BUSN 33946 & ECON 35101

International Macroeconomics and Trade

Jonathan Dingel

Autumn 2020, Week 10

The University of Chicago Booth School of Business

Today: Spatial Sorting of Skills and Sectors

My goal is to tackle three questions:

- Why should we care about spatial sorting?
- How should we characterize skills and sectors?
- What tools are relevant for building and estimating models?

We'll discuss three papers in some detail:

- Davis & Dingel - A Spatial Knowledge Economy
- Davis & Dingel - The Comparative Advantage of Cities
- Diamond - The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980-2000

Spatial distributions of skills and sectors

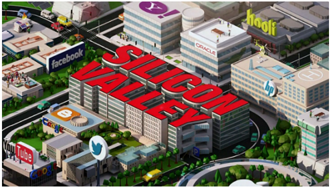
Why should we care about the spatial distributions of skills and sectors?

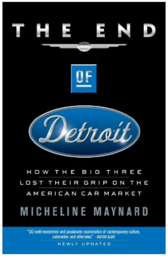
1. They vary a lot
2. They covary with city characteristics
3. They're often the basis for identification
4. They should help us understand how cities work

Spatial distributions of skills and sectors

- Public discussion describes US cities in terms of skills and sectors
- Ranking cities by educational attainment is a popular media exercise

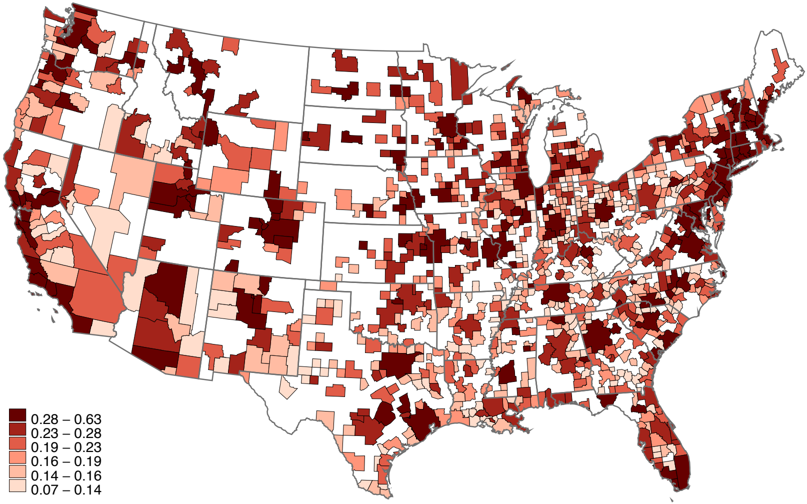
- Place names are shorthand for sectors





Educational attainment varies a lot across cities

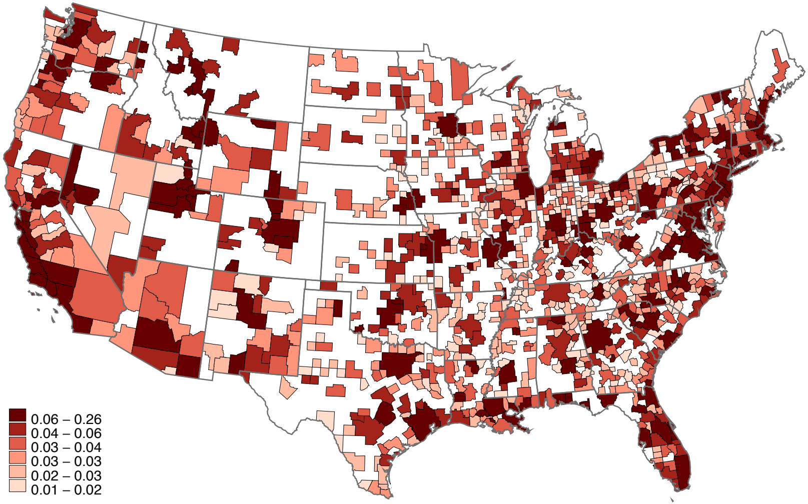
Share of population 25 and older with bachelor's degree or higher



DATA SOURCE: American Community Survey, 2005-2009, Series S1501 PLOT: CBSAs for maptile

Sectoral composition varies a lot across cities

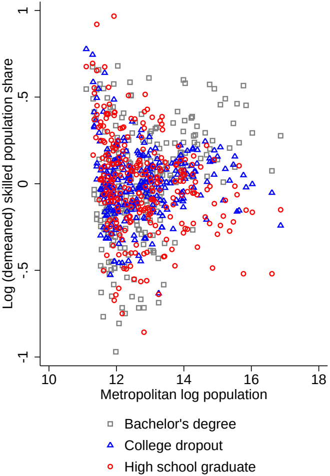
Employment share of Professional, Scientific, and Technical Services

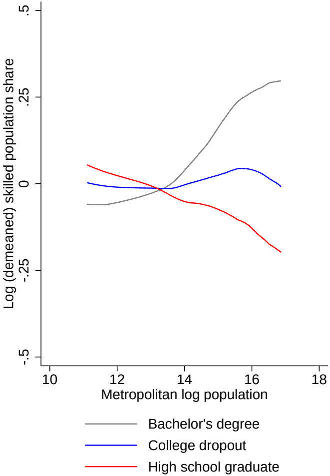


DATA SOURCE: County Business Patterns, 2009, NAICS 54   PLOT: CBSAs for maptile

They covary with city characteristics

Populations of three educational groups across US metropolitan areas

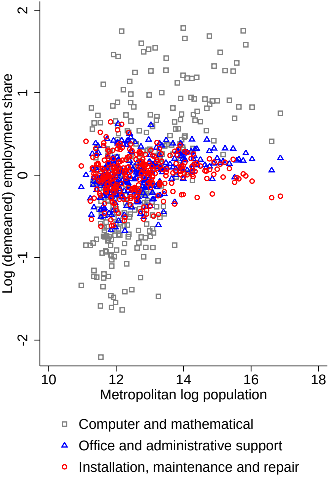


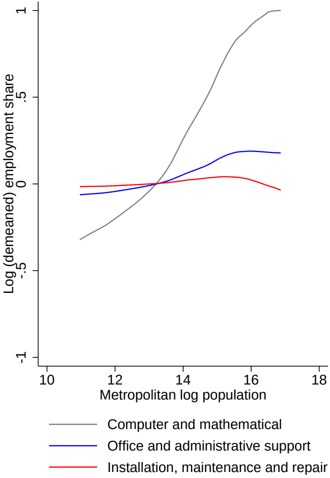


Data source: 2000 Census of Population microdata via IPUMS-USA

They covary with city characteristics

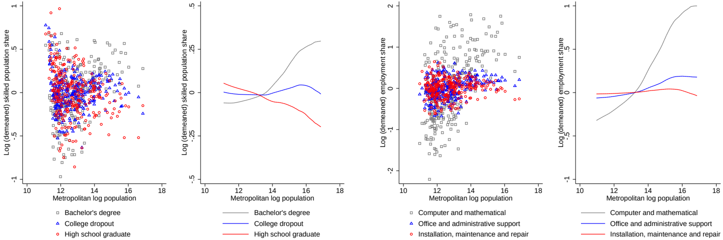
Employment in three occupations across US metropolitan areas





DATA SOURCE: Occupational Employment Statistics 2000

They covary with city characteristics



Skills and sectors are strongly linked to cities' sizes

(a) Confounds inference: Agglomeration benefits vs compositional effects

(b) Confounds counterfactuals: Making NYC 10x larger raises finance's share of national employment and GDP

They're often the basis for identification

Recent JMPs by Notowidigdo, Diamond, and Yagan

- Theory: all locations produce a homogeneous good
- Empirics: exploit variation in industrial composition to estimate model parameters via shifts in local labor demand
- Shift-share instrument: local composition $\times$ national changes

What variation does the instrument exploit?

- Skill mix vs industrial mix (e.g. endogenous local SBTC - Beaudry, Doms, Lewis 2010)
- City characteristics covarying with skills and sectors highlight exclusion-restriction assumptions

They should help us understand how cities work

- Why do different people and different businesses locate in different places?
- The answers should be crucial to understanding how cities work
- Which elements of the Marshallian trinity imply we'll find finance and dot-coms in big cities?
- Coagglomeration (Ellison Glaeser Kerr 2010) and heterogeneous agglomeration (Faggio, Silva, Strange 2015) can provide clues
- Theory is laggard: Most models of sectoral composition are polarized, with *specialized* cities that have only one tradable sector and *perfectly diversified* cities that have all the tradable sectors (Helsley and Strange 2014)

Spatial distributions of skills and sectors

How should we characterize skills and sectors?

- Important question for both theory and empirics

A richer depiction of firms and workers improves realism, but...

- more types threaten to make theoretical models intractable
- more types increase the burden of finding instruments

While the trade-offs are specific to the research question under investigation, we can start by asking: Are two skill groups enough?

Spatial equilibrium with two skill groups

A simple starting point

1. Two skill groups, $s \in \{L, H\}$
2. Spatial equilibrium: $U_s(A_c, w_{s,c}, p_c) = U_s(A_{c'}, w_{s,c'}, p_{c'}) \ \forall c, c' \ \forall s$
3. Homotheticity: $U_s(A_c, w_{s,c}, p_c) = \frac{w_{s,c}}{A_c p_c}$

These jointly imply that relative wages are spatially invariant

$$\frac{w_{H,c}}{A_c p_c} = \frac{w_{H,c'}}{A_{c'} p_{c'}} \quad \text{and} \quad \frac{w_{L,c}}{A_c p_c} = \frac{w_{L,c'}}{A_{c'} p_{c'}}$$

$$\Rightarrow \frac{w_{H,c}}{w_{L,c}} = \frac{w_{H,c'}}{w_{L,c'}} \ \forall c, c'$$

Spatial equilibrium with two skill groups

A simple starting point

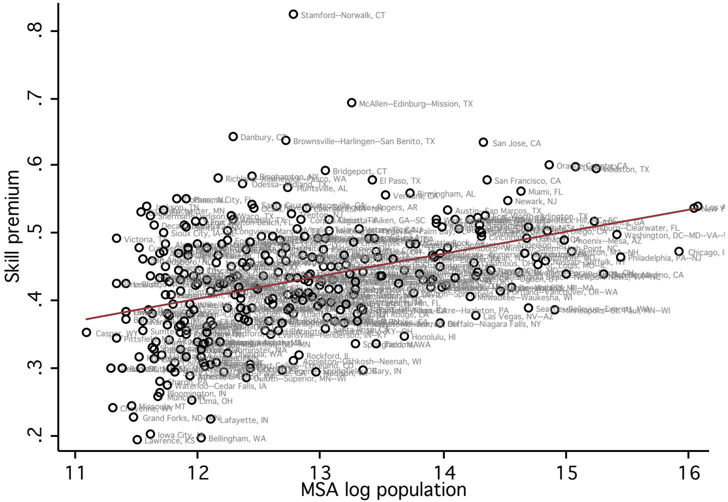
1. Two skill groups, $s \in \{L, H\}$
2. Spatial equilibrium: $U_s(A_c, w_{s,c}, p_c) = U_s(A_{c'}, w_{s,c'}, p_{c'}) \ \forall c, c' \ \forall s$
3. Homotheticity: $U_s(A_c, w_{s,c}, p_c) = \frac{w_{s,c}}{A_c p_c}$

These jointly imply that relative wages are spatially invariant

$$\frac{w_{H,c}}{A_c p_c} = \frac{w_{H,c'}}{A_{c'} p_{c'}} \quad \text{and} \quad \frac{w_{L,c}}{A_c p_c} = \frac{w_{L,c'}}{A_{c'} p_{c'}}$$
$$\Rightarrow \frac{w_{H,c}}{w_{L,c}} = \frac{w_{H,c'}}{w_{L,c'}} \ \forall c, c'$$

Spatial variation in skill premia

College wage premia are higher in larger cities

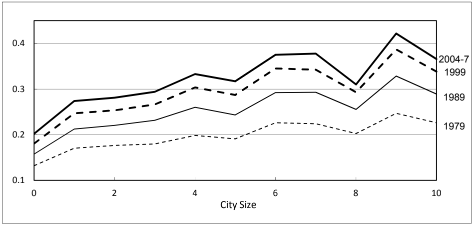


[Davis and Dingel, "A Spatial Knowledge Economy", 2013](#)

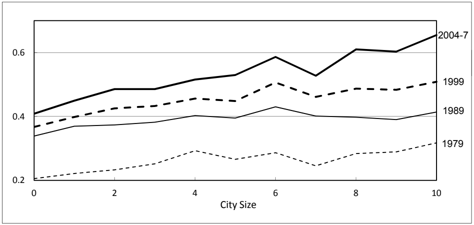
Spatial variation in skill premia

This pattern is getting stronger over time

Panel A: Fraction College or More by City Size



Panel B: College Log Wage Premium by City Size



Baum-Snow and Pavan, “Inequality and City Size”, 2013

How to proceed?

The data reject our simple model; skill premia are higher in larger cities

$$\frac{w_{H,c}}{w_{L,c}} \neq \frac{w_{H,c'}}{w_{L,c'}}$$

Three possible routes to take

1. Non-homothetic preferences
2. Upward-sloping local labor supplies
3. More than two skill groups

How to proceed?

The data reject our simple model; skill premia are higher in larger cities

$$\frac{w_{H,c}}{w_{L,c}} \neq \frac{w_{H,c'}}{w_{L,c'}}$$

Three possible routes to take

1. Non-homothetic preferences (Black, Kolesnikova, Taylor 2009)
2. Upward-sloping local labor supplies (Topel, Moretti, Diamond)
3. More than two skill groups (My focus)

How to proceed?

The data reject our simple model; skill premia are higher in larger cities

$$\frac{w_{H,c}}{w_{L,c}} \neq \frac{w_{H,c'}}{w_{L,c'}}$$

Three possible routes to take

1. Non-homothetic preferences (Black, Kolesnikova, Taylor 2009)

$\Rightarrow$ Do more skilled people find big cities less attractive for consumption? (Albouy, Ehrlich, Liu 2015, Handbury 2012)

2. Upward-sloping local labor supplies (Topel, Moretti, Diamond)
3. More than two skill groups (My focus)

How to proceed?

The data reject our simple model; skill premia are higher in larger cities

$$\frac{w_{H,c}}{w_{L,c}} \neq \frac{w_{H,c'}}{w_{L,c'}}$$

Three possible routes to take

1. Non-homothetic preferences (Black, Kolesnikova, Taylor 2009)

⇒ Do more skilled people find big cities less attractive for consumption? (Albouy, Ehrlich, Liu 2015, Handbury 2012)

2. Upward-sloping local labor supplies (Topel, Moretti, Diamond)

⇒ Relative prices and quantities imply higher relative demand for skilled in larger cities

3. More than two skill groups (My focus)

How to proceed?

The data reject our simple model; skill premia are higher in larger cities

$$\frac{w_{H,c}}{w_{L,c}} \neq \frac{w_{H,c'}}{w_{L,c'}}$$

Three possible routes to take

1. Non-homothetic preferences (Black, Kolesnikova, Taylor 2009)

⇒ Do more skilled people find big cities less attractive for consumption? (Albouy, Ehrlich, Liu 2015, Handbury 2012)

2. Upward-sloping local labor supplies (Topel, Moretti, Diamond)

⇒ Relative prices and quantities imply higher relative demand for skilled in larger cities

3. More than two skill groups (My focus)

Both 2 and 3 push us towards thinking about the complementarity between agglomeration and skills

A continuum of skills

Recent research works with a continuum of skills

- High-dimensional: Infinite types of individuals
- One-dimensional: Skills are ordered

A few reasons to take this route

1. Dichotomous results depend on dichotomous definitions
2. Broad categories miss important variation
3. Continuum case can be quite tractable

Two types in theory and practice

Two-type models can be simple – but what about two-type empirics?

- Omit types: Our plot of college wage premia was bachelor's degrees vs HS diplomas – use only 45% of population to test price prediction
- Convert quantities to "equivalents": "one person with some college is equivalent to a total of 0.69 of a high school graduate and 0.29 of a college graduate" (Katz & Murphy 1992, p.68)

Results may be sensitive to dichotomous definitions

- Diamond (2016): "A MSA's share of college graduates in 1980 is positively associated with larger growth in its share of college workers from 1980 to 2000"
- Baum-Snow, Freedman, Pavan (2015): "Diamond's result does not hold for CBSAs if those with some college education are included in the skilled group."

Two types in theory and practice

Two-type models can be simple – but what about two-type empirics?

- Omit types: Our plot of college wage premia was bachelor's degrees vs HS diplomas – use only 45% of population to test price prediction
- Convert quantities to "equivalents": "one person with some college is equivalent to a total of 0.69 of a high school graduate and 0.29 of a college graduate" (Katz & Murphy 1992, p.68)

Results may be sensitive to dichotomous definitions

- Diamond (2016): "A MSA's share of college graduates in 1980 is positively associated with larger growth in its share of college workers from 1980 to 2000"
- Baum-Snow, Freedman, Pavan (2015): "Diamond's result does not hold for CBSAs if those with some college education are included in the skilled group."

Dichotomous approach misses relevant variation

- In labor economics, the canonical two-skill model "is largely silent on a number of central empirical developments of the last three decades", such as wage polarization and job polarization (Acemoglu and Autor 2011)
- There is systematic variation across cities in terms of finer observable categories: population elasticities for high school graduates (.925), associate's degree (0.997), bachelor's degree (1.087), and professional degree (1.113) (Davis and Dingel 2020)

Do broad categories miss important variation?

Contrasting views

- "Workers in cities with a well-educated labor force are likely to have unobserved characteristics that make them more productive than workers with the same level of schooling in cities with a less-educated labor force. For example, a lawyer in New York is likely to be different from a lawyer in El Paso, TX." (Moretti 2004, p.2246)
- "Within broad occupation or education groups, there appears to be little sorting on ability" (de la Roca, Ottaviano, Puga 2014)

Data sources for "no sorting" evidence

- NLSY79: Longitudinal study of about 11,000 US individuals
- Spanish tax data 2004-2009: 150,375 workers (de la Roca and Puga 2017)

Do broad categories miss important variation?

National Longitudinal Survey of Youth 1979

- Bacolod, Blum, Strange (2009): "The mean AFQT scores do not vary much across [four] city sizes" within occupational categories
- BBS observe only one sales person in MSAs with 0.5m – 1.0m residents (10th and 90th percentiles of AFQT are equal) ▸ table
- Baum-Snow & Pavan (2012): Structural estimation of finite-mixture model implies "sorting on unobserved ability within education group...contribute little to observed city size wage premia."
- BSP use NLSY79 data on 1754 white men; 583 have bachelor's degree or more; college wage premia don't rise with city size ▸ table

Spanish tax data (de la Roca and Puga 2017)

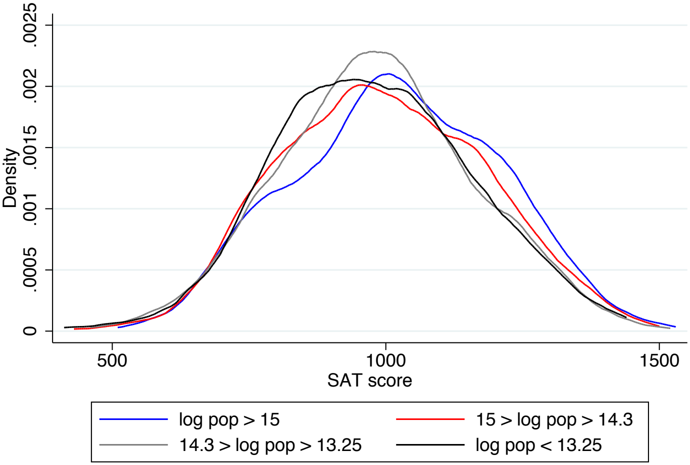
- 150,375 workers and 37,443 migrations
- Identification of sorting relies on random migration conditional on observables
- Little sorting within five educational categories

Bringing more data to bear on sorting

- Baccalaureate and Beyond tracks a cohort graduating from four-year colleges in 1993
- In 2003, look at 2300 white individuals who obtained no further education after bachelor's degree and now live in a PMSA
- Look at variation in SAT scores across cities – all variation is within the finest age-race-education cell in typical public data sets
- Mean SAT score in metros with more than 3.25m residents is 40 points higher than metros with fewer than 0.57m residents

Sorting within observable demographic cells

- Mean SAT score in metros with more than 3.25m residents is 40 points higher than metros with fewer than 0.57m residents
- Full distribution suggests stochastic dominance



The continuum case

Why work with a continuum?

- Evidence for sorting on characteristics that are typically not observed
- Need at least five types to capture sorting on observables in the sense of de la Roca and Puga (2017)
- Modeling a finite, particular number of types is potentially painful

Continuum case can be quite tractable

- Recent work: Behrens, Duranton, Robert-Nicoud (2014), Davis and Dingel (2019, 2020), Gaubert (2018), Behrens and Robert-Nicoud (*Handbook* 2015)
- These papers rely on tools from the assignment literature
- Assignments of individuals/firms to cities, with endogenous city characteristics determined in equilibrium
- Davis and Dingel (2020) speak to both skills and sectors

Davis & Dingel - A Spatial Knowledge Economy

We have models of:

- Knowledge spillovers as a pure externality (one interpretation of Henderson 1974, Black 1999, Lucas 2001)
- Endogenous exchange of ideas in a single (or symmetric) location(s) (Helsley and Strange 2004, Berliant, Reed III, and Wang 2006, Berliant and Fujita 2008, Lucas and Moll 2011)

Our contribution:

- Introduce a model of a system of cities in which costly idea exchange is the agglomeration force
- Our model replicates a broad set of established facts about the cross section of cities
- We provide a spatial-equilibrium explanation of why skill premia are higher in larger cities and how this emerges from symmetric fundamentals

Model summary

Our model's core components:

- Spatial equilibrium – zero mobility costs
- Heterogeneous workers – continuum of abilities
- Two sectors:
  - Tradables: Labor heterogeneity matters for productivity
  - Non-tradables: Homogeneous productivity
- Skilled tradables sector has local learning opportunities
  - Workers choose to spend time exchanging ideas
  - Gains from interactions increasing in own ability and peers' ability
- Congestion costs make housing more expensive in larger cities
- Workers choose locations, occupations, and time spent exchanging ideas

Preferences and congestion

- Preferences: Unit demand for housing and $\bar{n}$-unit demand for non-tradable:

$$V(p_{n,c}, p_{h,c}, y) = y - p_{n,c}\bar{n} - p_{h,c}.$$

- Each individual in a city of population L_c pays a net urban cost (in units of the numeraire) of

$$p_{h,c} = \theta L_c^{\gamma}$$

- Individuals are perfectly mobile across cities and jobs, so their locational and occupational choices maximize $V(p_{n,c}, p_{h,c}, y)$.

Production

- An individual can produce tradables (t) or non-tradables (n)
- An individual working in sector σ earns income equal to the value of her output, which is

$$y = \begin{cases} p_{n,c} & \text{if } \sigma = n \\ \tilde{z}(z, Z_c) & \text{if } \sigma = t \end{cases}$$

- Tradables production depends on own ability (z), time spent producing (β), time spent exchanging ideas ($1-\beta$), and local learning opportunities (Z_c):

$$\tilde{z}(z, Z_c) = \max_{\beta \in [0,1]} B(1-\beta, z, Z_c)$$

Idea exchange

Tradables production:

$$\tilde{z}(z, Z_c) = \max_{\beta \in [0,1]} B(1-\beta, z, Z_c)$$

- Scalar Z_c depends on time-allocation decisions of all agents in c.
- Denote idea-exchange time of ability z in city c by $1-\beta_{z,c}$
- Denote local ability distribution $\mu(z, c)$, where $\frac{\mu(z,c)}{\mu(z)}$ is the share of z in c.

$$Z_c = Z(\{1-\beta_{z,c}\}, \{\mu(z, c)\}).$$

- Denote total time devoted to learning by tradables producers in city c by M_c

$$M_c = L \int_{z:\sigma(z)=t} (1-\beta_{z,c})\mu(z, c) dz.$$

Idea exchange: General assumptions

- **Assumption 1**. The production function for tradables $B(1-\beta, z, Z_c)$ is continuous, strictly concave in $1-\beta$, strictly increasing in z, and increasing in Z_c. $B(1-\beta, z, 0) = \beta z$ and $B(0, z, Z_c) = z \; \forall z$.
- **Assumption 2**. Tradables output $\tilde{z}(z, Z_c)$ is supermodular and is strictly supermodular on $\otimes \equiv \{(z, Z) : \tilde{z}(z, Z) > z\}$.
- **Assumption 3**. The idea-exchange functional $Z(\{1-\beta_{z,c}\}, \{L \cdot \mu(z,c)\})$ is continuous, equal to zero if $M_c = 0$, and bounded above by $\sup\{z : 1-\beta_{z,c} > 0, \mu(z,c) > 0\}$. If $M_c > M_{c'}$ and $\{(1-\beta_{z,c})\mu(z,c)\}$ stochastically dominates $\{(1-\beta_{z,c'})\mu(z,c')\}$, then $Z(\{1-\beta_{z,c}\}, \{L \cdot \mu(z,c)\}) > Z(\{1-\beta_{z,c'}\}, \{L \cdot \mu(z,c')\})$.

Idea exchange: Special case

For some of our analysis, we focus on particular functional forms for $B(\cdot)$ and $Z(\cdot)$:

$$B(1-\beta, z, Z_c) = \beta z(1 + (1-\beta) A Z_c z)$$

$$Z(\{(1-\beta_{z,c}), \mu(z,c)\}) = (1 - \exp(-\nu M_c))\, \bar{z}_c$$

$$\bar{z}_c = \begin{cases} \int_{z:\sigma(z)=t} \frac{(1-\beta_{z,c})z}{\int_{z:\sigma(z)=t}(1-\beta_{z,c})\mu(z,c)dz} \mu(z,c)dz & \text{if } M_c > 0 \\ 0 & \text{otherwise} \end{cases}$$

- Random matching: Probability of encounter during each moment of time spent seeking idea exchanges is $(1 - \exp(-\nu M_c))$
- M_c is the total time devoted to idea exchange
- $\bar{z}_c$ is the average ability of the individuals encountered

Two lemmas

Lemma (Comparative advantage)

Suppose that Assumption 1 holds. There is an ability level z_m such that individuals of greater ability produce tradables and individuals of lesser ability produce non-tradables.

$$\sigma(z) = \begin{cases} t & \text{if } z > z_m \\ n & \text{if } z < z_m \end{cases}$$

Lemma (Spatial sorting of tradables producers engaged in idea exchange)

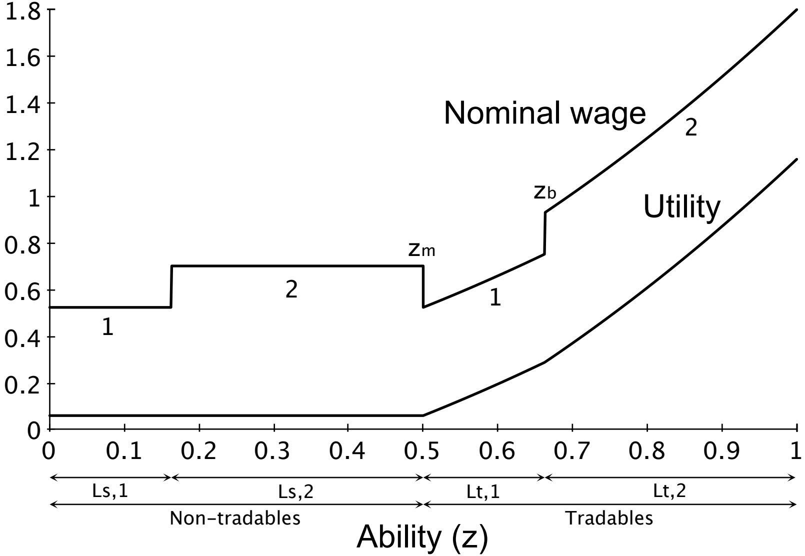
Suppose that Assumption 2 holds. For $z > z' > z_m$, if $\mu(z, c) > 0$, $\mu(z', c') > 0$, $\beta(z, Z_c) < 1$, and $\beta(z', Z_{c'}) < 1$, then $Z_c \geq Z_{c'}$.

Spatial equilibrium

Proposition (Heterogeneous cities' characteristics)

Suppose that Assumptions 1 and 2 hold. In any equilibrium, a larger city has higher housing prices, higher non-tradables prices, a better idea-exchange environment, and higher-ability tradables producers. If $L_c > L_{c'}$ *in equilibrium, then* $p_{h,c} > p_{h,c'}$, $p_{n,c} > p_{n,c'}$, $Z_c > Z_{c'}$, *and* $z > z' > z_m \Rightarrow \mu(z, c)\mu(z', c') \geq \mu(z, c')\mu(z', c) = 0$.

Spatial equilibrium: Two-city example



Differences in average wages

- Differences in tradables producers' wages are the sum of three components: composition, learning, and compensation effects
- Denote z_b the "boundary" ability of indifferent tradables producer
- Define inframarginal learning $\Delta(z,c,c') \equiv [\tilde{z}(z,Z_c) - \tilde{z}(z,Z_{c'})] - [\tilde{z}(z_b,Z_c) - \tilde{z}(z_b,Z_{c'})]$
- Define the density of tradables producers' abilities in city c by $\tilde{\mu}(z,c) \equiv \frac{\mu(z,c)}{\int_{z':\sigma(z')=t} \mu(z',c)dz'}$

$$\bar{w}_c - \bar{w}_{c'} \equiv \frac{\int_{z:\sigma(z)=t} \tilde{z}(z,Z_c)\mu(z,c)dz}{\int_{z:\sigma(z)=t} \mu(z,c)dz} - \frac{\int_{z:\sigma(z)=t} \tilde{z}(z,Z_{c'})\mu(z,c')dz}{\int_{z:\sigma(z)=t} \mu(z,c')dz}$$

$$= \underbrace{\int_{z_m}^{\infty} [\tilde{\mu}(z,c) - \tilde{\mu}(z,c')]\tilde{z}(z,Z_{c'})dz}_{\text{composition}} + \underbrace{\int_{z_m}^{\infty} \tilde{\mu}(z,c)\Delta(z,c,c')dz}_{\text{inframarginal learning}}$$

$$+ \underbrace{p_{n,c} - p_{n,c'}}_{\text{compensation}}$$

Skill premia

- Define a city's observed skill premium as its average tradables wage divided by its (common) non-tradables wage, $\frac{\bar{w}_c}{p_{n,c}}$
- When a tradables producer of ability z_b is indifferent between cities c and c', this skill premium is higher in c if and only if

$$\underbrace{\int_{z_m}^{\infty}[\tilde{\mu}(z,c)-\tilde{\mu}(z,c')]\tilde{z}(z,Z_{c'})dz}_{\text{composition}}+\underbrace{\int_{z_m}^{\infty}\tilde{\mu}(z,c)\Delta(z,c,c')dz}_{\text{inframarginal learning}}$$
$$\geq \underbrace{(p_{n,c}-p_{n,c'})\left(\frac{\bar{w}_{c'}}{p_{n,c'}}-1\right)}_{\text{relative compensation}}$$

- Helpful to define a production-function property:

Condition

The ability elasticity of tradable output, $\frac{\partial \ln \tilde{z}(z,Z_c)}{\partial \ln z}$, is non-decreasing in z and Z_c.

Larger cities have higher skill premia

Proposition (Skill premia)

Suppose that Assumptions 1 and 2 hold. In an equilibrium in which the smallest city has population L_1 and the second-smallest city has population $L_2 > L_1$,

1. *if the ability distribution is decreasing, $\mu'(z) \leq 0$, $\tilde{z}(z, Z_c)$ is log-convex in z, and $\tilde{z}(z, Z_c)$ is log-supermodular, then* $\frac{\bar{w}_2}{p_{n,2}} > \frac{\bar{w}_1}{p_{n,1}}$;
2. *if the ability distribution is Pareto, $\mu(z) \propto z^{-k-1}$ for $z \geq z_{\min}$ and $k > 0$, and the production function satisfies Condition 1, then* $\frac{\bar{w}_2}{p_{n,2}} > \frac{\bar{w}_1}{p_{n,1}}$;
3. *if the ability distribution is uniform, $z \sim U(z_{\min}, z_{\max})$, the production function satisfies Condition 1, and* $\frac{L_2 - L_1}{L_1^2} > \frac{1}{L} \frac{(1-\bar{n})(z_{\max} - z_{\min})}{z_{\min} + \bar{n}(z_{\max} - z_{\min})}$, *then* $\frac{\bar{w}_2}{p_{n,2}} > \frac{\bar{w}_1}{p_{n,1}}$.

Larger cities have higher skill premia

- The three cases in Proposition 2 trade off stronger assumptions about the production function with weaker assumptions about the ability distribution
- Paper contains numerical results for more than two cities for special case of
  $B(1-\beta, z, Z_c) = \beta z(1 + (1-\beta)A\,(1 - \exp(-\nu M_c))\,\bar{z}_c z)$
- Paper contains illustrative example with 275 cities that quantitatively matches Zipf's law, premia-population correlation, and size-invariant housing expenditure shares

Assignment models

Many markets concern assignment problems

- Who marries whom? (Becker)
- Which worker performs which job? (Roy)
- Which country makes which goods? (Ricardo)

If relevant objects are well ordered, we can use tools from mathematics of complementarity to characterize equilibrium prices and quantities

- Supermodularity (Topkis 1998)
- Log-supermodularity (Athey 2002)

Basis for today's introduction

- Sattinger - "Assignment Models of the Distribution of Earnings"
- Costinot & Vogel - "Beyond Ricardo: Assignment Models in International Trade"
- Davis & Dingel - "The Comparative Advantage of Cities"

Differentials rents model

In the spirit of Ricardo's analysis of rent, start with land and labor:

- A plot of land has fertility $\gamma \in \mathbb{R}$
- A farmer has skill $\omega \in \mathbb{R}$
- Profits are $\pi(\gamma, \omega) = p \cdot y(\gamma, \omega) - r(\gamma)$

Which farmer will use which plot of land?

- Farmers optimize: $\gamma^*(\omega) \equiv \arg\max_\gamma \pi(\gamma, \omega)$
- Equilibrium prices $r(\gamma)$ must support the equilibrium assignment of farmers to plots

Supermodularity

Definition (Supermodularity)

A function $g : \mathbb{R}^n \to \mathbb{R}$ is *supermodular* if $\forall x, x' \in \mathbb{R}^n$

$$g\left(\max\left(x, x'\right)\right) + g\left(\min\left(x, x'\right)\right) \geq g(x) + g(x')$$

where max and min are component-wise operators.

- Supermodularity means the arguments of $g(\cdot)$ are complements
- $g(x)$ is SM in (x_i, x_j) if $g(x_i, x_j; x_{-i,-j})$ is SM
- $g(x)$ is SM $\iff$ $g(x)$ is SM in (x_i, x_j) $\forall i, j$
- If g is C^2, $\frac{\partial^2 g}{\partial x_i \partial x_j} \geq 0 \iff g(x)$ is SM in (x_i, x_j)

Supermodularity implies PAM

Positive assortative matching:

- If $g(x,t)$ is supermodular in (x,t), then $x^*(t) \equiv \arg\max_{x \in X} g(x,t)$ is increasing in t
- If $y(\gamma,\omega)$ is strictly supermodular (fertility and skill are complements), then $\gamma^*(\omega)$ is increasing
- More skilled farmers are assigned to more fertile land

Why? Suppose not:

- Suppose $\exists \omega > \omega', \gamma > \gamma'$ where $\gamma' \in \gamma^*(\omega), \gamma \in \gamma^*(\omega')$
- $\gamma' \in \gamma^*(\omega) \Rightarrow p \cdot y(\gamma',\omega) - r(\gamma') \geq p \cdot y(\gamma,\omega) - r(\gamma) \ \forall \gamma$
- $\gamma \in \gamma^*(\omega') \Rightarrow p \cdot y(\gamma,\omega') - r(\gamma) \geq p \cdot y(\gamma',\omega') - r(\gamma') \ \forall \gamma'$
- Summing: $p \cdot (y(\gamma',\omega) + y(\gamma,\omega')) \geq p \cdot (y(\gamma,\omega) + y(\gamma',\omega'))$
- Would contradict strict supermodularity of $y(\cdot)$

Ricardian trade model

Costinot and Vogel (2015) survey Ricardo-Roy models

- Ricardo: Linear production functions
- Roy: Multiple factors of production (ω)

Output in sector σ in country c is

$$Q(\sigma, c) = \int_{\Omega} A(\omega, \sigma, c) L(\omega, \sigma, c) d\omega$$

Ricardo 1817: England $= c > c' =$ Portugal and cloth $= \sigma > \sigma' =$ wine

$$A(\sigma, c)/A(\sigma', c) \geq A(\sigma, c')/A(\sigma', c')$$

Log-supermodularity (1/2)

Definition (Log-supermodularity)

A function $g : \mathbb{R}^n \to \mathbb{R}^+$ is *log-supermodular* if $\forall x, x' \in \mathbb{R}^n$

$$g\left(\max\left(x, x'\right)\right) \cdot g\left(\min\left(x, x'\right)\right) \geq g(x) \cdot g(x')$$

where max and min are component-wise operators.

- Example: $A : \Sigma \times \mathbb{C} \to \mathbb{R}^+$, where $\Sigma \subseteq \mathbb{R}$ and $\mathbb{C} \subseteq \mathbb{R}$, with $\sigma > \sigma'$ and $c > c'$

$$A(\sigma, c)A(\sigma', c') \geq A(\sigma', c)A(\sigma, c')$$

- $g(x)$ is LSM in (x_i, x_j) if $g(x_i, x_j; x_{-i,-j})$ is LSM
- $g(x)$ is LSM $\iff$ $g(x)$ is LSM in (x_i, x_j) $\forall i, j$
- $g > 0$ and g is $C^2 \Rightarrow \frac{\partial^2 \ln g}{\partial x_i \partial x_j} \geq 0 \iff g(x)$ is LSM in (x_i, x_j)

Log-supermodularity (2/2)

Three handy properties:

1. If $g, h : \mathbb{R}^n \to \mathbb{R}^+$ are log-supermodular, then gh is log-supermodular.
2. If $g : \mathbb{R}^n \to \mathbb{R}^+$ is log-supermodular, then $G(x_{-i}) \equiv \int g(x_i, x_{-i}) dx_i$ is log-supermodular.
3. If $g : \mathbb{R}^n \to \mathbb{R}^+$ is log-supermodular, then $x_i^*(x_{-i}) \equiv \arg\max_{x_i \in \mathbb{R}} g(x_i, x_{-i})$ is increasing in x_{-i}.

Assignments with factor endowments (Costinot 2009)

Primitives:

- Technologies $A(\omega, \sigma, c) = A(\omega, \sigma) \; \forall c$
- Endowments $L(\omega, \gamma_{L,c})$

Profit maximization by firms:

$$p(\sigma) \leq \min_{\omega \in \Omega} \{w(\omega, c)/A(\omega, \sigma)\}$$

$$\Omega(\sigma, c) \equiv \{\omega \in \Omega : L(\omega, \sigma, c) > 0)\} \subseteq \arg\min_{\omega \in \Omega} \{w(\omega, c)/A(\omega, \sigma)\}$$

$A(\omega, \sigma)$ is strictly log-supermodular in $(\omega, \sigma) \Rightarrow$

- $\Omega(\sigma, c)$ is increasing in σ by property 3 of LSM
- High-ω factors are employed in high-σ activities

Equilibrium:

- FPE $w(\omega, c) = w(\omega)$
- Continuum $\Rightarrow \Sigma(\omega, c) = \Sigma(\omega)$ singleton

Output quantities (Costinot 2009)

Labor market clearing:

$$\int_{\Sigma} L(\omega, \sigma, c) d\sigma = L(\omega, \gamma_{L,c}) \quad \forall \omega, c$$

$L(\omega, \gamma_{L,c})$ is strictly log-supermodular: High-$\gamma_{L,c}$ locations are relatively abundant in high-ω factors

$$\begin{aligned} Q(\sigma, c) &= \int_{\Omega} A(\omega, \sigma) L(\omega, \sigma, c) d\omega \\ &= \int_{\Omega(\sigma)} A(\omega, \sigma) L(\omega, \gamma_{L,c}) d\omega \qquad \text{by } \Sigma(\omega, c) \text{ singleton} \end{aligned}$$

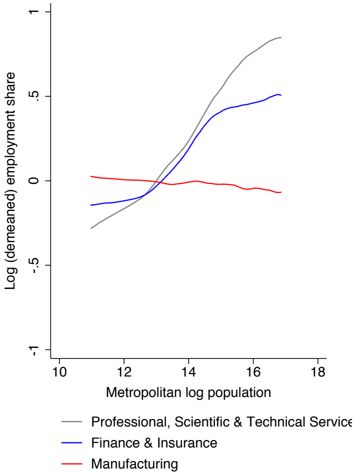
Rybczynski: $A(\omega, \sigma)$ and $L(\omega, \gamma_{L,c})$ SLSM $\Rightarrow Q(\sigma, \gamma_{L,c})$ SLSM by properties 1 and 2 of LSM

Comparative Advantage of Cities: Theory

- Davis and Dingel (2020) describe comparative advantage of cities as jointly governed by individuals' comparative advantage and locational choices
- Cities endogenously differ in TFP due to agglomeration
- More skilled individuals are more willing to pay for more attractive locations
- Larger cities are skill-abundant in equilibrium
- By individuals' comparative advantage, larger cities specialize in skill-intensive activities
- Under a further condition, larger cities are larger in all activities

Comparative Advantage of Cities: Empirics (1/2)

- Use US data on skills and sectors
- Characterize the comparative advantage of cities with two tests
- Elasticity test of variation in relative population/employment
  - Compare elasticities of different skills, sectors
  - Steeper slope in log-log plot is higher elasticity
  - Elasticities may be positive for all sectors



Comparative Advantage of Cities: Empirics (2/2)

Pairwise comparison test (LSM)

- The function $f(\omega, c)$ is log-supermodular if

$$c > c', \omega > \omega' \Rightarrow f(\omega, c) f(\omega', c') \geq f(\omega', c) f(\omega, c')$$

- Our theory says skill distribution $f(\omega, c)$ and sectoral employment distribution $f(\sigma, c)$ are log-supermodular
- For example, population of skill ω in city c is $f(\omega, c)$. Check whether, for $c > c', \omega > \omega'$,

$$\frac{f(\omega, c)}{f(\omega', c)} \geq \frac{f(\omega, c')}{f(\omega', c')}$$

Are larger cities larger in all sectors?

- Check if $c > c' \Rightarrow f(\sigma, c) \geq f(\sigma, c')$

Model components

Producers

- Skills: Continuum of skills indexed by ω (educational attainment)
- Sectors: Continuum of sectors σ (occupations, industries)
- Goods: Freely traded intermediates assembled into final good
- All markets are perfectly competitive

Places

- Cities are *ex ante* identical
- Locations within cities vary in their desirability
- TFP depends on agglomeration of "scale and skills"

$$A(c) = J\left(L, \int_{\omega \in \Omega} j(\omega) f(\omega, c) d\omega\right)$$

Individual optimization

Perfectly mobile individuals simultaneously choose

- A sector σ of employment
- A city with total factor productivity $A(c)$
- A location τ (distance from ideal) within city c

The productivity of an individual of skill ω is

$$q(c, \tau, \sigma; \omega) = A(c)T(\tau)H(\omega, \sigma)$$

Utility is consumption of the numeraire final good, which is income minus locational cost:

$$\begin{aligned} U(c, \tau, \sigma; \omega) &= q(c, \tau, \sigma; \omega)p(\sigma) - r(c, \tau) \\ &= A(c)T(\tau)H(\omega, \sigma))p(\sigma) - r(c, \tau) \end{aligned}$$

Sectoral choice

- Individuals' choices of locations and sectors are separable:

$$\arg\max_{\sigma} \underbrace{A(c)T(\tau)}_{\text{locational}} \underbrace{H(\omega,\sigma)p(\sigma)}_{\text{sectoral}} - r(c,\tau) = \arg\max_{\sigma} H(\omega,\sigma)p(\sigma)$$

- $H(\omega, \sigma)$ is log-supermodular in ω, σ and strictly increasing in ω
- Comparative advantage assigns high-ω individuals to high-σ sectors
- Absolute advantage makes more skilled have higher incomes ($G(\omega) = \max_\sigma H(\omega, \sigma)p(\sigma)$ is increasing)

Locational choice

- A location's attractiveness $\gamma = A(c)T(\tau)$ depends on c and τ
- $T'(\tau) < 0$ may be interpreted as commuting to CBD, proximity to productive opportunities, or consumption value
- More skilled are more willing to pay for more attractive locations
- Equally attractive locations have same rental price and skill type
- Location in higher-TFP city is farther from ideal desirability

$$\gamma = A(c)T(\tau) = A(c')T(\tau')$$
$$A(c) > A(c') \Rightarrow \tau > \tau'$$

- Locational hierarchy: A smaller city's locations are a subset of larger city's in terms of attractiveness: $A(c)T(0) > A(c')T(0)$

Equilibrium distributions

- Skill and sectoral distributions reflect distribution of locational attractiveness: Higher-γ locations occupied by higher-ω individuals who work in higher-σ sectors
- Locational hierarchy $\Rightarrow$ hierarchy of skills and sectors
- The distributions $f(\omega, c)$ and $f(\sigma, c)$ are log-supermodular if and only if the supply of locations with attractiveness γ in city c, $s(\gamma, c)$, is log-supermodular

$$s(\gamma, c) = \begin{cases} \frac{1}{A(c)} V\left(\frac{\gamma}{A(c)}\right) & \text{if } \gamma \leq A(c)T(0) \\ 0 & \text{otherwise} \end{cases}$$

where $V(z) \equiv -\frac{\partial}{\partial z} S\left(T^{-1}(z)\right)$ is the supply of locations with innate desirability τ such that$T(\tau) = z$

When is $s(\gamma, c)$ log-supermodular?

Proposition (Locational attractiveness distribution)

The supply of locations of attractiveness γ in city c, $s(\gamma, c)$, is log-supermodular if and only if the supply of locations with innate desirability $T^{-1}(z)$ within each city, $V(z)$, has a decreasing elasticity.

- Links each city's exogeneous distribution of locations, $V(z)$, to endogenous equilibrium locational supplies $s(\gamma, c)$
- Informally, ranking relative supplies is ranking elasticities of $V(z)$

$$s(\gamma, c) \propto V\left(\frac{\gamma}{A(c)}\right) \Rightarrow \frac{\partial \ln s(\gamma, c)}{\partial \ln \gamma} = \frac{\partial \ln V\left(\frac{\gamma}{A(c)}\right)}{\partial \ln z}$$

- Satisfied by the canonical von Thünen/monocentric geography

The Comparative Advantage of Cities

Corollary (Skill and employment distributions)

If $V(z)$ has a decreasing elasticity, then $f(\omega, c)$ and $f(\sigma, c)$ are log-supermodular.

- Larger cities are skill-abundant in equilibrium (satisfies Assumption 2 in Costinot 2009)
- Locational productivity differences are Hicks-neutral in equilibrium (satisfies Definition 4 in Costinot 2009)
- $H(\omega, \sigma)$ is log-supermodular (Assumption 3 in Costinot 2009)

Corollary (Output and revenue distributions)

If $V(z)$ has a decreasing elasticity, then sectoral output $Q(\sigma, c)$ and revenue $R(\sigma, c) \equiv p(\sigma)Q(\sigma, c)$ are log-supermodular.

The Comparative Advantage of Cities

Corollary (Skill and employment distributions)

If $V(z)$ has a decreasing elasticity, then $f(\omega, c)$ and $f(\sigma, c)$ are log-supermodular.

- Larger cities are skill-abundant in equilibrium (satisfies Assumption 2 in Costinot 2009)
- Locational productivity differences are Hicks-neutral in equilibrium (satisfies Definition 4 in Costinot 2009)
- $H(\omega, \sigma)$ is log-supermodular (Assumption 3 in Costinot 2009)

Corollary (Output and revenue distributions)

If $V(z)$ has a decreasing elasticity, then sectoral output $Q(\sigma, c)$ and revenue $R(\sigma, c) \equiv p(\sigma)Q(\sigma, c)$ are log-supermodular.

When are bigger cities bigger in everything?

We identify a sufficient condition under which a larger city has a larger supply of locations of a given attractiveness

Proposition

For any $A(c) > A(c')$, if $V(z)$ has a decreasing elasticity that is less than -1 at $z = \frac{\gamma}{A(c)}$, $s(\gamma, c) \geq s(\gamma, c')$.

Now apply this result to the least-attractive locations, so larger cities are larger in all skills and sectors

Corollary

If $V(z)$ has a decreasing elasticity that is less than -1 at $z = \frac{K^{-1}(\underline{\omega})}{A(c)} = \frac{\underline{\gamma}}{A(c)}$, $A(c) > A(c')$ implies $f(\omega, c) \geq f(\omega, c')$ and $f(M(\omega), c) \geq f(M(\omega), c')$ $\forall \omega \in \Omega$.

Empirical tests

Our theory says $f(\omega, c)$ and $f(\sigma, c)$ are log-supermodular.

Two tests to describe skill and sectoral employment distributions:

- Elasticities test:
  - Compare population elasticities estimated via linear regression
  - More skilled types should have higher population elasticities
  - More skill-intensive sectors should have higher population elasticities
- Pairwise comparisons test:
  - Compare any two cities and any two skills/sectors
  - Relative population of more skilled should be higher in larger city: $c > c', \omega > \omega' \Rightarrow \frac{f(\omega,c)}{f(\omega',c)} \geq \frac{f(\omega,c')}{f(\omega',c')}$
  - Relative employment of more skill-intensive sector should be higher in larger city: $c > c', \sigma > \sigma' \Rightarrow \frac{f(\sigma,c)}{f(\sigma',c)} \geq \frac{f(\sigma,c')}{f(\sigma',c')}$
  - "Bin" together cities ordered by size and compare bins similarly

Data: Skills

- Proxy skills by educational attainment, assuming $f(edu, \omega, c)$ is log-supermodular in edu and ω (Costinot and Vogel 2010)
- Following Acemoglu and Autor (2011), we use a minimum of three skill groups.

Skill (3 groups)	Population share	Share US-born	Skill (9 groups)	Population share	Share US-born
High school or less	.37	.78	Less than high school	.04	.29
			High school dropout	.08	.73
			High school graduate	.25	.88
Some college	.31	.89	College dropout	.23	.89
			Associate's degree	.08	.87
Bachelor's or more	.32	.85	Bachelor's degree	.20	.86
			Master's degree	.08	.84
			Professional degree	.03	.81
			Doctorate	.01	.72

NOTES: Sample is individuals 25 and older in the labor force residing in 270 metropolitan areas. Data source: 2000 Census of Population microdata via IPUMS-USA

Data: Sectors

- 19 industrial categories (2-digit NAICS, 2000 County Business Patterns)
- 22 occupations (2-digit SOC, 2000 BLS Occupational Employment Statistics)
- Infer sectors' skill intensities from average years of schooling of workers employed in them

SOC	Occupational category	Skill intensity	NAICS	Industry	Skill intensity
45	Farming, Fishing & Forestry	8.7	11	Forestry, fishing, hunting & agriculture support	10.5
37	Cleaning & Maintenance	10.8	72	Accommodation & food services	11.8
35	Food Preparation & Serving Related	11.5	23	Construction	11.9
47	Construction & Extraction	11.5	56	Admin, support, waste mgt, remediation	12.2
51	Production	11.5	48	Transportation & warehousing	12.6
29	Healthcare Practitioners & Technical	15.6	52	Finance & insurance	14.1
21	Community & Social Services	15.8	51	Information	14.1
25	Education, Training & Library	16.3	55	Management of companies & enterprises	14.6
19	Life, Physical & Social Science	17.2	54	Professional, scientific & technical services	15.3
23	Legal Occupations	17.3	61	Educational services	15.6

Data source: 2000 Census of Population microdata via IPUMS-USA

Empirical results: Three skill groups

	(1)	(2)	Population	Share
Dependent variable: $\ln f(\omega, c)$	All	US-born	share	US-born
$\beta_{\omega 1}$ High school or less $\times$ log population	0.954	0.895	.37	.78
	(0.0108)	(0.0153)		
$\beta_{\omega 2}$ Some college $\times$ log population	0.996	0.969	.31	.89
	(0.0105)	(0.0122)		
$\beta_{\omega 3}$ Bachelor's or more $\times$ log population	1.086	1.057	.32	.85
	(0.0153)	(0.0162)		

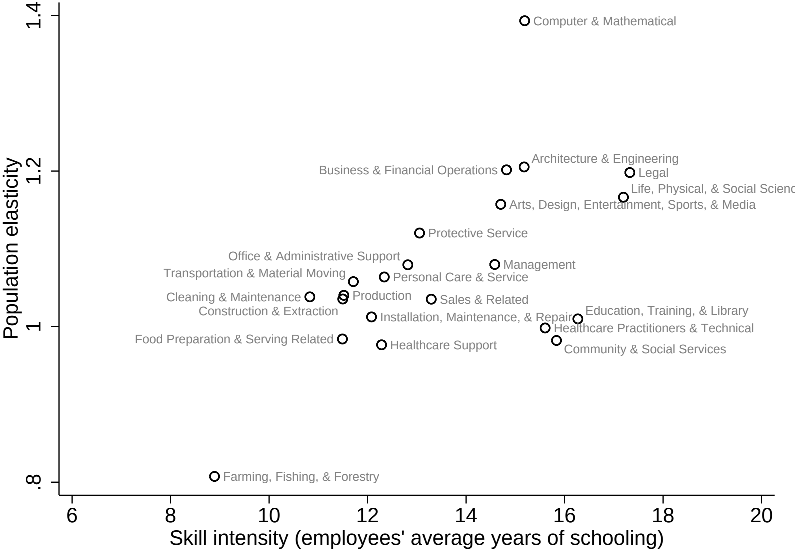
Empirical results: Nine skill groups

Dependent variable: $\ln f(\omega, c)$	(1) All	(2) US-born	Population share	Share US-born
$\beta_{\omega 1}$ Less than high school $\times$ log population	1.089	0.858	.04	.29
	(0.0314)	(0.0239)		
$\beta_{\omega 2}$ High school dropout $\times$ log population	1.005	0.933	.08	.73
	(0.0152)	(0.0181)		
$\beta_{\omega 3}$ High school graduate $\times$ log population	0.925	0.890	.25	.88
	(0.0132)	(0.0163)		
$\beta_{\omega 4}$ College dropout $\times$ log population	0.997	0.971	.23	.89
	(0.0111)	(0.0128)		
$\beta_{\omega 5}$ Associate's degree $\times$ log population	0.997	0.965	.08	.87
	(0.0146)	(0.0157)		
$\beta_{\omega 6}$ Bachelor's degree $\times$ log population	1.087	1.059	.20	.86
	(0.0149)	(0.0164)		
$\beta_{\omega 7}$ Master's degree $\times$ log population	1.095	1.063	.08	.84
	(0.0179)	(0.0181)		
$\beta_{\omega 8}$ Professional degree $\times$ log population	1.113	1.082	.03	.81
	(0.0168)	(0.0178)		
$\beta_{\omega 9}$ PhD $\times$ log population	1.069	1.021	.01	.72
	(0.0321)	(0.0303)		

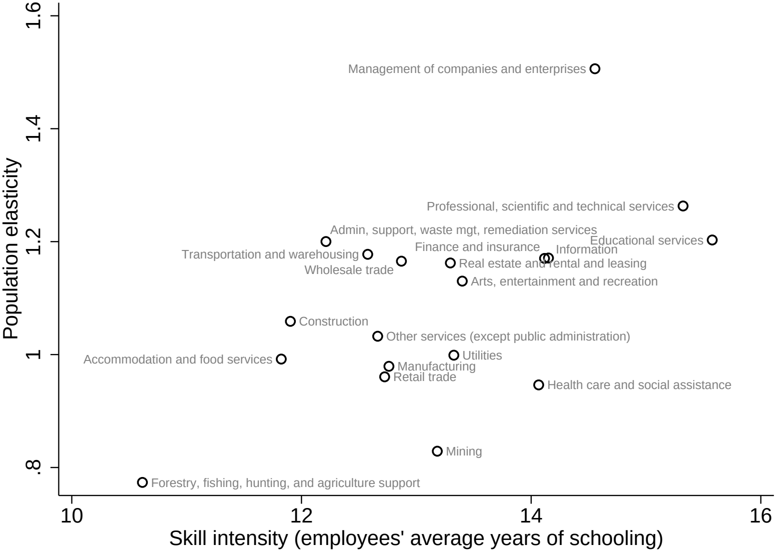
Spatial distribution of skills in 1980

Dependent variable: $\ln f(\omega, c)$	(1) All	(2) US-born	Population share	Share US-born
$\beta_{\omega 1}$ Less than high school × log population	0.975	0.892	.09	.72
	(0.0236)	(0.0255)		
$\beta_{\omega 2}$ High school dropout × log population	1.006	0.983	.12	.91
	(0.0157)	(0.0179)		
$\beta_{\omega 3}$ Grade 12 × log population	0.989	0.971	.33	.93
	(0.00936)	(0.0111)		
$\beta_{\omega 4}$ 1 year college × log population	1.047	1.033	.10	.94
	(0.0144)	(0.0151)		
$\beta_{\omega 5}$ 2-3 years college × log population	1.095	1.076	.12	.91
	(0.0153)	(0.0155)		
$\beta_{\omega 6}$ 4 years college × log population	1.091	1.073	.12	.92
	(0.0153)	(0.0157)		
$\beta_{\omega 7}$ 5+ years college × log population	1.113	1.093	.12	.90
	(0.0202)	(0.0196)		

Occupations' elasticities and skill intensities



Industry population elasticities and skill intensities



Diamond (2016)

Since 1980, college graduates have been concentrating in US cities with faster wage and housing-price growth. Questions:

- Why are they doing so?
- Is this spatial divergence associated with greater welfare inequality?

By reading Diamond (2016), you'll learn about a bunch of relevant concepts and tools:

- Great Divergence (Moretti's *New Geography of Jobs*)
- Endogenous amenities
- Inferring welfare with multiple types
- "Bartik (1991)" shift-share instruments
- Housing supply elasticities

Great Divergence

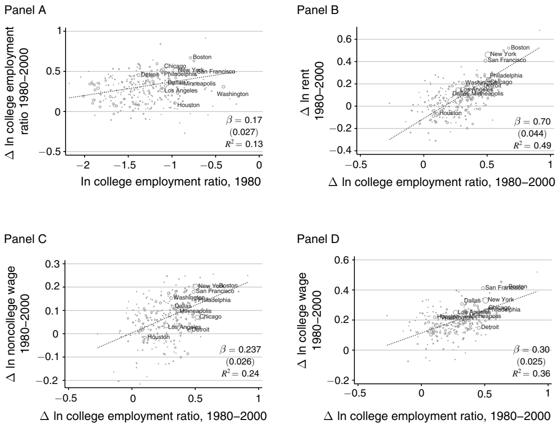


FIGURE 1. CHANGES IN WAGES, RENTS, AND COLLEGE EMPLOYMENT RATIOS, 1980–2000

Notes: Weighted by 1980 population. Largest 15 MSAs in 1980 labeled.

These facts are in Berry & Glaeser (2005)

Moretti (2013) raises welfare questions by pointing out housing price growth for college-abundant cities shrinks nominal wage gap

Endogenous amenities

- Welfare isn't just nominal wages and housing prices
- Infer compensating differential from spatial-indifference condition
- Amenities: Exogenous sunshine vs endogenous crime
- Diamond's amenities: "all characteristics of a city which could influence the desirability of a city beyond local wages and prices"
- In reality, "retail amenities" are private goods with prices and "schooling amenities" are govt expenditures paid by local taxes
- Inferring amenities much harder with multiple types (Roback 1988) and endogeneity

NEWS IN BRIEF

Neighborhood Starting To Get Too Safe For Family To Afford

8/28/15 12:00am • SEE MORE: LOCAL

The "Bartik" instrument for local labor demand

- "The idea is to isolate shifts in local labor demand that come only from national shocks in each sector of the economy, thereby purging potentially endogenous local demand shocks driving variation in employment or wages" (Baum-Snow and Ferreira 2015)
- "A host of papers make use of such instruments for identification"
- "The main source of identifying variation in Bartik instruments comes from differing base year industry compositions across local labor markets. Therefore, validity of these instruments relies on the assertion that neither industry composition nor unobserved variables correlated with it directly predict the outcome of interest conditional on controls."
- There is suddenly an econometrics literature on this. See Adao, Kolesar, Morales for inference. See Goldsmith-Pinkham, Sorkin, Swift and Borusyak, Jaravel, Hull for consistency/validity.

Housing supply elasticities

- If housing is supplied elastically, a local labor demand shock mostly shows up in increased population (quantities)
- If housing is inelastic, wages and prices increase instead
- Housing is durable, so expansion and contraction are asymmetric
- Housing supply depends on exogenous features (hills, water) and on endogenous regulatory regime (Saiz *QJE* 2010)

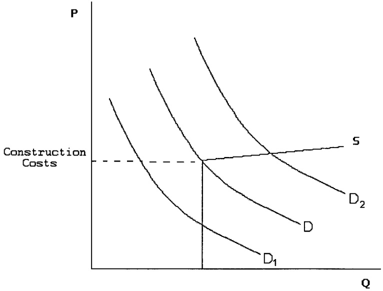
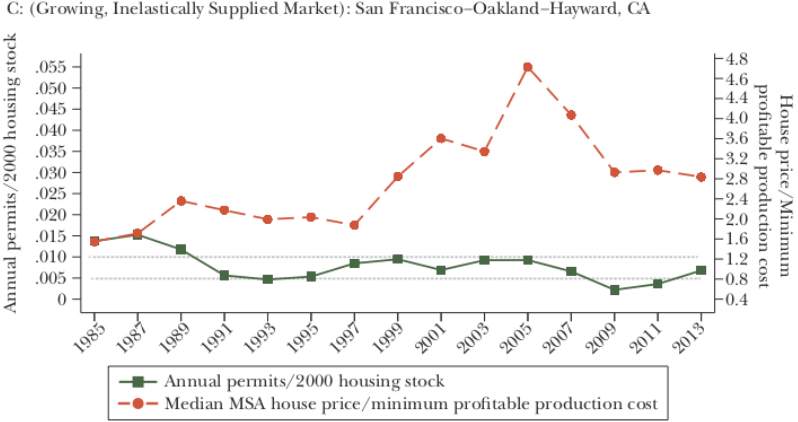


FIG. 1.—The nature of housing supply and construction costs

Glaeser & Gyourko (JPE 2005)

Housing supply elasticities – prices vs quantities

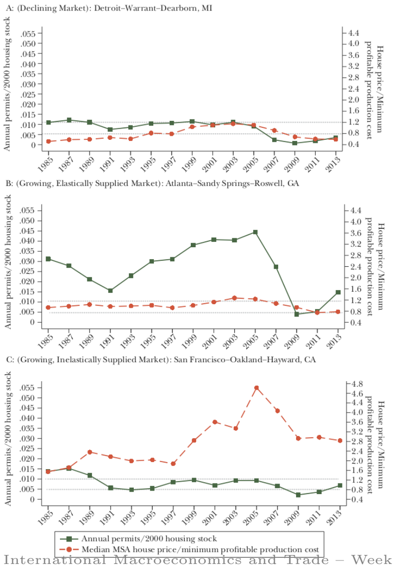


Glaeser & Gyourko (JEP 2018)

Housing supply elasticities – prices vs quantities

Figure 2

New Housing Supply and House Prices (Relative to Costs)

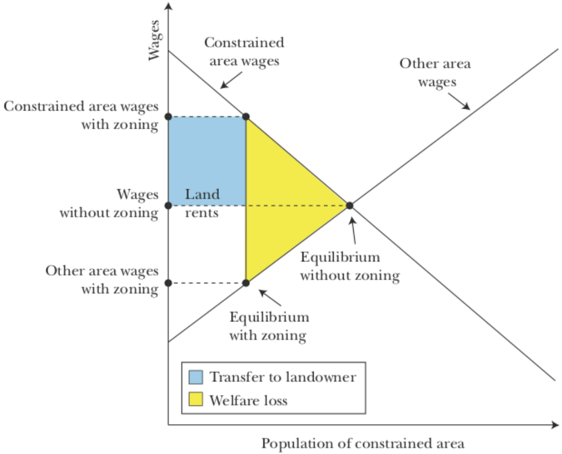


Source: House prices come from American Housing Survey micro data.
Note: Minimum Profitable Production Costs (MPPC) are for 1800 square feet, economy class, 20 percent

Incidence of building restrictions

Figure 4

Welfare Consequences of Restricting Development in a Productive Market



Glaeser & Gyourko (JEP 2018)

Diamond (2016) empirical implementation

- Write a spatial-equilibrium model with two skill types and elasticities to be estimated
- "In the first step, a maximum likelihood estimator is used to identify how desirable each city is to each type of worker, on average, in each decade, controlling for workers' preferences to live close to their state of birth."
- "The second step of estimation uses a simultaneous equation nonlinear generalized method of moments (GMM) estimator. Moment restrictions on workers' preferences are combined with moments identifying cities' labor demand, housing supply, and amenity supply curves."
- Housing supply elasticities are identified by the response of housing rents to the Bartik shocks across cities.
- The interactions of the Bartik productivity shocks with cities' housing markets identify the labor demand elasticities.

Labor demand

$$Y_{djt} = N_{djt}^{\alpha} K_{djt}^{1-\alpha} \tag{1}$$

$$N_{djt} = \left(\theta_{jt}^{L} L_{djt}^{\rho} + \theta_{jt}^{H} H_{djt}^{\rho}\right)^{1/\rho}$$

$$\theta_{jt}^{L} = f_L \left(H_{jt}, L_{jt}\right) \exp\left(\epsilon_{jt}^{L}\right) \tag{2}$$

$$\theta_{jt}^{H} = f_H \left(H_{jt}, L_{jt}\right) \exp\left(\epsilon_{jt}^{H}\right) \tag{3}$$

K isn't interesting since interest rate assumed national

$$w_{jt}^{H} = \ln W_{jt}^{H} = c_t + (1-\rho)\ln N_{jt} + (\rho - 1)\ln H_{jt} + \ln\left(f_H(H_{jt}, L_{jt})\right) + \epsilon_{jt}^{H}$$

$$w_{jt}^{L} = \ln W_{jt}^{L} = c_t + (1-\rho)\ln N_{jt} + (\rho - 1)\ln H_{jt} + \ln\left(f_H(H_{jt}, L_{jt})\right) + \epsilon_{jt}^{L}$$

$$w_{jt}^{H} = g_H \left(H_{jt}, L_{jt}\right) + \epsilon_{jt}^{H} \tag{7}$$

$$\approx \gamma_{HH} \ln H_{jt} + \gamma_{HL} \ln L_{jt} + \epsilon_{jt}^{H} \tag{9}$$

$$w_{jt}^{L} = g_L \left(H_{jt}, L_{jt}\right) + \epsilon_{jt}^{L} \tag{8}$$

$$\approx \gamma_{LH} \ln H_{jt} + \gamma_{LL} \ln L_{jt} + \epsilon_{jt}^{L} \tag{10}$$

Labor supply

- Logit preferences
- Common component: Cobb-Douglas preference over freely traded homogeneous good with price P_t and local housing with rent R_{jt}.
- Augmented by amenity vector $\mathbf{A}_{jt}$, which has race-specific valuations
- Race-specific valuations of birthplace dummies
- Type 1 extreme-value error term

$$H_{jt} = \sum_{i \in \mathcal{H}_t} \frac{\exp(\delta_{jt}^{z_i} + \mathbf{x}_j^{st}\mathbf{st}_i \beta^{st}\mathbf{z}_i + \mathbf{x}_j^{\text{div}}\text{div}_i \beta^{\text{div}}\mathbf{z}_i)}{\sum_k^J \exp(\delta_{kt}^{z_i} + \mathbf{x}_k^{st}\mathbf{st}_i \beta^{st}\mathbf{z}_i + \mathbf{x}_k^{\text{div}}\text{div}_i \beta^{\text{div}}\mathbf{z}_i)}$$

$$L_{jt} = \sum_{i \in \mathcal{L}_t} \frac{\exp(\delta_{jt}^{z_i} + \mathbf{x}_j^{st}\mathbf{st}_i \beta^{st}\mathbf{z}_i + \mathbf{x}_j^{\text{div}}\text{div}_i \beta^{\text{div}}\mathbf{z}_i)}{\sum_k^J \exp(\delta_{kt}^{z_i} + \mathbf{x}_k^{st}\mathbf{st}_i \beta^{st}\mathbf{z}_i + \mathbf{x}_k^{\text{div}}\text{div}_i \beta^{\text{div}}\mathbf{z}_i)}.$$

Housing supply and amenity supply

- The elasticity of housing supply depends on geographic and regulatory components from Saiz (*QJE* 2010)
- Amenities are an endogenous function of the $\frac{H}{L}$ ratio
- You might find it interesting to read Tom Davidoff's "Supply Constraints Are Not Valid Instrumental Variables for Home Prices Because They Are Correlated With Many Demand Factors" (*Critical Finance Review* 2016)

Estimation: Two-step GMM

Instruments:

$$\Delta Z_{jt} \in \left\{ \Delta B_{jt}^{H}, \Delta B_{jt}^{L}, \Delta B_{jt}^{H} x_j^{reg}, \Delta B_{jt}^{H} x_j^{geo}, \Delta B_{jt}^{L} x_j^{reg}, \Delta B_{jt}^{L} x_j^{geo} \right\}$$

"the level of land-unavailability and land-use regulation are uncorrelated with unobserved local productivity changes [$\Delta\tilde{\epsilon}_{jt}^{H}$ and $\Delta\tilde{\epsilon}_{jt}^{L}$, which are uncorrelated with the Bartik local labor demand shocks]"

$$E\left(\Delta\tilde{\epsilon}_{jt}^{H}\Delta Z_{jt}\right) = 0 \quad E\left(\Delta\tilde{\epsilon}_{jt}^{L}\Delta Z_{jt}\right) = 0$$

"Bartik labor demand shocks are uncorrelated with changes in local construction costs" ($\ln CC_{jt}$ are unobserved factors driving housing prices):

$$E\left(\Delta \ln(CC_{jt})\Delta Z_{jt}\right) = 0$$

"housing supply elasticity characteristics are independent of changes in local exogenous amenities" ($\Delta\xi_{jt}^{z} \equiv \beta^{A}\mathbf{z}\Delta\mathbf{x}_{jt}^{A}$):

$$E\left(\Delta\xi_{jt}\Delta Z_{jt}\right) = 0$$

"these instruments are uncorrelated with unobserved exogenous changes in the city's local amenities which make up the amenity index":

$$E\left(\Delta\epsilon_{jt}^{a}\Delta Z_{jt}\right) = 0$$

Estimation results

- "My results suggest that endogenous local amenity changes are an important mechanism driving workers' migration responses to local labor demand shocks."
- "the positive aggregate labor demand elasticities for college workers suggests that the endogenous productivity effects of college workers on college workers' productivity may be large and could overwhelm the standard forces leading to downward-sloping labor demand"
- "an increase in well-being inequality between college and high school graduates which was significantly larger than would be suggested by the increase in the college wage gap alone"

Summary

- Spatial distributions of skills and sectors are prominent in public discussion of cities, exploited for identification in empirical work, and potentially key to understanding agglomeration processes
- We need models with more than two skills groups and more than perfectly specialized/diversified cities
- Recent research exploits tools from assignment literature to characterize spatial sorting of skills and sectors
- Assignment mechanisms can be used in quantitative work via assumptions on components observed and unobserved by the econometrician – Fréchet distribution is most popular

Thank you

Bacolod, Blum, Strange on AFQT scores

Table 5
Agglomeration and the AFQT and Rotter scores: Distributions for selected occupations and city size categories.

	Panel A. 10th & 90th Percentiles of AFQT Score				*Panel B. 10th & 90th Percentiles of Rotter Score*			
	MSA Size				MSA Size			
Occupation	Small	Medium	Large	Very Large	Small	Medium	Large	Very Large
Managers	51.99	42.02	36.37	24.6	0.47	0.46	0.43	0.37
	69.65	64.81	82.29	91.72	0.55	0.52	0.65	0.68
Engineers	62.92	79.22	62.95	49.67	0.47	0.49	0.42	0.41
	79.22	86.96	87.59	94.93	0.53	0.53	0.58	0.63
Therapists	60.75	70.92	44.98	41.62	0.57	0.6	0.49	0.42
	60.9	72.93	60.03	82.56	0.57	0.6	0.62	0.62
College Professors	74.1	59.79	70.4	45.13	0.45	0.47	0.46	0.4
	81.43	81.77	88.25	93.61	0.49	0.6	0.55	0.6
Teachers	60.32	63.82	50.88	34.51	0.51	0.45	0.43	0.38
	68.81	75.67	81.96	86.44	0.54	0.52	0.62	0.62
Sales Persons	69.74	82.27	62.92	66.41	0.49	0.42	0.44	0.42
	81.45	82.27	86.18	96.12	0.56	0.42	0.5	0.59
Food Services	47.48	21.05	27.21	10.71	0.53	0.49	0.42	0.38
	58.01	54.9	64.57	80.6	0.58	0.64	0.66	0.7
Mechanics	39.73	29.72	24.13	12.71	0.51	0.45	0.41	0.38
	57.01	61.59	67.99	74.14	0.56	0.55	0.62	0.68
Construction Workers	42.4	26.8	15.22	8.89	0.46	0.48	0.46	0.39
	51.75	42.58	63.56	68.33	0.51	0.58	0.7	0.69
Janitors	34.54	35.99	11.83	5.55	0.52	0.48	0.43	0.4
	45.41	55.4	53.21	64.15	0.55	0.63	0.67	0.72
Natural Scientists	75.67	53.53	47.25	63.06	0.52	0.45	0.47	0.44
	75.67	77.7	58.03	92.92	0.52	0.51	0.49	0.6
Nurses	57.33	61.02	61.97	51.23	0.53	0.48	0.46	0.41
	58.88	65.34	76.31	83.92	0.54	0.51	0.59	0.57
Social Workers	38.52	54.14	57.37	34.1	0.49	0.52	0.53	0.4
	52.54	57.04	69.24	77.37	0.5	0.54	0.58	0.63
Technicians	67.28	52.01	46.84	30.44	0.47	0.42	0.42	0.38
	79.89	81.6	85.74	93.88	0.55	0.61	0.62	0.67
Administrative Support	34.18	37.9	34.05	14.65	0.49	0.45	0.41	0.37
	55.98	70.32	75.89	83.85	0.6	0.62	0.62	0.7
Personal Services	60.54	34.46	19.58	14.74	0.51	0.5	0.44	0.39
	68.11	57.92	65.6	73.21	0.56	0.59	0.67	0.68
Total	56.78	52.77	44.92	33.86	0.5	0.48	0.45	0.4
	66.61	69.49	74	84.39	0.54	0.56	0.6	0.65

Notes. The first row reports the 10th percentile, while the second row reports the 90th percentile. Small MSA size: population between 100,000 and 500,000; Medium: between 500,000 and 1 million; Large: between 1 million and 4 million; Very Large: more than 4 million.

College wage premia in NLSY vs Census

	Baum-Snow & Pavan Table 1, column 1	2000 Census PMSA	2000 Census CMSA
Non-Hispanic white males with fewer than 15 years of work experience			
Medium-city college wage premium	.09	0.0978	0.0937
		(0.00578)	(0.00613)
Large-city college wage premium	.05	0.145	0.154
		(0.00565)	(0.00551)
N	17991	301326	301326
Individuals observed	1257	301326	301326
R^2		0.197	0.202
p-value for equal premia		0	0

Robust standard errors in parentheses

NOTES: This table describes full-time, full-year employees ages 18-55. Following BSP, "college graduate" means anyone with a bachelor's degree or greater educational attainment. Large means population greater than 1.5m. Medium means population .25m to 1.5m. Small includes rural areas. The premia in the first column are obtained by differencing the numbers for high-school and college graduates' log wages in the first column of BSP's Table 1. Note that they report results for temporally deflated panel data, while we report cross-sectional results. BSP assign individual to metropolitan statistical areas using the 1999 boundary definitions, but they do not specify whether they use consolidated MSAs or primary MSAs for large cities. Hence we report both.