ECON G6905
Topics in Trade

Jonathan Dingel

Spring 2025, Week 10

COLUMBIA UNIVERSITY
IN THE CITY OF NEW YORK

Today: Skill-Biased Agglomeration

Economic geography of human capital: Where do skilled workers live and why?

My goal today is to tackle four questions:

- Why should we care about the spatial distributions of skills and sectors?
- How do we know that agglomeration is skill-biased?
- How should we characterize skills when modeling cities?
- What tools are relevant for building and estimating models?

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 are often used as exogenous variation
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

The 25 Most Educated Cities In America

Alyson Penn
Sep. 19, 2014, 11:39 AM 36,046 4

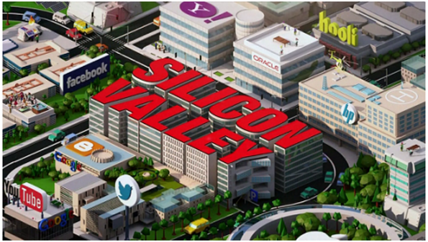
FACEBOOK LINKEDIN TWITTER

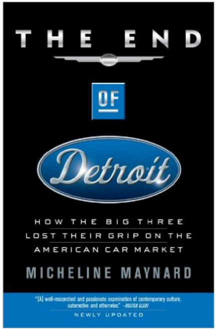
The 10 smartest cities in America

By MarketWatch
Published: July 28, 2015 2:33 p.m. ET

11,657 271 91

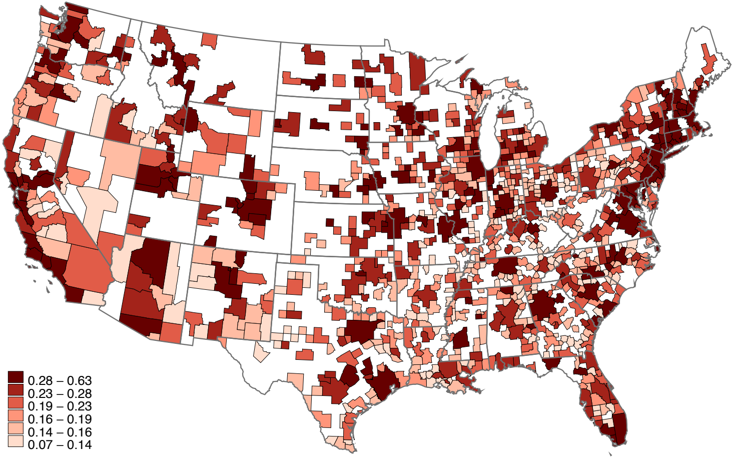
- Place names are shorthand for sectors





Educational attainment varies a lot across cities

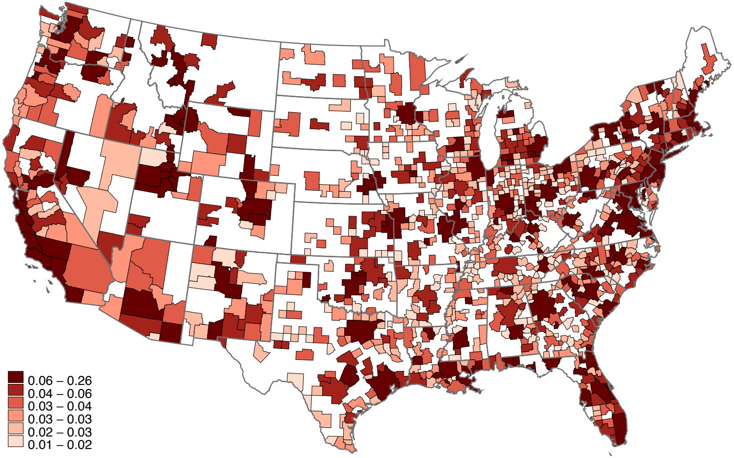
Share of population 25 and older with bachelor's degree or higher



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

Sectoral composition varies a lot across cities

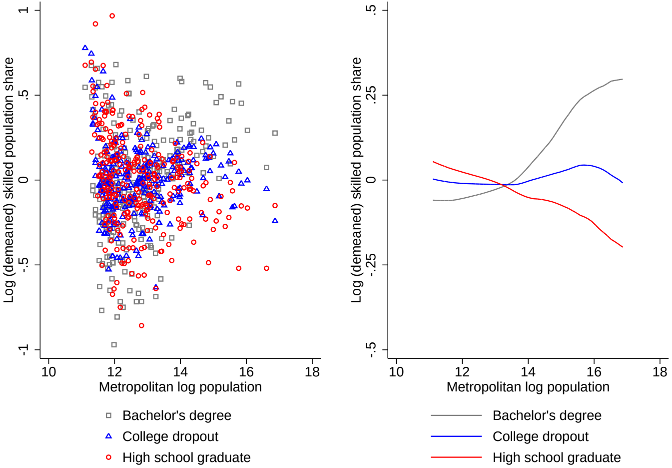
Employment share of Professional, Scientific, and Technical Services



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

They covary with city characteristics

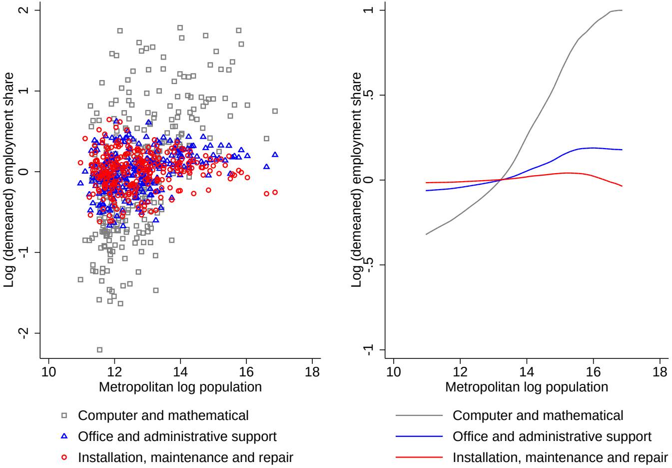
Populations of three educational groups across US metropolitan areas



DATA SOURCE: 2000 Census of Population microdata via IPUMS-USA

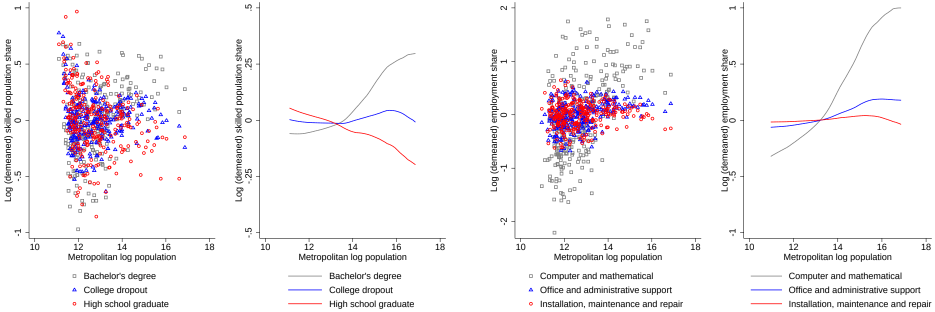
They covary with city characteristics

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 are often used as exogenous variation

It is common to see the following theory-empirics pairing:

- Model: all locations produce a homogeneous good
- Estimation by shift-share design: exogenous shifts in local labor demand via local industrial composition $\times$ national changes in industrial employment

What variation does this shift-share design exploit?

- GPSS: employment shares are IVs measuring exposure to shifts
- AKM & BHJ: sectoral shifters are randomly assigned and independent
- Skill mix vs industry mix (e.g., endogenous local SBTC of Beaudry, Doms, Lewis 2010)
- City characteristics covarying with skills and sectors highlight exclusion-restriction assumptions

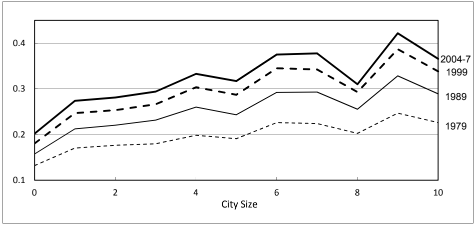
They should help us understand how cities work

- Why do different people and different businesses locate in different places?
- The answers should be crucial to understanding how cities work
- Lucas (1988): "the 'force' we need to postulate account for the central role of cities in economic life is of exactly the same character as the 'external human capital' I have postulated as a force to account for certain features of aggregative development."
- 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
- Eckert, Ganapati, Walsh (2024): Wage growth since 1980 has been faster in larger cities and it's all in business services

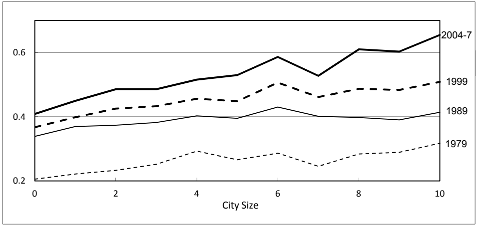
An introduction to skill-biased agglomeration

- Central fact: Larger cities have higher college-graduate shares (Berry & Glaeser 2005, Moretti 2012) and higher college wage premia (Baum-Snow & Pavan 2013, Davis & Dingel 2019)

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 explain spatial variation in relative prices and relative quantities?
- Start from "canonical model" with two skill types and spatial variation in relative supply and relative demand

Spatial equilibrium with two skill groups

A simple starting point

1. Two skill groups, $s \in \{L, H\}$
2. Elastic labor supply: $U_s(A_c, w_{s,c}, p_c) = U_s(A_{c'}, w_{s,c'}, p_{c'}) \; \forall c, c' \; \forall s$
3. Homothetic preferences: $U_s(A_c, w_{s,c}, p_c) = \frac{w_{s,c} A_c}{p_c}$

These jointly imply that relative wages are spatially invariant:

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

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

Spatial equilibrium with two skill groups

A simple starting point

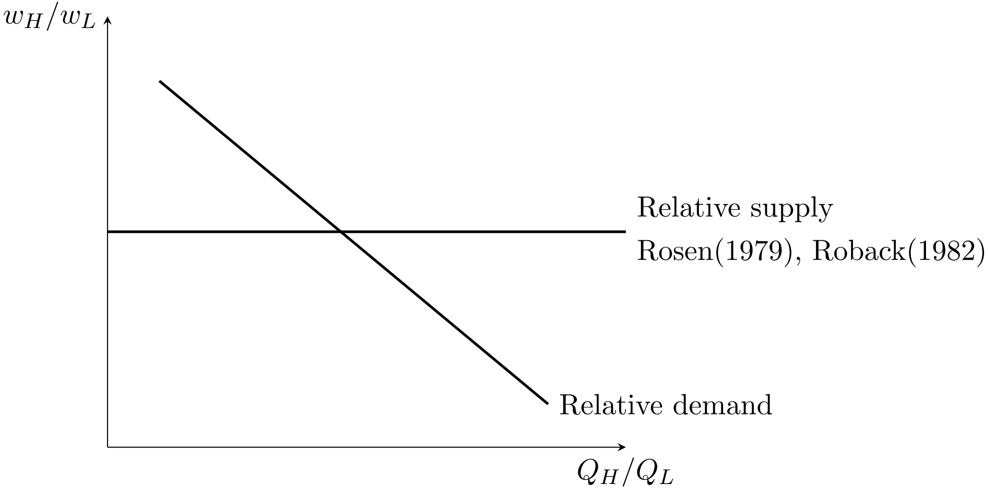
1. Two skill groups, $s \in \{L, H\}$
2. Elastic labor supply: $U_s(A_c, w_{s,c}, p_c) = U_s(A_{c'}, w_{s,c'}, p_{c'}) \ \forall c, c' \ \forall s$
3. Homothetic preferences: $U_s(A_c, w_{s,c}, p_c) = \frac{w_{s,c} A_c}{p_c}$

These jointly imply that relative wages are spatially invariant:

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

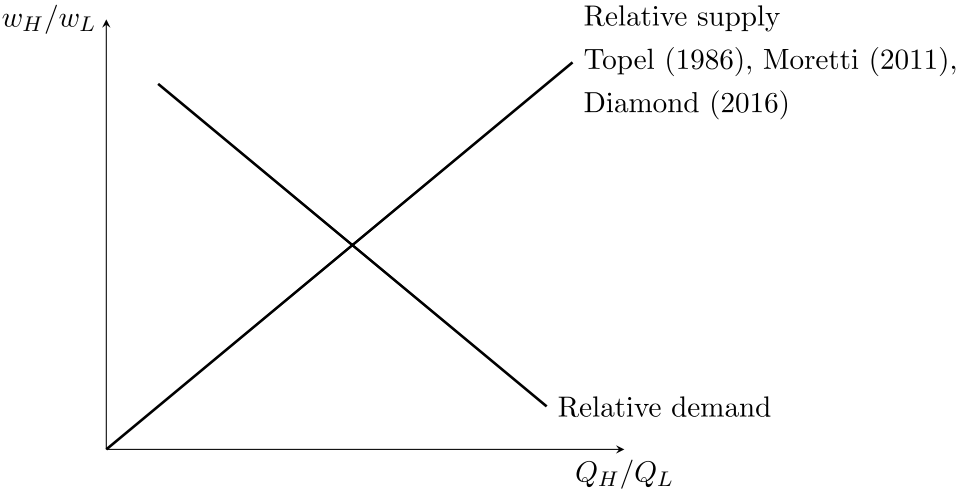
$$\implies \frac{w_{H,c}}{w_{L,c}} = \frac{w_{H,c'}}{w_{L,c'}} \ \forall c, c'$$

Spatial equilibrium and skill premia (two types; homothetic prefs)

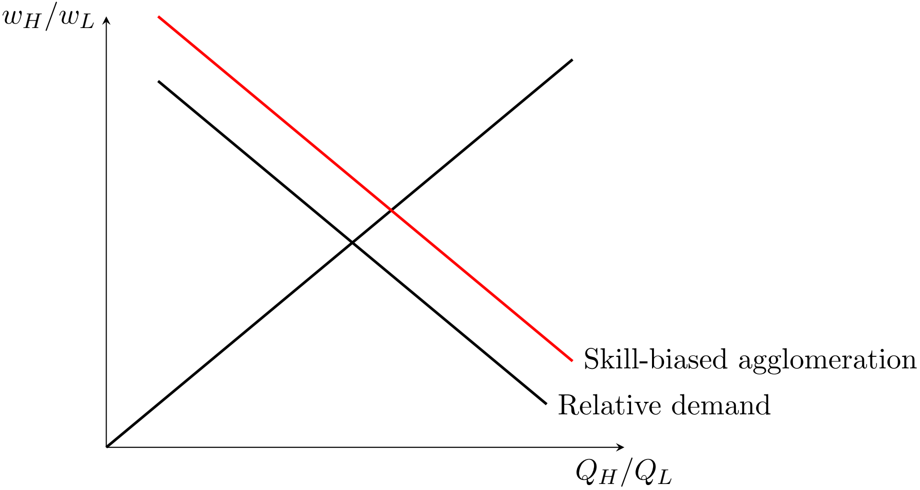


Differences in productivity "tend to show up exclusively in changes in quantities of skilled people, not in different returns to skilled people across space" (Glaeser 2008)

Spatial equilibrium and skill premia (two types; homothetic prefs)



Spatial equilibrium and skill premia (two types; homothetic prefs)



Why not a story about relative supply?

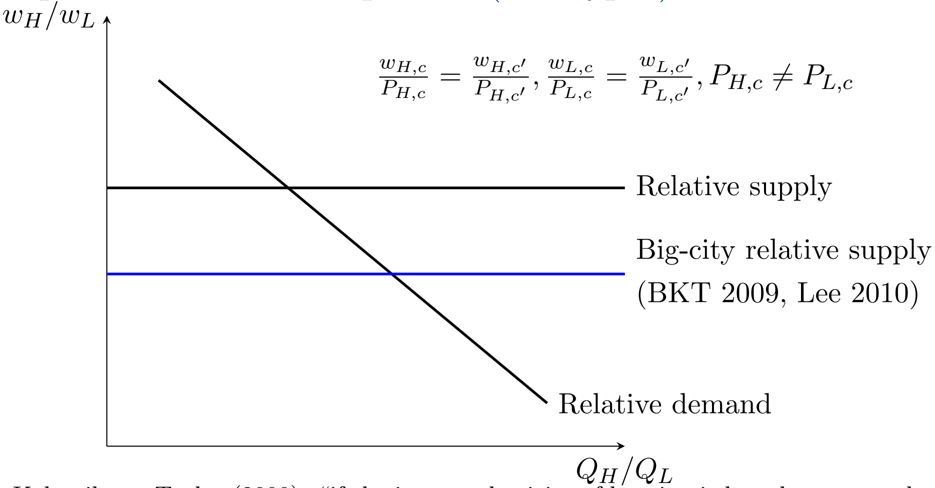
Preferences are not homothetic. What happens if we relax that assumption?

Relative prices of income-elastic goods are lower in larger cities:

- Income elasticity of housing demand is less than one:
  Albouy, Ehrlich, and Liu (2016), Finlay and Williams (2023)
- Larger cities specialize in producing income-elastic tradable goods:
  Dingel (2017), Handbury (2021), Onoda (2023)

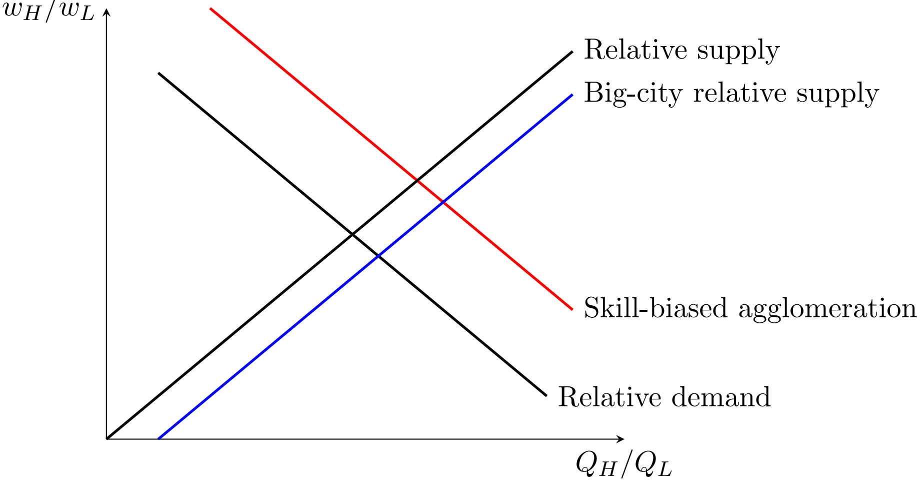
These are empirically relevant forces, but they increase the relative supply of skill in larger cities

Spatial equilibrium and skill premia (two types)



Black, Kolesnikova, Taylor (2009): "if the income elasticity of housing is less than one. . . the return to education is lower in cities that are more expensive"

Spatial equilibrium and skill premia (two types)



With two types, the explanation for spatial variation in relative prices and relative quantities lies in skill-biased agglomeration shifting relative demand.

What is skill-biased agglomeration?

The canonical model is one way of interpreting the central fact

- In two-type model, larger cities have higher relative demand for skill
- With more than two underlying types, these relative quantities and relative prices may reflect compositional differences (i.e., spatial sorting)
- I read the available empirical evidence as saying two types are not enough

Two types in theory and practice

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

- Omit types: Davis & Dingel (2019) plot of college wage premia shows bachelor's 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: Davis & Dingel (2019) plot of college wage premia shows bachelor's 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)
- In the urban context, 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)

How much should we worry about spatial sorting?

Contrasting views. Suggestions of sorting:

- "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 occupations, job postings in larger cities require more interactive tasks and using newer technologies, especially for college graduates (Atalay, Sotelo, Tannenbaum 2021)

Claim of little sorting (e.g., de la Roca, Ottaviano, Puga 2023) stems from:

- Cognitive test scores in NLSY79 (11,000 US individuals)
- Estimated finite-mixture model using NLSY79 (Baum-Snow & Pavan 2012)
- Individual fixed effects from AKM-style regressions

Evidence of little spatial sorting from US NLSY79

National Longitudinal Survey of Youth 1979 (NLSY79)

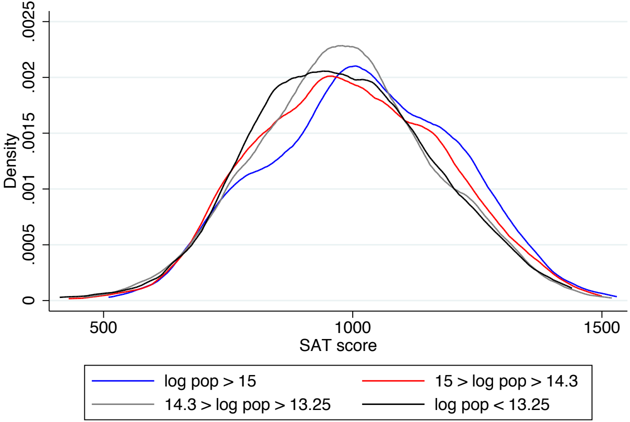
- Bacolod, Blum, Strange (2009): "The mean AFQT scores do not vary much across [four] city sizes" within occupational categories
- BBS observe only one sales person in MSAs with 0.5m – 1.0m residents (10th and 90th percentiles of AFQT are equal) ▸ table
- Baum-Snow & Pavan (2012): Estimated finite-mixture 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

Bringing more US data to bear on sorting

- Baccalaureate and Beyond tracks a cohort graduating from four-year colleges in 1993
- In 2003, look at 2,300 white individuals who obtained no further education after bachelor's degree and now live in a MSA
- 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



Evidence from wage regressions with worker fixed effects

de la Roca and Puga (2017) use Spanish tax data on 2004-2009 earnings:

- 157,113 workers and 40,809 cross-city moves
- Assume random migration conditional on observables
- e_{ijt} is worker i's experience in city j through time t
- Estimating equation:

$$\ln w_{ict} = \sigma_c + \mu_i + \sum_{j=1}^{C} \delta_{jc} e_{ijt} + \beta x_{it} + \epsilon_{ict}$$

- If one only estimates a static specification

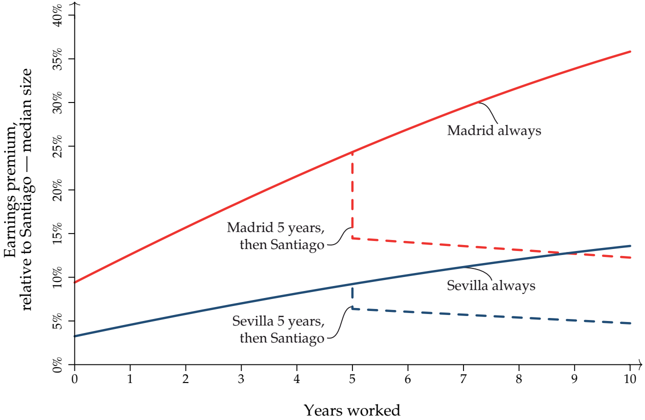
$$\ln w_{ict} = \sigma_c + \mu_i + \beta x_{it} + \zeta_{ict}$$

$\hat{\sigma}_c$ may be biased if $\mathrm{Cov}\left((\iota_{ict} - \bar{\iota}_{ic}), \sum_j^C \delta_{jc}(e_{ijt} - \bar{e}_{ij})\right) \neq 0$ (if workers' experience is higher when they are in c)

de la Roca and Puga (2017) on dynamics and sorting

- Experience in larger cities is more valuable and portable

  Earnings of equivalent workers in Madrid (largest) and Sevilla (fourth largest) relative to Santiago de Compostela (median)



Panel (a) Profiles allowing for learning benefits of bigger cities

de la Roca and Puga (2017) on dynamics and sorting

- Experience in larger cities is more valuable and portable
- Big-city experience is more valuable for workers with higher ability

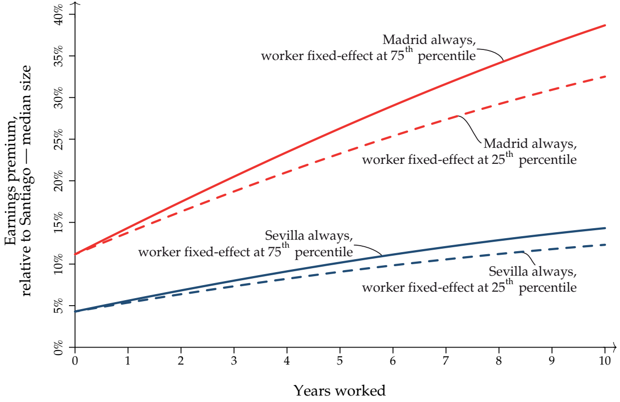


FIGURE 7

Earnings profiles relative to median-sized city, high- and low-ability worker

de la Roca and Puga (2017) on dynamics and sorting

- Experience in larger cities is more valuable and portable
- Big-city experience is more valuable for workers with higher ability
- There is sorting across five occupational categories

TABLE 5

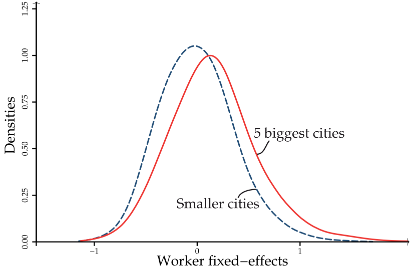
Comparison of occupational groups across cities of different sizes

	Occupational groups (%)				
	Very-high-skilled	High-skilled	Medium-high skilled	Medium-low skilled	Low-skilled
First to second biggest cities	10.9	13.8	24.2	41.7	9.4
Third to fifth biggest cities	6.3	10.9	21.0	48.2	13.8
Other cities	3.5	7.9	18.4	54.0	16.1

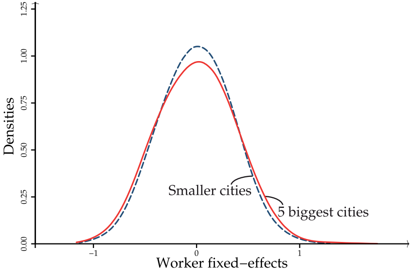
Notes: Employers assign workers into one of ten social security categories which we regroup into five occupational skill categories. Shares are averages of monthly observations in the sample.

de la Roca and Puga (2017) on dynamics and sorting

- Experience in larger cities is more valuable and portable
- Big-city experience is more valuable for workers with higher ability
- There is sorting across five occupational categories
- Little sorting within these categories when using full dynamic specification



Panel (c)
Fixed-effects, static premium



Panel (a)
Fixed-effects, heterogeneous dynamic and static premium

Norweigan wage regressions with worker fixed effects

Carlsen Rattsø & Stokke (2016) use 2003-2010 data on Norway:

- College-educated workers have higher return to big-city experience
- The city wage premium trajectories depend on job tenure

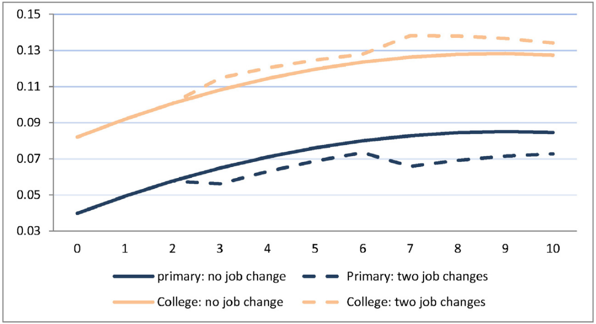
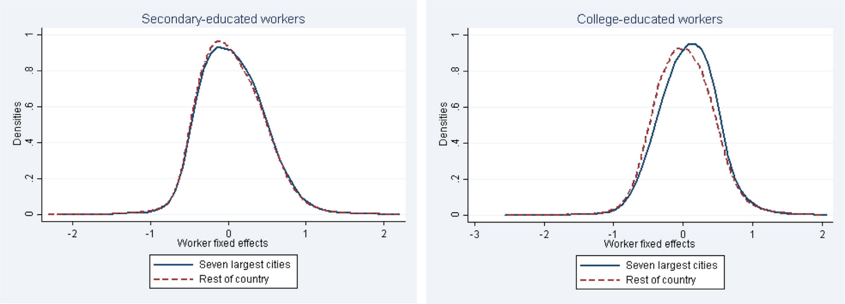


Fig. 1. Urban wage premium trajectories for primary- and college-educated workers, years after move to Oslo.

Norweigan wage regressions with worker fixed effects

Carlsen Rattsø & Stokke (2016) use 2003-2010 data on Norway:

- College-educated workers have higher return to big-city experience
- The city wage premium trajectories depend on job tenure
- Sorting on unobserved abilities is driven by the college educated



- Sorting by young: no sorting on worker fixed effects among older workers

US wage regressions with worker fixed effects

Card, Rothstein, Yi (2023) use Longitudinal Employer-Household Dynamics data:

- Abowd-Kramarz-Margolis wage regression with worker FEs, establishment FEs, and controls for age and time
- Commuting Zone effect is average of establishment FEs in the CZ (specification with worker FEs and CZ FEs biased by "hierarchy effect")
- "We find some evidence of the kind of dynamic returns to 'big city' experience highlighted by de la Roca and Puga (2017) but the addition of this channel has little impact on the static returns to different CZs"
- Half of the observed large-CZ earnings premium reflects worker sorting
- 2/3 of spatial variation in college wage premia is variation in relative skills
- Much of the remaining 1/3 is because of enhanced sorting of college-educated workers to high-wage industries in larger CZs

Modeling a spatial economy with a continuum of skills

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

- See Behrens, Duranton, Robert-Nicoud (2014), Davis and Dingel (2019, 2020), 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 (2019)

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 2014)

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 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 $Z(\cdot)$:

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

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

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

- 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)

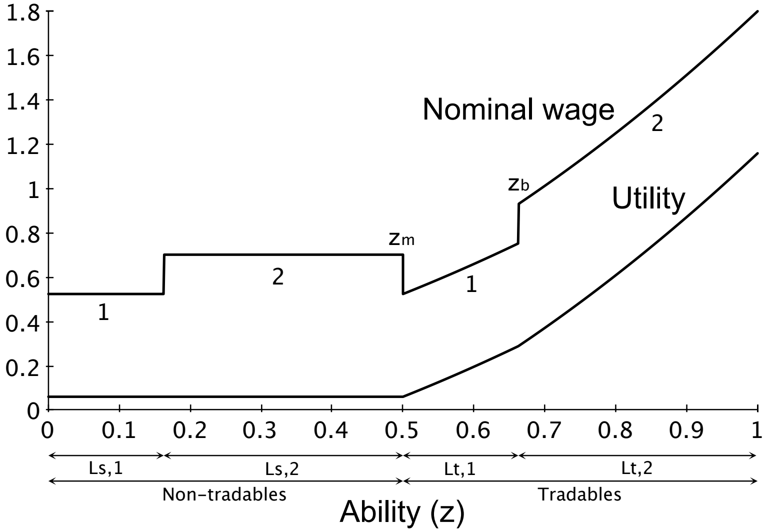
Suppose that Assumption 2 holds. For $z > z' > z_m$*, if* $\mu(z, c) > 0$*,* $\mu(z', c') > 0$*,* $\beta(z, Z_c) < 1$*, and* $\beta(z', Z_{c'}) < 1$*, then* $Z_c \geq Z_{c'}$*.*

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 $B(1-\beta, z, Z_c) = \beta z(1 + (1-\beta)A\left(1 - \exp(-\nu M_c)\right)\bar{z}_c z)$
- Paper contains illustrative example with 275 cities that quantitatively matches Zipf's law, premia-population correlation, and size-invariant housing expenditure shares
- Numerical comparative static: 10% increase in A leaves the power-law exponent virtually unchanged and increases both the economy-wide average skill premium and the population elasticity of skill premia by 7%–8%.

Davis and Dingel (2019) summary

A microfounded account of skill-biased agglomeration that matches the facts:

- Cities facilitate idea exchange, more skilled value learning opportunities more, and more skilled are more valuable idea-exchange partners
- Spatial sorting of skilled explains spatial variation in skill premia
- Nests Black, Kolesnikova, Taylor (2009) two-type logic

Related to recent job-market papers on spatial sorting of heterogeneous agents

- Lhuillier (2024): Spatial sorting based on human capital accumulation
  - ⋆ Absent migration costs, model with supermodular learning function and $t+1$ payoff works just like model with instantaneous supermodular benefits
- Oh (2024): Cities are too large when worker-firm matching opportunities are rationed by congestion costs (cities as platforms)
  - ⋆ Normative analysis of model akin to worker-worker matching when $\nu \to \infty$ so idea-exchange benefits depend only on composition, not scale

Skill-biased agglomeration around the world

Most evidence is for United States or Western Europe.

Dingel, Miscio, Davis (JUE 2021):

- Use lights at night in satellite imagery to define metropolitan areas for Brazil, China, and India
- In all three developing economies, larger cities are skill-abundant (using years of schooling as proxy for skills)
- In Brazil, college wage premia are higher in more populous cities, consistent with developed-economy patterns

Agglomeration has become more skill-biased since 1980

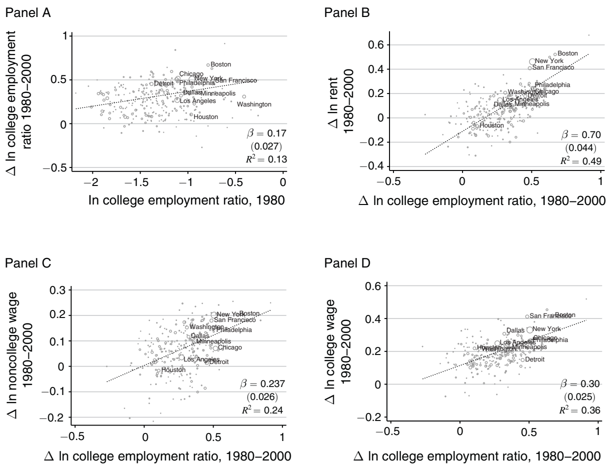


FIGURE 1. CHANGES IN WAGES, RENTS, AND COLLEGE EMPLOYMENT RATIOS, 1980–2000

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

Davis and Dingel (2019): spatially neutral SBTC ($\uparrow A$) raises skill premia in larger cities, consistent with 1990-2007 increase

Two-type touchstone: Diamond (AER 2016)

Since 1980, college graduates have been concentrating in more skilled US cities.

Those cities have faster wage and housing-price growth. Questions:

- Why are college graduates choosing already-skilled cities?
- Is this spatial divergence associated with greater welfare inequality?

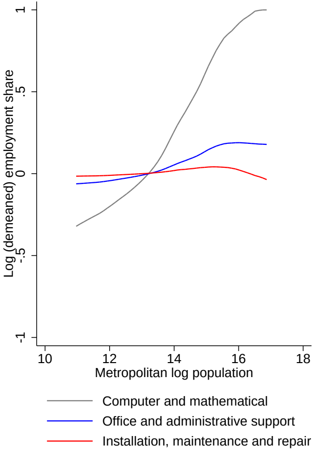
Read Diamond (2016) to learn about a bunch of relevant concepts and tools:

- Great Divergence (Berry Glaeser 2005, Moretti's *New Geography of Jobs*)
- Endogenous amenities (Milena Almagro's 2023 summer school slides)
- Inferring welfare with multiple types (c.f. Moretti 2013)
- "Bartik (1991)" shift-share instruments (AKM 2019, BJH 2022, GPSS 2020)
- Housing supply elasticities (Glaeser Gyourko 2005, Saiz 2010)

This paper uses shift-share IVs to estimate parameters of a one-sector model

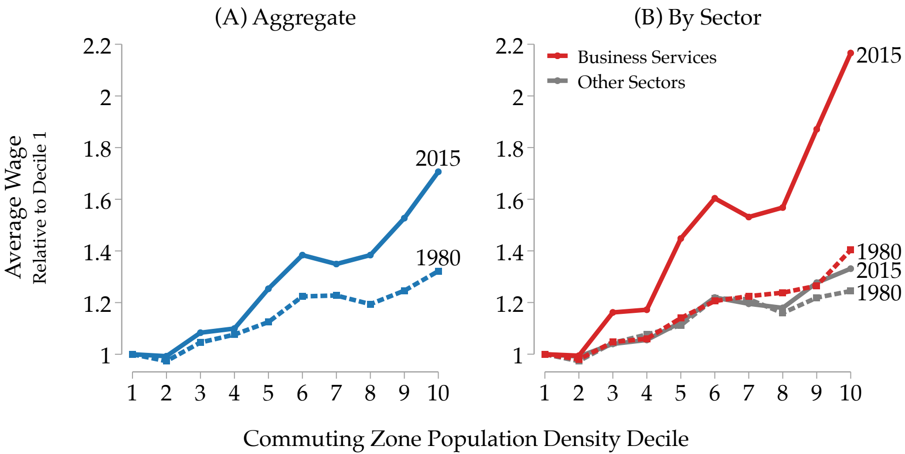
Explaining sectors: The Comparative Advantage of Cities

- Davis and Dingel (2020) describe comparative advantage of cities as jointly governed by individuals' comparative advantage and locational choices
- City-level TFP is endogenous outcome of agglomeration economies and locations within cities vary in their desirability
- 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



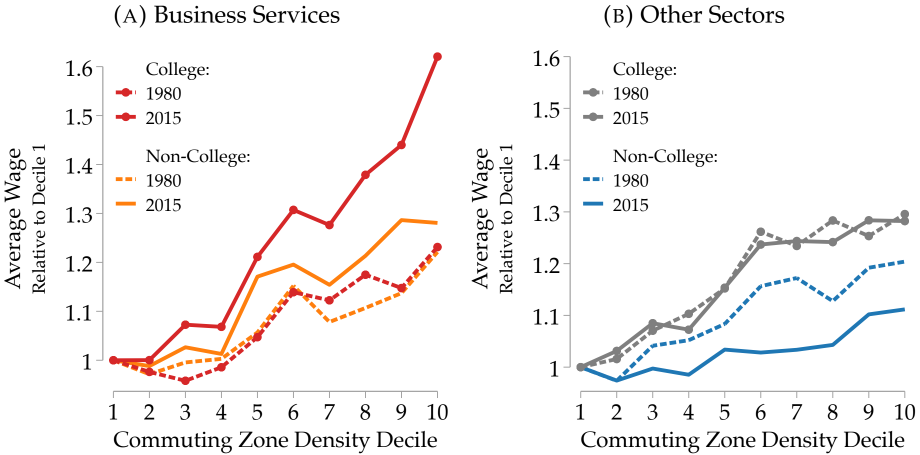
One key sector: Business services (Eckert, Ganapati, Walsh 2024)

FIGURE 1: THE US WAGE-DENSITY GRADIENT IN 1980 AND 2015

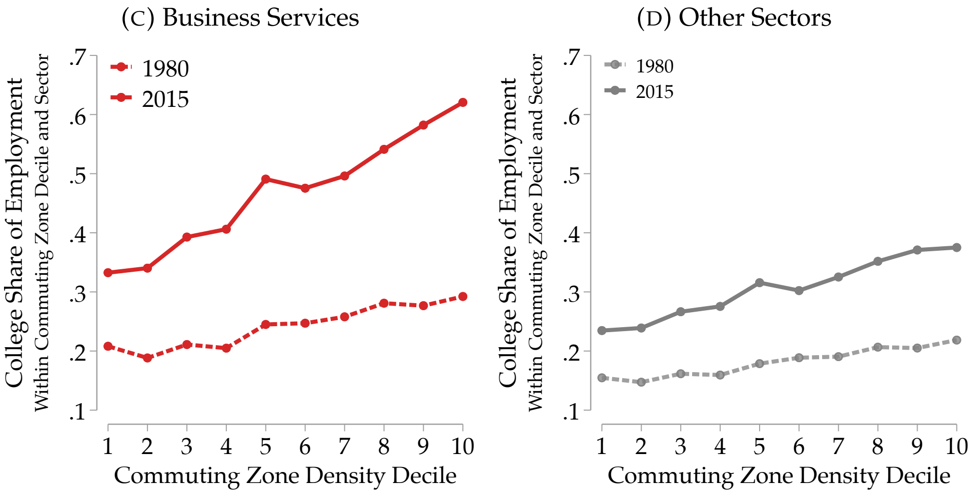


One key sector: Business services (Eckert, Ganapati, Walsh 2024)

Figure 6: The Role of Education in Urban-Biased Growth



One key sector: Business services (Eckert, Ganapati, Walsh 2024)



Eckert, Ganapati, Walsh (2024)

Empirical patterns:

- Business Services sector accounts for almost all urban-biased wage growth
- Large (> 100 emp) Business Services establishments account for 70%
- IT investment is concentrated in large, urban Business Services firms

Quantitative model:

- Scale effect from complementarity between firm size and capital
- Investment-specific technical change reduces price of IT
- Growth accounting exercise: observed decline in IT capital prices alone explains most urban-biased growth since 1980
- Absent scale effect, very little urban-biased growth

See also Chen, Novy, Perroni, Wong (2023) on urban-biased structural change

Skill-biased agglomeration and dynamics

My (static) account with continuum of types:

- Davis & Dingel (2019): Endogenous process is idea exchange between heterogenous agents with skill complementarity
- Davis & Dingel (2020): Larger cities are skill-abundant and specialize in more skill-intensive industries and occupations

Some of the benefits are dynamic (Glaeser & Maré 2001, Carlsen Rattsø & Stokke 2016, de la Roca & Puga 2017, Card Rothstein & Yi 2024)

- Large cities have faster wage growth and human capital is portable

Models of human-capital dynamics with heterogeneous agents in recent JMPs:

- Paolo Martellini: Life-cycle model of spatial sorting
- Hugo Lhuillier: Spatial variation in peer effects on human-capital growth
- Levi Crews: Aggregate growth rate depends on spatial distribution

Agglomeration and dynamics

Cities as engines of growth in human capital and human-capital-driven growth:

- Lucas (1988): "But from the viewpoint of a technology through which the average skill level of a group of people is assumed to affect the productivity of each individual within the group, a national economy is a completely arbitrary unit to consider."
- Estimates from dlRP, CRS, CRY all say that workers learn more in big cities
- Models of knowledge diffusion as source of growth (Lucas 2009; Lucas, Moll 2014; Perla, Tonetti 2014; Buera, Lucas 2018) neglect cities
- Optimal spatial policy should account for growth effects

Other life-cycle components of spatial choices I'll neglect today:

- Kids and suburbs: delayed childbearing explains half of urban revival (Moreno-Maldonado, Santamaria 2024)
- Retirees value amenities and prices, not wages (Komissarova 2022; c.f. Albert, Monras 2022)

Crews (2023) "Dynamic Spatial Knowledge Economy": Overview

Theory: Local human capital externalities → agglomeration & growth

- system of cities
- heterogeneous workers **learn & migrate** over the life cycle
- human capital process drives **both** agglomeration and growth
  - learn from others in your city, more if bigger or more skilled (*local externalities*)
  - learning → human capital dist. shifts right → output grows
- characterize "cities drive growth": **growth rate = f(spatial distribution)**

Solves the **hard problem of regional econ** (Breinlich, Ottaviano, Temple 2014)

- "How to model growth and agglomeration as outcomes of a joint process"
- Agents must know *distribution* of economic activity over *time & space* → high-dimensional
- *how?* **Equilibrium is a mean field game** (Achdou et al 2022) → can track distribution

Crews (2023) "Dynamic Spatial Knowledge Economy": Overview

Theory: Local human capital externalities → agglomeration & growth

- system of cities
- heterogeneous workers **learn & migrate** over the life cycle
- human capital process drives **both** agglomeration and growth
  - learn from others in your city, more if bigger or more skilled (*local externalities*)
  - learning → human capital dist. shifts right → output grows
- characterize "cities drive growth": **growth rate** $=$ f(**spatial distribution**)

Solves the **hard problem of regional econ** (Breinlich, Ottaviano, Temple 2014)

- "How to model growth and agglomeration as outcomes of a joint process"
- Agents must know *distribution* of economic activity over *time & space* → high-dimensional
- *how?* **Equilibrium is a mean field game** (Achdou et al 2022) → can track distribution

Crews (2023) “Dynamic Spatial Knowledge Economy”: Overview

2. **Quantitative: Using US data, jointly rationalize...**

   - **...urban cross-section**:
     - match city size distribution
     - big cities more productive, more expensive, more skilled on avg.
   - **...worker panels**:
     - life-cycle of human capital investment (Ben-Porath 1967; Huggett, Ventura, Yaron 2006
     - migration driven by expected income; young & educated move more (Kennan, Walker 2011)
     - city size wage premium = higher wage level + **faster wage growth w/ permanent value**
   - **...aggregate growth**: 2% per year on BGP

Crews (2023) "Dynamic Spatial Knowledge Economy": Overview

3. **Long-run effects of place-based policy**
   - policy: relax land-use regulations in NYC and SF to US median
   - outcome: aggregate growth **increases by 13bp**
   - **through what channel?**
     - *not* syphoning skill from elsewhere
     - instead, stronger dynamic spillover → **faster human capital accumulation**

   Spatial **policy** → Δ spatial **distribution** → Δ **growth** in two (complementary) ways:

   - by *attracting* more skilled workers to particular cities (e.g., push *skilled* to NYC)
   - by *producing* more skilled workers for whole economy (e.g., push *young* to NYC)

For details of model (setup, equilibrium, balanced-growth path, and main result), see slides on Levi's webpage

Crews (2023) "Dynamic Spatial Knowledge Economy": Overview

3. **Long-run effects of place-based policy**
    - policy: relax land-use regulations in NYC and SF to US median
    - outcome: aggregate growth **increases by 13bp**
    - **through what channel?**
        - *not* syphoning skill from elsewhere
        - instead, stronger dynamic spillover → **faster human capital accumulation**

    Spatial **policy** → Δ spatial **distribution** → Δ **growth** in two (complementary) ways:

    - by *attracting* more skilled workers to particular cities (e.g., push *skilled* to NYC)
    - by *producing* more skilled workers for whole economy (e.g., push *young* to NYC)

For details of model (setup, equilibrium, balanced-growth path, and main result), see slides on Levi's webpage

Crews (2023) "Dynamic Spatial Knowledge Economy" summary

- **A classic hypothesis** (Jacobs 1969, Lucas 1988, Glaeser 2011) **...**

	human capital spillovers	$\rightarrow$ agglomeration (cities!)
+	human capital accumulation	$\rightarrow$ growth
	human capital accumulation s.t. local spillovers	$\rightarrow$ **"cities drive growth"**

  ... but no models $\rightarrow$ no testing, no counterfactuals, no optimal policy

- Forward-looking dynamics + place-specific conditions $\rightarrow$ *high-dimensional*

- **Crews (2023)**: Introduce new tools/model + apply to U.S. data & policy
  1. characterize "cities drive growth": growth rate = f(spatial distribution)
  2. rationalize patterns in U.S. data: worker panels, city cross-section, aggregate BGP trend
  3. policy counterfactual: relax LURs in NYC and SF $\rightarrow \uparrow$ aggregate growth 13bp

Summary

- Spatial distributions of skills and sectors are prominent in public discussion of cities, exploited for variation in empirical work, and potentially key to understanding agglomeration processes
- Agglomeration has skill-biased productivity benefits: larger cities have higher relative quantities and relative wages for skilled
- We need models with more than two skills groups and more than perfectly specialized/diversified cities to understand empirical patterns
- Agglomeration has become more skill-biased since 1980; business services seem important driver of recent decades
- Dynamics: Still much to investigate in the story of Lucas (1988, §6)

More references

See the summer 2020 *JEP* symposium

- Gilles Duranton and Diego Puga - “The Economics of Urban Density”
- Stuart S. Rosenthal and William C. Strange - “How Close Is Close? The Spatial Reach of Agglomeration Economies”
- William R. Kerr and Frederic Robert-Nicoud - “Tech Clusters”
- Gaetano Basso and Giovanni Peri - “Internal Mobility: The Greater Responsiveness of Foreign-Born to Economic Conditions”

Great Divergence

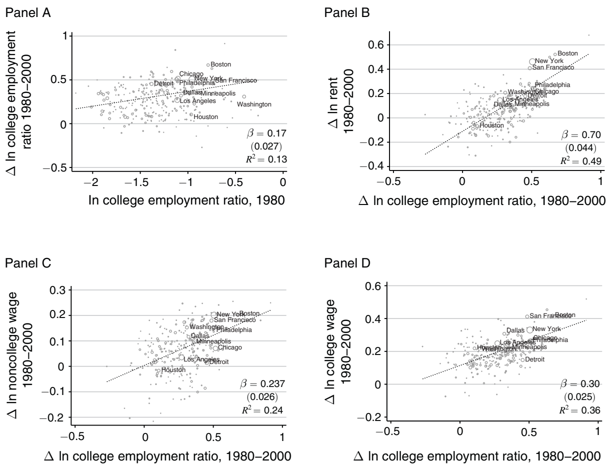


Figure 1. Changes in Wages, Rents, and College Employment Ratios, 1980–2000

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 (1/2)

- “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.

The “Bartik” instrument for local labor demand (2/2)

BJH:

> *The claim for instrument validity in shift-share instrumental variable (SSIV) regressions must rely on some assumptions about the shocks, exposure shares, or both.*

GPSS say shares are exogenous.

> *Once we think about the shares as the instruments, the implied empirical strategy is an exposure research design, where the industry shares measure the differential exogenous exposure to the common shock*

AKM and BJH say shocks are exogenous. BJH on ADH example:

> *exogeneity of industry employment shares is difficult to justify a priori since unobserved industry shocks (e.g., automation or innovation trends) are likely to affect regional outcomes through the same mixture of exposure shares. Our approach, in contrast, allows researchers to specify a set of shocks that are plausibly uncorrelated with such unobserved factors.*

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)

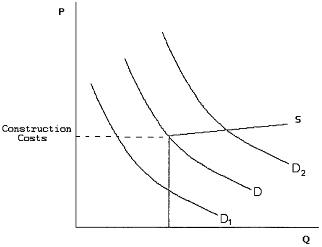
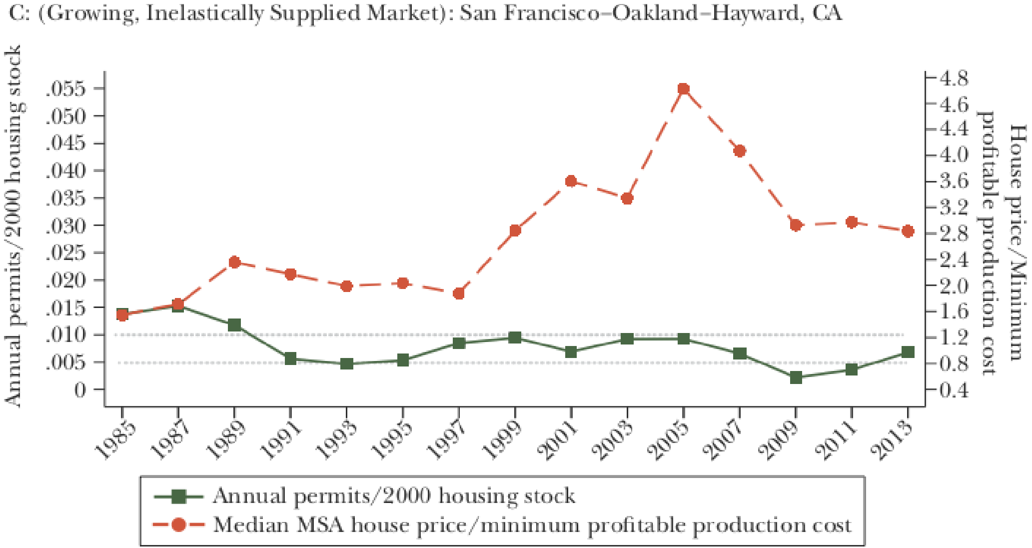


Fig. 1.—The nature of housing supply and construction costs

Glaeser & Gyourko (JPE 2005)

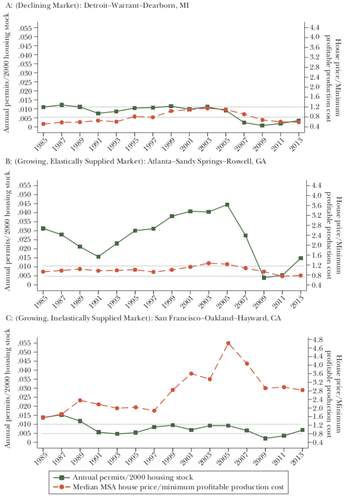
Housing supply elasticities – prices vs quantities



Housing supply elasticities – prices vs quantities

Figure 2

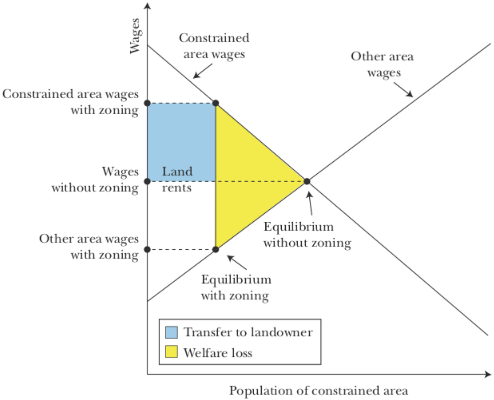
New Housing Supply and House Prices (Relative to Costs)



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 land share, and 17 percent gross margin homes.

Incidence of building restrictions

Figure 4
Welfare Consequences of Restricting Development in a Productive Market



Glaeser & Gyourko (JEP 2018)

Housing is durable $\rightarrow$ filtering

How do you build housing for poor people? You build housing for rich people in the past. Rosenthal (*AER* 2014):

> *While filtering has long been considered the primary mechanism by which markets supply low-income housing, direct estimates of that process have been absent. This has contributed to doubts about the viability of markets and to misplaced policy. I fill this gap by estimating a "repeat income" model using 1985-2011 panel data. Real annual filtering rates are faster for rental housing (2.5 percent) than owner-occupied (0.5 percent), vary inversely with the income elasticity of demand and house price inflation, and are sensitive to tenure transitions as homes age. For most locations, filtering is robust which lends support for housing voucher programs.*

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

$$w_{jt}^{H} = \ln W_{jt}^{H} = c_t + (1-\rho) \ln N_{jt} + (\rho - 1) \ln H_{jt} + \ln\left(f_H(H_{jt}, L_{jt})\right) + \epsilon_{jt}^{H} \tag{4}$$

$$w_{jt}^{L} = \ln W_{jt}^{L} = c_t + (1-\rho) \ln N_{jt} + (\rho - 1) \ln H_{jt} + \ln\left(f_H(H_{jt}, L_{jt})\right) + \epsilon_{jt}^{L} \tag{5}$$

$$w_{jt}^{H} = g_H\left(H_{jt}, L_{jt}\right) + \epsilon_{jt}^{H} \approx \gamma_{HH} \ln H_{jt} + \gamma_{HL} \ln L_{jt} + \epsilon_{jt}^{H} \qquad (7) \approx (9)$$

$$w_{jt}^{L} = g_L\left(H_{jt}, L_{jt}\right) + \epsilon_{jt}^{L} \approx \gamma_{LH} \ln H_{jt} + \gamma_{LL} \ln L_{jt} + \epsilon_{jt}^{L} \qquad (8) \approx (10)$$

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:

$$\Delta Z_{jt} \in \left\{ \Delta B^H_{jt}, \Delta B^L_{jt}, \Delta B^H_{jt} x^{reg}_j, \Delta B^H_{jt} x^{geo}_j, \Delta B^L_{jt} x^{reg}_j, \Delta B^L_{jt} x^{geo}_j \right\}$$

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

$$E\left(\Delta\tilde{\epsilon}^H_{jt}\Delta Z_{jt}\right) = 0 \quad E\left(\Delta\tilde{\epsilon}^L_{jt}\Delta Z_{jt}\right) = 0$$

"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" ($\Delta\xi^z_{jt} \equiv \beta^A \mathbf{z} \Delta \mathbf{x}^A_{jt}$):

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

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

$$E\left(\Delta\epsilon^a_{jt}\Delta Z_{jt}\right) = 0$$

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"
- Immigrants more attracted to higher nominal wage and less deterred by higher nominal rents (c.f. Albert and Monras 2022)

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 (0.25–1.5m) college wage premium	.09	0.0978	0.0937
		(0.00578)	(0.00613)
Large-city (>1.5m) 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

NOTES: This table describes full-time, full-year employees ages 18-55. Following BSP, "college graduate" means a bachelor's degree or greater educational attainment. 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. Robust standard errors in parentheses.