Task Specialization in U.S. Cities 1880-2000

Guy Michaels
LSE

Ferdinand Rauch
Oxford

Stephen Redding
Princeton

Motivation

- As a large and growing share of the world's population concentrates in cities, worker's employment outcomes are increasingly determined in urban labor markets
  - What kinds of jobs are available in urban areas?
  - How do these differ from the jobs available in rural areas?
  - Does the task composition of employment vary by city size?
  - Are these patterns stable over time?
  - If they have changed, what explains the changes?
  - What are the implications of the evolving task composition of employment for the sources of agglomeration?
- Combine Census of Population and Dictionary of Occupations (DOTs) data to provide novel evidence on the task composition of employment in urban and rural areas in the United States from 1880-2000

Motivation

- Understanding the task composition of employment is central to evaluating alternative theories of agglomeration
- Traditional emphasis on costs of moving goods and people
  - New economic geography literature (e.g. Fujita, Krugman and Venables 1999)
  - Canonical models of urban economies (e.g. Alonso 1964, Muth 1968, Mills 1967)
- More recent research on costs of moving ideas
  - Human capital externalities (e.g. Moretti 2004, Davis and Dingel 2013)
  - Costs of exchanging ideas (e.g. Davis and Dingel 2012)
- Externalities for the movement of ideas could differ from those for the movement of goods and people
- Urbanization is likely to lead to substantial changes in the relative demand for different occupations and skills

This Paper

- Measure multiple production tasks undertaken by workers within occupations using 3,000 verbs from around 12,000 occupational descriptions in DOTs.
- Dynamic rather than static view of urban and rural areas, in which relative importance of different sources of agglomeration has changed over time.
  - In 1880, tasks concentrated in metro areas were "Braid," "Sew," "Stretch" and "Thread"
  - By 2000, tasks concentrated in metro areas were "Analyze," "Advise," "Confer" and "Report"
  - The correlation between thesaurus categories that were concentrated in metro areas in 1880 and 2000 is negative
- Increased importance of "interactive" tasks in cities (thought, communication and intersocial activity)
- Consistent with increased importance of exchange of ideas
- Show related to communication and transport technology

Related Literature

- Theoretical and empirical literature on formation of cities
  - Duranton & Puga (2004), Helpman (1998), Henderson (1974)
  - Rosenthal & Strange (2004)

- Research on sectoral versus functional specialization, human capital and skills, and the division of labor
  - Baumgardner (1988), Duranton (1998), Gaspar & Glaeser (1998), Otta & Fujita (1993), Glaeser & Saiz (2004), Duranton & Jayet (2011), Duranton & Puga (2005), Rossi-Hansberg, Sarte & Owens (2009), Bacolod, Blum & Strange (2009), Duranton & Jayet (2011), Helsley & Strange (2007), Lin (2011), Davis & Dingel (2012)

- Research on tasks, technology and offshoring
  - Autor, Levy & Murname (2003), Autor & Dorn (2012), Blinder (2005), Becker, Ekholm & Muendler (2009), Blinder & Krueger (2009), Firpo et al. (2011), Grossman & Rossi-Hansberg (2008, 2012), Holmes & Mitchell (2008), Jensen (2011)

Outline

- Data
- Task specialization in metro and non-metro areas
- Robustness
- Theoretical framework
- Explanations

Data

- IPUMS decennial census micro data
  - Exclude agriculture to ensure that the results are not driven by the decline in agricultural employment
  - Results robust to including agriculture
  - Eight two-digit occupations and over 150 three-digit occupations
  - Eleven two-digit sectors and over 100 three-digit sectors
  - Time-varying boundaries of metro areas based on Metropolitan Statistical Areas (MSAs)
- Dictionary of Occupational Titles (DOTs)
  - Contemporary (1991)
  - Historical (earliest from 1939)
- Roget's Thesaurus (1911)
- Phones and highways
  - Residence phones by county (AT&T 1935)
  - Highway maps, US Department of Transportation

Outline

- Data
- Task specialization in metro and non-metro areas
- Robustness
- Theoretical framework
- Explanations

Specialization Across Occupations & Sectors

- Estimate the following regression for each year t separately using data across occupations o and sectors s:

$$\text{MetroShare}_{ost} = \mu_{ot} + \eta_{st} + \varepsilon_{ost}$$

- MetroShare_{ost} is the share of employment in metro areas in occupation o, sector s and year t
- Observations are weighted by person weights
- μ_{ot} are occupation-year fixed effects
- η_{st} are sector-year fixed effects
- ε_{ost} is a stochastic error

Specialization Across Occupations & Sectors

TABLE 1. Metro area specialization for aggregate occupations and sectors.

	Coefficient 1880	Standard Error 1880	Rank 1880	Coefficient 2000	Standard Error 2000	Rank 2000
Panel A: Two-digit occupation						
Clerical and Kindred	0.15	0.08	1	0.04	0.01	4
Craftsmen	0.09	0.06	2	−0.01	0.01	6
Operatives	0.06	0.07	3	−0.05	0.01	7
Sales workers	0.01	0.07	4	0.05	0.01	2
Service Workers	0.00	0.08	5	0.00	0.01	5
Managers, Officials, and Proprietors	−0.03	0.08	6	0.05	0.01	3
Professional, Technical	−0.07	0.08	7	0.07	0.01	1
Laborers	−0.2	0.18	8	−0.15	0.07	8
Panel B: Two-digit sector						
Entertainment and Recreation Services	0.29	0.08	1	0.04	0.01	4
Wholesale and Retail Trade	0.13	0.05	2	0.02	0.01	6
Finance, Insurance, and Real Estate	0.13	0.06	3	0.06	0.01	2
Manufacturing	0.06	0.05	4	−0.01	0.01	10
Personal Services	0.01	0.06	5	0.03	0.01	5
Transportation, Communication, and Other Utilities	0.01	0.04	6	0.05	0.01	3
Public Administration	−0.03	0.07	7	0.01	0.01	7
Professional and Related Services	−0.03	0.06	8	0.00	0.01	9
Business and Repair Services	−0.12	0.08	9	0.08	0.01	1
Construction	−0.14	0.08	10	0.00	0.01	8
Mining	−0.31	0.05	11	−0.27	0.03	11

Measuring Tasks Within Occupations

- Introduce a new methodology by characterizing the tasks undertaken for each occupation
- DOTs lists over 12,700 detailed occupations and their descriptions
- Use time-invariant occupational descriptions to abstract from changes in word use over time
  - Baseline specification uses 1991 DOTs
  - Robustness using 1939 DOTs
  - Examine changes in task content of occupations over time
- We use the verbs in each occupation's description to characterize the tasks undertaken by that occupation
  - Verbs capture the performance of tasks because they correspond to an action (bring, read), an occurrence (happen, become), or a state of being (exist, stand)

Verbs from Occupation Descriptions

- Start with a list of over 3000 English verbs from "Writing English", a company offering English language consulting
- Search each occupational description in 1991 DOTs
  - All lowercase or uppercase appearances of first person singular (e.g. (I) talk), third person singular (e.g. (she) talks) or present participle (e.g. (he is) talking) versions of each verb
- Robustness test using 1939 DOTs

Search for Verbs, adjusting Tense and Capitals

ECONOMIST: Plans, designs, and conducts research to aid in interpretation of economic relationships and in solution of problems arising from production and distribution of goods and services: Studies economic and statistical data in area of specialization, such as finance, labor, or agriculture. Devises methods and procedures for collecting and processing data, utilizing knowledge of available sources of data and various econometric and sampling techniques. Compiles data relating to research area, such as employment, productivity, and wages and hours. Reviews and analyzes economic data in order to prepare reports detailing results of investigation, and to stay abreast of economic changes ...

Type 1: **Nouns** / **Adjectives** Misinterpreted as Verbs
Type 2: Unrecognized Verbs

ECONOMIST: Plans, designs, and conducts research to aid in interpretation of economic relationships and in solution of problems arising from production and distribution of goods and services: Studies economic and statistical data in area of specialization, such as finance, labor, or agriculture. Devises methods and procedures for collecting and processing data, utilizing knowledge of available sources of data and various econometric and sampling techniques. Compiles data relating to research area, such as employment, productivity, and wages and hours. Reviews and analyzes economic data in order to prepare reports detailing results of investigation, and to stay abreast of economic changes ...

Verbs and Occupations

- Match DOT occupations (over 12,700) to (stable) IPUMS 1950 occupations (over 150)
- For each IPUMS 1950 occupation o and verb v, we calculate the frequency with which each verb used for each occupation

$$\text{VerbFreq}_{vo} = \frac{\text{Appearances of verb } v \text{ matched to } o}{\text{Appearances of all verbs matched to } o}$$

- Estimate a separate regression for each verb v and year t using data across occupations o and sectors s

$$(\text{MetroShare})_{ost} = \alpha_{vt}\text{VerbFreq}_{vo} + \eta_{st} + \epsilon_{ost}$$

- A rise in α_{vt} implies that employment in occupations using that verb is increasingly concentrating in metro areas within sectors over time

Verbs Most Strongly Correlated with Metro Areas

1991 DOTs

- Top 10 verbs, ranked by estimated effect of one standard deviation change in the frequency of verb use on the fraction of workers in metro areas
- Controlling for sector fixed effects by year

Rank	1880	1900	1920	1940	1960	1980	2000
Panel A: Verbs most strongly correlated with metro area employment shares							
1	Thread	Thread	File	File	Document	Identify	Develop
2	Stretch	Stitch	Distribute	Bill	Schedule	Document	Determine
3	Interfere	Telephone	Record	Take	File	Advise	Analyze
4	Hand	Sew	Notice	Compile	Record	Concern	Factor
5	Ravel	Hand	Telephone	Distribute	Distribute	Report	Review
6	Sew	Assist	Bill	Pay	Compile	Schedule	Confer
7	Braid	Visit	Envelope	Letter	Notice	Develop	Advise
8	Visit	Describe	Document	Notice	Identify	Analyze	Report
9	Receive	Number	Learn	Record	Send	Determine	Concern
10	Sack	Stamp	Number	Send	Notify	Notify	Plan

Verbs Least Strongly Correlated with Metro Areas

1991 DOTs

- Bottom 10 verbs, ranked by estimated effect of one standard deviation change in the frequency of verb use on the fraction of workers in metro areas
- Controlling for sector fixed effects by year

Panel B: Verbs least strongly correlated with metro area employment shares

1821	Conduct	Abstract	Counsel	Recur	Accord	Power	Restrain
1822	Teach	Tread	Discuss	Enlist	Feed	Pour	Cut
1823	Channel	Pinch	Hear	Labor	Escape	Erect	Power
1824	Sound	Assign	Assign	Tread	Hook	Clean	Massage
1825	Rule	Settle	Teach	Assign	Traverse	Massage	Remove
1826	Matter	Matter	Matter	Approve	Tread	Pump	Feed
1827	Drill	Tunnel	Consolidate	Extract	Loosen	Cut	Clean
1828	Tread	Sound	Rule	Tunnel	Range	Feed	Pump
1829	Tunnel	Rule	Tunnel	Malt	Activate	Move	Move
1830	Pinch	Sole	Sound	Establish	Turn	Turn	Turn

Quantifying Task Specialization

- Develop a quantitative measure of the tasks undertaken within occupations using the meanings of verbs
- Classify the meanings of verbs using Roget's Thesaurus
  - Seminal reference for English language use
  - Enumerates multiple possible uses of the same word
  - Classes, Divisions, Sections and Categories

CLASS I	Abstract Relations
CLASS II	Space
CLASS III	Matter
CLASS IV	Intellect
CLASS V	Volition
CLASS VI	Emotion, Religion and Morality

Thesaurus Subdivisions and Occupations

- Measure the meaning of each verb v using the frequency with which it appears in each subdivision k of Roget's Thesaurus

$$\text{ThesFreq}_{vk} = \frac{\text{Appearances of verb } v \text{ in subdivision } k \text{ of thesaurus}}{\text{Total appearances of verb } v \text{ in thesaurus}},$$

- Frequency with which occupations use concepts from each subdivision of the thesaurus

$$\text{TaskContent}_{ko} = \sum_{v \in V} \text{VerbFreq}_{vo} \times \text{ThesFreq}_{vk}.$$

- Estimate a separate regression for each thesaurus subdivision k and year t using data across occupations o and sectors s

$$\text{MetroShare}_{ost} = \beta_{kt} \text{TaskContent}_{ko} + \eta_{kst} + \varepsilon_{ost},$$

- A rise in β_{kt} implies that employment in occupations using that subdivision of the thesaurus is increasingly concentrating in metro areas within sectors over time

Thesaurus Sections

TABLE 3. Ranking of thesaurus sections by concentration in metro areas in 1880 and 2000.

Thesaurus Class (C), Division (D), and Section (S)	Rank Section 1880	Rank Section 2000	Difference
C 1, Abstract relations, S I. EXISTENCE	15	12	3
C 1, Abstract relations, S II. RELATION	6	15	−9
C 1, Abstract relations, S III. QUANTITY	1	34	−33
C 1, Abstract relations, S IV. ORDER	23	9	14
C 1, Abstract relations, S V. NUMBER	24	10	14
C 1, Abstract relations, S VI. TIME	3	23	−20
C 1, Abstract relations, S VII. CHANGE	34	11	23
C 1, Abstract relations, S VIII. CAUSATION	26	22	4
C 2, Space, S I. SPACE IN GENERAL	10	32	−22
C 2, Space, S II. DIMENSIONS	4	36	−32
C 2, Space, S IV. MOTION	19	27	−8
C 3, Matter, S I. MATTER IN GENERAL	2	31	−29
C 3, Matter, S II. INORGANIC MATTER	7	37	−30
C 3, Matter, S III. ORGANIC MATTER	11	38	−27
C 4, Intellect, D I, S I. OPERATIONS OF INTELLECT IN GENERAL	21	14	7
C 4, Intellect, D I, S II. PRECURSORY CONDITIONS & OPERATIONS	16	19	−3
C 4, Intellect, D I, S III. MATERIALS FOR REASONING	25	7	18
C 4, Intellect, D I, S IV. REASONING PROCESSES	35	4	31
C 4, Intellect, D I, S V. RESULTS OF REASONING	33	5	28
C 4, Intellect, D I, S VI. EXTENSION OF THOUGHT	8	3	5
C 4, Intellect, D I, S VII. CREATIVE THOUGHT	38	21	17
C 4, Intellect, D II, S I. NATURE OF IDEAS COMMUNICATED.	27	1	26
C 4, Intellect, D II, S II. MODES OF COMMUNICATION	28	17	11
C 4, Intellect, D II, S III. MEANS OF COMMUNICATING IDEAS	32	18	14
C 5, Will, D I, S I. VOLITION IN GENERAL	14	29	−15
C 5, Will, D I, S II. Prospective Volition 1	29	20	9
C 5, Will, D I, S III. VOLUNTARY ACTION	20	33	−13
C 5, Will, D I, S IV. ANTAGONISM	30	16	14
C 5, Will, D II, S I. GENERAL INTERSOCIAL VOLITION	31	13	18
C 5, Will, D II, S II. SPECIAL INTERSOCIAL VOLITION	37	2	35
C 5, Will, D II, S III. CONDITIONAL INTERSOCIAL VOLITION	9	30	−21
C 5, Will, D II, S IV. POSSESSIVE RELATIONS	13	8	5
C 5, Will, S V. RESULTS OF VOLUNTARY ACTION	22	25	−3
C 6, Emotion, Religion, Morality, S I. AFFECTIONS IN GENERAL	5	35	−30
C 6, Emotion, Religion, Morality, S II. PERSONAL AFFECTIONS	12	28	−16
C 6, Emotion, Religion, Morality, S III. SYMPATHETIC AFFECTIONS	18	26	−8
C 6, Emotion, Religion, Morality, S IV. MORAL AFFECTIONS	36	6	30
C 6, Emotion, Religion, Morality, S V. RELIGIOUS AFFECTIONS	17	24	−7

Interactiveness

- Sharp change in the relative ranking of thesaurus sections for external world (Classes I-III) and internal world (Classes IV-V)

Internal world		
CLASS IV	Intellect	
Division 1	Formation of Ideas	Thought
Division 2	Communication of Ideas	Communication
CLASS V	Volition	
Division 1	Individual Volition	
Division 2	Intersocial Volition	Intersocial Activity

- Baseline measure – Classes IV and V of thesaurus
- We define this combination of tasks as "interactiveness," because captures essence of human interaction: the generation ("thought") and transmission ("communication") of ideas to other humans ("intersocial activity")

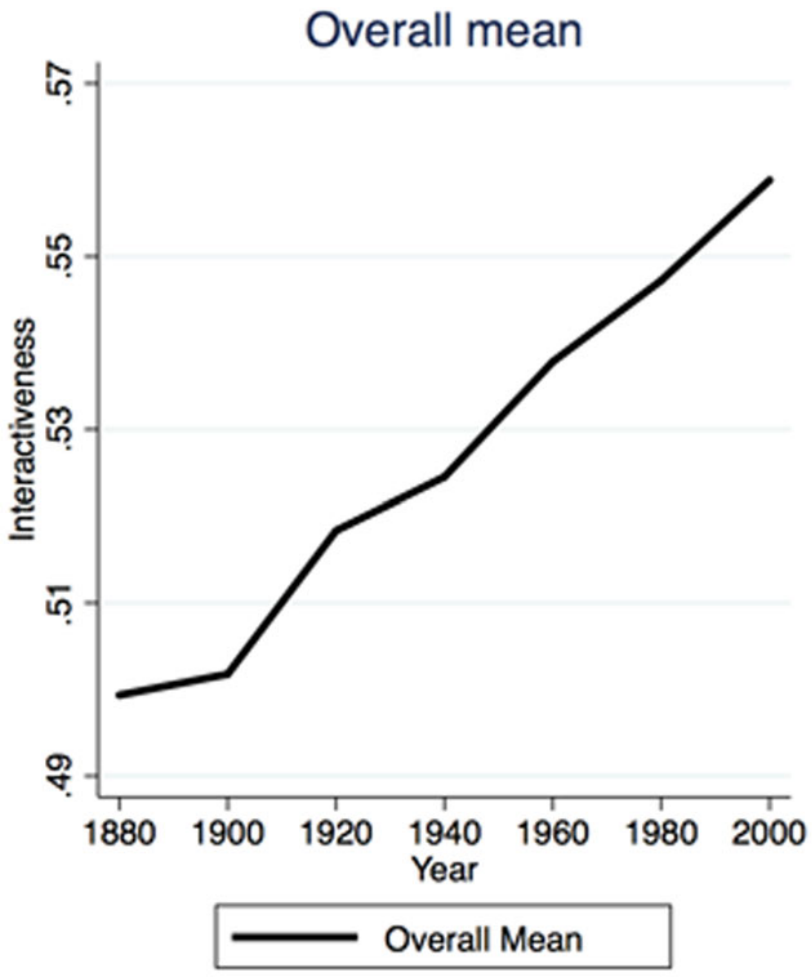
10 Most & Least Interactive Occupations

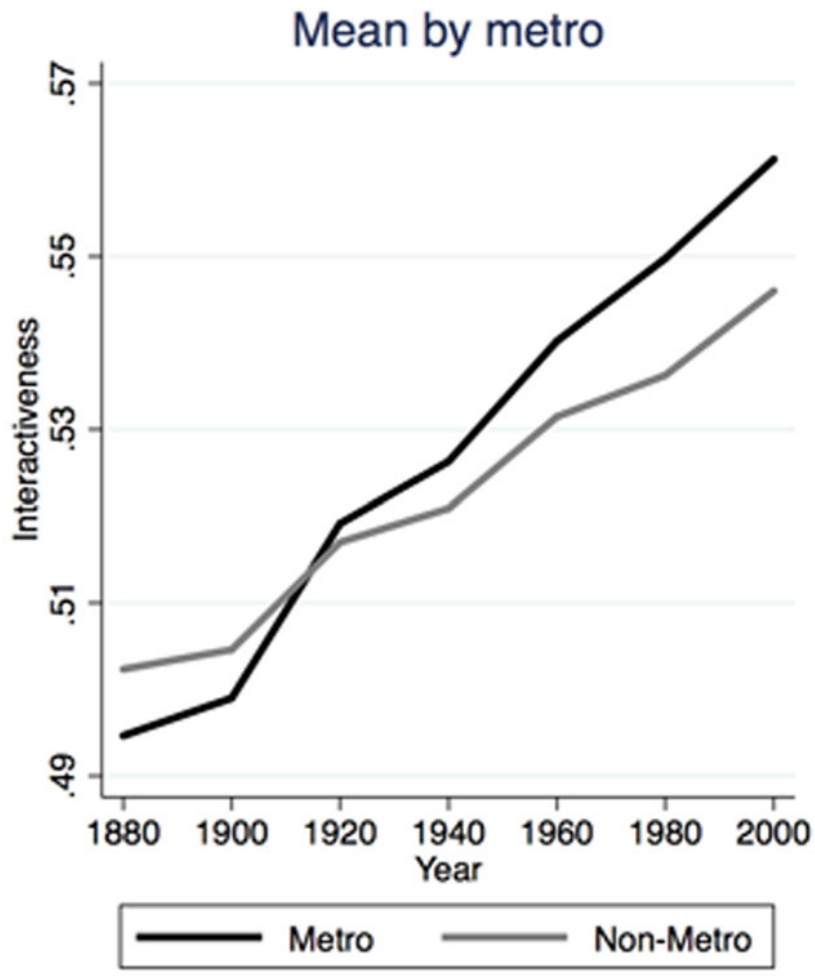
IPUMs 1950 Classification

TABLE 4. Most and least interactive occupations.

Panel A: Top ten interactive occupations
Economists
Nurses, professional
Pharmacists
Clergymen
Religious workers
Accountants and auditors
Postmasters
Buyers and dept heads, store
Aeronautical-Engineers
Statisticians and actuaries
Panel B: Bottom ten interactive occupations
Brickmasons, stonemasons, and tile setters
Attendants, auto service, and parking
Painters, except construction or maintenance
Plumbers and pipe fitters
Upholsterers
Asbestos and insulation workers
Welders and flame cutters
Blasters and powdermen
Dressmakers and seamstresses except factory
Roofers and slaters

Interactiveness Over Time





Outline

- Data
- Task specialization in metro and non-metro areas
- Robustness
- Theoretical framework
- Explanations

Robustness

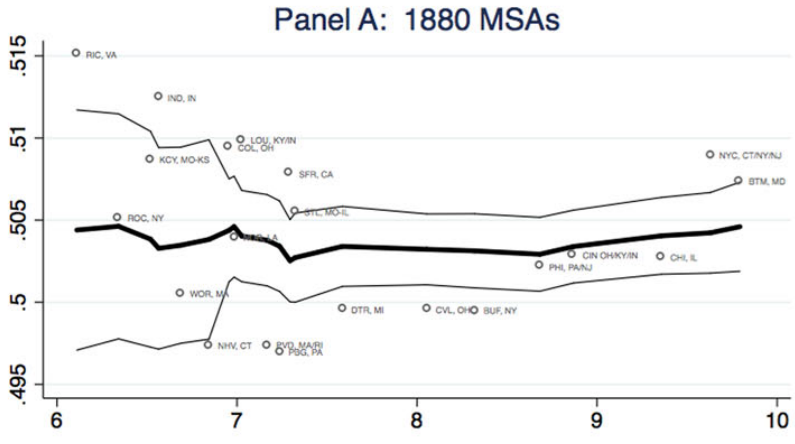
- 1939 Dictionary of Occupations (DOTs)
- Variation across metro areas

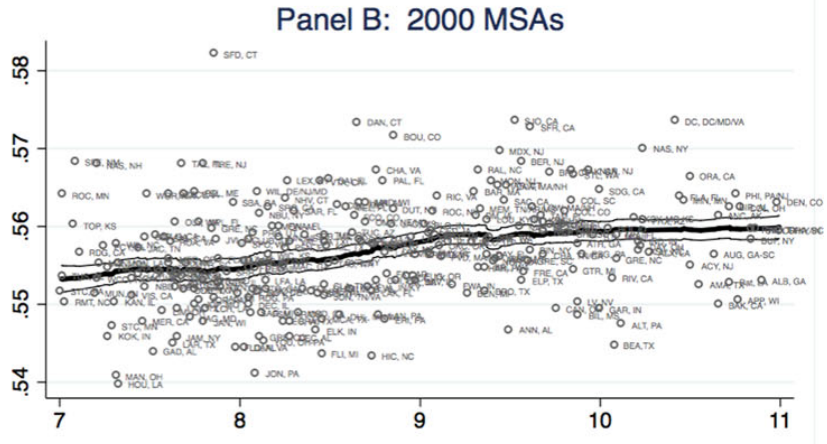
1939 DOTS

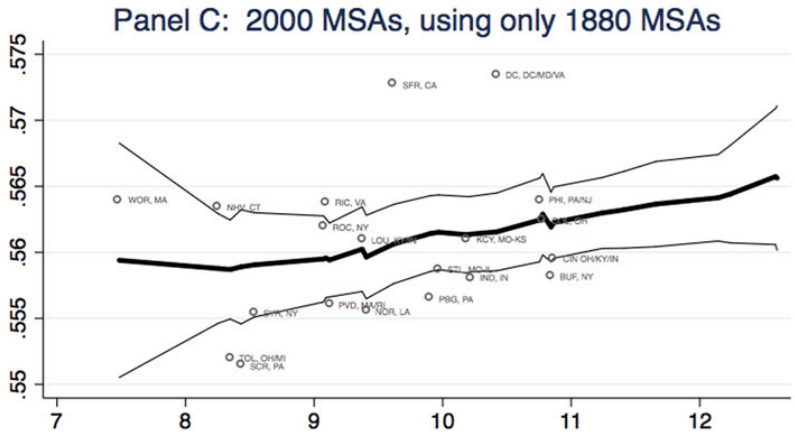
Panel A: Verbs Most Strongly Correlated with Metro Area Employment Shares							
Rank	1880	1900	1920	1940	1960	1980	2000
1	Retouch	Permeate	Permeate	Bounce	Accrue	Estimate	Advise
2	Flounce	Flounce	Fake	Blaze	Kid	Advise	Sell
3	Permeate	Retouch	Hum	Reserve	Undercharge	Calculate	Estimate
4	Lure	Initiate	Seep	Converge	Seep	Appraise	Investigate
5	Abut	Report	Kid	Favor	Prompt	Investigate	Prefer
6	Highlight	Enamel	Undercharge	Mail	Converge	Question	Appraise
7	Solidify	Solidify	Smash	Kid	Necessitate	Accrue	Quote
8	Glow	Refund	Accrue	Undercharge	Document	Wage	Display
9	Trawl	Enlarge	Necessitate	Lobby	Allocate	Adjudicate	Bid
10	Finish	Identify	Overload	Seep	Doff	Inform	Jail
Panel A: Verbs Least Strongly Correlated with Metro Area Employment Shares							
Rank	1880	1900	1920	1940	1960	1980	2000
1682	Dovetail	Lift	Demand	Program	Lounge	Flout	Narrow
1683	Sail	Rain	Discuss	Transact	Spoon	Heft	Line
1684	Overturn	Pink	Resist	Seam	Encounter	Sinter	Transport
1685	Extend	Finish	Induce	Grapple	Sinter	Hang	Truck
1686	Attain	Sew	Spoon	Board	Smoke	Hook	Drive
1687	Late	Top	Snarl	Pick	Back	Truck	Remove
1688	Embroider	Notch	Resume	Back	Pile	Bolt	Screw
1689	Fudge	Embroider	Intersperse	Flicker	Hand	Remove	Hook
1690	Foul	Offset	Top	Resume	Boat	Line	Bolt
1691	Offset	Scrutinize	Recede	Recede	Anchor	Hand	Hand

Variation Across Metro Areas

- Mean interactiveness across metro areas in 1880 and 2000







Outline

- Data
- Task specialization in metro and non-metro areas
- Robustness
- Theoretical Framework
- Explanations

Model Outline

- Develop a theoretical model of the distribution of employment across occupations, sectors and locations
- Key predictions of the model are comparative statics with respect to the costs of trading the tasks produced by each occupation and the final goods produced by each sector
- When these costs are large, all locations have similar employment structures across sectors, and all tasks within each sector are undertaken where the final good is produced
- As costs of trading final goods and tasks fall, locations specialize across sectors and across occupations within sectors
- If densely-populated urban locations have a comparative advantage in interactive tasks relative to sparsely-populated rural locations, the model predicts that a fall in the costs of trading tasks leads to an increase in the interactiveness of employment within sectors in urban relative to rural areas

Preferences

- Workers' preferences are defined over a goods consumption index (C_n) and residential land use (H_n)

$$U_n = \left(\frac{C_n}{\alpha}\right)^{\alpha} \left(\frac{H_n}{1-\alpha}\right)^{1-\alpha}, \qquad 0 < \alpha < 1$$

- Goods consumption index (C_n) is defined over sectors (e.g. Manufacturing, Services) indexed by $s \in S$:

$$C_n = \left[\sum_{s \in S} C_{ns}^{\frac{\beta-1}{\beta}}\right]^{\frac{\beta}{\beta-1}}$$

- Consumption index for each sector is defined over a continuum of goods (e.g. Motor Vehicles) indexed by $j \in [0, 1]$:

$$C_{ns} = \left[\int_0^1 c_{ns}(j)^{\frac{\sigma_s-1}{\sigma_s}} dj\right]^{\frac{\sigma_s}{\sigma_s-1}}$$

Production Technology

- Goods in each sector are produced by workers from different occupations who perform a continuum of tasks
- Cost to a consumer in location n of purchasing one unit of good j within sector s from location i is:

$$p_{nis}(j) = \frac{d_{nis} G_{is}(j)}{z_{is}(j)}$$

- where d_{nis} are iceberg goods trade costs
- Final goods productivity for each good, sector and location is assumed to be drawn independently from a Fréchet distribution:

$$F_{is}(z) = e^{-T_{is} L_{is}^{\eta_s} z^{\theta_s}},$$

Occupational Structure

- Final goods in each sector produced using occupations $o \in O_s$:

$$y_{is}(j) = \left[\sum_{o \in O_s} X_{iso}(j)^{\frac{\mu_s - 1}{\mu_s}} \right]^{\frac{\mu_s}{\mu_s - 1}}$$

- Within each occupation, worker perform tasks $t \in [0, 1]$:

$$X_{iso}(j) = \left[\int_0^1 x_{iso}(j, t)^{\frac{\nu_{so} - 1}{\nu_{so}}} \, dt \right]^{\frac{\nu_{so}}{\nu_{so} - 1}}$$

- Cost to firm in location n of sourcing a task t from location i within occupation o and sector s is:

$$g_{niso}(j, t) = \frac{\tau_{niso} w_i}{a_{iso}(j, t)}$$

- Task productivity drawn from a Fréchet distribution:

$$\mathcal{F}_{iso} = e^{-U_{iso} L_{iso}^{\chi_{so}} a^{-\epsilon_{so}}}$$

Trade in Goods and Tasks

- Trade in tasks
  - Share of firm costs in location n accounted for by tasks sourced from location i within occupation o and sector s (λ_{niso}):

$$\lambda_{niso} = \frac{U_{iso} L_{iso}^{\chi_{so}} \left(\tau_{niso} w_i\right)^{-\epsilon_{so}}}{\sum_{k \in N} U_{kso} L_{kso}^{\chi_{so}} \left(\tau_{nkso} w_k\right)^{-\epsilon_{so}}}.$$

- Trade in final goods
  - Share of location n's expenditure on final goods produced in location i within sector s (π_{nis}):

$$\pi_{nis} = \frac{T_{is} L_{is}^{\eta_s} \left(d_{nis} \Phi_{is} w_i\right)^{-\theta_s}}{\sum_{k \in N} T_{ks} L_{ks}^{\eta_s} \left(d_{nks} \Phi_{ks} w_k\right)^{-\theta_s}},$$

Falls in Communication and Transport Costs

- Comparative advantage across occupations
  - Location i specializes more in occupation o relative to occupation m compared to another location k when it has lower production costs and lower bilateral costs of trading tasks (τ_{niso})

$$\frac{\lambda_{niso}/\lambda_{nkso}}{\lambda_{nism}/\lambda_{nksm}} = \frac{\left[U_{iso}L_{iso}^{\chi_{so}}\left(\tau_{niso}w_i\right)^{-\epsilon_{so}}\right] / \left[U_{kso}L_{kso}^{\chi_{so}}\left(\tau_{nkso}w_k\right)^{-\epsilon_{so}}\right]}{\left[U_{ism}L_{ism}^{\chi_{so}}\left(\tau_{nism}w_i\right)^{-\epsilon_{sm}}\right] / \left[U_{ksm}L_{ksm}^{\chi_{so}}\left(\tau_{nksm}w_k\right)^{-\epsilon_{sm}}\right]}.$$

- Comparative advantage across sectors
  - Location i specializes more in sector s relative to sector r compared to another location k when it has lower production costs and lower bilateral costs of trading goods (d_{nis}).

$$\frac{\pi_{nis}/\pi_{nks}}{\pi_{nir}/\pi_{nkr}} = \frac{\left[T_{is}L_{is}^{\eta_s}\left(d_{nis}\Phi_{is}w_i\right)^{-\theta_s}\right] / \left[T_{ks}L_{ks}^{\eta_s}\left(d_{nks}\Phi_{ks}w_k\right)^{-\theta_s}\right]}{\left[T_{ir}L_{ir}^{\eta_r}\left(d_{nir}\Phi_{ir}w_i\right)^{-\theta_r}\right] / \left[T_{kr}L_{kr}^{\eta_r}\left(d_{nkr}\Phi_{kr}w_k\right)^{-\theta_r}\right]}.$$

- Falls in communication and transport costs lead to increased specialization by comparative advantage

Outline

- Data
- Task specialization in metro and non-metro areas
- Robustness
- Theoretical Framework
- Explanations

Explanations

- Decomposing interactiveness
- Variation within and between sectors
- Communication technology

Decomposing Changes in Interactiveness

- Overall interactiveness is the employment-weighted average of interactiveness for each two-digit-sector-occupation cell z:

$$I_{jt} = \sum_{z \in \Omega} \sum_{o \in \Omega_z} \frac{E_{ojt}}{E_{jt}} I_o, \qquad j \in \{M, N\},$$

- Change in overall interactiveness can be decomposed as:

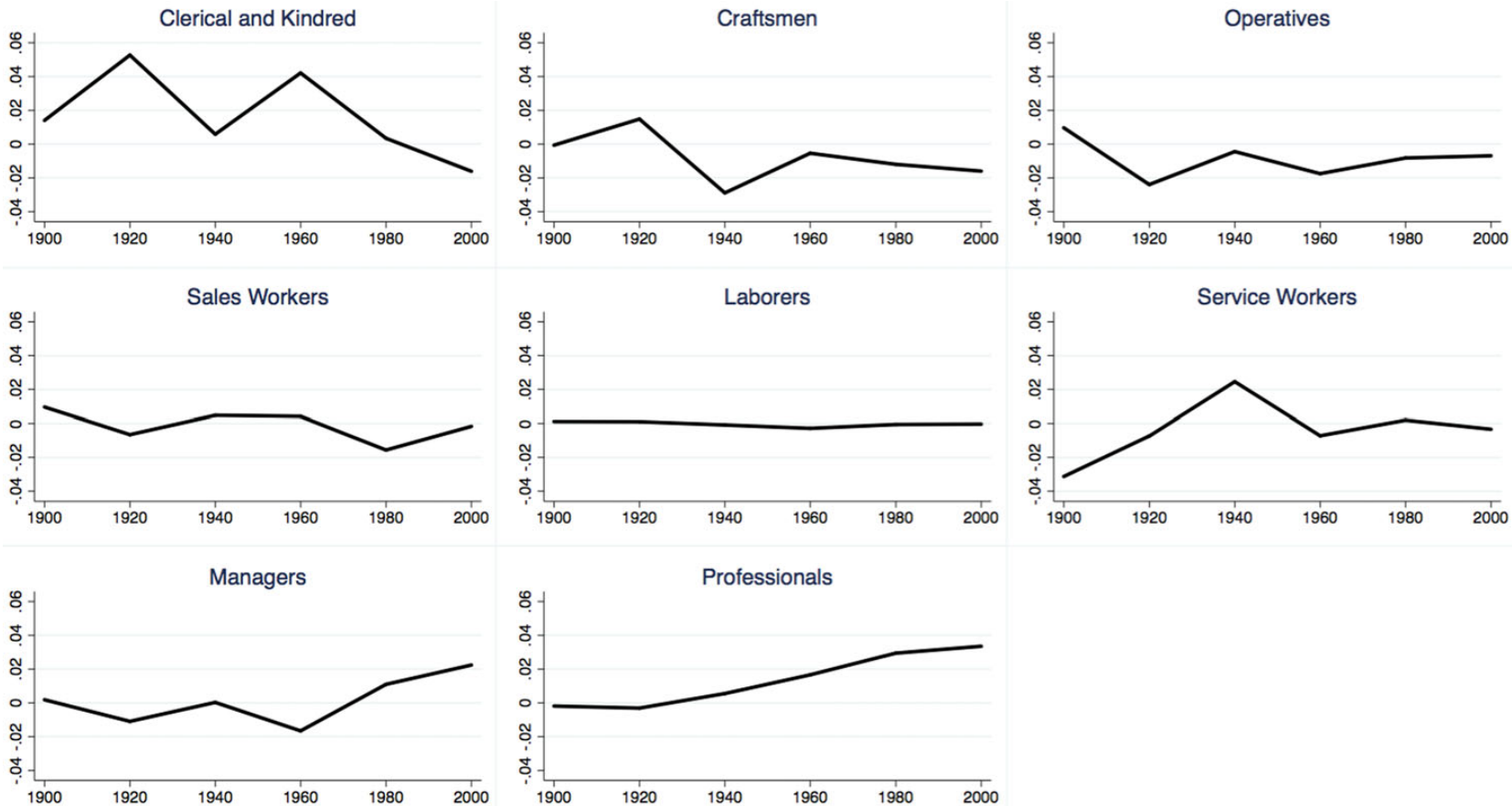
$$\triangle I_{jt} = \sum_{z \in \Omega} \sum_{o \in \Omega_z} \left[\triangle \left(\frac{E_{ojt}}{E_{jt}} \right) \right] I_o, \qquad j \in \{M, N\},$$

- Difference between metro and non-metro areas in the change in overall interactiveness can be decomposed as:

$$\triangle I_{Mt} - \triangle I_{Nt} = \sum_{z \in \Omega} \sum_{o \in \Omega_z} \left[\triangle \frac{E_{oMt}}{E_{Mt}} - \triangle \frac{E_{oNt}}{E_{Nt}} \right] I_o,$$

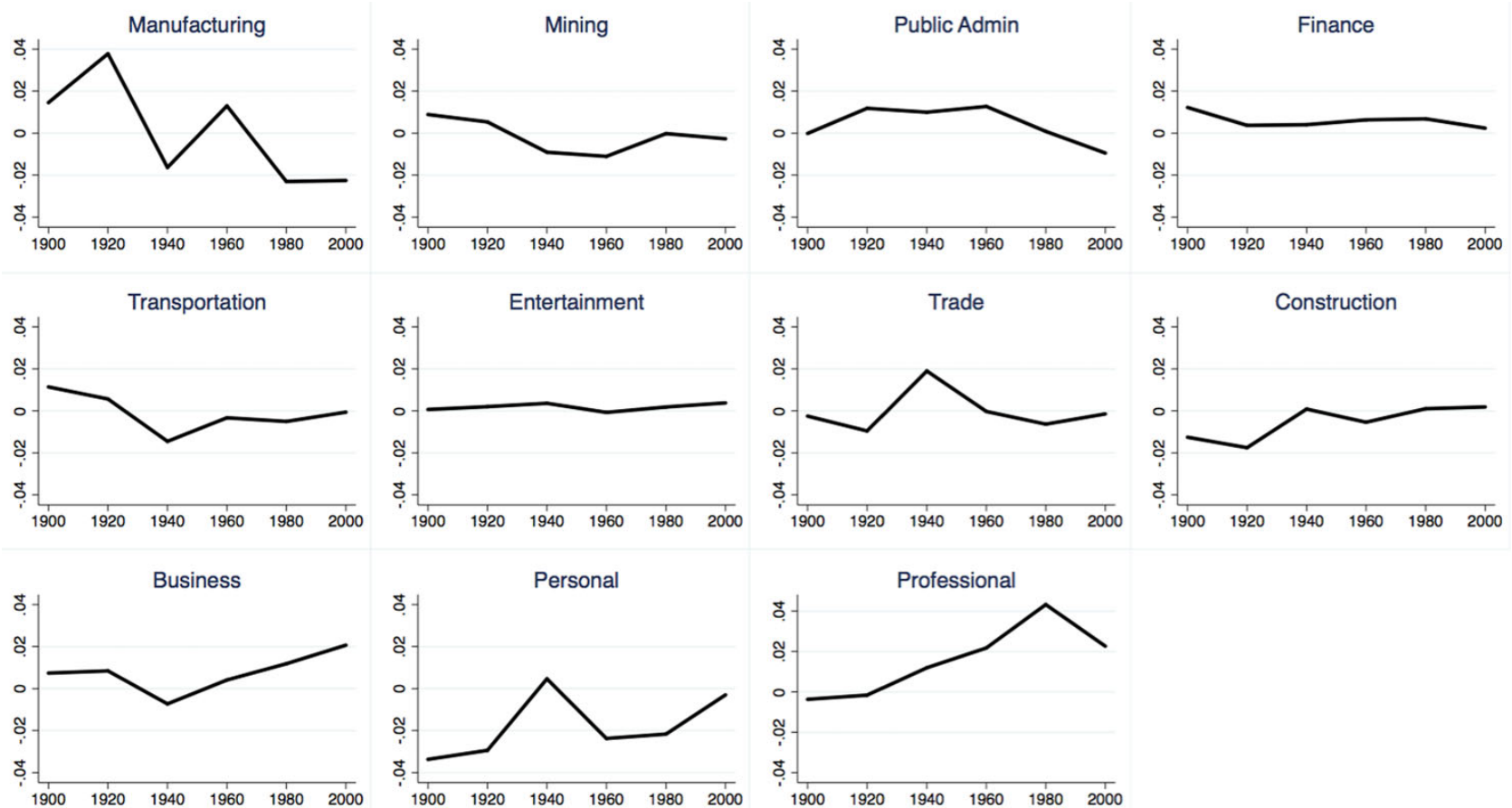
Decomposition

- Decomposition of difference in change in interactiveness between metro and non-metro areas by occupation



Decomposition

- Decomposition of difference in change in interactiveness between metro and non-metro areas by sector



Variation Between Sectors

- Define sector interactiveness as the employment-weighted mean of the interactiveness of each occupation

$$\text{Interactive}_{st} = \sum_{o} \frac{E_{ost}}{E_{st}} \text{Interactive}_{o}$$

- Using this measure, we run a regression across sectors of the share of a sector's employment in metro areas (MetroShare_{st}) on its interactiveness (Interactive_{st}) for each year separately:

$$\text{MetroShare}_{st} = \alpha_t \text{Interactive}_{st} + \varepsilon_{st}$$

Variation Within Sectors

- Estimate separate regression for each year of occupation-sector metro employment share on occupation interactiveness

$$\text{MetroShare}_{ost} = \alpha_t \text{Interactive}_o + \eta_{st} + \epsilon_{ost}$$

- where η_{st} are sector-year fixed effects

Variation Within & Between Sectors

TABLE 5. Metro employment and wagebill shares and interactiveness.

LHS	Measure	1880	1900	1920	1940	1960	1980	2000
Panel A: Between sectors								
Employment	Interactiveness	−0.130	−0.132	0.258	0.556	0.728***	0.901***	0.814***
		(0.267)	(0.239)	(0.419)	(0.405)	(0.267)	(0.200)	(0.182)
Employment	Thought	−0.722***	−1.293***	−1.806***	−0.622	0.190	0.788***	1.202***
		(0.260)	(0.261)	(0.357)	(0.487)	(0.310)	(0.274)	(0.237)
Employment	Communication	−0.459***	−0.582***	−0.645***	−0.220	0.210	0.360*	0.530**
		(0.146)	(0.151)	(0.186)	(0.266)	(0.193)	(0.208)	(0.233)
Employment	Intersocial	−0.351**	−0.481***	−0.599***	−0.117	0.101	0.268**	0.342***
		(0.135)	(0.135)	(0.165)	(0.209)	(0.133)	(0.122)	(0.109)
Employment	Individual volition	−0.157***	−0.195***	−0.268***	−0.212***	−0.115**	0.019	0.085
		(0.051)	(0.054)	(0.079)	(0.059)	(0.054)	(0.054)	(0.062)
Wagebill	Interactiveness				0.557	0.557*	0.814***	0.733***
					(0.366)	(0.283)	(0.215)	(0.201)
Panel B: Within sectors								
Employment	Interactiveness	−0.410***	−0.261**	−0.104	−0.036	0.190***	0.274***	0.317***
		(0.120)	(0.119)	(0.119)	(0.119)	(0.064)	(0.051)	(0.040)
Employment	Thought	−0.340**	−0.411***	−0.299***	−0.145	0.153***	0.227***	0.246***
		(0.134)	(0.132)	(0.093)	(0.095)	(0.049)	(0.037)	(0.039)
Employment	Communication	−0.041	−0.042	0.025	0.118	0.183***	0.168***	0.140***
		(0.144)	(0.118)	(0.098)	(0.079)	(0.036)	(0.032)	(0.038)
Employment	Intersocial	−0.030	−0.081	−0.017	0.0197	0.105***	0.0652*	0.046
		(0.130)	(0.078)	(0.058)	(0.049)	(0.032)	(0.034)	(0.048)
Employment	Individual volition	−0.095*	−0.058	−0.021	−0.016	0.006	0.015	0.027**
		(0.056)	(0.070)	(0.054)	(0.039)	(0.025)	(0.016)	(0.013)
Wagebill	Interactiveness				0.0430	0.207***	0.281***	0.311***
					(0.0874)	(0.0529)	(0.0433)	(0.0374)
Sector-year fixed effects		Yes	Yes	Yes	Yes	Yes	Yes	Yes

Panel Specification

- Regress occupation-sector metro employment share on interactions between occupation interactiveness and time

$$\text{MetroShare}_{ost} = \alpha_t \left(\text{Interactive}_o \times \text{Year}_t\right) + \mu_o + \eta_s + \delta_t + \epsilon_{ost}$$

- 1880 is the excluded category

Panel Specification

TABLE 6. Metro area employment shares and interactiveness, within-sector and within-occupation.

	Metro employment share							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Interactiveness × 1900	−0.261**	0.104	−0.001	0.260	0.076	0.126	−0.032	0.119
	(0.117)	(0.162)	(0.121)	(0.224)	(0.159)	(0.198)	(0.092)	(0.158)
Interactiveness × 1920	−0.104	0.187	0.019	0.428	0.328	0.214	−0.012	0.218
	(0.118)	(0.218)	(0.198)	(0.273)	(0.198)	(0.234)	(0.102)	(0.203)
Interactiveness × 1940	−0.04	0.321	0.177	0.534*	0.424**	0.405	0.012	0.409
	(0.119)	(0.235)	(0.231)	(0.286)	(0.210)	(0.251)	(0.124)	(0.157)
Interactiveness × 1960	0.190**	0.485***	0.331*	0.756***	0.578***	0.563***		
	(0.064)	(0.185)	(0.180)	(0.243)	(0.200)	(0.215)		
Interactiveness × 1980	0.274***	0.560***	0.449***	0.777***	0.658***	0.651***	0.295***	0.634***
	(0.052)	(0.174)	(0.168)	(0.231)	(0.191)	(0.210)	(0.090)	(0.160)
Interactiveness × 2000	0.317***	0.596***	0.478***	0.798***	0.794***	0.697***	0.339***	0.684***
	(0.040)	(0.174)	(0.169)	(0.233)	(0.196)	(0.227)	(0.091)	(0.157)
Observations	56,760	56,760	50,180	42,460	23,189	31,133	38,647	42,653
Occupation fixed effects		Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sector-year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Married only sample			Yes					
Single only sample				Yes				
Manufacturing only					Yes			
Services only						Yes		
No more skilled MSAs							Yes	
No less skilled MSAs								Yes

Communication & Transport Technology

- Combine data on employment by occupation, sector and county for 1880 and 1930 with information on the spatial diffusion of the telephone and the paved highway network
  - Telephone and paved highways were virtually non-existent in 1880 and diffused rapidly from 1880-1930
  - This is the period for which we observe the largest increase in the relative interactiveness of metro areas
  - 1930 is the last year for which county identifiers are available in IPUMS and hence for which we can measure county interactiveness in IPUMs

Communication & Transport Technology

- Telephones and highways are unlikely to be randomly assigned
- Develop instruments based on institutional features of the development of the telephone and highway network

$$\triangle \text{Interactive}_c = \alpha_P \ln(\text{Phonepc}_c) + \alpha_H \text{Highwaypa}_c + X_c\alpha_X + u_c,$$

$$\ln(\text{Phonepc}_c) = \beta_P Z_{Pc} + \beta_H Z_{Hc} + X_c\beta_X + \varepsilon_c,$$

$$\text{Highwaypa}_c = \gamma_P Z_{Pc} + \gamma_H Z_{Hc} + X_c\gamma_X + \omega_c,$$

- $\ln(\text{Phonepc}_c)$ is log county 1935 residence telephones divided by 1930 population
- Highwaypa_c is 1931 highway length per county area
- X_c are controls for log county 1880 population and log area
- Z_{Pc} and Z_{Hc} are our instruments: proximity to AT&T's long distance network and Pershing highway length per county area

Communication & Transport Technology

TABLE 7. Interactiveness and improvements in communication and transport technologies.

	(1) Change in interactiveness 1880–1930	(2) Change in interactiveness 1880–1930	(3) Log phones per capita 1935	(4) Highways per km 1931	(5) Change in interactiveness 1880–1930
Highways per km	0.007	0.086***			
	(0.004)	(0.028)			
Log phones per capita	0.022***	0.083***			
	(0.002)	(0.019)			
Log area	0.007***	0.010***	−0.013**	−0.030***	0.007***
	(0.001)	(0.001)	(0.005)	(0.003)	(0.001)
Log population 1880	0.004***	0.002*	0.006*	0.016***	0.004***
	(0.001)	(0.001)	(0.003)	(0.002)	(0.007)
Pershing highways per km			−0.113**	0.274***	0.015**
			(0.055)	(0.032)	(0.005)
Log remoteness from long distance outlet			−0.063***	0.008**	−0.005***
			(0.009)	(0.004)	(0.001)
Observations	2467	2467	2467	2509	2509
R-squared	0.12	0.19	0.02	0.19	0.09
Estimation	OLS	2SLS	OLS	OLS	OLS
Specification	Second-stage	Second-stage	First-stage	First-stage	Reduced-form
F-statistic instruments			26.35	38.4	14.05
Underidentification test (Kleibergen–Paap LM statistic)		35.63			
Weak identification test (Kleibergen–Paap *F*-statistic)		18.61			

Conclusions

- We develop a new methodology for measuring the production tasks undertaken in urban and rural areas
- We provide the first evidence on task specialization in urban and rural areas in the United States from 1880-2000.
- Results suggest a dynamic view of cities in which the nature of agglomeration changes over time
  - 1880 tasks: “Braid,” “Sew,” “Stretch” and “Thread”
  - 2000 tasks: “Analyze,” “Advise, “Confer” and “Report”
- Evidence of a secular change in the organization of economic activity within industries
- Increased importance of “interactive” tasks in cities (thought, communication and intersocial activity)
- Related to new communication and transport technologies (telephone and paved highways)

Thank You