The Growing Importance of Social Skills in the Labor Market

by David J. Deming

The Quarterly Journal of Economics, 2017, pp. 1593-1640

James J. Heckman

Econ 350, Winter 2021

I. Introduction

FIGURE I

Change in Relative Employment for Cognitive Occupations, 2000–2012

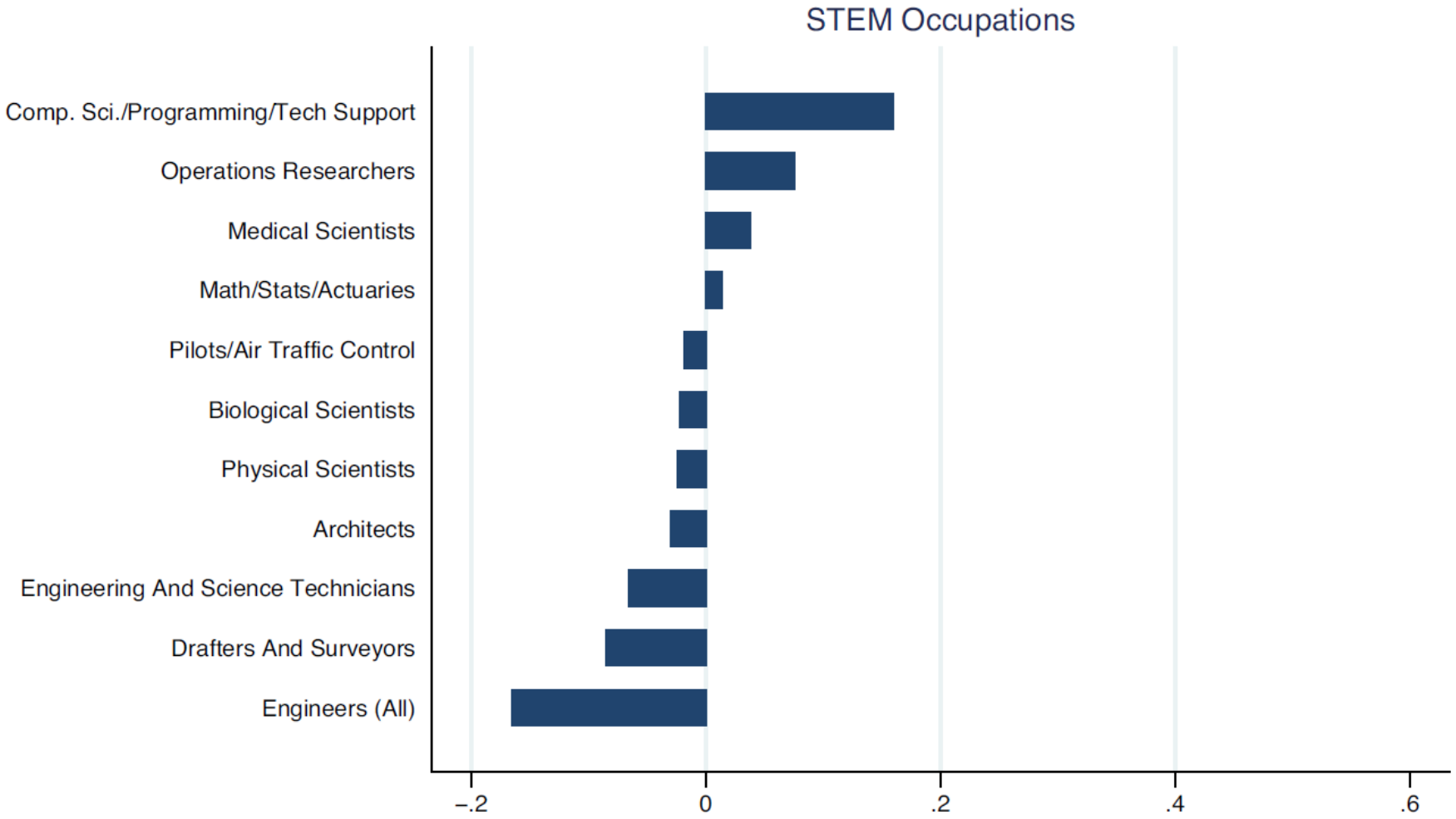
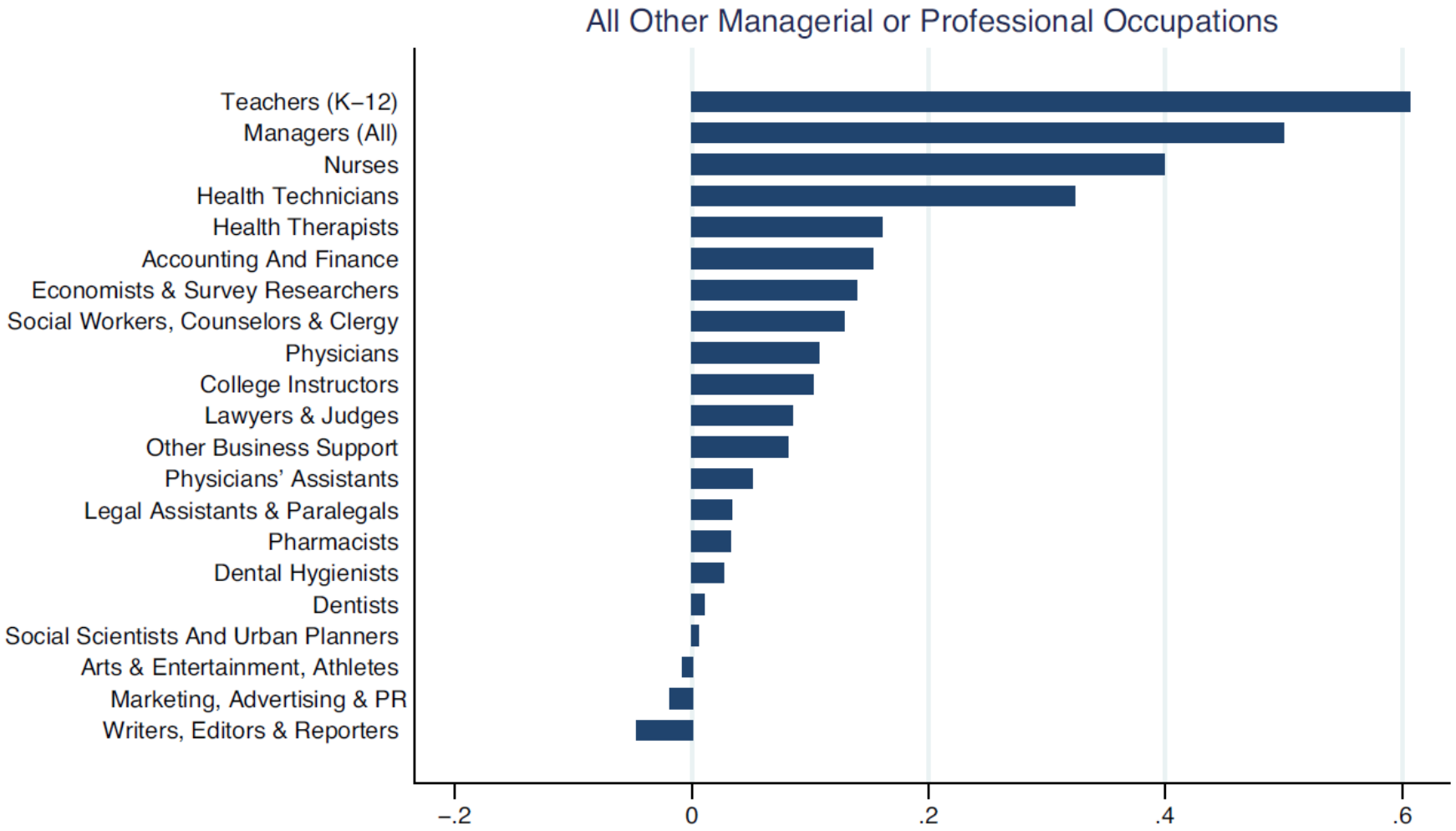


FIGURE I

Change in Relative Employment for Cognitive Occupations, 2000–2012



II. The Model

In a standard human capital model, worker skill takes a simple factor-augmenting form, where the output of worker j is increasing in some measure of skill (such as cognitive ability or education) A_j times L_j, the quantity of labor supplied:

$$y_j = A_j L_j. \tag{1}$$

$$y_j(i) = A_j \alpha_j(i) l_j(i), \tag{2}$$

where $y_j(i)$ specifies the production function for task i as worker j's cognitive skill A_j (still taking the factor-augmenting form) times a task-specific productivity parameter $\alpha_j(i)$ times labor supplied to task i.

Any job can be separated into an infinite number of discrete tasks that must be performed jointly to produce some final good Y. Following Acemoglu and Autor (2011), I assume that workers perform a continuum of tasks indexed over the unit interval according to a Cobb-Douglas technology:

$$Y_j = \exp[\int_0^1 \ln y_j(i) di]. \tag{3}$$

For simplicity, I assume that each worker supplies one unit of labor inelastically:

$$\int_0^1 l_j(i) di = L_j = 1. \tag{4}$$

Because the order of tasks over the unit interval is arbitary, it is convenient to index tasks in order of decreasing comparative advantage for worker 1 (i.e., $\frac{\alpha_1(0)}{\alpha_2(0)} > \cdots > \frac{\alpha_1(i)}{\alpha_2(i)} > \cdots > \frac{\alpha_1(1)}{\alpha_2(1)}$). Define the comparative advantage schedule over tasks as:

$$\gamma(i) \equiv \frac{A_1 \alpha_1(i)}{A_2 \alpha_2(i)}, \tag{5}$$

with $\gamma'(i) < 0$ by assumption.

For concreteness, I assume that the comparative advantage schedule takes the form:

$$\gamma(i) = \bar{A}\exp(\theta(1-2i)), \tag{6}$$

with $\bar{A} = \frac{A_1}{A_2}$. This functional form for $\gamma(i)$ can be derived from an underlying process where worker productivity in task i is drawn from a log-normal distribution with a mean that is increasing in cognitive skill A_j, and a variance that is increasing in θ.[10]

Each worker maximizes output by obtaining tasks from the lowest-cost producer, including herself. Workers trade tasks with each other at "prices" defined by efficiency units of labor, with a budget equal to each worker's labor supply constraint in equation (4). The worker-specific price of task i is:

$$p_j(i) = \frac{w_j}{A_j \alpha_j(i)}, \tag{7}$$

where w_j is the endogenously determined wage paid to worker j for a unit of labor.

$$\omega = \gamma(i^*), \tag{8}$$

where $\omega = \frac{w_1}{w_2}$. Worker 1 will perform all tasks in the interval [0, i^*] and worker 2 will perform all tasks in the interval [i^*, 1].

$$\omega = \frac{i^*}{1 - i^*}. \tag{9}$$

The relative wage ω is clearly increasing in the task threshold—for example, if $A_1 = A_2$, then $i^* = \frac{1}{2}$ and $\omega = 1$. Equilibrium wages for worker 1 are given by:

$$w_1 = P^* A_1^{i^*} (A_2 \omega)^{1-i^*} \exp\left[\int_0^{i^*} \ln\alpha_1(i)di + \int_{i^*}^{1} \ln\alpha_2(i)di\right]. \tag{10}$$

$$p_1(i) < p_2^S(i)$$

$$\frac{w_1}{A_1\alpha_1(i)} < \frac{w_2}{S^* A_2\alpha_2(i)}$$

$$\omega < \frac{\gamma(i)}{S^*}. \tag{11}$$

Likewise, worker 2 will produce her own tasks if $\omega > S^*\gamma(i)$. Thus in equilibrium there will be two task thresholds, defined by:

$$\gamma(i^H) = S^*\omega \tag{12}$$

$$\gamma(i^L) = \frac{\omega}{S^*}. \tag{13}$$

Since $\gamma'(i) < 0$, it is clear that $i^H > i^* > i^L$ when $S^* < 1$.

Figure II shows that—all else equal—the size of the non-traded zone $[i^L, i^H]$ is decreasing in θ. This can also be demonstrated by solving equations (12) and (13) for ω, which yields:

$$i^H - i^L = -\frac{\ln S^*}{\theta}. \tag{14}$$

As in the case of costless trade, equilibrium can be obtained by solving for the intersection between the two comparative advantage schedules in equations (12) and (13) and the demand for tasks, which is given simply by:

$$\omega = \frac{i^L}{1 - i^H}. \tag{15}$$

FIGURE II

Equilibrium Task Thresholds with Different Values of Theta

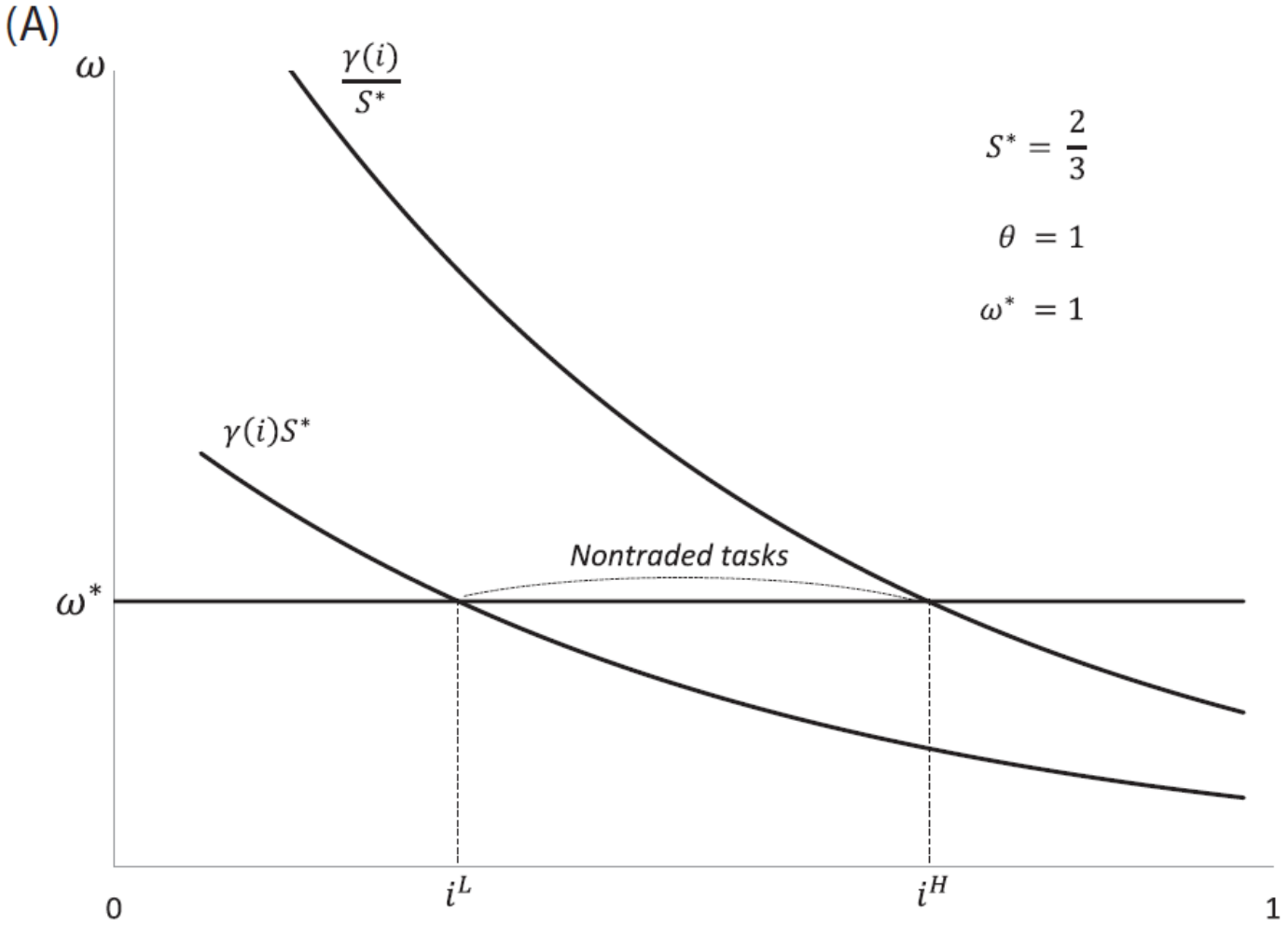
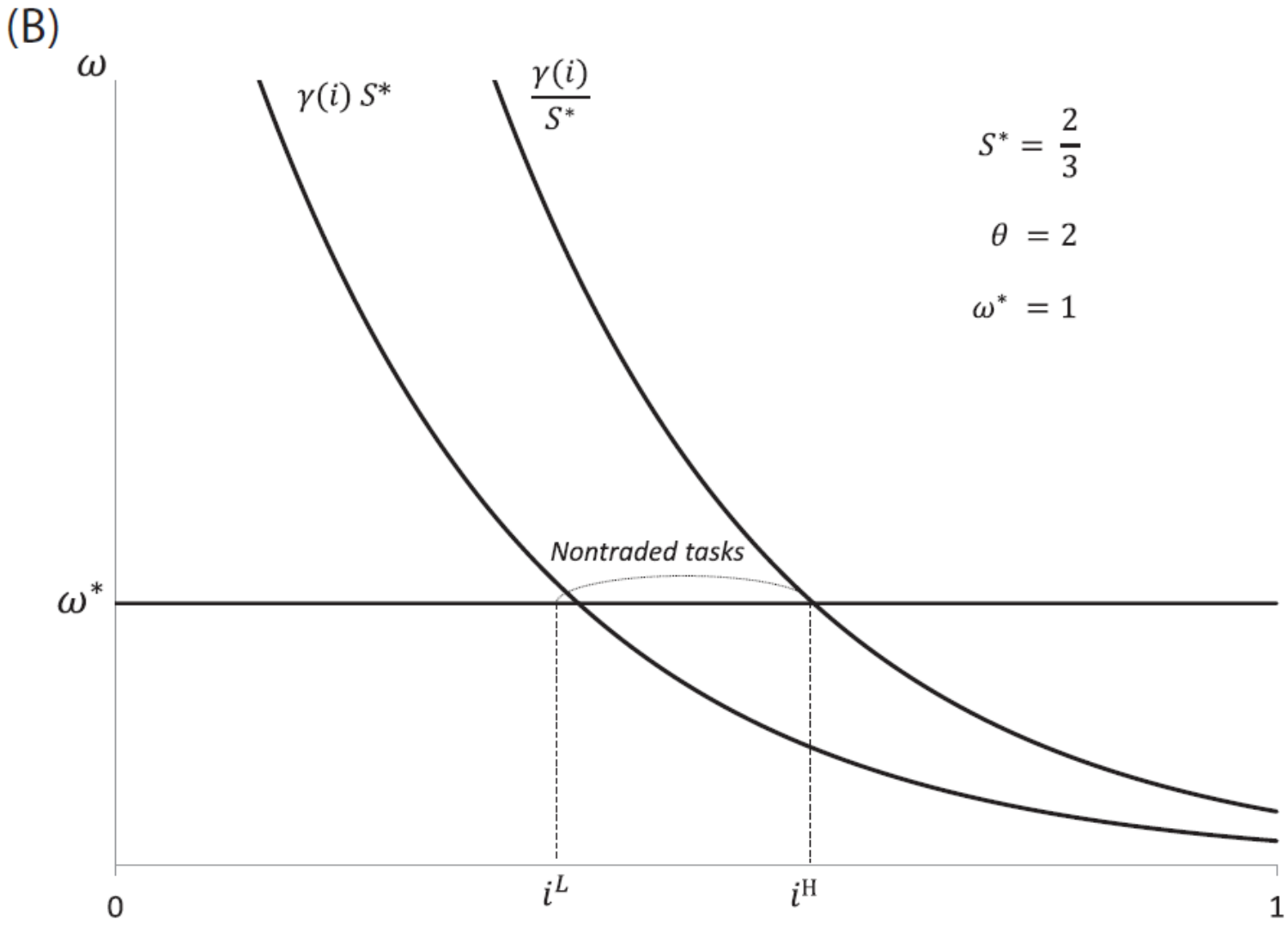


FIGURE II

Equilibrium Task Thresholds with Different Values of Theta



Finally, equilibrium wages for workers 1 and 2 are given by:

$$(16)\quad w_1 = P^* A_1^{i^H} (S^* A_2 \omega)^{1-i^H} \exp\left[\int_0^{i^H} \ln\alpha_1(i)di + \int_{i^H}^1 \ln\alpha_2(i)di\right],$$

$$(17)\quad w_2 = P^* A_2^{1-i^L} (S^* A_1 \omega^{-1})^{i^L} \exp\left[\int_0^{i^L} \ln\alpha_1(i)di + \int_{i^L}^1 \ln\alpha_2(i)di\right].$$

III. Data

*III.A. O*NET and Census / ACS Data*

I study changes in the the task content of work using data from O*NET. O*NET is a survey administered by the U.S. Department of Labor to a random sample of U.S. workers in each occupation. The O*NET survey began in 1998 and is updated periodically. I use the 1998 O*NET to most accurately reflect the task content of occupations in earlier years, although results with later versions of O*NET are generally similar.

I focus on changes in three key indicators of task content. First, I measure an occupation's routine task intensity as the average of the following two questions: (i) "how automated is the job?" and (ii) "how important is repeating the same physical activities (e.g. key entry) or mental activities (e.g. checking entries in a ledger) over and over, without stopping, to performing this job?"[18]

Second, I closely follow Autor, Levy, and Murnane (2003) and define nonroutine analytical (math) task intensity as the average of three O*NET variables that capture an occupation's mathematical reasoning requirements.[19] Third, I define an occupation's social skill intensity as the average of the four items in the O*NET module on "social skills": (i) coordination, (ii) negotiation, (iii) persuasion, and (iv) social perceptiveness.[20]

I use respondents' standardized scores on the Armed Forces Qualifying Test (AFQT) to proxy for cognitive skill, following many other studies (e.g., Neal and Johnson 1996; Altonji, Bharadwaj, and Lange 2012). Altonji, Bharadwaj, and Lange (2012) construct a mapping of the AFQT score across NLSY waves that is designed to account for differences in age-at-test, test format, and other idiosyncrasies. I take the raw scores from Altonji, Bharadwaj, and Lange (2012) and normalize them to have mean 0 and standard deviation 1.

Several psychometrically valid and field-tested measures of social skills exist, but none are used by the NLSY or other panel surveys of adult workers. As an alternative, I construct a premarket measure of social skills using the following four variables:

i. Self-reported sociability in 1981 (extremely shy, somewhat shy, somewhat outgoing, extremely outgoing)
ii. Self-reported sociability in 1981 at age 6 (retrospective)
iii. The number of clubs in which the respondent participated in high school
iv. Participation in high school sports (yes/no)

I normalize each variable to have a mean of 0 and a standard deviation of 1. I then take the average across all four variables and restandardize so that cognitive skills and social skills have the same distribution. The results are not sensitive to other reasonable choices, such as dropping any one of the four measures or constructing a composite using principal component analysis.

To account for possible bias from unmeasured ability differences, I control for completed years of education in addition to AFQT in some specifications. I also construct a measure of "noncognitive" skills using the normalized average of the Rotter Locus of Control and the Rosenberg Self-Esteem Scale—which are also used by Heckman, Stixrud, and Urzua (2006). This "noncognitive" skill measure is modestly positively correlated with both AFQT (0.30) and the social skills composite (0.20). To the extent that my measure of social skills is an imperfect or even poor proxy for the underlying construct, the results may understate its relative importance.

IV. NSLY79 Results

$$\ln\left(wage_{ijt}\right) = \alpha + \beta_1 COG_i + \beta_2 SS_i + \beta_3 COG_i * SS_i + \gamma X_{ijt} + \delta_j + \zeta_t + \epsilon_{ijt}. \quad (18)$$

$$\ln\left(wage_{ijt}\right) = \beta_1 COG_i * T_{ijt} + \beta_2 SS_i * T_{ijt} + \beta_3 COG_i * SS_i * T_{ijt} + \gamma X_{ijt} + \eta_i + \delta_j + \zeta_t + \epsilon_{ijt}. \quad (19)$$

TABLE I

LABOR MARKET RETURNS TO COGNITIVE SKILLS AND SOCIAL SKILLS IN THE NLSY79

Outcome is log hourly wage (in 2012 dollars)	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Cognitive skills (AQT, standardized)		0.206***	0.206***	0.189***	0.126***	0.190***	0.126***
		[0.007]	[0.007]	[0.007]	[0.008]	[0.007]	[0.008]
Social skills (standardized)	0.107***	0.055***	0.049***	0.043***	0.029***	0.044***	0.029***
	[0.006]	[0.006]	[0.006]	[0.006]	[0.006]	[0.006]	[0.006]
Cognitive * Social			0.019***	0.019***	0.011*	0.017***	0.010*
			[0.006]	[0.006]	[0.006]	[0.006]	[0.006]
Noncognitive skills (standardized)				0.048***	0.040***	0.046***	0.040***
				[0.006]	[0.006]	[0.006]	[0.006]
Cognitive * Noncognitive						0.008	0.001
						[0.006]	[0.006]
Demographics and age/year fixed effects		X	X	X	X	X	X
Years of completed education					X		X
R-squared	0.300	0.343	0.344	0.347	0.359	0.347	0.359
Observations	126,251	126,251	126,251	126,191	126,191	126,191	126,191

Notes. Each column reports results from an estimate of equation (18), with real log hourly wages as the outcome and person-year as the unit of observation. The data source is the National Longitudinal Survey of Youth 1979 cohort (NLSY79). Cognitive skills are measured by each NLSY79 respondent's score on the Armed Forces Qualifying Test (AFQT), and are normalized to have a mean of 0 and a standard deviation of 1. I use the AFQT score crosswalk developed by Altonji, Bharadwaj and Lange (2012). Social skills is a standardized composite of four variables (i) sociability in childhood, (ii) sociability in adulthood, (iii) participation in high school clubs, and (iv) participation in team sports; see the text for details on construction of the social skills measure. My measure of noncognitive skills is the normalized average of the Rotter and Rosenberg scores in the NLSY. The regression also controls for race-by-gender indicator variables, age, year, census region, and urbanicity fixed effects, plus additional controls as indicated. Standard errors are in brackets and are clustered at the individual level. ***$p < .01$, **$p < .05$, *$p < .10$.

TABLE II

OCCUPATIONAL SORTING ON SKILLS IN THE NLSY79

	Routine		Social skills	
Outcomes are O*NET task measures	(1)	(2)	(3)	(4)
Cognitive skills (AQT, standardized)	−0.055*	0.161***	0.345***	−0.044**
	[0.030]	[0.032]	[0.028]	[0.019]
Social skills (standardized)	−0.188***	−0.149***	0.208***	0.119***
	[0.022]	[0.024]	[0.020]	[0.014]
Cognitive * Social	−0.058***	−0.054**	0.014	0.013
	[0.021]	[0.023]	[0.019]	[0.014]
Demogs, age/year, education fixed effects	X	X	X	X
Controls for O*NET cognitive tasks		X		X
Observations	133,599	133,599	133,599	133,599
R-squared	0.204	0.237	0.305	0.668

Notes. Each column reports results from an estimate of equation (18), with the indicated 1998 O*NET task intensity of an occupation as the outcome and person-year as the unit of observation. The task measures are percentiles that range from 0 to 10 and are weighted by labor supply to conform to the 1980 occupation distribution. The additional O*NET cognitive task measures are nonroutine analytical, number facility, inductive/deductive reasoning, and analyze/use information. See the text and Online Appendix for details on the construction of each O*NET task measure. The data source is the National Longitudinal Survey of Youth 1979 cohort (NLSY79). Cognitive skills are measured by each NLSY79 respondent's score on the Armed Forces Qualifying Test (AFQT), and are normalized to have a mean of 0 and a standard deviation of 1. I use the AFQT score crosswalk developed by Altonji, Bharadwaj and Lange (2012). Social skills is a standardized composite of four variables: (i) sociability in childhood, (ii) sociability in adulthood, (iii) participation in high school clubs, and (iv) participation in team sports (see the text for details on construction of the social skills measure). My measure of noncognitive skills is the normalized average of the Rotter and Rosenberg scores in the NLSY. The regression also controls for race-by-gender indicator variables, age, year, census region, and urbanicity fixed effects, plus additional controls as indicated. Standard errors are in brackets and are clustered at the individual level. ***$p < .01$, **$p < .05$, *$p < .10$.

TABLE III

RETURNS TO SKILLS BY OCCUPATION TASK INTENSITY IN THE NLSY79

Outcome is log hourly wage (in 2012 dollars)	(1)	(2)	(3)
Routine task intensity	0.0136***		0.0212***
	[0.0012]		[0.0014]
Cognitive * Routine task intensity	−0.0034***		0.0005
	[0.0013]		[0.0015]
Social skills * Routine task intensity	−0.0025**		−0.0008
	[0.0013]		[0.0015]
Cognitive * Social * Routine task intensity	−0.0008		−0.0011
	[0.0012]		[0.0014]
Social skill task intensity		0.0039***	0.0176***
		[0.0013]	[0.0016]
Cognitive * Social skill task intensity		0.0113***	0.0112***
		[0.0015]	[0.0018]
Social skills * Social skill task intensity		0.0050***	0.0041**
		[0.0015]	[0.0018]
Cognitive * Social * Social skill task intensity		0.0021	0.0011
		[0.0015]	[0.0023]
Worker fixed effects	X	X	X
Observations	126,251	126,251	126,251
Number of individuals	11,050	11,050	11,050

V. The Growing Importance of Social Skills

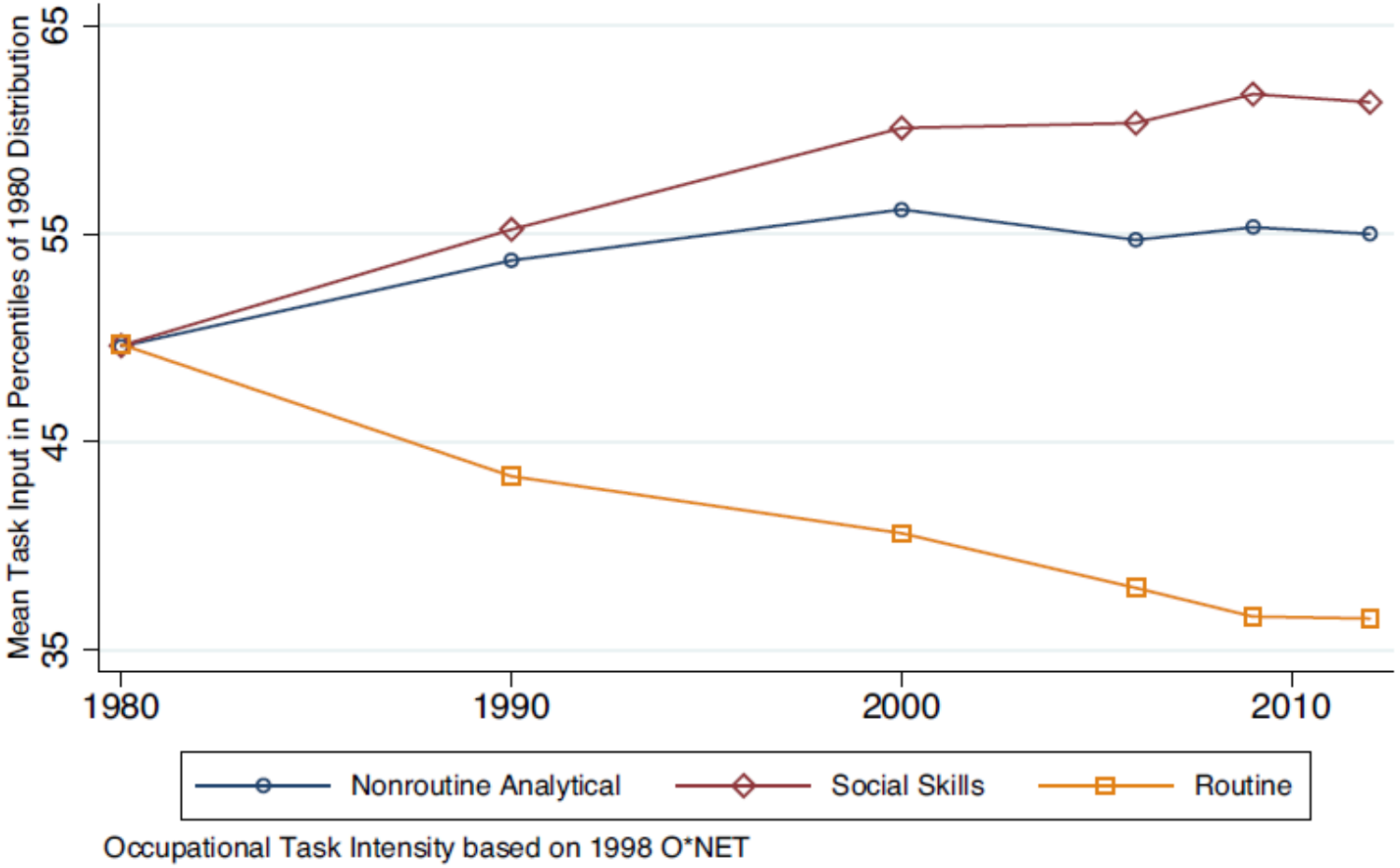


FIGURE III

Worker Tasks in the U.S. Economy, 1980–2012 (Update of Autor, Levy, and Murnane (2003) Figure I)

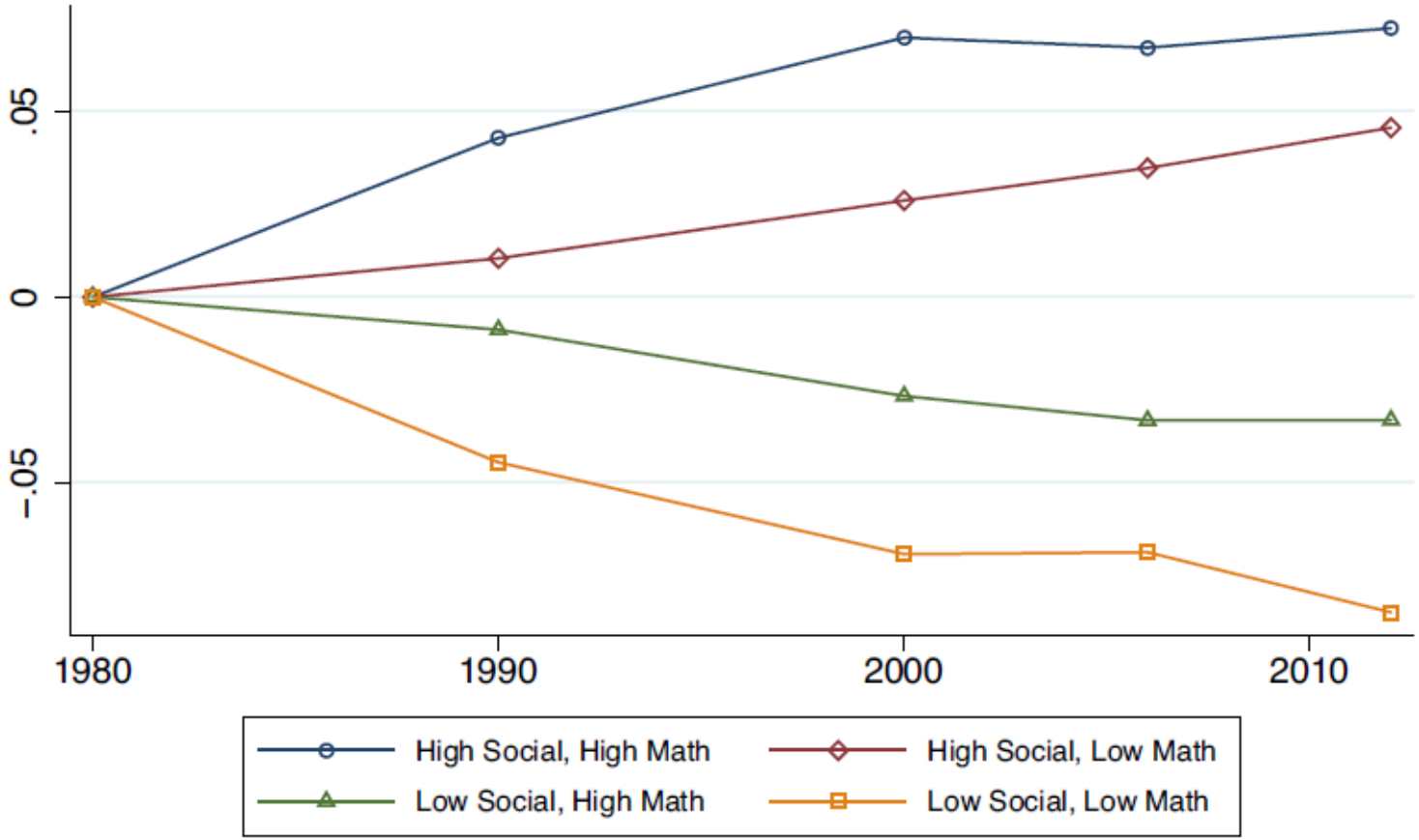


FIGURE IV

Cumulative Changes in Employment Share by Occupation Task Intensity, 1980–2012

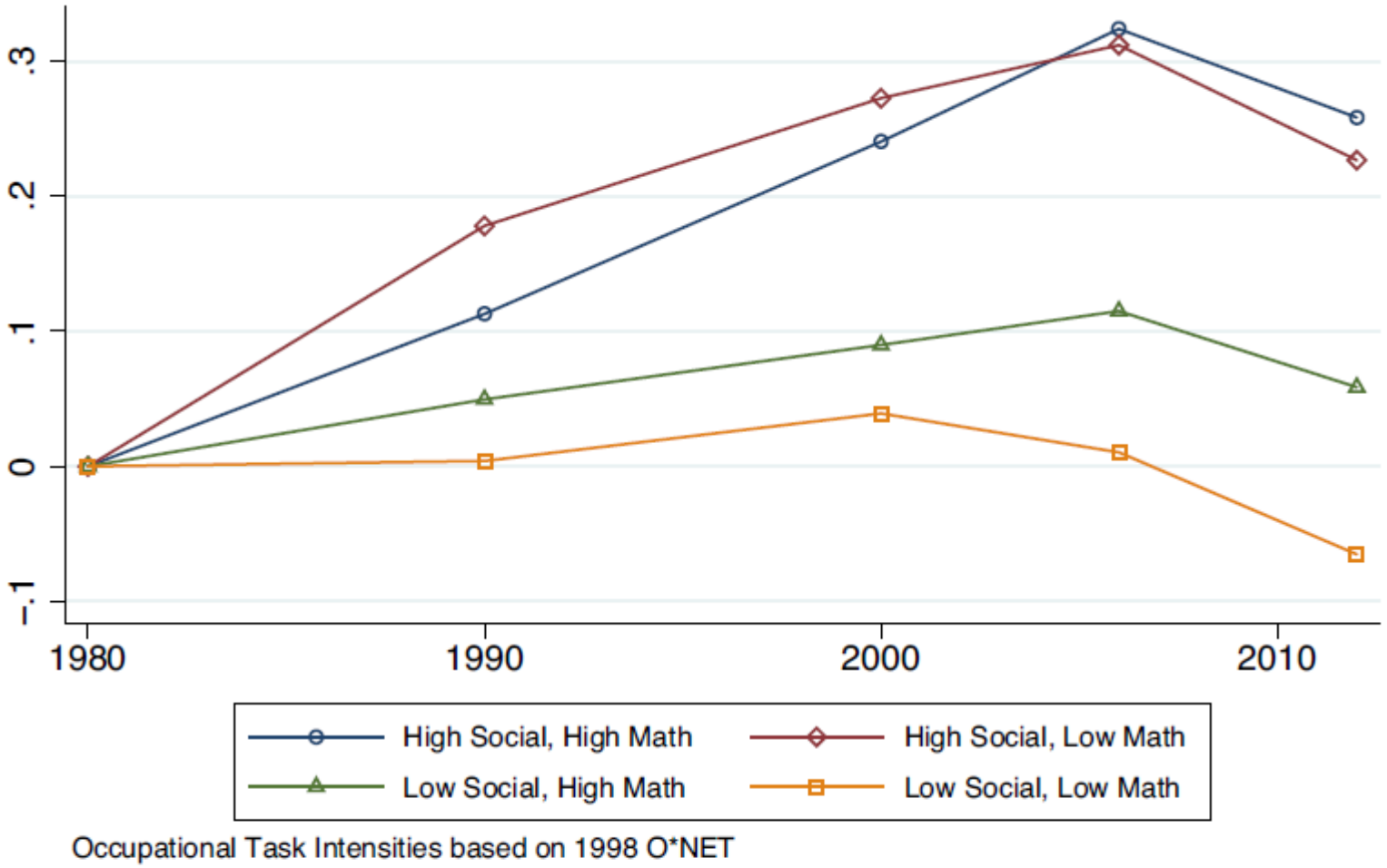


FIGURE V

Cumulative Changes in Real Hourly Wages by Occupation Task Intensity, 1980–2012

$$y_{ijt} = \alpha + \sum_{s=1}^{S} \left[\beta_s SKILL_i + \gamma_s \left(SKILL_i * NLSY97_i \right) \right] + \zeta X_{ijt} + \delta_j + \zeta_t + \epsilon_{ijt}. \quad (20)$$

TABLE IV.

LABOR MARKET RETURNS TO SKILLS IN THE NLSY79 VERSUS NLSY97

	Full-time employment			Log real hourly wage		
	(1)	(2)	(3)	(4)	(5)	(6)
Cognitive skills (AQT, standardized)	0.068***	0.042***	0.040***	0.203***	0.129***	0.116***
	[0.003]	[0.003]	[0.003]	[0.005]	[0.006]	[0.006]
Cognitive skills * NLSY97	0.008*	0.005	0.009*	−0.052***	−0.061***	−0.045***
	[0.004]	[0.004]	[0.005]	[0.008]	[0.008]	[0.008]
Social skills (standardized)	0.007***	0.005**	0.004*	0.020***	0.015***	0.011**
	[0.002]	[0.002]	[0.002]	[0.004]	[0.004]	[0.004]
Social skills * NLSY97	0.023***	0.021***	0.019***	0.017**	0.019**	0.016**
	[0.004]	[0.004]	[0.004]	[0.008]	[0.008]	[0.008]
Cognitive * Social	−0.007***	−0.006**	−0.007**	0.006	0.003	0.002
	[0.003]	[0.003]	[0.003]	[0.004]	[0.004]	[0.004]
Cognitive * Social * NLSY97	−0.006	−0.006	−0.006	−0.004	−0.004	−0.002
	[0.004]	[0.004]	[0.004]	[0.008]	[0.008]	[0.008]
Noncognitive skills (standardized)			0.008**			0.041***
			[0.003]			[0.005]
Noncognitive skills * NLSY97			0.013***			0.016*
			[0.005]			[0.009]
Demographics and age/year FE	X	X	X	X	X	X
Years of completed education		X	X		X	X
R-squared	0.081	0.096	0.097	0.309	0.333	0.337
Observations	104,613	104,252	104,206	77,845	77,631	77,599

Notes. Each column reports results from an estimate of equation (20), with an indicator for being employed full-time as the outcome in columns (1)–(3), real log hourly wages as the outcome in columns (4)–(6), and person-year as the unit of observation. The data are a pooled sample of two cohorts of youth: the National Longitudinal Survey of Youth 1979 (NLSY79) and 1997 (NLSY97) waves. I restrict the age range to 25–33, which allows for a comparison of NLSY respondents at similar ages across survey waves. Cognitive skills are measured by each NLSY respondent's score on the Armed Forces Qualifying Test (AFQT), and are normalized to have a mean of 0 and a standard deviation of 1. I use the AFQT score crosswalk developed by Altonji, Bharadwaj, and Lange (2012), which adjusts for differences across survey waves in age-at-test and test format. Social skills is a standardized composite of two variables that measure extroversion in both the NLSY79 (sociability in childhood and sociability in adulthood) and in the NLSY97 (two items from the Big 5 personality inventory that measure extroversion). The noncognitive skill measures are a normalized average of the Rotter and Rosenberg scores in the NLSY79, and two items from the NLSY97 that measure the Big 5 personality factor conscientiousness. The regression also controls for an indicator for whether the respondent was in the NLSY97 wave, race-by-gender indicator variables, age, year, census region, and urbanicity fixed effects, plus additional controls as indicated. Standard errors are in brackets and are clustered at the individual level. ***$p < .01$, **$p < .05$, *$p < .10$.

$$\ln\left(wage_{ijt}\right) = \sum_{s=1}^{S}\left[\beta_s\left(SKILL_i * T_{ijt}\right) + \vartheta_s\left(T_{ijt} * NLSY97_i\right)\right.$$
$$\left. + \gamma_s\left(SKILL_i * T_{ijt} * NLSY97_i\right)\right]$$
$$+ \zeta X_{ijt} + \eta_i + \delta_j + \phi_t + \epsilon_{ijt}. \tag{21}$$

TABLE V

RETURNS TO SKILLS BY OCCUPATION TASK INTENSITY IN THE NLSY79 VERSUS NLSY97

Outcome is log hourly wage (in 2012 dollars)	(1)	(2)	(3)	(4)
Social skill task intensity	0.0004	−0.0096***	−0.0095***	−0.0096***
	[0.0013]	[0.0015]	[0.0015]	[0.0015]
Social skill task intensity * NLSY97	0.0210***	0.0253***	0.0217***	0.0225***
	[0.0036]	[0.0041]	[0.0040]	[0.0040]
Math task intensity		0.0175***	0.0177***	0.0177***
		[0.0015]	[0.0015]	[0.0015]
Math task intensity * NLSY97		−0.0082**	−0.0085**	−0.0099***
		[0.0035]	[0.0034]	[0.0034]
Cognitive skill * Social skill task intensity			0.0069***	0.0074***
			[0.0013]	[0.0016]
Cognitive skill * Social skill task intensity * NLSY97			0.0114***	0.0047
			[0.0036]	[0.0044]
Social skill * Social skill task intensity			0.00008	0.0011
			[0.0013]	[0.0016]
Social skill * Social skill task intensity * NLSY97			0.0040	0.0069*
			[0.0032]	[0.0038]
P (Social skill * Social skill intensity in NLSY97 > 0)			0.108	0.023
P (All skills * Social skill intensity in NLSY97 > 0)			0.000	0.001
P (All skills in NLSY97 > All skills in NLSY79)				0.000
Observations	77,845	77,845	77,845	77,845
Number of individuals	14,998	14,998	14,998	14,998

Notes. Each column reports results from an estimate of equation (21), with real log hourly wages as the outcome and person-year as the unit of observation. The data are a pooled sample of two cohorts of youth: the National Longitudinal Survey of Youth 1979 (NLSY79) and 1997 (NLSY97) waves. I restrict the age range to 25–33, which allows for a comparison of NLSY respondents at similar ages across survey waves. Cognitive skills are measured by each NLSY respondent's score on the Armed Forces Qualifying Test (AFQT), and are normalized to have a mean of 0 and a standard deviation of 1. I use the AFQT score crosswalk developed by Altonji, Bharadwaj and Lange (2012), which adjusts for differences across survey waves in age-at-test and test format. Social skills is a standardized composite of two variables that measure extroversion in both the NLSY79 (sociability in childhood and sociability in adulthood) and in the NLSY97 (two items from the Big 5 personality inventory that measure extroversion). The regression also controls for age, year, census region, and urbanicity fixed effects, plus additional controls as indicated. The interactions between cognitive/social skills and 1998 O*NET task intensities measure whether the returns to skills vary with the task content of the worker's occupation. The task measures are percentiles that range from 0 to 10 and are weighted by labor supply to conform to the 1980 occupation distribution. See the text and Online Appendix for details on the construction of each O*NET task measure. Standard errors are in brackets and are clustered at the individual level. ***$p < .01$, **$p < .05$, *$p < .10$.

VI. Conclusion