Appendix Results

NOT FOR PUBLICATION

Table A1: Change in Skill Requirements by Occupation Category, 2007-2017

Panel A	Social	Cognitive	Character	Creativity	Writing	Management	Finance
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Management	0.073	0.039	0.056	0.011	-0.009	-0.009	*
STEM	0.043	0.037	0.019	0.023	0.008	0.005	*
Business	0.037	0.038	0.037	*	*	0.020	-0.016
Social Science / Service	0.083	0.037	0.046	0.009	*	*	-0.012
Art/Design/Media	0.034	0.023	0.035	0.040	*	-0.016	-0.018
Health	0.092	-0.005	0.029	-0.002	*	0.019	*
Sales and Admin	0.027	0.008	-0.022	0.027	-0.011	0.026	0.038

Panel B	Business Systems	Customer Service	Office Software	Technical Support	Data Analysis	Specialized Software	ML and AI
Management	0.011	0.018	*	-0.046	0.003	-0.007	0.008
STEM	-0.004	-0.007	-0.039	-0.087	0.035	-0.006	0.104
Business	0.038	-0.015	-0.026	-0.066	0.009	0.005	0.015
Social Science / Service	*	0.030	-0.016	-0.028	*	*	0.005
Art/Design/Media	*	-0.014	-0.049	-0.082	0.012	*	0.014
Health	0.004	0.149	0.027	*	*	0.020	0.004
Sales and Admin	0.035	*	-0.056	-0.054	-0.005	*	*

Notes: Each cell in this table presents the coefficient on an indicator for the 2017 year from a separate vacancy-level regression of the frequency of each skill requirement on the 2017 indicator, the total number of skills listed in each vacancy, education and experience requirements, and occupation-city-employer fixed effects. The occupations are grouped based on 2010 Standard Occupation Classification (SOC) codes. Data come from online job vacancies collected by Burning Glass Technologies in 2007 and 2017. Stars indicate a coefficient that was NOT statistically different from zero at the 5 percent level or less. See the Data Appendix for detailed descriptions of how each skill category is constructed.

Table A2: Occupations with the Highest and Lowest Rates of Task Change - Alternative Method of Calculation

Panel A: Fastest-Changing Professional Occupations (3-digit)

SOC code	Occupation Title	Rate of Task Change
231	Lawyers, Judges and Related Workers	0.239
171	Architects and Surveyors	0.217
192	Physical Scientists	0.207
191	Life Scientists	0.203
172	Engineers	0.197
173	Drafters and Engineering Technicians	0.197
152	Mathematical Scientists	0.184
113	Operations Specialties Managers	0.173
254	Librarians, Curators and Archivists	0.172
232	Legal Support Workers	0.165

Panel B: Fastest-Changing Professional Occupations (6-digit)

SOC code	Occupation Title	Rate of Task Change
173013	Mechanical Drafters	0.404
151131	Computer Programmers	0.355
173011	Architectural and Civil Drafters	0.344
151133	Software Developers, Systems Software	0.301
112011	Advertising and Promotions Managers	0.282
172081	Environmental Engineers	0.281
132053	Insurance Underwriters	0.281
291051	Pharmacists	0.281
173012	Electrical and Electronics Drafters	0.274
152011	Actuaries	0.244

Panel C: Slowest-Changing Professional Occupations (3-digit)

SOC code	Occupation Title	Rate of Task Change
193	Social Scientists and Related Workers	0.099
291	Health Diagnosing and Treating Practitioners	0.101
252	Pre-K, Primary and Secondary School Teachers	0.104
259	Other Education, Training and Library Occupations	0.105
253	Other Teachers and Instructors	0.109
292	Health Technologists and Technicians	0.117
111	Managers and Executives	0.122
194	Life, Physical and Social Science Technicians	0.126
299	Other Healthcare Practitioners	0.128
211	Counselors and Social Workers	0.131

Panel D: Slowest-Changing Professional Occupations (6-digit)

SOC code	Occupation Title	Rate of Task Change
291065	Pediatricians	0.030
291171	Nurse Practitioners	0.050
291021	Dentists	0.052
193031	Clinical Psychologists	0.054
291131	Veterinarians	0.054
292052	Pharmacy Technicians	0.056
252059	Special Education Teachers, All Other	0.056
291066	Psychiatrists	0.057
291069	Physicians and Surgeons, All Other	0.063
291151	Nurse Anesthetists	0.067

Notes: This table uses online job vacancy data from Burning Glass Technologies (BG) to calculate the rate of task change between 2007 and 2017 for each 3- and 6-digit Standard Occupational Classification (SOC) code. The task change measure ranges between 0 and 1, with zero indicating that the tasks demanded by employers in the occupation in 2007 were exactly the same as in 2017, and 1 indicating that the job has a completely different set of task demands. The average value of the task change measure is 0.13 - see the text for details. Professional Occupations are SOC codes where the first digit begins with a 1 or a 2.

Table A3: Occupations with the Highest and Lowest Rates of Software-Related Task Change

Panel A: Fastest-Changing Professional Occupations (3-digit)

SOC code	Occupation Title	Software Task Change
171	Architects and Surveyors	0.062
151	Computer Occupations	0.058
173	Drafters and Engineering Technicians	0.030
172	Engineers	0.030
152	Mathematical Scientists	0.029
271	Art and Design Workers	0.028
273	Media and Communications Workers	0.025
274	Media and Communications Equipment Workers	0.024
131	Business Operations Specialists	0.022
132	Financial Specialists	0.019

Panel B: Fastest-Changing Professional Occupations (6-digit)

SOC code	Occupation Title	Software Task Change
151131	Computer Programmers	0.140
151133	Software Developers, Systems Software	0.133
151134	Web Developers	0.098
173011	Architectural and Civil Drafters	0.094
173013	Mechanical Drafters	0.094
271014	Multimedia Artists and Animators	0.076
151142	Network and Computer Systems Administr	0.069
271024	Graphic Designers	0.067
151141	Database Administrators	0.065
151141	Software Developers, Applications	0.065

Panel C: Slowest-Changing Professional Occupations (3-digit)

SOC code	Occupation Title	Software Task Change
291	Health Diagnosing and Treating Practitioners	0.002
252	Pre-K, Primary and Secondary School Teachers	0.004
253	Other Teachers and Instructors	0.004
292	Health Technologists and Technicians	0.004
211	Counselors and Social Workers	0.008
251	Postsecondary Teachers	0.009
259	Other Education, Training and Library Occupations	0.010
299	Other Healthcare Practitioners	0.010
191	Life Scientists	0.010
193	Social Scientists	0.010

Panel D: Slowest-Changing Professional Occupations (6-digit)

SOC code	Occupation Title	Software Task Change
291065	Pediatricians	0.000
291151	Nurse Anesthetists	0.000
292021	Dental Hygienists	0.001
291131	Veterinarians	0.001
291171	Nurse Practitioners	0.001
291066	Psychiatrists	0.001
252011	Preschool Teachers	0.001
291021	Dentists	0.001
292055	Surgical Technologists	0.001
291062	Family and General Practitioners	0.001

Notes: This table uses online job vacancy data from Burning Glass Technologies (BG) to calculate the rate of change for software-related tasks between 2007 and 2017 for each 3- and 6-digit Standard Occupational Classification (SOC) code. The task change measure ranges between 0 and 1, with zero indicating that the tasks demanded by employers in the occupation in 2007 were exactly the same as in 2017, and 1 indicating that the job has a completely different set of task demands. We restrict the set of occupations in the table to those with at least 10,000 total vacancies posted in both 2007 and 2017. The average value of the software task change measure is 0.016 - see the text for details. Professional Occupations are SOC codes where the first digit begins with a 1 or a 2.

Table A4: Selection into Graduate School in the NLSY

Outcome is an indicator for graduate education	(1)	(2)	(3)
Cognitive Skill (AFQT, standardized)	0.099***	0.099***	0.131***
	[0.022]	[0.022]	[0.032]
STEM Major	-0.034	-0.110***	-0.013
	[0.032]	[0.041]	[0.073]
NLSY Wave	0.103***	0.063**	0.089*
	[0.027]	[0.031]	[0.051]
STEM Major * NLSY Wave		0.181***	0.162
		[0.064]	[0.121]
AFQT * NLSY Wave			-0.028
			[0.044]
STEM Major * AFQT			-0.100*
			[0.062]
STEM Major * AFQT * NLSY Wave			0.027
			[0.100]
R-squared	0.030	0.035	0.038
Number of Observations	1,360	1,360	1,360

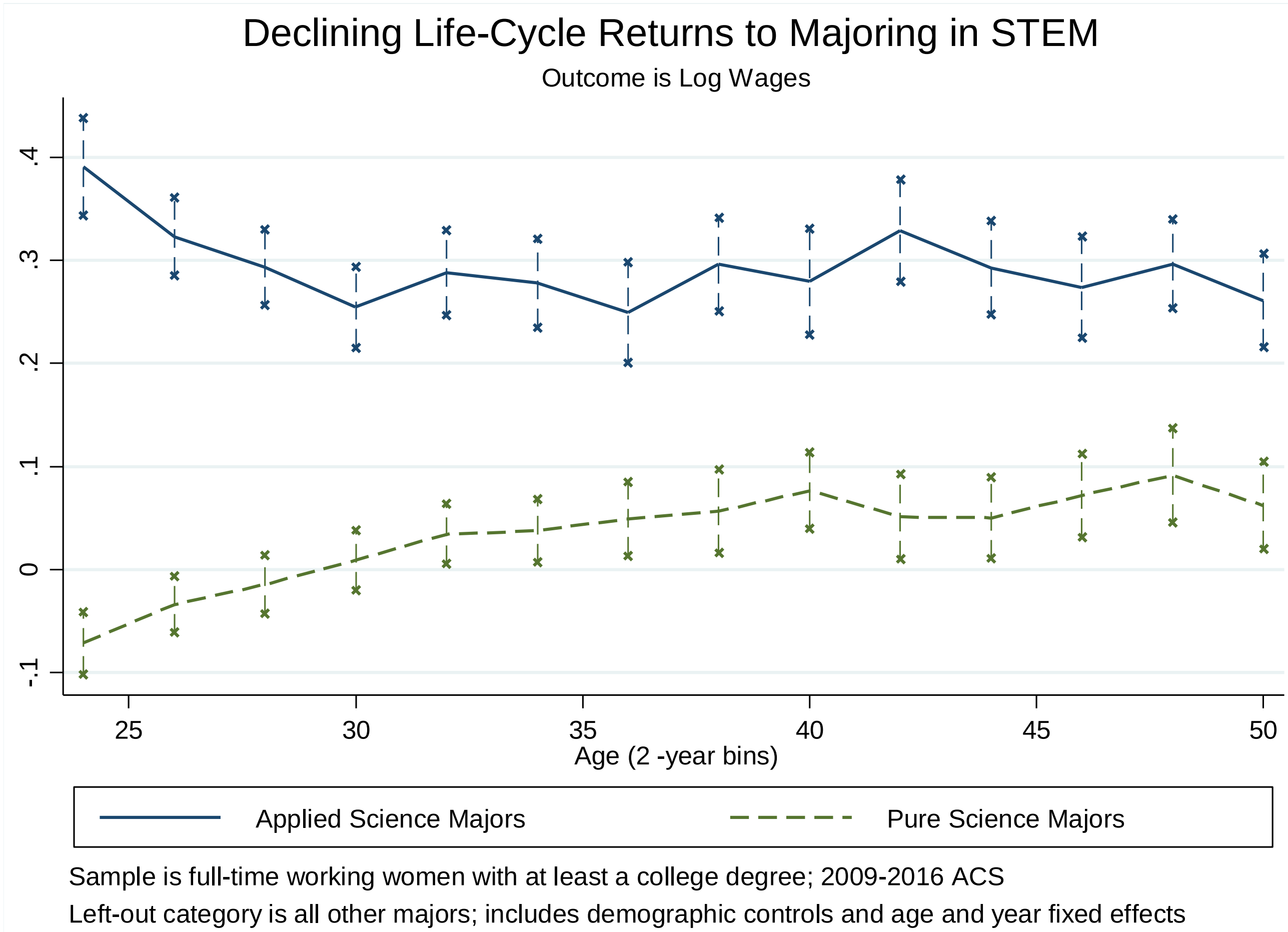
Notes: Each column reports results from a regression of an indicator for graduate school attendance on the Armed Forces Qualifying Test (AFQT) score, indicators for college major, an indicator for whether the respondent is in the National Longitudinal Survey of Youth (NLSY) 1997 survey wave, and other variables as shown. The regression also includes controls for race. The sample pools the 1979 and 1997 NLSY waves together and is restricted to men with at least a college degree. STEM majors are defined following Peri, Shih and Sparber (2015). "Pure" Science majors include biology, chemistry, physics, mathematics and statistics, while "Applied" Science includes engineering and computer science. We normalize scores across NLSY waves using the crosswalk developed by Altonji, Bharadwaj and Lange (2012). The sample is restricted to ages 23-34 to maximize comparability across survey waves. *** p<0.01, ** p<0.05, * p<0.10.

Table A5: Ability Sorting into STEM Majors in the NLSY

Outcome is AFQT Score (standardized)	(1)	(2)	(3)
STEM Major	0.083**	0.076	
	[0.039]	[0.051]	
NLSY 97 Wave	0.048	0.044	0.051
	[0.053]	[0.056]	[0.055]
STEM Major * NLSY Wave		0.019	
		[0.079]	
Applied Science Major			0.083
			[0.056]
Applied Science Major * NLSY Wave			-0.021
			[0.089]
R-squared	0.217	0.217	0.217
Number of Observations	1,360	1,360	1,360

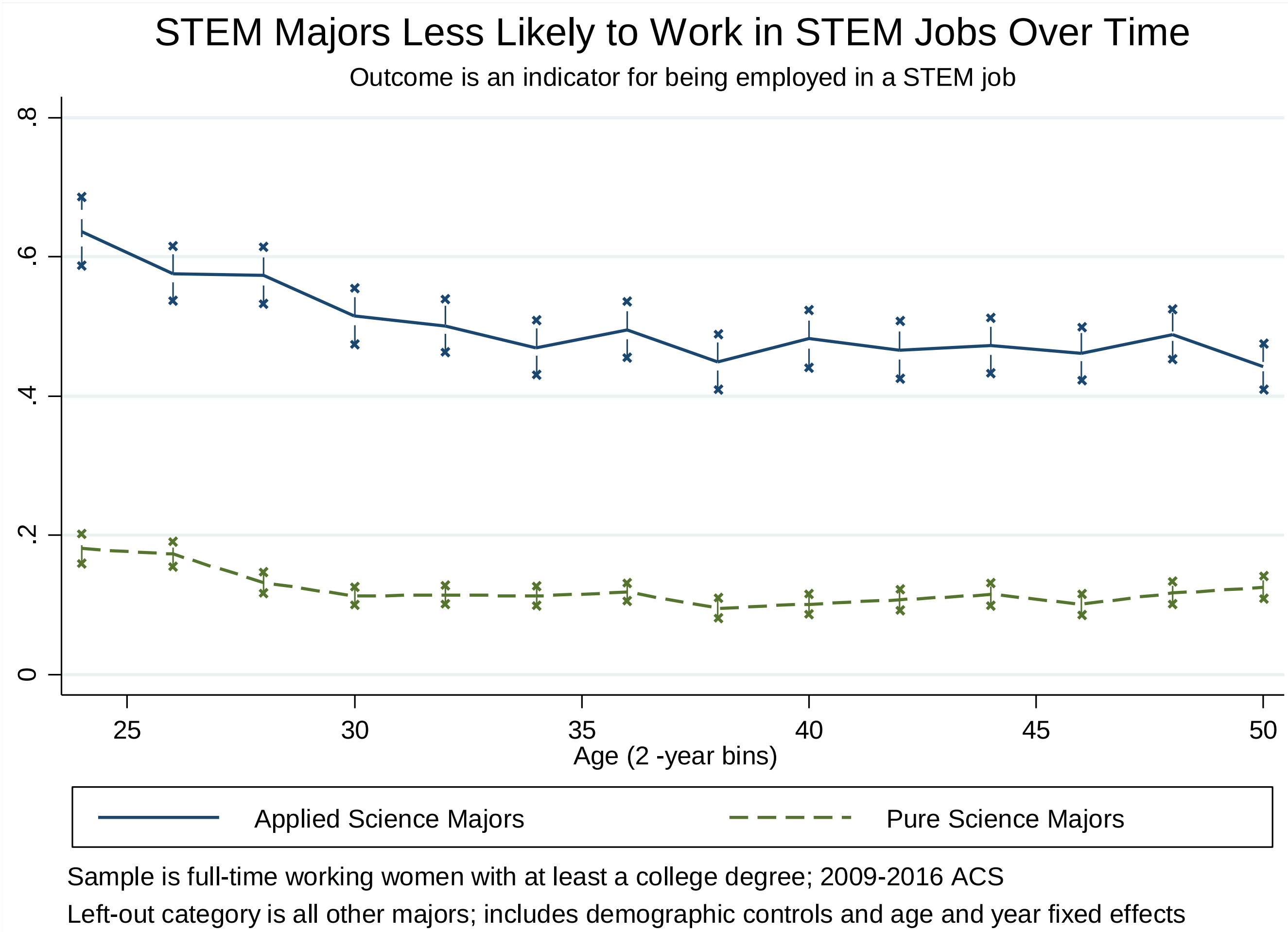
Notes: Each column reports results from a regression of the Armed Forces Qualifying Test (AFQT) score on indicators for college major and an indicator for whether the respondent is in the National Longitudinal Survey of Youth (NLSY) 1997 survey wave. The regression also includes controls for race and the age at which the test was taken. The sample pools the 1979 and 1997 NLSY waves together and is restricted to men with at least a college degree. STEM majors are defined following Peri, Shih and Sparber (2015). "Pure" Science majors include biology, chemistry, physics, mathematics and statistics, while "Applied" Science includes engineering and computer science. We normalize scores across NLSY waves using the crosswalk developed by Altonji, Bharadwaj and Lange (2012). The sample is restricted to ages 23-34 to maximize comparability across survey waves. *** p<0.01, ** p<0.05, * p<0.10.

Figure A1



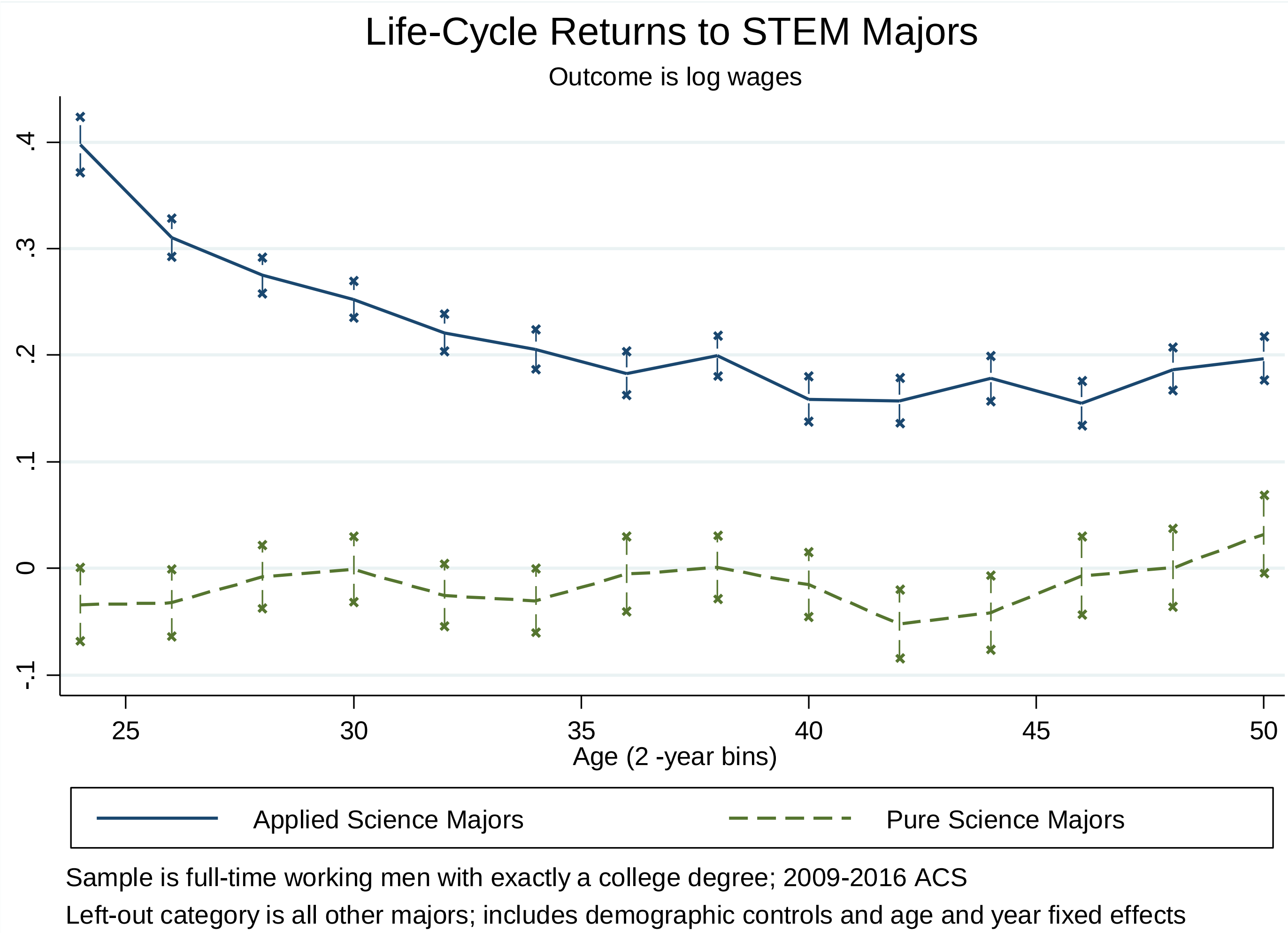
Notes: The figure plots coefficients and 95 percent confidence intervals from an estimate of the returns to majors over time for women, following equation (10) in the paper. "Pure" Science includes biology, chemistry, physics, mathematics and statistics, while "Applied" Science includes engineering and computer science.

Figure A2



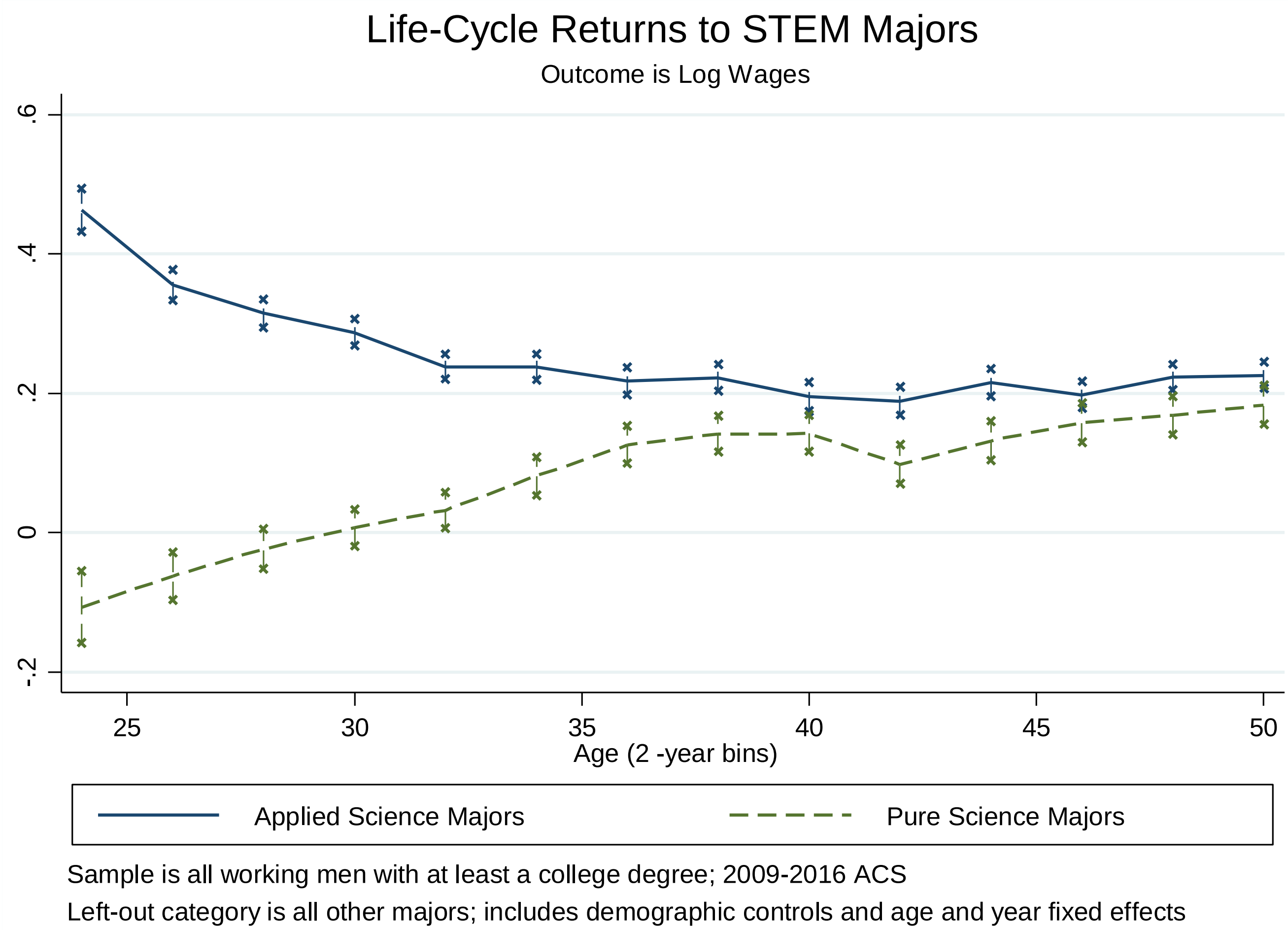
Notes: The figure plots coefficients and 95 percent confidence intervals from an estimate of equation (10) in the paper, except with an indicator for working in a STEM occupation as the outcome. "Pure" Science includes biology, chemistry, physics, mathematics and statistics, while "Applied" Science includes engineering and computer science. STEM occupations are defined using the 2010 Census Bureau classification.

Figure A3



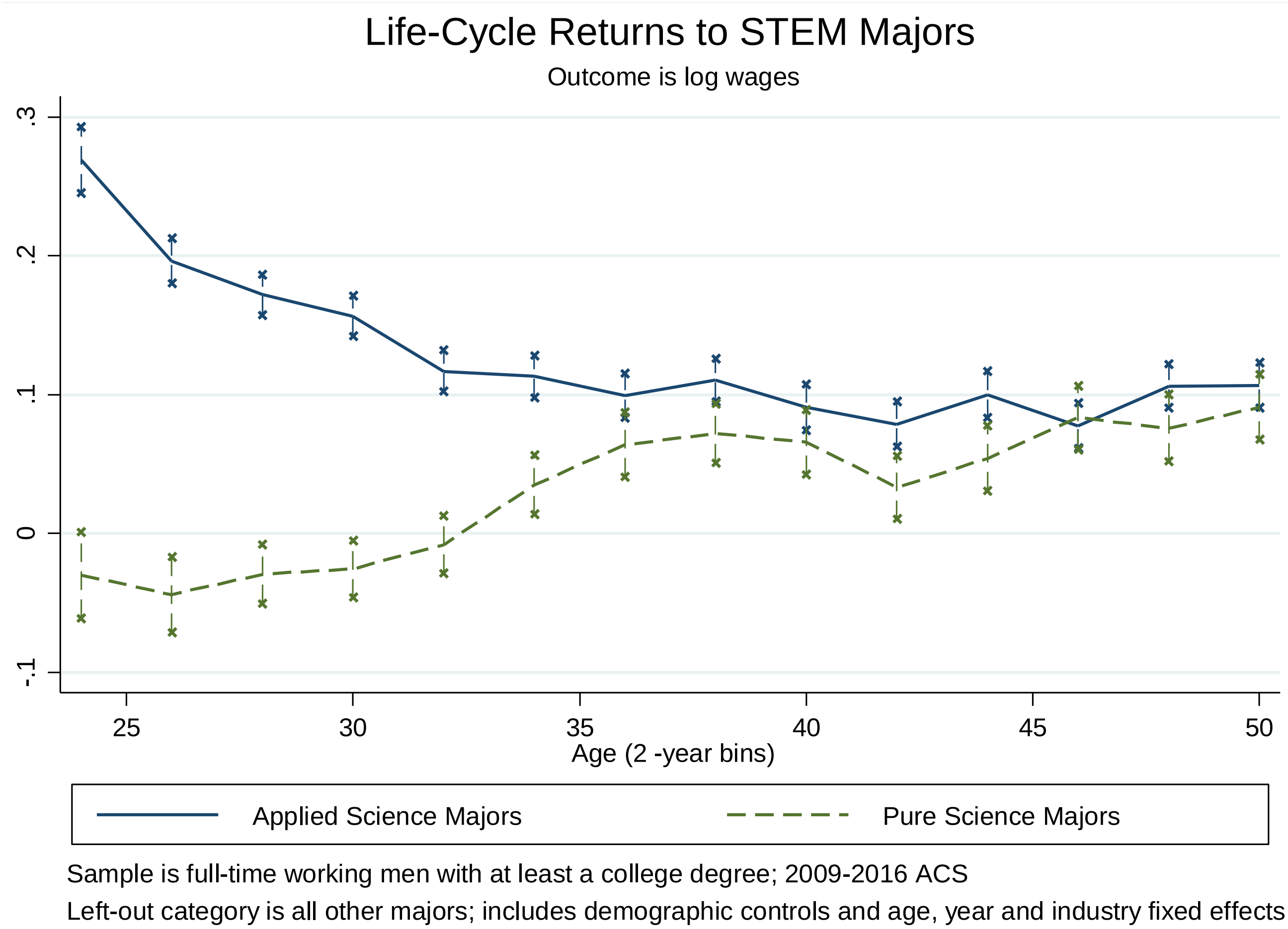
Notes: The figure plots coefficients and 95 percent confidence intervals from an estimate of equation (10) in the paper, except the sample is restricted to full-time working men with exactly a bachelor's degree. "Pure" Science includes biology, chemistry, physics, mathematics and statistics, while "Applied" Science includes engineering and computer science.

Figure A4



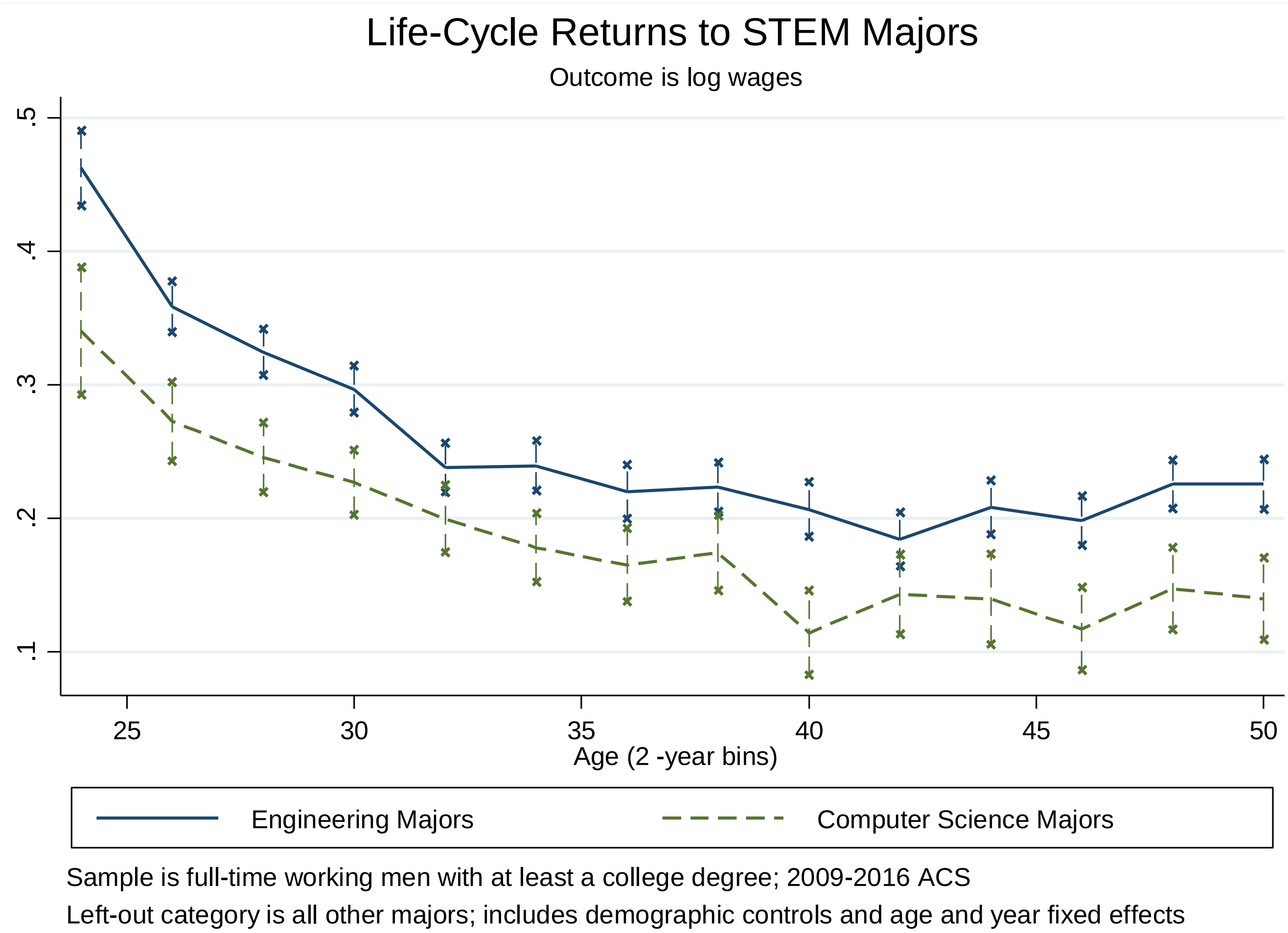
Notes: The figure plots coefficients and 95 percent confidence intervals from an estimate of equation (10) in the paper, except the sample includes all working men (not just full-time). "Pure" Science includes biology, chemistry, physics, mathematics and statistics, while "Applied" Science includes engineering and computer science.

Figure A5



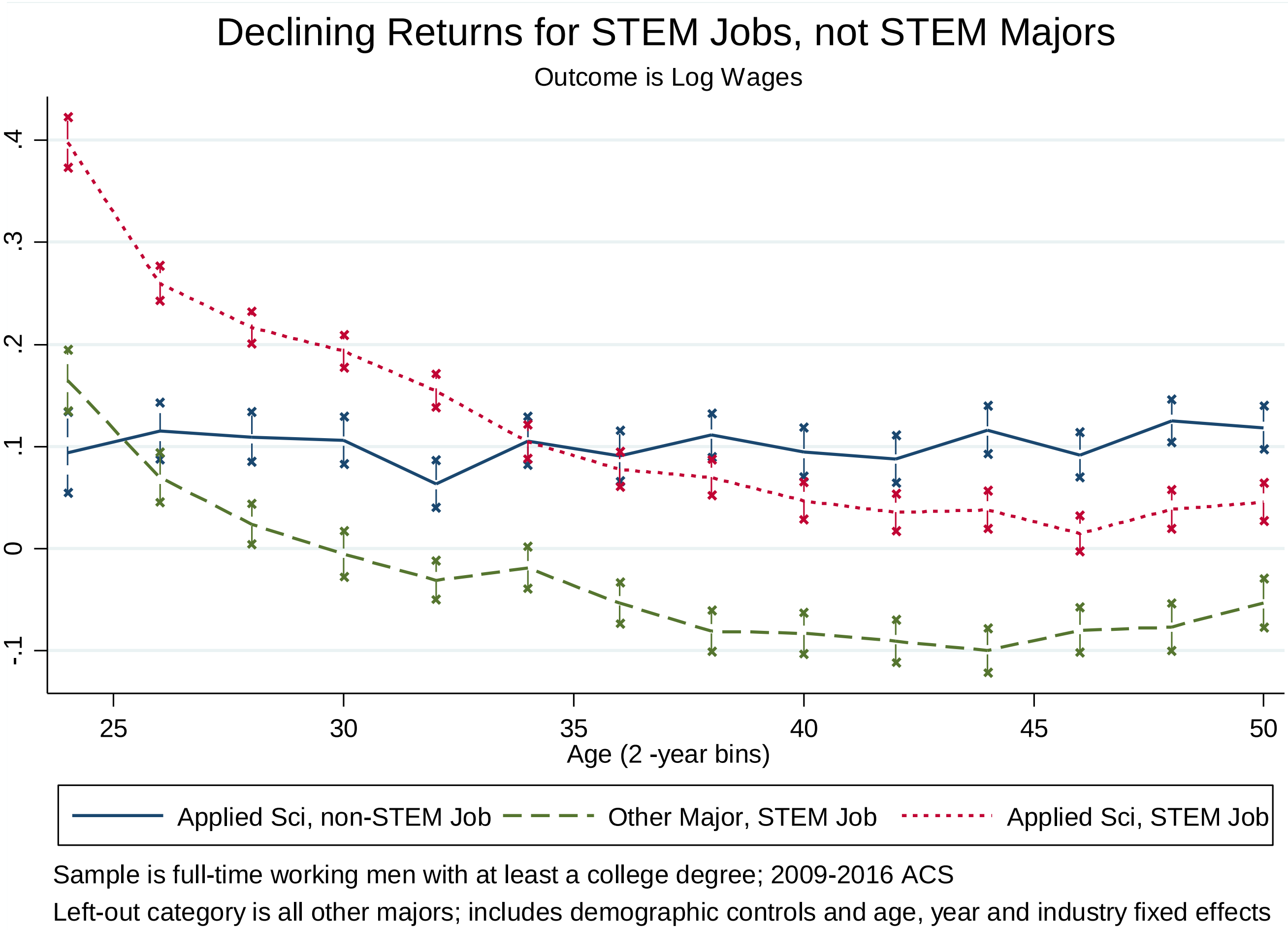
Notes: The figure plots coefficients and 95 percent confidence intervals from an estimate of equation (10) in the paper, except with industry fixed effects also included in the regression. "Pure" Science includes biology, chemistry, physics, mathematics and statistics, while "Applied" Science includes engineering and computer science.

Figure A6



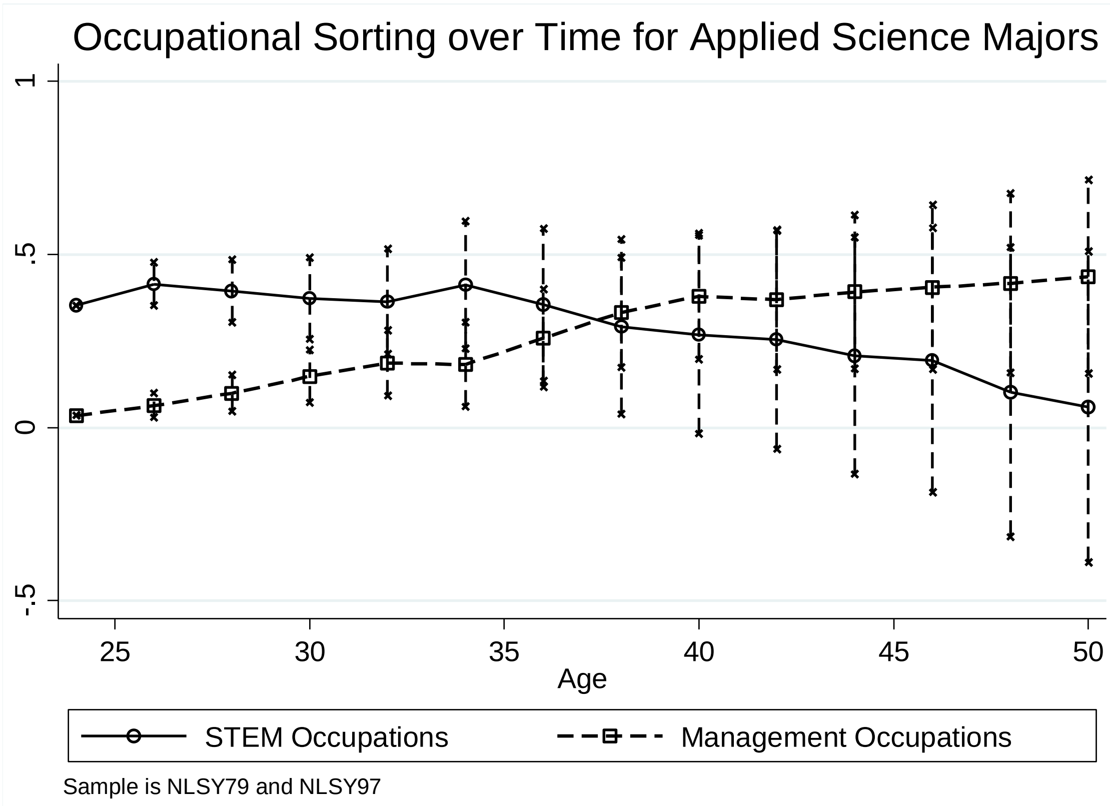
Notes: The figure plots coefficients and 95 percent confidence intervals from an estimate of equation (10) in the paper, except with Engineering and Computer Science majors estimated separately.

Figure A7



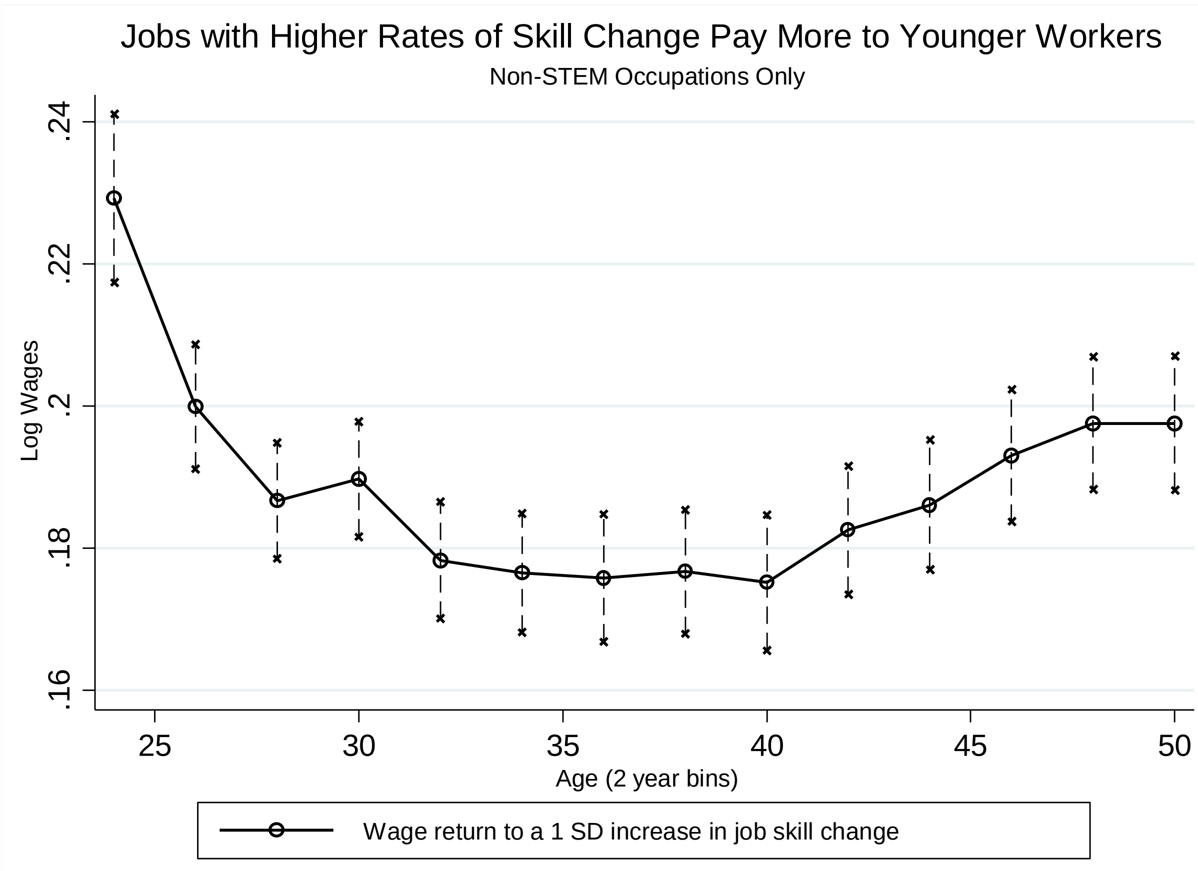
Notes: The figure plots coefficients and 95 percent confidence intervals from an estimate of the returns to majors over time, following equation (10) in the paper, but adding occupation and major interactions as well as industry fixed effects. "Applied" Science majors include engineering and computer science.

Figure A8



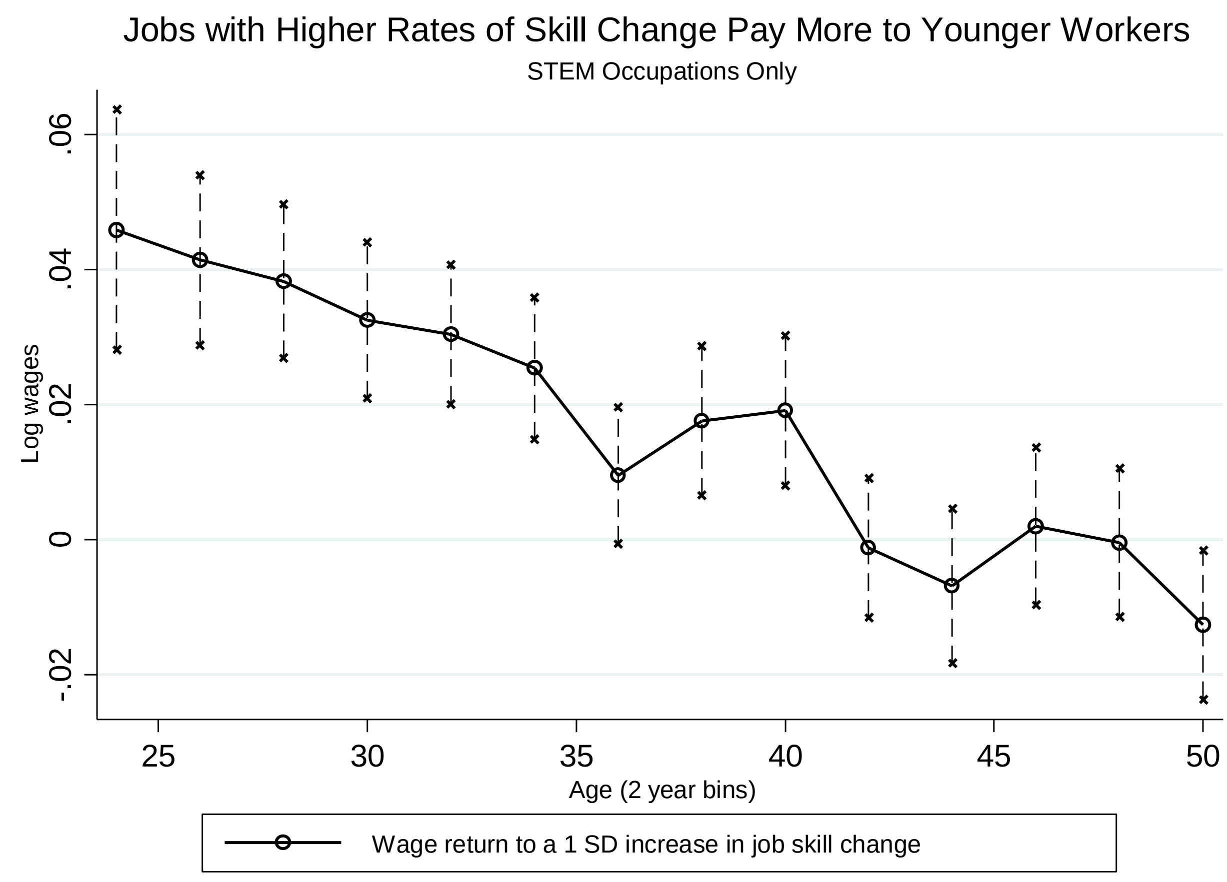
Notes: The figure plots coefficients and 95 percent confidence intervals from two separate estimates of equation (10) in the paper, restricting the sample to Applied Science majors and with indicators for working in STEM and management occupations as the outcome variables. STEM occupations are defined using the 2010 Census Bureau classification. Includes demographic controls, age and year fixed effects, and controls for cognitive and non-cognitive skills.

Figure A9



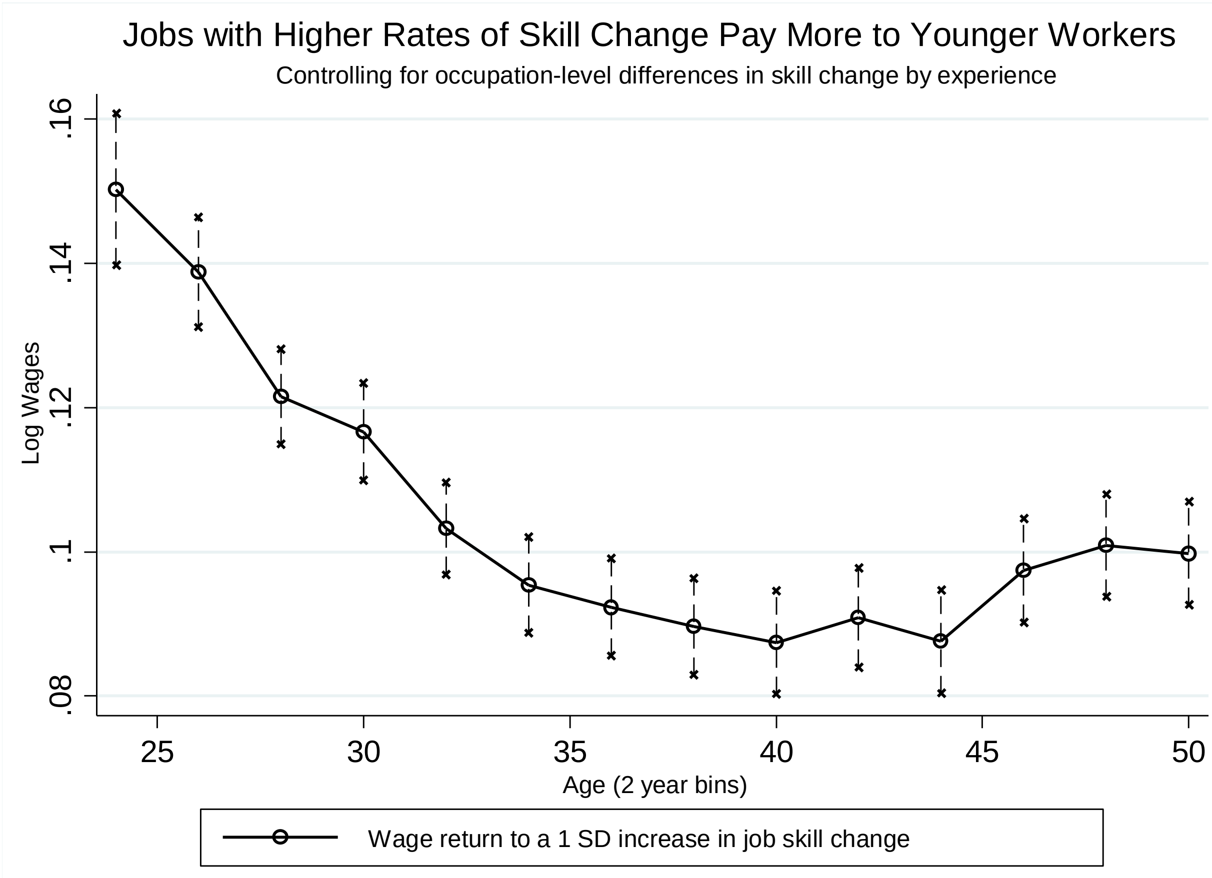
Notes: The figure plots coefficients and 95 percent confidence intervals from an estimate of equation (11) in the paper, a regression of log wages on interactions between two-year age bins and the skill change measure Δ_j that is estimated using 2007-2017 online job vacancy data from Burning Glass Technologies. The sample is restricted to non-STEM occupations only. STEM occupations are defined using the 2010 Census Bureau classification. See the text for details.

Figure A10



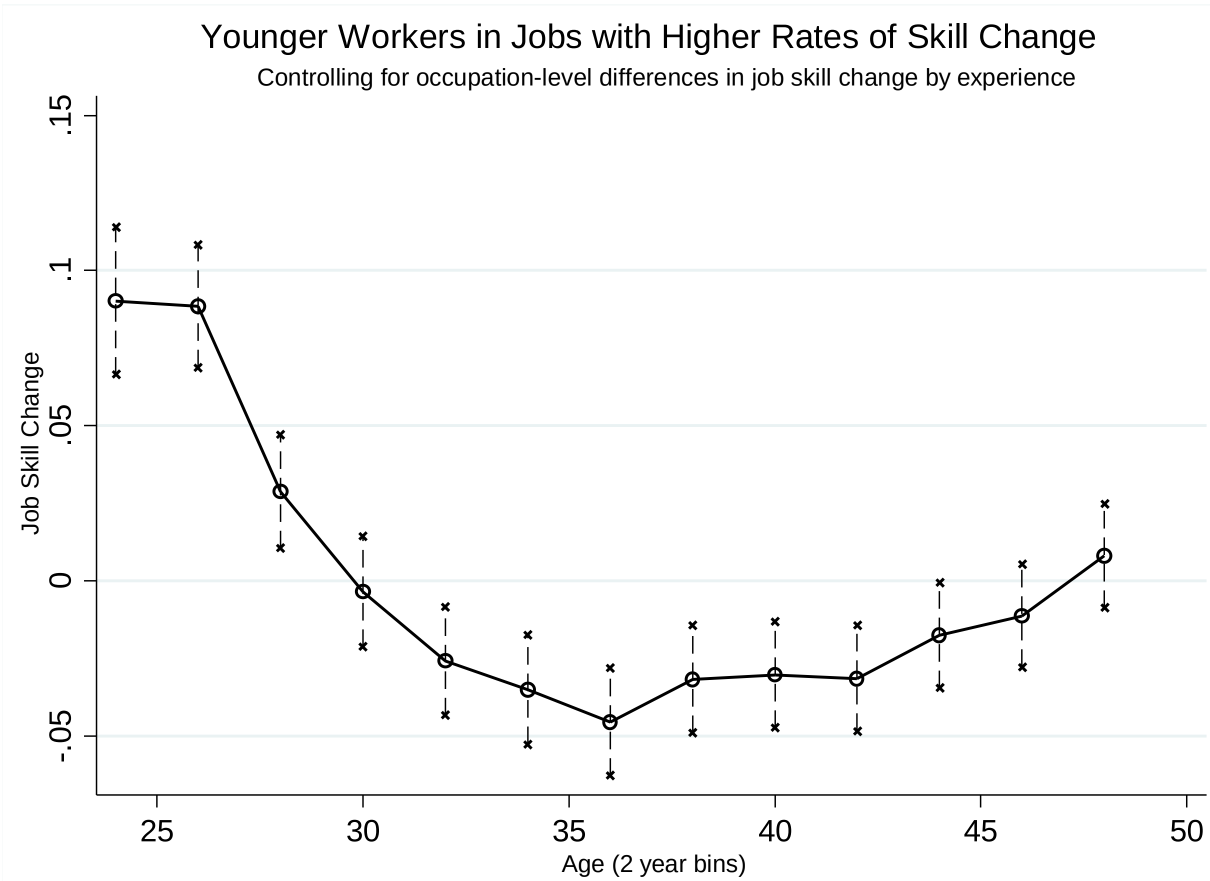
Notes: The figure plots coefficients and 95 percent confidence intervals from an estimate of equation (11) in the paper, a regression of log wages on interactions between two-year age bins and the skill change measure Δ_j that is estimated using 2007-2017 online job vacancy data from Burning Glass Technologies. The sample is restricted to STEM occupations only. STEM occupations are defined using the 2010 Census Bureau classification. See the text for details.

Figure A11



Notes: The figure plots coefficients and 95 percent confidence intervals from an estimate of equation (11) in the paper, a regression of log wages on interactions between two-year age bins and the skill change measure Δ_j that is estimated using 2007-2017 online job vacancy data from Burning Glass Technologies. The regression also controls for an alternative measure of Δ_j that is calculated within-occupation but across experience requirements. STEM occupations are defined using the 2010 Census Bureau classification. See the text for details.

Figure A12



Notes: The figure plots coefficients and 95 percent confidence intervals from an estimate of equation (12) in the paper, a regression of the task change measure Δ_j (which is constructed using 2007-2017 online job vacancy data from Burning Glass Technologies) on occupation by age group interactions. The regression also controls for an alternative measure of Δ_j that is calculated within-occupation but across experience requirements. STEM occupations are defined using the 2010 Census Bureau classification. See the text for details.

Data Appendix to "STEM Careers and the Changing Skill Content of Work"

June 2019

This paper uses data from Burning Glass Technologies (BG), an employment analytics and labor market information firm that scrapes job vacancy data from more than 40,000 online job boards and company websites. BG applies an algorithm to the raw scraped data that removes duplicate postings and parses the data into a number of fields, including job title and six digit Standard Occupational Classification (SOC) code, industry, firm, location, and education and work experience. BG also codes key words and phrases into a large number of unique skill requirements.

The BG database only covers job vacancies that are posted on the Internet. Rothwell (2014) compares the distribution of occupations in an extract of BG data to state vacancy surveys for select metropolitan areas for which data are available. He finds that computer, management, and business occupations are overrepresented relative to the state vacancy surveys, while health care support, transportation, maintenance, sales, and food service workers are underrepresented.

Carnevale, Jayasundera, and Repnikov (2014) show that the occupation-industry composition of the BG data are similar to another database of online job vacancies, the Help-Wanted Online (HWOL) Index collected by the Conference Board. Carnevale, Jayasundera, and Repnikov (2014) also compare a sample of job postings in the BG database to the actual text of the postings and find a high degree of accuracy for verifiable measures such as occupation and education and experience requirements. Additionally, BG has refined its algorithm over time to increase accuracy relative to the early extract studied by Carnevale, Jayasundera, and Repnikov (2014)

Hershbein and Kahn (2018) and Deming and Kahn (2018) provide more detail on the representation of vacancies and occupations in BG data compared to other external sources such as JOLTS, OES and CPS. The bottom line is that while the BG data do have a higher share of technical, STEM jobs than other external sources, this relative representation has not changed over time. Similarly, the BG data underrepresent blue-collar and low-paid service jobs in fields such as food preparation and serving, production, and construction, although this has also not changed very much over time.

One of the most novel features of the BG data is the information available on job skills. BG use a parsing algorithm to identify key words and phrases and code them up as a set of skill requirements. BG regularly update the algorithm to pick up new skills, but then they apply the new algorithm to all years of data retrospectively. More than 93 percent of all job ads have at least one skill requirement, and the average number is 9. There are 13,544 unique skills in our analysis dataset.

We further refine the list of skills by creating a set of common categories that capture major features of the BG data. The table below lists the most common skill strings that we use to create our measures of skills in the paper. For the full list, please see the replication file.

Job Skills	**Keywords and Phrases**
Social	Communication, Collaboration, Negotiation, "Team", Persuasion, Listening, Presentation
Cognitive	Solving, Research, "Analy", Decision, Thinking, Math, "Statistic", Calculation
Character	Organizational Skills, Time Management, Detail-Oriented, Meeting Deadlines, Multi-Tasking, Energetic, Self-Starter, Initiative, Self-Motivation
Creativity	"Creativ"
Writing	Writing, Editing, Preparing Reports, Preparing Proposals
Management	Supervisory, Leadership, Mentoring, Staff Supervision/Development, Performance/Personnel Management
Finance	"Financ", Budgeting, Accounting, Cost
Business Systems	Systems Development/Integration/Architecture, Business Intelligence/Systems/Planning/Strategy, Six Sigma, KPIs
Customer Service	Customer, Sales, Patient, Client
Office Software	Microsoft Word/Excel/Outlook/PowerPoint/Office/Windows, Computer Literacy, Basic Internet Skills
Technical Support	Computer Installation/Repair/Maintenance/Troubleshooting, Web Development/Site Design, Software Installation, Help Desk Support
Data Analysis	Data Analysis/Analytics/Engineering/Modeling/Visualization/Mining/Science, Predictive Analytics/Models, Spreadsheets, Tableau
Specialized Software	Specific software that is tracked by BG and not otherwise categorized (e.g. SQL, Javascript, Adobe)
ML and AI	Artificial Intelligence, Machine Learning, Decision Trees, Apache Hadoop, Python, Bayesian Networks, Automation Tools, Neural Networks, Support Vector Machines (SVM), Decision Trees, Supervised Learning, TensorFlow, MapReduce, Splunk, Convolutional Neural Network (CNN), Cluster Analysis

Model Appendix to "STEM Careers and the Changing Skill Content of Work"

June 2019

1 Wage Growth across Careers as a function of Δ_j

The first prediction of the model is that wage growth over time is lower in careers with higher rates of task change Δ_j. This is equivalent to showing that the derivative of the difference in wages between year t and year zero with respect to Δ ($\frac{\partial(W_{j0}-W_{jt})}{\partial\Delta}$) is positive for $t \geq 1$.

Proposition 1. $\frac{\partial(W_{j0}-W_{jt})}{\partial(\Delta)} \geq 0 \; \forall t \geq 1$

Proof. Rearranging equation (10) - the expression for wages W_{jt} in section 3.4 - and taking the difference between year 0 and year t, we obtain:

$$\begin{aligned} W_{j0} - W_{jt} = FS(1-(1-\Delta)^t)) + a[1-(1-\Delta)^t(t+1) \\ - \sum_{v=1}^{t} \Delta(1-\Delta)^{t-v}(t-v+1). \end{aligned} \tag{1}$$

The derivative of this expression is

$$\frac{\partial(W_{j0}-W_{jt})}{\partial(\Delta)} = tFS(1-\Delta)^{t-1} + at(t+1)(1-\Delta)^{t-1} + \sum_{v=1}^{t}(t-v+1)(1-\Delta)^{t-v-1}[\Delta(t-v+1)-1], \tag{2}$$

$$\frac{\partial(W_{j0}-W_{jt})}{\partial(\Delta)} = tFS(1-\Delta)^{t-1} + at(t+1)(1-\Delta)^{t-1} + \sum_{x=0}^{t-1}(x+1)(1-\Delta)^{x-1}[\Delta(x+1)-1], \tag{3}$$

$$\frac{\partial(W_{j0}-W_{jt})}{\partial(\Delta)} = tFS(1-\Delta)^{t-1} + at(t+1)(1-\Delta)^{t-1} - \left(\frac{(\Delta^2x^2(1-\Delta)^{t-1} + \Delta^2(1-\Delta)^{t-1} + 2\Delta^2x(1-\Delta)^{t-1}}{\Delta^2}\right) - \left(\frac{\Delta(t-1)(1-\Delta)^{t-1} + (1-\Delta)^{t-1}-1)}{\Delta^2}\right). \tag{4}$$

Where the final expression substitutes in the closed form solution for the summation component. We show that this derivative is positive in two steps. First, we demonstrate that it is monotonically increasing in t for all relevant values of the parameters. Then, we show that this derivative is positive for all values of Δ when $t = 1$. Thus we conclude that the derivative is always positive on the parameter space.

To begin, assume that $F = 0$. Because the first term is always positive, increasing F will only tend to make $\frac{\partial(W_{j0}-W_{jt})}{\partial(\Delta)}$ larger. So if the derivative is positive for $F = 0$, it is also positive for $F > 0$. Furthermore, set $a = 1$, the lowest possible value. Taking the derivative of the remaining expression with respect to t yields:

$$\frac{\partial^2(W_{j0}-W_{jt})}{\partial(\Delta)\partial(t-1)} = \frac{((\Delta-1)(1-\Delta)^{(t-1)}((\Delta(t-1)+\Delta+1)log(1-\Delta)+\Delta))}{\Delta^2}$$

Note that the first term in the numerator will be negative, so for this expression to be positive we need

$$(\Delta(t-1)+\Delta+1)log(1-\Delta)+\Delta < 0$$

or

$$t-1 > -\frac{1}{log(1-\Delta)} - \frac{1}{\Delta} - 1.$$

Because the right hand side of this inequality is always negative, we know that $t \geq 1 \implies$

$\frac{\partial^2(W_{j0}-W_{jt})}{\partial(\Delta)\partial(t-1)} > 0$. Thus, it suffices to demonstrate that $\frac{\partial(W_{j0}-W_{jt})}{\partial(\Delta)} > 0$ when $t = 1$. If we evaluate the derivative at $t = 1$ we get $2a - 1$ which is greater than zero for all $a > 1/2$. $\square$

2 Showing Career Choice and Selection on Ability in a Three Period Model

A clear way to see the predictions of this model are with a simple three-period and two-industry example. Let the only industries be STEM ($j = T$) and non-STEM ($j = N$). Assume STEM education is more technical than the non-STEM education. Specifically, assume that $S = 1$ in for STEM and $S = 0$ for non-STEM. Furthermore, assume that the rate of task change is higher in STEM than in non-STEM (i.e. $\Delta_T > \Delta_N$) and the initial productivity in STEM is higher than in non-STEM (i.e. $F_T > F_N$). Finally, assume that the cost function is $C(s, a, u) = s(c - a - u)$ where $c \in \mathbb{R}^+$ is a constant such that, $\forall a$ and u, $(c - a - u) > 0$.

We solve this problem in two parts. First, we find the optimal choice of schooling ($S = 0$ or $S = 1$ in this simplified case) given each chosen career path. Second, we solve for the payoffs of each career choice and characterize the conditions under which workers of a given (a, u) type will make each choice. There are eight possible career paths in the three-period case, so we begin by finding the optimal choice of S for each of them. We will refer to these eight cases in shorthand as $TTT, TTN, TNN, TNT, NTT, NTN, NNT$and NNN, with the order of letters representing the order of periods.

If we restrict industry choice to STEM in all periods, the optimization problem becomes

$$\begin{aligned} &\max_s \,(F_T(1-|S-s|)+a)+\big((1-\Delta_T)(F_T(1-|S-s|)+2a)+\Delta_T a\big) \\ &\quad +(1-\Delta_T)^2[F_T(1-|S-s|)+3a]+\Delta_T(1-\Delta_T)2a+\Delta_T a - s(c-a-u) \qquad (5) \\ &\max_s \,(F_T(s)+a)+\big((1-\Delta_T)(F_T(s)+2a)+\Delta_T a\big) \\ &\quad +(1-\Delta_T)^2[F_T(s)+3a]+\Delta_T(1-\Delta_T)2a+\Delta_T a - s(c-a-u) \qquad (6) \end{aligned}$$

Where equation (6) is equivalent to equation (5) because $S = 1$ and $0 \leq s \leq 1$.[1] Taking the derivative of equation (6) with respect to s and setting it equal to zero yields:

$$\underbrace{F_T + (1-\Delta_T)F_T + (1-\Delta_T)^2 F_T}_{\text{MB of Technical Education}} - \underbrace{(c-a-u)}_{\text{MC of Technical Education}} = 0.$$

[1]We suppress discount factors for simplicity - including them does not change any of the model's qualitative predictions.

This first order condition can be easily separated into two interpretable sections. The first section is the marginal benefit of an additional unit of technical education, the second section is the marginal cost of an additional unit of technical education. Satisfaction of this first order condition implies a solution where an individual is indifferent between all levels of technical education. Because the likelihood that this condition holds exactly is zero, we ignore this case.

Thus, individuals will always select completely technical coursework (STEM) or completely non-technical coursework (non-STEM). We have the following individual demand curve for technical education

$$s^* = \begin{cases} 1 & \text{if } F_T + (1-\Delta_T)F_T + (1-\Delta_T)^2 F_T - (c-a-u) > 0 \\ 0 & \text{if } F_T + (1-\Delta_T)F_T + (1-\Delta_T)^2 F_T - (c-a-u) < 0 \end{cases}$$

where s^* represents the optimal level of technical education. To avoid the trivial case where individuals never choose any technical education, assume that, for every type, the marginal benefit of technical education, conditional on choosing strictly a STEM career, is larger in magnitude than the marginal cost. From these assumptions we derive our first proposition.

Proposition 2. *Individuals that choose to work in STEM for all three periods* (TTT) *will allocate all education time to technical fields* $(s^* = 1)$.

Using the analogous assumptions, we can derive the individual demand curve for technical education if we restrict industry choice to non-STEM in both periods. Specifically, we get

$$s^* = \begin{cases} 1 & \text{if } -F_N - (1-\Delta_N)F_N - (1-\Delta_N)^2 F_N - (c-a-u) > 0 \\ 0 & \text{if } -F_N - (1-\Delta_N)F_N - (1-\Delta_N)^2 F_N - (c-a-u) < 0. \end{cases}$$

As long as wages in non-STEM fields are greater than zero, case one (complete technical education) is never optimal. Thus, we arrive at our second proposition

Proposition 3. *Individuals that choose to work in non-STEM for all three periods* (NNN) *will allocate all education time to non-technical fields* $(s^* = 0)$.

There are six ways individuals can choose to split their career between STEM and non-STEM occupations over their life cycle. As long as $\Delta_T > \Delta_N$ and $F_T > a$, the cases where workers switch from non-STEM to STEM (e.g. TNT, NTT, NTN, NNT) are never optimal. This is because switching into a career with a higher rate of task change will diminish both the value of what was learned in school and the value of future accumulated learning. Furthermore, it can be shown that it never makes sense to switch from STEM to non-STEM in the last period (TTN). This is because there is always an immediate loss from switching associated with losing the value previous learning.[2]

Because of this, the only relevant alternative to consider other than full specialization (TTT or NNN) is when a worker chooses STEM initially but switches to non-STEM in the final two periods (TNN). The maximization problem for the TNN career path is:

$$\begin{aligned} &\underset{s}{max}\ (F_T(1-|S-s|)+a)+(1-\Delta_N)(F_N(1-|S-s|))+a \\ &\qquad +(1-\Delta_N)^2(F_N(1-|S-s|))+(1-\Delta_N)2a+\Delta_N a+-s(c-a-u) \\ &\underset{s}{max}\ (F_T(s)+a)+(1-\Delta_N)(F_N(1-s))+a \\ &\qquad +(1-\Delta_N)^2(F_N(1-s))+(1-\Delta_N)2a+\Delta_N a+-s(c-a-u) \end{aligned}$$

Proposition 4. *Workers on the TNN career path will allocate all education time to technical fields ($s^*=1$).*

Proof. The proof above shows that individuals will always choose either $s^*=0$ or $s^*=1$. Any individual who chooses $s^*=0$ will earn more by choosing NNN rather than working in STEM in the first period (TNN), because $F_N(1)+a > F_t(0)+a = a$ and $(1-\Delta_N)(F_N(1)+2a)+\Delta_N a+(1-\Delta_N)^2[F_N(1)+3a]+\Delta_N(1-\Delta_N)2a+\Delta_N a > (1-\Delta_N)(F_N(1))+a+(1-\Delta_N)^2(F_N(1)+(1-\Delta_N)2a+\Delta_N a$. Thus, any worker who chooses the TNN career path must have chosen $s^*=1$. □

Now that we have determined the optimal education choice conditional on each career path, we can move to the second part of the problem: given the total utility of each option, which individuals should select which career path? This is equivalent to a discrete choice problem where each relevant career path is the choice variable. Mathematically, individuals solve the following maximization problem

$$Max(W^{TTT}-C(1,u,a),W^{NNN}-C(0,u,a),W^{TNN}-C(1,u,a))$$

[2] We exclude these proofs but can reproduce them upon request.

where

$$\begin{aligned}W^{TTT} - C(1, u, a) =& F_T + a + (1 - \Delta_T)(F_T + 2a) + \Delta_T a \\ &+ (1 - \Delta_T)^2(F_T + 3a) + \Delta_T(1 - \Delta_T)2a + \Delta_t a - (c - a - u) \quad (7)\\ W^{NNN} - C(0, u, a) =& F_N + a + (1 - \Delta_N)(F_N + 2a) + \Delta_N a \\ &+ (1 - \Delta_N)^2(F_N + 3a) + \Delta_N(1 - \Delta_N)2a + \Delta_N a \quad (8)\\ W^{TNN} - C(1, u, a) =& F_T + a + (1 - \Delta_N)(a) + \Delta_N a \\ &+ (1 - \Delta_N)^2(2a) + \Delta_N(1 - \Delta_N)2a + \Delta_N a - (c - a - u). \quad (9)\end{aligned}$$

To figure out exactly which types of individuals will select into each career track, we can compare the payoffs for each track and derive indifference conditions that allow us to split the type space by preference orderings over the three options. First, we compare the TNN career path to the TTT career path.

To figure out exactly who will choose the TNN career path, we derive the conditions under which an individual would choose TNN over TTT. The gains from switching can be obtained by subtracting the payoff for TNN (equation 9) from the payoff for TTT (equation 7):

$$\begin{aligned}W^{TNN} - W^{TTT} \equiv& (1 - \Delta_N)a + \Delta_N a + (1 - \Delta_N)^2 2a + (1 - \Delta_N)\Delta_N(2a) \\ &- (1 - \Delta_T)(F_T + 2a) - \Delta_T a - (1 - \Delta_T)^2(F_T + 3a) - \Delta_T(1 - \Delta_T)2a - \Delta_T a\end{aligned} \quad (10)$$

Workers will major in STEM, work in STEM initially and then switch in the second period (e.g. choose $s^* = 1$ and then the TNNcareer path) when the expression above is greater than zero. This results in a clear prediction for ability sorting:

Proposition 5. *For every $\Delta_N \in [0, 1]$, there exists a minimum difference in rates of change across sectors $\Delta_T - \Delta_N > k_{\Delta_N}$ such that workers who choose TNN will have higher ability that workers who choose TTT. Furthermore, $\Delta_T - \Delta_N > k_{\Delta_N}$ is a necessary condition for $W^{TNN} - W^{TTT} > 0$.*

Proof. Differentiating the expression above with respect to ability yields:

$$\frac{\partial \left(W^{TNN} - W^{TTT}\right)}{\partial a} = (3 - \Delta_N) - (5 - 4\Delta_T + \Delta_T^2). \quad (11)$$

This is greater than zero if $\Delta_T > 2 - \sqrt{2 - \Delta_N}$. Thus, as long as Δ_T is sufficiently large relative to Δ_N, workers who choose TNNwill have higher ability than workers who choose

TTT.[3] For $\Delta_T, \Delta_N \in [0,1]$, $k_{\Delta_N} \in (0, .59]$ and is monotonically decreasing in Δ_N.

To prove the second part of proposition 5, we need to show that $W^{TNN} - W^{TTT} > 0$ only if $\Delta_T - \Delta_N > k_{\Delta_N}$. We can use equation (10) to represent the the switching returns inequality as follows.

$$W^{TNN} - W^{TTT} = (3 - \Delta_N)a - (5 - 4\Delta_T + \Delta_T^2)a - \Delta_T(1 - \Delta_T)F_T > 0. \tag{12}$$

If the inequality in equation (11) does not hold, then $(3-\Delta_N)a-(5-4\Delta_T+\Delta_T^2)a$ is negative, and because $\Delta_T(1-\Delta_T)F_T$ is positive, it is impossible for the inequality in equation (12) to hold. Thus if the gap between Δ_T and Δ_N is not large enough, workers simply won't switch at all and the only career paths will be TTT and NNN. This means that any career switching we observe from T to N over time will always be among higher-ability workers. □

We can also solve for the ability threshold $a_{TTT,TNN}$ at which workers would choose TNN over TTT by setting the expressions for the gains from switching equal to zero. With some simplification, we obtain the indifference condition:

$$a_{TTT,TNN} = \frac{F_T(2 - 3\Delta_T + \Delta_T^2)}{-2 - \Delta_N + 4\Delta_T - \Delta_T^2}.$$

By a similar logic, we can generate indifference conditions for other career trajectories assuming that workers make optimal education choices. Specifically, we solve for u as a function of a, which yields indifference curves along the (a, u) type space:

$$u_{TTT,NNN}(a) = \underbrace{Y_N(3 - \Delta_N + \Delta_N^2) - Y_T(3 - \Delta_T + \Delta_T^2) + C}_{k_1} + a\underbrace{[4(\Delta_S - \Delta_N) - 1 - (\Delta_T^2 - \Delta_N^2)]}_{k_2} \tag{13}$$

$$u_{NNN,TNN}(a) = \underbrace{Y_N(3 - 3\Delta_N + \Delta_N^2) - Y_T + C}_{k_3} + a\underbrace{[-1 - 5\Delta_N + \Delta_N^2]}_{k_4} \tag{14}$$

where $u_{TTT,NNN}(a)$ defines the indifference threshold between TTT and NNN - a fully STEM vs. non-STEM career - and $u_{TTT,TNN}(a)$ defines the indifference threshold between TTT and TNN. To reduce notational clutter, we abbreviate the subscripts for each indiference condition such that $u_{TTT,NNN}(a) = u_{T,N}(a)$, $u_{NNN,TNN}(a) = u_{N,TNN}(a)$, and $a_{TTT,TNN} = a_{switch}$.

[3] We have also analyzed this result for models with four or more periods. While the exact values that satisfy the $\Delta_T >> \Delta_N$ condition vary somewhat depending on the number of periods, the basic structure remains unchanged.

These indifference curves allow us to divide the two-dimensional type-space into six sections that represent all of the possible preference combinations over TTT,NNN and TNN.

Figure M.A1 shows a visual example of the optimal career choice for individuals of each (a, u) type, under the assumption that $\Delta_T - \Delta_N > k_{\Delta_N}$, so there will be career switching.

Sections 1 and 2 represent types who choose NNN, sections 3 and 4 represent types who choose TTT, and sections 5 and 6 represent types who choose TNN (we call these types "switchers" as shorthand).

Since we have assumed u and a are joint uniformly distributed, we can solve analytically for the average ability in each career path by using the indifference conditions as a weighting function:

$$\bar{a}_{TTT} = \frac{\int_1^{a_{T,TNN}} (u_{max} - u_{T,N}(a)) * a \; da}{\int_1^{a_{T,TNN}} (u_{max} - u_{T,N}(a)) \; da} \tag{15}$$

$$\bar{a}_{NNN} = \frac{\int_1^{a_{T,TNN}} u_{T,N}(a) * a \; da + \int_{a_{T,TNN}}^{u_{T,N}^{-1}(0)} u_{N,TNN}(a) * a \; da}{\int_{1,}^{a_{T,TNN}} u_{TN}(a) \; da + \int_{a_{T,TNN}}^{u^{-1}(0)} u_{N,TNN}(a) \; da} \tag{16}$$

$$\bar{a}_{TNN} = \frac{\int_{a_{T,TNN}}^{u_{T,N}^{-1}(0)} (u_{max} - u_{N,TNN}(a)) * a \; da + \int_{u_{T,N}^{-1}(0)}^{a_{max}} u_{max} * a \; da}{\int_{a_{T,TNN}}^{u_{T,N}^{-1}(0)} (u_{max} - u_{N,TNN}(a)) \; da + \int_{u_{T,N}^{-1}(0)}^{a_{max}} u_{max} \; da} \tag{17}$$

where $\bar{a}_{TTT}$ is the average ability of those who select into a STEM career, $\bar{a}_{NNN}$ is the average ability for those who select into a non-STEM career, and $\bar{a}_{TNN}$ is the average ability of those who select into a STEM occupation the first period and switch into non-STEM afterward.

As long as $\Delta_T - \Delta_N$ is large enough to cause ability selection out of STEM, then $\bar{a}_{TNN} > \bar{a}_{TTT}$ because $min(a_{TNN}) > max(a_{TTT})$. Furthermore, as long as $u_{N,TNN}(a)$ slopes downward, $\bar{a}_W > \bar{a}_N$. This is because if $u_{N,TNN}(a)$ slopes downward, the weighting function in equation 17 is monotonically increasing in a over $[a_{T,TNN}, a_{max}]$, meaning that $\bar{a}_{TNN} > \frac{a_{max}+a_{T,TNN}}{2}$. Furthermore, if we overestimate $\bar{a}_{NNN}$ by limiting our calculation those whose abilities are greater than $a_{T,TNN}$, we know that this overestimate is less than $\frac{a_{max}+a_{T,TNN}}{2}$ because the weighting function $u_{N,TNN}(a)$ is monotonically decreasing in a.

Thus $\bar{a}_{NNN} < \frac{a_{max}+a_{T,TNN}}{2} < \bar{a}_{TNN} \implies \bar{a}_{NNN} < \bar{a}_{TNN}$. Also, in this case, we can verify from equation 14 that $u_{N,TNN}(a)$ always slopes downward because $k_4 < 0$ for $\Delta_N \in [0, 1]$. Thus, under fairly general conditions, non-STEM occupations will attract higher ability workers over time.