Enghin Atalay, Phai Phongthiengtham, Sebastian Sotelo, Daniel Tannenbaum (2018) "The Evolving U.S. Occupational Structure", *Working Paper*

Presented by Bao Hoang Nguyen

ECON8800 Presentation, 2019

Motivation

- The considerable increase in labour income inequality in the U.S. between 1960-2000.
- The dramatic changes in the composition of U.S. employment: the employment share of occupations intensive in routine tasks has shrunk, while the share of occupations emphasizing nonroutine tasks, social and cognitive skills, has grown.
- How about the content of occupations themselves?
- If there have been job task shifts within occupations, what are the implications for the distribution of earning?

$\Rightarrow$ The current paper aims at answering the above questions by utilizing a newly-created data set of occupational content from 1960 to 2000 using text from newspaper job ads.

New Data Set of Occupational Characteristics

- Processing the Newspaper Text Files
  - The primary data set is raw text files, purchased from ProQuest and originally published in the *News York Times* (1940 - 2000), *Wall Street Journal*(1940-1998), *Boston Globe* (1960-1983).
  - Subsets of text from ads can be isolated but:
    - Cannot distinguish between job ads and other types of ads (Latent Dirichlet Allocation model)
    - Cannot indicate where one job ad ends and another one begins
    - Cannot identify job title of each job ad
- Grouping Occupations by Standard Occupational Classification code
  - Apply a continuous bag of words (CBOW) model to identify the ad's job SOC code from list of job titles, using ONET's Sample of Reported Titles and list of Alternative Titles.
- Eliciting Skill and Task Related Information
  - Main classification of job tasks follows Spitz-Oener (2006): *nonroutine analytic*, *nonroutine interactive*, *nonroutine manual*, *routine cognitive*, *routine manual*
  - Within the body of job ads, map similar words to a common task. List of words related to each of five tasks are created as follows:
    - Fist, used the list of words described in Spitz-Oener (2006).
    - Second, augment the list with words whose meaning are similar to those in the original list (similarity is again determined by CBOW model)

Trends in occupational tasks from 1960 to 2000

Figure 1: Trends in Keyword Frequencies

	Total	Between	Within	Within Share	Total	Between	Within	Within Share
	A. Nonroutine Analytic				B. Nonroutine Interactive			
1960 Level	3.43				4.95			
	(0.08)				(0.05)			
1960-1970	0.26	0.14	0.13	0.48	-0.15	0.17	-0.32	2.09
	(0.07)	(0.01)	(0.07)	(0.23)	(0.05)	(0.01)	(0.05)	(0.93)
1970-1980	0.54	0.21	0.34	0.62	0.81	0.30	0.51	0.63
	(0.04)	(0.01)	(0.04)	(0.04)	(0.03)	(0.01)	(0.03)	(0.01)
1980-1990	0.77	0.12	0.64	0.84	1.02	0.10	0.93	0.90
	(0.04)	(0.02)	(0.04)	(0.03)	(0.04)	(0.02)	(0.05)	(0.02)
1990-2000	1.04	-0.26	1.31	1.25	0.93	-0.20	1.12	1.21
	(0.11)	(0.15)	(0.20)	(0.13)	(0.12)	(0.15)	(0.24)	(0.15)
1960-2000	2.61	0.20	2.41	0.92	2.61	0.37	2.24	0.86
	(0.14)	(0.15)	(0.23)	(0.06)	(0.15)	(0.16)	(0.27)	(0.06)
	C. Nonroutine Manual				D. Routine Cognitive			
1960 Level	1.00				1.23			
	(0.03)				(0.02)			
1960-1970	0.03	0.03	0.00	0.13	-0.25	0.04	-0.30	1.17
	(0.02)	(0.00)	(0.03)	(1.56)	(0.02)	(0.00)	(0.02)	(0.02)
1970-1980	-0.06	-0.11	0.05	-0.91	-0.13	-0.01	-0.12	0.96
	(0.01)	(0.01)	(0.01)	(0.60)	(0.01)	(0.00)	(0.01)	(0.02)
1980-1990	-0.08	-0.03	-0.05	0.61	-0.04	-0.04	0.00	-0.10
	(0.02)	(0.01)	(0.02)	(0.14)	(0.01)	(0.01)	(0.02)	(0.54)
1990-2000	-0.02	0.03	-0.06	2.38	-0.05	-0.14	0.09	-1.80
	(0.04)	(0.03)	(0.05)	(5.21)	(0.06)	(0.14)	(0.20)	(18.44)
1960-2000	-0.13	-0.08	-0.05	0.38	-0.47	-0.14	-0.33	0.70
	(0.06)	(0.04)	(0.07)	(0.90)	(0.06)	(0.15)	(0.20)	(0.41)
	E. Routine Manual							
1960 Level	0.78							
	(0.03)							
1960-1970	-0.20	-0.03	-0.17	0.88				
	(0.03)	(0.00)	(0.03)	(0.03)				
1970-1980	-0.23	-0.05	-0.17	0.77				
	(0.01)	(0.00)	(0.01)	(0.02)				
1980-1990	-0.20	0.02	-0.23	1.11				
	(0.01)	(0.01)	(0.01)	(0.04)				
1990-2000	-0.09	0.03	-0.13	1.35				
	(0.02)	(0.02)	(0.01)	(0.37)				
1960-2000	-0.72	-0.02	-0.70	0.97				
	(0.03)	(0.02)	(0.03)	(0.03)				

Decomposition using reduced form statistical model - Fortin, Lemieux, and Firpo (2011) model

- The basic of decomposition is to model workers' earnings as a function of their observed characteristics and occupational characteristics.

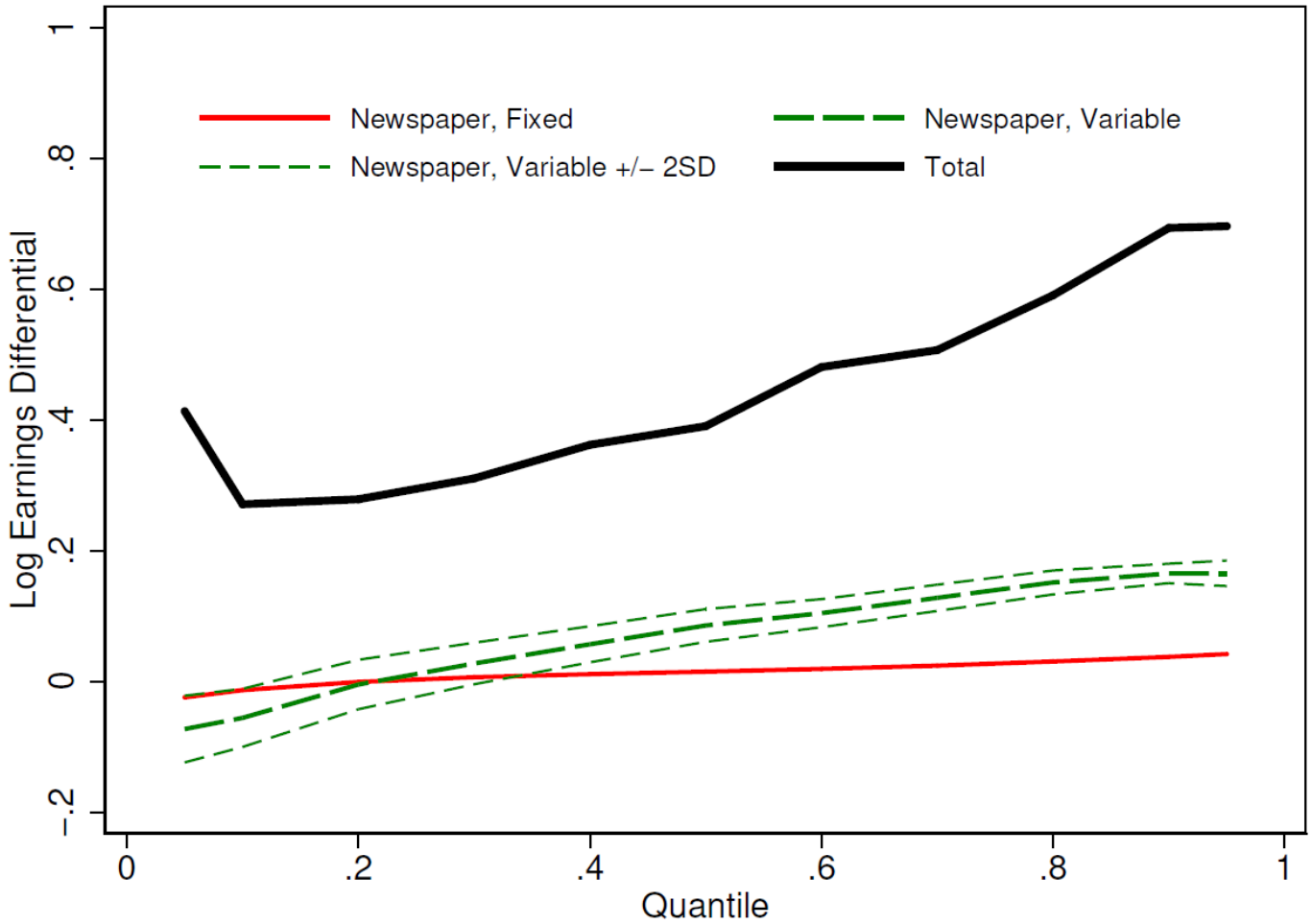
$$\log(W_{ijt}) = \log(\pi_{0t}) + \sum_{h=1}^{H} T_{hjt} \log \pi_{ht} + \sum_{k=1}^{K} \alpha_k \tilde{S}_{kg} + \log \epsilon_{ijt} \qquad (1)$$

- In any quantile of the earnings distribution, break down earning changes overtime into the contribution of observable changes in worker and occupational characteristics (called the composition effect) and the implicit change in wage structure. Then, decompose the composition effect further to examine the contribution of occupational characteristics. [1]
- Employ two sets of decompositions to highlight the importance of time-varying occupational measure:
  - First, fix task measure overtime for a given occupation (i.e. replacing T_{hjt} by its weighted average over time)
  - Second, allow task measure to vary freely across time within occupation.

[1] To do decomposition, they need to do Recentred Influence Function (RIF) estimation to derive the relation between real log earnings and occupational and worker characteristics in each quantile

Decomposition using reduced form statistical model - Results

Figure 2: Decomposition of Real Log Earnings



Decomposition using equilibrium model of occupational choice - Burstein, Morales, and Vogel (2015)

- Closed economy, many time periods, a continuum of workers.
- Each worker belongs to one of a finite number of group, $g = 1, ..., G$. The mass of workers in group g is L_{gt}
- Workers type g are endowed with a set of skills: $S_{gh}, h = 1, ..., H$.
- Worker i produces task h by combining skills and time: $q_{iht} = S_{gh} l_{iht}$ $(\sum_h l_{iht} = 1)$
- Worker i if she works in occupation j, produce

$$V_{ijt} = \epsilon_{ijt} \prod_{h=1}^{H} \left(\frac{q_{iht}}{T_{hjt}} \right)^{T_{hjt}} \tag{2}$$

- Assume that the wage equal value of output produced, says: $W_{ijt} = P_{jt} V_{ijt}$.
- Take P_{jt} as given, workers choose amount of time spent on each task to maximize the value of her output, implying the optimal allocation of time:

$$l_{ijt} = T_{hjt} \tag{3}$$

Decomposition using equilibrium model of occupational choice - Burstein, Morales, and Vogel (2015) (Cont.)

- Sub. (3) into (2) and also use $W_{ijt} = P_{jt} V_{ijt}$, we have

$$\log W_{ijt} = \log P_{jt} + \sum_{h=1}^{H} T_{hjt} \log S_{gh} + \log \epsilon_{ijt} \tag{4}$$

- Workers are assumed to freely choose occupation to maximize their wage. ϵ is assumed to be Frechet distribution with shape parameter θ
- The probability a worker of type g sort into occupation j

$$\lambda_{gjt} = \frac{\left(P_{jt} \prod_{h=1}^{H} (S_{gh})^{T_{ijt}}\right)^{\theta}}{\sum_{j'} \left(P_{j't} \prod_{h=1}^{H} (S_{gh})^{T_{ij't}}\right)^{\theta}} \tag{5}$$

- The average wage of group g workers equal

$$\bar{W}_{gt} = \Gamma\left(1 - \frac{1}{\theta}\right) \left(\sum_{j} \left(P_{jt} \prod_{h=1}^{H} (S_{gh})^{T_{ijt}}\right)^{\theta}\right)^{\frac{1}{\theta}} \tag{6}$$

Decomposition using equilibrium model of occupational choice - Burstein, Morales, and Vogel (2015) (Cont.)

- Assuming representative consumer has CES preference

$$V_t = \left[\sum_j (\xi_j)^{\frac{1}{\rho}} (V_j t)^{\frac{\rho-1}{\rho}} \right]^{\frac{\rho-1}{\rho}} \tag{7}$$

- The market clearing condition for occupation j output

$$\underbrace{\frac{\xi_j (P_{jt})^{1-\rho}}{\sum_{j'} \xi_{j'} (P_{j't})^{1-\rho}}}_{\text{(A)}} \underbrace{\sum_{g=1}^{G} \bar{W}_{gt} L_{gt}}_{\text{(B)}} = \underbrace{\sum_{g=1}^{G} \Gamma\left(1-\frac{1}{\theta}\right) \left(\sum_j \left(P_{jt} \prod_{h=1}^{H} (S_{gh})^{T_{ijt}} \right)^{\theta} \right)^{\frac{1}{\theta}}}_{\text{(C)}} \underbrace{\lambda_{gjt} L_{gt}}_{\text{(D)}} \tag{8}$$

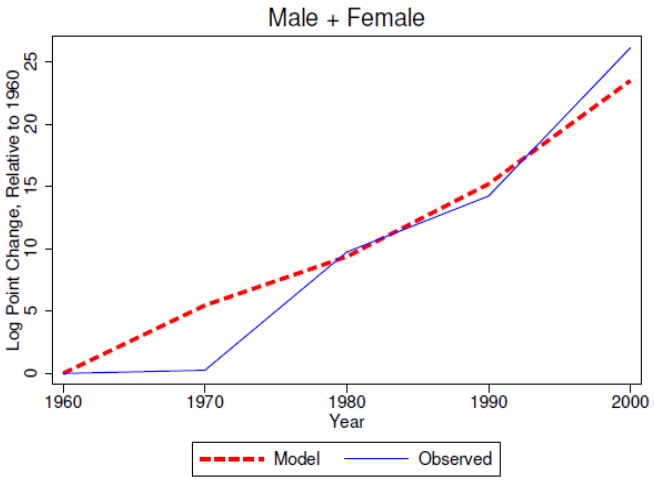
- An equilibrium for economy is the solution to $\{\lambda_{gj}\}$ and $\{P_j\}$ as determined by equation (5) and (8)

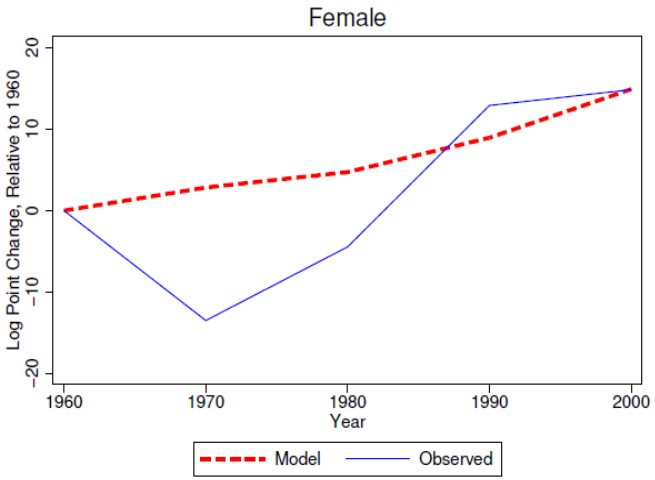
Decomposition using equilibrium model of occupational choice - Counterfactual change in wage distribution

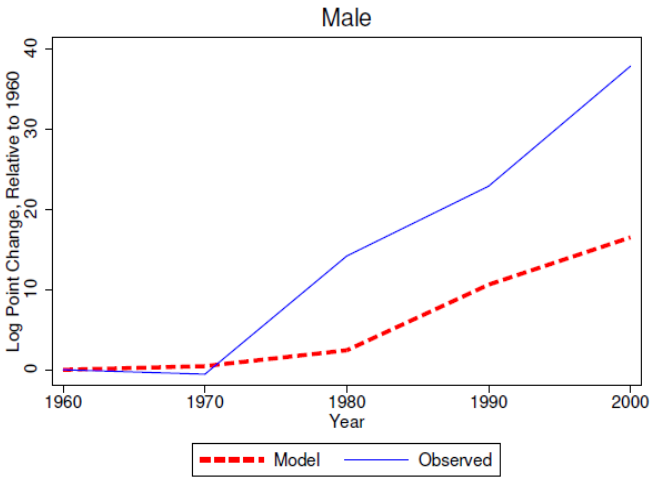
- First, use data from a single period (1960) with equation (5) and (6) to estimate demographic groups' skills (i.e. S_{gh}) in producing tasks.
- Fixing $\log S_{gh}$ to the estimated values, but allowing T_{hjt} to change over time, we can compute the couterfactual change in distribution of wages that is due only to changes in demand for tasks

Decomposition using equilibrium model of occupational choice - Results

Figure 3: Counterfactual and Observed 90-10 Inequality, 1960-2000







Summary

- Changes in task composition of workforce has occurred within rather between occupations.
- Within-occupation task changes are fundamentally important in the examination of evolution of income inequality.

THANK YOU!!!