New Technologies and the Labor Market

Enghin Atalay [1] Phai Phongthiengtham[2]
Sebastian Sotelo[3] Daniel Tannenbaum[4]

[1,2]University of Wisconsin-Madison, [3]University of Michigan-Ann Arbor,

[4]University of Nebraska-Lincoln

January, 2019

Motivation

- Over the last half century:
  - Dramatic decline in the price of computing power (Nordhaus, 2007).
  - Increased usage of information and communication technologies (ICT) on the job.
- Skill-biased technological change
  - Better, cheaper computers $\rightarrow$ Reduction the demand for routine tasks

Research questions

1. How do individual pieces of ICTs reshape the types of tasks which workers perform on the job?
2. What are the macroeconomic implications of ICTs:
    - on the allocation of workers across occupations, and
    - on the skill premium (through their effect on changes in the relative value of workers output)?

What we do

1. Construct a new data set measuring jobs task content and technology adoption in the U.S.
    - Using the text of 4.2 million job ads posted between 1960 and 2000, we measure the adoption of 48 unique technologies. (e.g., Microsoft Excel, Word, PowerPoint, WordPerfect, Fortran, Unix)
2. Develop and estimate a general equilibrium model of workers technology adoption decisions and occupational choice
    - Examine counterfactual equilbria without the introduction of ICTs

Main results

1. Relationships between tasks and technologies, at the ad level:
    - For most technologies: new technologies are associated with increased mentions of nonroutine analytic tasks, fewer mentions of other groups of tasks.
    - *Exceptions*: Microsoft Office software is associated with nonroutine interactive tasks, networking ICTs with routine cognitive tasks.
2. Without the introduction of ICTs, the College-HS skill premium would have been lower by 17 percent.
    - Workers with college education have a comparative advantage in producing nonroutine analytic tasks.
    - From (1): New technologies increase the labor demand in nonroutine analytic tasks.

Literature

1. New technologies, skill prices, and the wage distribution
   - Krusell, Ohanian, Rios-Rull, Violante (2000); Autor, Levy, Murnane (2003); Autor and Dorn (2013); Acemoglu and Autor (2011); Michaels, Natraj, Van Reenen (2014); Burstein, Morales, Vogel (2015)
2. Impact of the adoption of new technologies on...
   - Firm organization: Levy and Murnane (1996); Bartel, Ichniowski, Shaw (2007)
   - Individual workers wages: Entorf and Kramarz (1998); Autor, Katz, Krueger(1998)
3. Using vacancy posting text to learn about the labor market
   - Deming and Kahn (2016), Hershbein and Kahn (2016), Marinescu and Wolthoff (2016), Atalay, Phongthiengtham, Tannenbuam, Sotelo (2017)

Roadmap

1. Introduction
2. A new data set measuring ICT adoption
3. Task and technology complementarity
4. The macroeconomic implications of ICTs
5. Conclusion

Processing newspaper text files

- ProQuest has produced the raw data using an algorithm that converts images of scanned newspaper pages into text files.
  - Display Ads and Classified Ads from New York Times (1960-2000), Wall Street Journal (1960-1998) and Boston Globe (1960-1983).
- We proceed to construct the dataset as follows:
  1. Classifying vacancy postings from other advertisements
  2. Detecting the boundaries between vacancy postings
  3. Extracting task and ICT usage information
  4. Classifying occupation code

Example of raw text

SOFiWARE ENGINEERS - Modal Software Develop air-to-surface modal software, including design, code, unit test, integration and test, and documentation. Requires 5+ years software engineering experience with a BSEE/CS or Computer Engineering. Software development for real-time, multi-tasking/multi-processor, embedded systems experience a must. 3+ years C programming experience in a Unix environment and familiarity with modern software design methodologies essential. Knowledge of radar design principles a plus. Joint STARS The premiere ground surveillance system far the U.S. and allied forces. The DoD has authorized the full production of Joint STARS. In addition, significant activity on Joint STARS upgrades is underway. SENIOR SYSTEMS ENGINEERS Design and develop advanced, high-resolution radar imaging systems, including ultra-high resolution SAR and Moving Target Imaging Systems in real-time or near real-time environments. Represent the engineering organization ta senior technical management, potential partners and customers in industry and government; plan/coordinate R&D program activities; lead a team of hardware/soare/systems engineers; develop and test complex signal processing modes and algorithms in a workstation environment; support development with analyses, reports, documentation and technical guidance. Requires an MS or PhD in Engineering, Physics or Mathematics with experience in specification, Imaging anss and testing of Advanced Coherent Radar High-Resolution Must have strong math, physics and signal processing skills, C/C++ and ,AN programming expertise, plus familiarity with workstations and analytical tools such as The following require knowledge oF emulators, debuggers, and logic ana/. Knowledge of Ada, Unix, VxWorks, DigitalAlpha Processor and assembly language desirable. Radar systems experience plus. SOFTWARE ENGINEERS Define requirements and develop software far RCU or Intel microprocessor-based RSEs. Help define software requirements far LRU ECPs and the Contractor Logistics software program, including design, code, integration and test, and documentation. BSCS/EE preferred with 3-5 years real-time software development experience using Ada and/or FORTRAN programming languages. U IS- * SOFiWARE

Classifying vacancy postings from other advertisements

- We use *Latent Dirichlet Allocation*, an unsupervised machine learning model, to cluster each newspaper page into topics.
- Before we estimate this LDA model, we:
  1. remove stop words (e.g., common words like "a", "the", "and"), numerals, and words that are not contained in the English dictionary.
  2. stem words such that words in different forms: singular nouns, plural nouns, verbs, adjectives, and adverbs, are grouped as one.

Table 1: Predictive Word Stems from New York Times Classified Ads

1	resum seek call must work exp excel new salari send
2	new home owner acr call car hous den area ask
3	build ave new park call studio east avail fee fir

Detecting the boundaries between vacancy postings

Certain phrases at the beginning or end of individual help wanted ads allow us to identify the boundaries between ads

- **Addresses**: Most job vacancy posts have the employers addresses at the end.
- **Ending phrases**: Some phrases indicate that a job vacancy post is ending, for example: send [...] resume, submit [...] resume, in confidence to, affirmative action employer,equal opportunity; or equal opportunities.
- **Job titles**: A job vacancy post usually starts with a job title, which stands alone within a few lines and is uppercase.

List of task catagories (Spitz-Oener, 2006)

nonroutine analytic	analyze, analyzing, design, designing, devising rule, evaluate, evaluating, interpreting rule, plan, planning, research, researching, sketch, sketching
nonroutine interactive	advertise, advertising, advise, advising, buying, coordinate, coordinating, entertain, entertaining, lobby, lobbying, managing, negotiate, negotiating, organize, organizing, presentation, presentations, presenting, purchase, sell, selling, teaching
nonroutine manual	accommodate, accommodating, accommodation, renovate, renovating, repair, repairing, restore, restoring, service, serving
routine cognitive	bookkeeping, calculate, calculating, correcting, corrections, measurement, measuring
routine manual	control, controlling, equip, equipment, equipping, operate, operating

[1] We include, for each of these words, synonyms based on machine-learning semantic similarity.

List of technologies

Database Management Systems	CICS, DB2, DOS, EDP, FoxPro IMS (Information Managment System), MVS, Oracle, Sybase, UNIX, VAX, VMS
Hardware and Mainframe Software	BAL, IBM 360, IBM 370, IBM 5520 IBM RPG, JCL, UNIVAC
Networking Hardware and Software	LAN, Novell, TCP, TSO (TCP segmentation offloading)
Microsoft Office and Competitors	MS Excel, MS PowerPoint, MS Word, WordPerfect, Lotus 123, Lotus Notes
General Purpose Software / Others	APL, CAD, COBOL, C++, Fortran, HTML, Java, Pascal, Quark, SQL, Visual Basic

Example of processed text

engineers|- modal software develop air-to-surface modal software , including design , code , unit test , integration and test , and documentation . requires 5+ years software engineering experience with a b see cs or computer engineering . software development for real-time , multitasking multiprocessor , embedded systems experience a must . 3+ years c programming experience in a UNIX environment and familiarity with modern software design methodologies essential . knowledge of radar design principles a plus . joint stars the premiere ground surveillance system far the u . s . and allied forces . the DOD has authorized the full production of joint stars . in addition , significant activity on joint stars upgrades is underway .

senior system engineer| design and develop advanced , high-resolution radar imaging systems , including ultra-high resolution sear and moving target imaging systems in real-time or near real-time environments . represent the engineering organization ta senior technical management , potential partners and customers in industry and government ; plan coordinater ; d program activities ; lead a team of hardware soared systems engineers ; develop and test complex signal processing modes and algorithms in a workstation environment ; support development with analysis , reports , documentation and technical guidance . requires an ms or PhD in engineering , physics or mathematics with experience in specification , imaging ans and testing of advanced coherent radar high-resolution must have strong math , physics and signal processing skills , c c and , an programming expertise , plus familiarity with workstations and analytical tools such as the following require knowledge of emulators , debuggers , and logic Ana . knowledge of Ada , UNIX , vxworks , digital alpha processor and assembly language desirable . radar systems experience plus.

software engineers|define requirements and develop software far r cu or Intel microprocessor-based rs es . help define software requirements far lr u e cps and the contractor logistics software program , including design , code , integration and test , and documentation . bscs ee preferred with 3-5 years real-time software development experience using Ada and or FORTRAN programming languages . u is- software

Classifying vacancy's occupation code

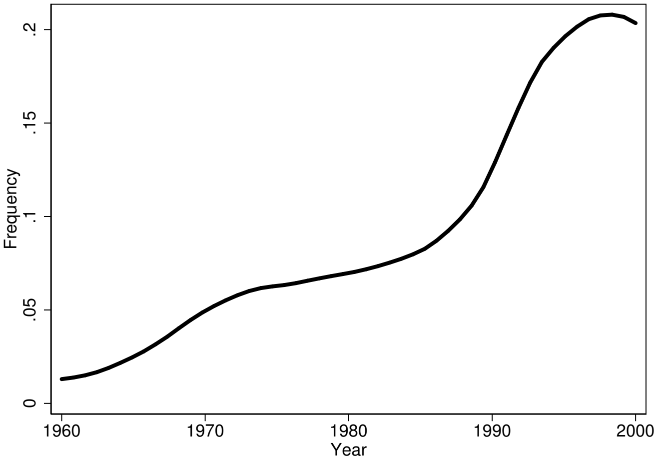
- We use the list of "Sample of Reported Titles" provided by the U.S. Department of Labor (O*NET database).
- We use *Continuous Bag of Words*, an unsupervised machine learning model, to vectorize each word.
  - We match each job title in the ad to the closest job title from the list of and assign its corresponding occupation code.
  - We are able to match, for example, *"customer relation specialist"* to an O*NET occupation code: 43-4051.00 (Customer Service Representatives).
- We then merge our dataset with the U.S. decennial census.

Word vector illustration

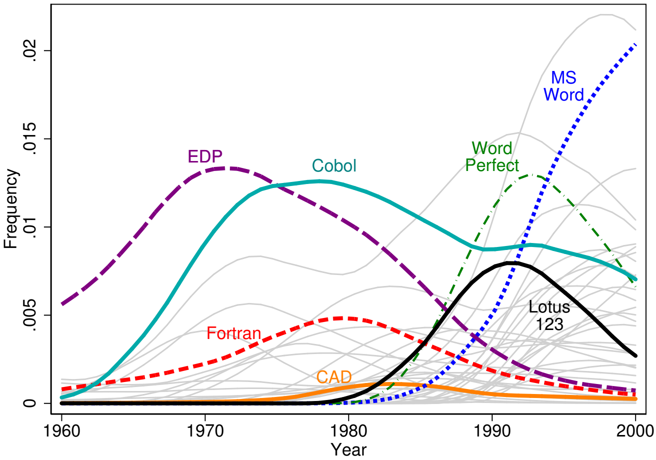
$$W = \begin{array}{c} \\ \text{cat} \\ \text{dog} \\ \text{python} \\ \text{stata} \\ \\ \end{array} \begin{array}{c} \begin{array}{ccc} \text{walks} & \text{eats} & \text{programmer} \end{array} \\ \begin{bmatrix} 0.012 & 0.011 & 0 & \dots \\ 0.014 & 0.008 & 0 & \\ 0 & 0.004 & 0.02 & \\ 0 & 0 & 0.014 & \\ \vdots & & & \ddots \end{bmatrix}_{VxV} \end{array}$$

- There are V words in the corpus.
- Word embedding matrix, W, represents actual text usage.
- For example: $W_{12} = 0.011 = Prob(w_{t-1} = \text{cat} | w_t = \text{eats})$
- Cosine similarity: $cos(\theta) = \frac{v_1 \cdot v_2}{|v_1||v_2|}$
- The word *"dog"* is closer to *"cat"* than *"stata"*.

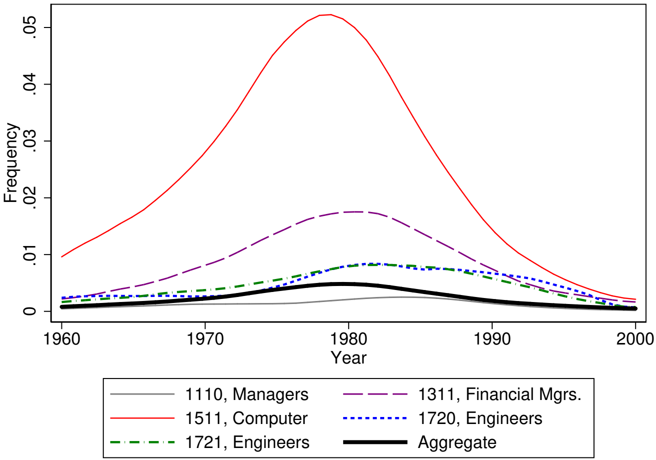
Overall mentions of technologies



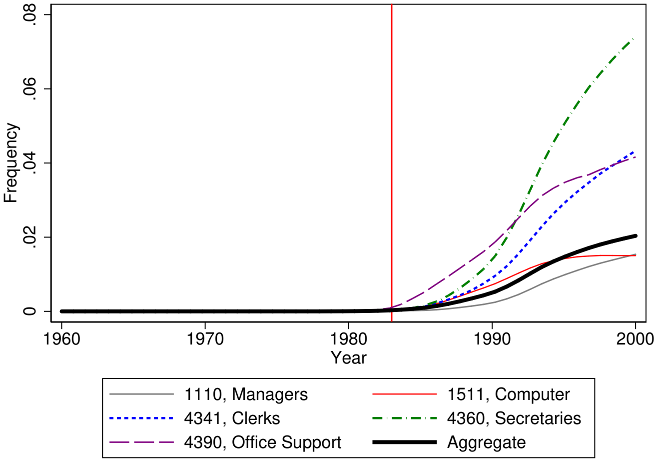
Individual mentions of technologies



Mentions of FORTRAN in the top 5 occupations



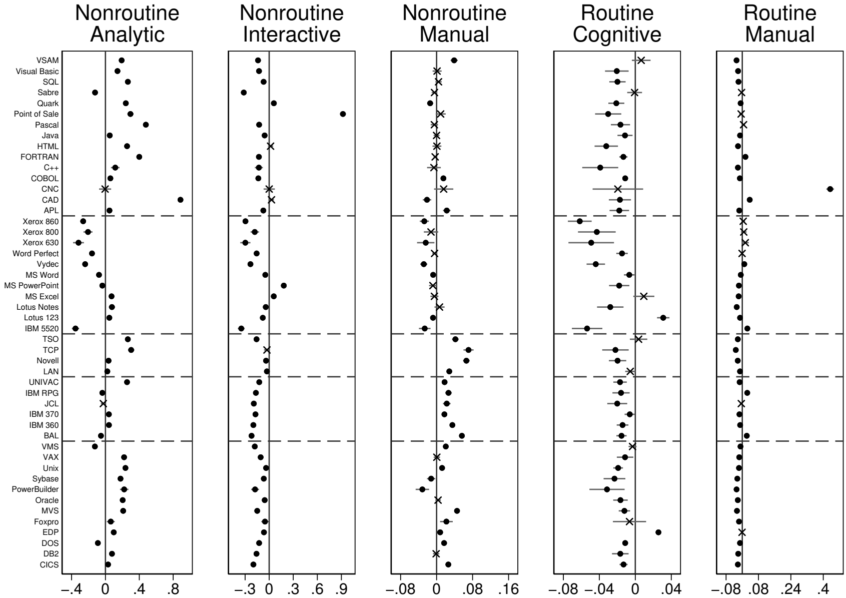
Mentions of MS Word in the top 5 occupations



Task and technology complementarity

$$\text{task}^h_{ajt} = \beta_{hk} \cdot \text{tech}_{ajkt} + f_h(\text{words}_{ajt}) + \iota_{jh} + \iota_{th} + \epsilon_{ahjkt} \qquad (1)$$

- task^h_{ajt}: mentions of task h $\in$ nonroutine (analytic, interactive, manual) and routine (cognitive and manual) tasks in ad a
- tech_{ajkt}: mentions of technology k in ad a
- ι_{jh} and ι_{th}: occupation (j) and year (t) fixed effects by task h
- $f_h(\text{words}_{ajt})$: quartic polynomial, controlling for differences in ad length



Nonroutine Analytic
Nonroutine Interactive
Nonroutine Manual
Routine Cognitive
Routine Manual
VSAM
Visual Basic
SQL
Sabre
Quark
Point of Sale
Pascal
Java
HTML
FORTRAN
C++
COBOL
CNC
CAD
APL
Xerox 860
Xerox 800
Xerox 630
Word Perfect
Vydec
MS Word
MS PowerPoint
MS Excel
Lotus Notes
Lotus 123
IBM 5520
TSO
TCP
Novell
LAN
UNIVAC
IBM RPG
JCL
IBM 370
IBM 360
BAL
VMS
VAX
Unix
Sybase
PowerBuilder
Oracle
MVS
Foxpro
EDP
DOS
DB2
CICS
–.4 0 .4 .8
–.3 0 .3 .6 .9
–.08 0 .08 .16
–.08 –.04 0 .04
–.08 .08 .24 .4

Relationship between tasks and ICTs

- Additional technology-related mention:
  - $\uparrow$ 0.09 std nonroutine analytic task mentions per job ad
  - $\downarrow$ 0.18 std nonroutine interactive tasks
  - $\downarrow$ 0.07 std routine cognitive tasks
  - $\downarrow$ 0.07 std routine manual tasks
  - No clear pattern for nonroutine manual tasks
- *Exception*: Office software (MS Suite) $\uparrow$ interactive $\downarrow$ analytic
- Confirm the near univerally accepted hypothesis that technology is complementary with human capital (as measured by cognitive or analytic skill).
  - There is, however, a substantial heterogeneity.

The macroeconomic implications of ICTs

- How do ICTs change the task content of occupations and the demand for workers?
- Key features of the model:
  1. There are many worker groups: by gender, education and experience. Each group has a different set of skill endowment.
  2. Workers jointly choose their occupation, whether to adopt one of the ICTs and time allocation to each of the five task catagories.
  3. The arrival of ICTs is depicted as an exogenous decrease in the price.
- These technologies are different from the standard TFP framework.

Quantitative GE model - Environment

- Time: t
- Task space: $h \in \{1, \dots, H\}$
- Worker groups: $g \in \{1, \dots, G\}$
  - Every worker is endowed with one unit of labor.
  - Each worker in group g is endowed with skill S_{gh}.
  - Exogenous mass L_{gt}
- ICT: $k \in \{0, 1, \dots, K\}$
  - $k = 0$ corresponds to no ICT adoption.
  - Price of ICT, c_{kt}, is exogenous.
- Occupations: $j \in \{1, \dots, J\}$
- A representative firm

Quantitative GE model - Preferences

The representative consumer has constant elasticity of substitution preferences across outputs of each of the J occupations:

$$U = \left(\sum_{j=1}^{J} a_j^{1/\rho} Y_j^{\frac{\rho-1}{\rho}} \right)^{\frac{\rho}{\rho-1}}$$

- Y_j : the sum of the production of individual workers who work in occupation j
- ρ : the elasticity of substitution
- Prices of occupational output $\{p_j\}$ are determined in equilibrium.

Quantitative GE model - Technology

Output of a worker from group g, if working in occupation j and using κ units of technology k:

$$\tilde{V}_{gjk}(\epsilon) = \epsilon^{\bar{\alpha}_k} \cdot \left(\prod_{h=1}^{H} \left[\frac{q_{hgjk}(\epsilon)}{\alpha_{hjk}} \right]^{\alpha_{hjk}} \right) \cdot \left(\frac{\kappa_{gjk}}{1-\bar{\alpha}_k} \right)^{1-\bar{\alpha}_k}$$

- ϵ_{igjt}: idiosyncratic efficiency shock – assume Frchet distribution $\Pr[\epsilon < x] = \exp\left(-x^{-\theta}\right)$
- q_{hgjk}: units of task h produced by the worker:

$$q_{hgjk} = S_{gh} \cdot \ell_{hgjk}$$

- α_{hjk}: the importance of task h in occupation-ICT pair (j, k)
- κ_{gjk}: units of ICT k used un production ($k = 0$: no ICT)
- Each unit of ICT k costs c_k units of the final good to produce.

Quantitative GE model - Equilibrium

- Each worker jointly chooses: occupation j, time allocation to each task, ℓ_{hgjk}, and ICT adoption, κ_{gjk} - to maximize her wage (paid competitively).
- Workers payment (per efficiency unit of labor) in occupation j:

$$w_{gjk} = p_j^{\frac{1}{\bar{\alpha}_k}} (c_k)^{-\frac{1-\bar{\alpha}_k}{\bar{\alpha}_k}} \prod_{h=1}^{H} S_{gh}^{\frac{\alpha_{hjk}}{\bar{\alpha}_k}}$$

- Given the price of ICTs $\{c_k\}$, an equilibrium is given by prices of occupational output $\{p_j\}$ and ICT uses κ_{gjk} such that:
    1. occupational-output markets clear
    2. ICT markets clear

Quantitative GE model - Equilibrium

- The fraction of workers in group g that sorts into occupation j and technology k:

$$\lambda_{gjk} = \frac{w_{gjk}^{\theta}}{\sum_{k'} \sum_{j'} w_{gj'k'}^{\theta}} \tag{2}$$

- Total payments to workers:

$$\bar{W}_g = \Gamma\left(1 - 1/\theta\right) \cdot \left(\sum_{j} \sum_{k} w_{gjk}^{\theta}\right)^{1/\theta} \tag{3}$$

- where $\Gamma(\cdot)$ is the Gamma function.

Quantitative GE model - Estimation

- Parameterize skills:

$$\log S_{gh} = a_{h,gender} \cdot D_{gender,g} + a_{h,edu} \cdot D_{edu,g} + a_{h,exp} \cdot D_{exp,g} \quad (4)$$

- Calibrate $\theta = \rho = 1.78$ (Burstein et al., 2015)
- Method of moments estimation:

Share of employment: $$\tilde{\lambda}_{gj} = \sum_{k=1}^{K} \left[\frac{w_{gjk}^{\theta}}{\sum_j w_{gjk'}^{\theta}} \right]$$

ICT adoption by occupation: $$\tilde{\pi}_{jk} = \sum_{g} \frac{\lambda_{gjk} \tilde{L}_{gj}}{\sum_{g'} \tilde{L}_{g'j}}$$

Average earnings: $$\tilde{\bar{W}}_g = \Gamma\left(1 - 1/\theta\right) \cdot \left(\sum_j \sum_k w_{gjk}^{\theta} \right)^{1/\theta}$$

- Parameters: a_h, p_j and c_k.

Estimates of skills

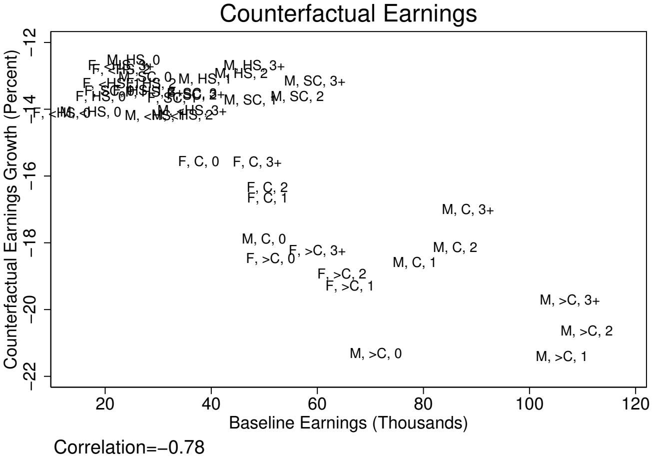
	Nonroutine Analytic	Nonroutine Interactive	Routine Cognitive	Routine Manual	Nonroutine Manual
Gender					
Male			omitted		
Female	-0.520	-0.241	2.168	-5.964	-1.434
Education					
$<$ HS	-2.351	0.015	1.413	3.189	1.534
High School	-1.210	-0.051	-0.376	2.231	1.098
Some College			omitted		
College	1.878	0.515	-0.589	-9.018	-2.444
Post-Graduate	2.466	0.698	-1.915	-20.127	0.588
Experience					
0-9 Years	-0.941	-0.232	-0.229	-1.469	-0.451
10-19 Years	-0.190	-0.102	0.005	-0.410	0.076
20-29 Years			omitted		
30+ Years	-0.135	0.135	0.300	-0.380	0.080

Estimates of skills

	Nonroutine Analytic	Nonroutine Interactive	Routine Cognitive	Routine Manual	Nonroutine Manual
Gender					
Male			omitted		
Female	-0.520	-0.241	**2.168**	-5.964	-1.434
Education					
$<$ HS	**-2.351**	0.015	1.413	3.189	1.534
High School	**-1.210**	-0.051	-0.376	2.231	1.098
Some College			omitted		
College	**1.878**	0.515	-0.589	-9.018	-2.444
Post-Graduate	**2.466**	0.698	-1.915	-20.127	0.588
Experience					
0-9 Years	-0.941	-0.232	-0.229	-1.469	-0.451
10-19 Years	-0.190	-0.102	0.005	-0.410	0.076
20-29 Years			omitted		
30+ Years	-0.135	0.135	0.300	-0.380	0.080

Quantitative GE model - Counterfactual

- What would happen if all ICTs were exogenously unavailable?
  - Begin with equilibrium in 2000, re-compute equilibrium with extremely high c_k
- Reduce earnings across the board (15% on average)
- Reduce inequality: college premium $\downarrow$ by 17%



The first character - "M" or "F" describes the gender - "<HS," "HS," "SC," "C," or ">C" describes the educational attainment; and - "0" for 0-9, "1" for 10-19, "2" for 20-29, "3+" for $\geq$ 30 - describes the number of years of potential experience. The correlation is weighted by the number of people in each demographic group.

Conclusion

This paper:

- New measurements of occupations adoption of a wide variety of ICTs.
- Job ads which mention ICTs are relatively intensive in nonroutine analytic tasks. There is, however, substantial heterogeneity.
- Overall, ICTs increase output but also inequality.

Looking ahead...

- More recent technologies (i.e., Artificial Intelligence)
- The role of (different types of) capital, especially when new technologies are more likely to be skill-biased.