Agglomeration Economies

Urban Economics

Ignacio Sarmiento-Barbieri

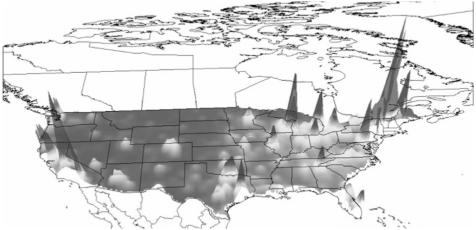
Universidad de los Andes

September 25, 2024

Agenda

Agglomeration Economies

- Why do we see such a remarkable clustering of human activity in a small number of urban areas?



Evidence of Agglomeration Economies

- Three strategies to identify agglomeration economies

  1. Show there is too much spatial concentration to be random (Duranton and Overman, 2005)
  2. Compare productivity over space (Greenstone et al., 2010)
  3. Compare wages and rents across space (Quantitative Spatial Models, Ahlfeldt et al, 2015)

Agenda

Spatial Concentration

Extremes of Localization and Dispersion

(c) Other Agricultural and Forestry Machinery (SIC2932)

(d) Machinery for Textile, Apparel and Leather Production (SIC2954)

Spatial Concentration

Ambiguous Cases

(a) Basic Pharmaceuticals (SIC2441)

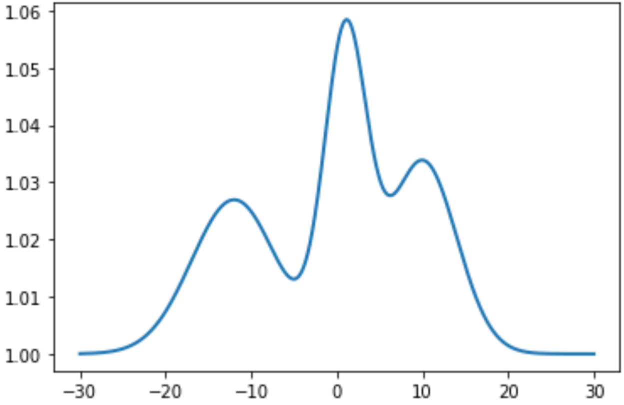
(b) Pharmaceutical Preparations (SIC2442)

Duranton and Overman

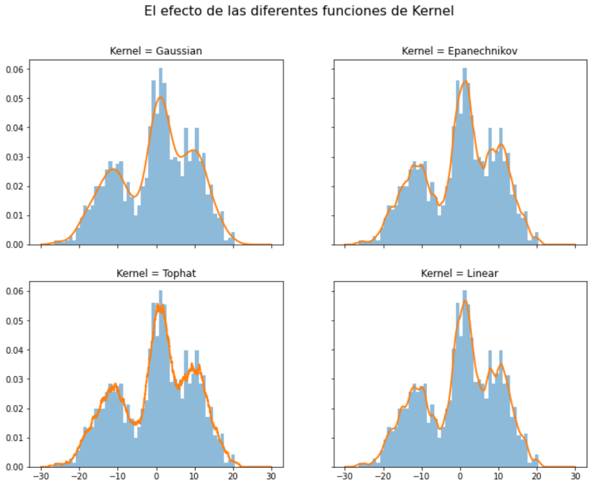
1. **Select Relevant Establishments:**
   - Choose establishments based on industry and size.
   - Consider different thresholds to assess robustness (e.g., include only those contributing to 90% of employment).
2. **Compute Bilateral Distances:**
   - Calculate Euclidean distances between all pairs of establishments.
   - Use Kernel Density Estimation (KDE) to estimate the density of these distances.

Spatial Concentration

K Density Estimates

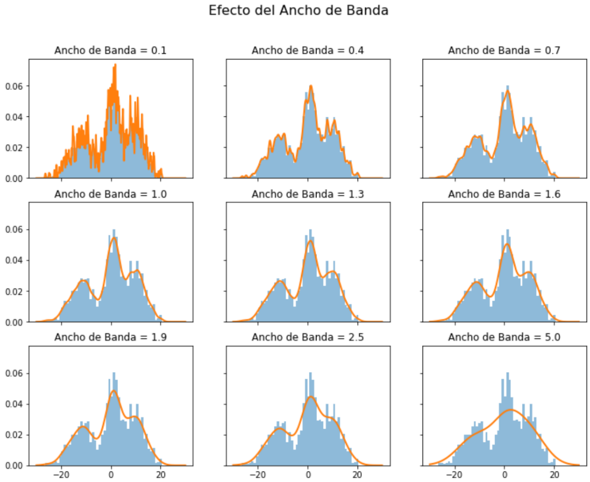


Spatial Concentration
K Density Estimates



Spatial Concentration

K Density Estimates



Duranton and Overman

Density estimates

$$\hat{K}(d) = \frac{1}{n(n-1)h} \sum_{i=1}^{n-1} \sum_{j=1+1}^{n} K\left(\frac{d - d_{i,j}}{h}\right) \tag{1}$$

- $d_{i,j}$ the observed distance between establishments i and j
- Gaussian kernel
- Section 3.4.2 of Silverman (1986): $h = 1.059\sigma n^{-1/5}$; $h = 0.79 IQR \sigma n^{-1/5}$; $h = 0.79 min\{\sigma, IQR/1.34\} n^{-1/5}$

Duranton and Overman

3. **Generate Counterfactuals:**
   - Randomly assign establishments to locations, maintaining the number of establishments and industrial concentration.
   - Create 1,000 simulations to construct a baseline for comparison.

(a) Basic Pharmaceuticals (SIC2441) (b) Pharmaceutical Preparations (SIC2442)

Duranton and Overman

4 Statistical Significance and Localization Metrics

Local confidence intervals are constructed for each distance d by comparing the observed K-density, $\hat{K}(d)$, with the distribution of simulated K-densities from random assignments.

$$\bar{K}_A(d) = \text{5th percentile of } \hat{K}_{\text{sim}}(d), \tag{2}$$

$$\underline{\mathrm{K}}_A(d) = \text{95th percentile of } \hat{K}_{\text{sim}}(d), \tag{3}$$

Interpretation:

- If $\hat{K}(d) > \bar{K}_A(d)$, it indicates significant **localization** at distance d.
- If $\hat{K}(d) < \underline{\mathrm{K}}_A(d)$, it indicates significant **dispersion** at distance d.

Duranton and Overman

4 Statistical Significance and Localization Metrics

Local confidence intervals are constructed for each distance d by comparing the observed K-density, $\hat{K}(d)$, with the distribution of simulated K-densities from random assignments.

We can also define an index of localization:

$$\gamma_A(d) \equiv \max\left(\hat{K}_A(d) - \bar{K}_A(d), 0\right), \tag{4}$$

as well as an index of dispersion:

$$\psi_A(d) \equiv \max\left(\bar{K}_A(d) - \hat{K}_A(d), 0\right). \tag{5}$$

Interpretation:

- To reject the hypothesis of randomness at distance d because of localization (dispersion), we only need $\gamma_A(d) > 0$ ($\psi_A(d) > 0$).
- The exact value of these two indices does not matter. However, the indices do indicate how much localization and dispersion there is at any level of distance.

Duranton and Overman

4 **Statistical Significance and Localization Metrics**

Global confidence bands to make statements about the overall location patterns of an industry.

- Denote $\bar{\bar{K}}_A(d)$ **the upper confidence band of industry** A.
- This band is hit by 5% of our simulations between 0 and 180 km.
- If

$$\hat{K}_A(d) > \bar{\bar{K}}_A(d)$$

for at least one $d \in [0, 180]$, this industry is said to exhibit **global localization** (at a 5% confidence level).

Duranton and Overman

4 **Statistical Significance and Localization Metrics**

Global confidence bands to make statements about the overall location patterns of an industry.

- An industry which is very localized at short distances can show dispersion at larger distances.
- Then an industry is said to exhibit **global dispersion** (at a 5% confidence level) when

$$\hat{K}_A(d) < \underline{\underline{K}}_A(d)$$

for at least one $d \in [0, 180]$ **and** the industry does not exhibit localization.

Duranton and Overman

4 **Statistical Significance and Localization Metrics**

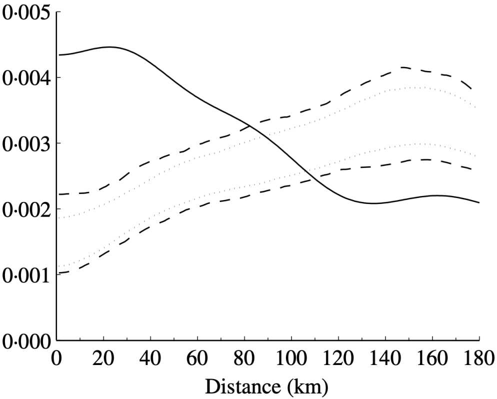
We can also define an index of global localization:

$$\gamma_A(d) \equiv \max\left(\hat{K}_A(d) - \bar{\bar{K}}_A(d), 0\right), \tag{6}$$

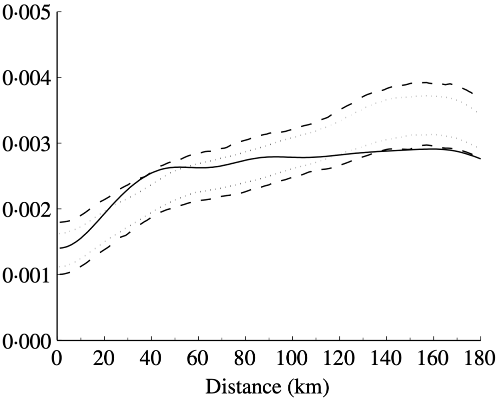
as well as an index of global dispersion:

$$\psi_A(d) \equiv \begin{cases} \max\left(\underline{\underline{K}}_A(d) - \hat{K}_A(d), 0\right) \text{ if } \sum_{d=0}^{d=180} = 0 \\ 0 \text{ o.w.} \end{cases} \tag{7}$$

Spatial Concentration



(a) Basic Pharmaceuticals (SIC2441)



(b) Pharmaceutical Preparations (SIC2442)

Agenda

1 Motivation

2 Duranton and Overman

3 Quantitative Spatial Models

Quantitative Spatial Models

- One of the most striking economic features of our world is the uneven distribution of economic activity across space.
- These rich patterns of the concentration of economic activity can be explained by a three-way interaction between natural advantages, agglomeration forces, and dispersion forces.
- The complexity of modeling these forces in spatial equilibrium has meant that the traditional theoretical literature on cities focused on stylized settings, such as a monocentric city or the Rosen Roback framework.
  - These models, however, **cannot** capture the rich internal variation in patterns of economic activity within and between real world locations

Quantitative Spatial (Urban) Models

- Although traditional models in urban economics explain certain features of the data, their simplifying assumptions olimit their usefulness for empirical work.
- These simplifying assumptions abstract from empirically relevant differences in natural advantage across locations, such as access to natural harbors or green parks.
- To address these limitations, recent quantitative spatial (urban) models allow for empirically relevant differences in natural advantage while also incorporating agglomeration forces.
- These models are designed to connect directly to observed data on cities.

Introduction to a basic quantitative spatial model

- Consider a city (or coutry), embedded in a wider economy.
- The city/country consists of a set of discrete locations (blocks/cities).
- These locations are populated by workers, who are mobile between locations and the larger economy.
- Workers have idiosyncratic preferences for living and working in different locations within the city/country).
- They consider all the personal, work-related, or amenity-related reasons and pick the locations that yields the highest utility.

Introduction to a basic quantitative spatial model

- We begin with a twist to Rosen-Roback; between cities
- Rosen–Roback key insight is that any local shock to the demand or supply of labor in a city is, in equilibrium, fully capitalized in the price of land.
- As a consequence, shocks to a local economy do not affect worker welfare.
- This rule's outs some interesting questions

Introduction to a basic quantitative spatial model

- The goal of the model is to clarify what happens to wages, costs of housing and worker utility when a local economy experiences a shock to labor demand or labor supply.
- An example of a shock
  - to labor demand is an increase in productivity.
  - to labor supply is an increase in amenities.

Introduction to a basic quantitative spatial model

- We'll work through a 2 location case to develop intuition, but it can be easily extended to n locations
- We assume that workers and firms are mobile across cities, but worker mobility is not necessarily infinite, because workers have **idiosyncratic preferences for certain locations**.

Rosen-Roback model: Exogenous Prices

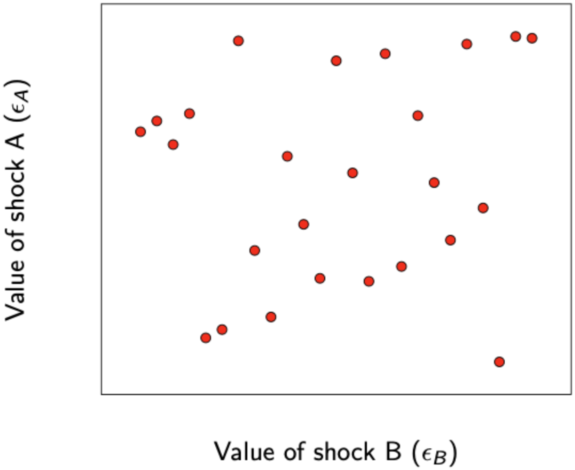
- Assume wages, rents, amenities are exogenous
- Person i's indirect utility of being in A:

$$V_A^i = w_A - r_A + A_A + \epsilon_A^i \tag{8}$$

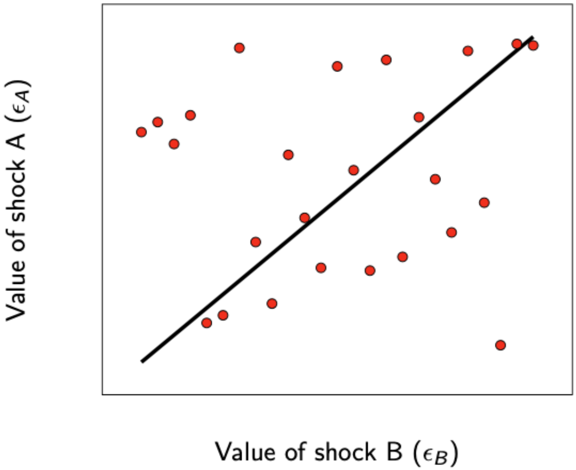
- Person i's indirect utility of being in B:

$$V_B^i = w_B - r_B + A_B + \epsilon_B^i \tag{9}$$

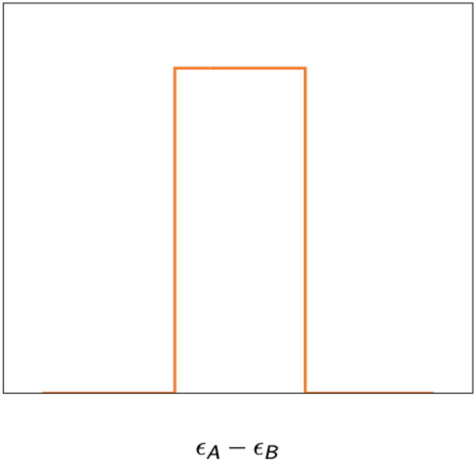
Rosen-Roback model: Exogenous Prices



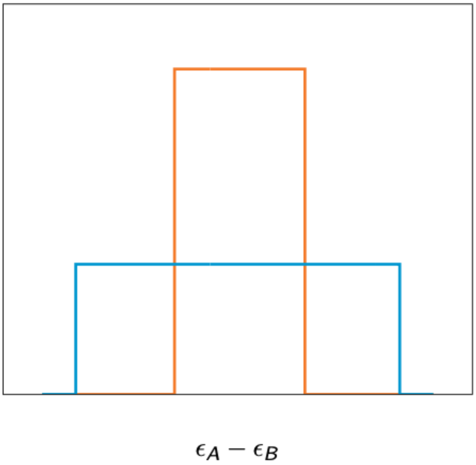
Rosen-Roback model: Exogenous Prices



Rosen-Roback model: Exogenous Prices



Rosen-Roback model: Exogenous Prices



Firms

- The production function for firms in city c is Cobb–Douglas with constant returns to scale, so that

$$\ln y_c = X_c - hN_c + (1-h)K_c \tag{10}$$

- where X_c is a city especific productivity shifter.
- there is an international capital market, and that capital is infinitely supplied at a given price i

Construction

- The supply of housing is

$$r_c = z + k_c N_c \tag{11}$$

- number of housing units in city c is assumed to be equal to the number of workers.
- k_c is the elasticity of the supply of housing

Equilibrium

- Local labor market supply

$$w_B = w_A + (r_B - r_A) + (A_A - A_B) + s\frac{(N_B - N_A)}{N} \tag{12}$$

- Labor demand

$$w_c = X_c - (1-h)N_c + (1-h)K_c + ln(h) \tag{13}$$

- Local housing demand

$$r_B = r_A + (w_B - w_A) + (A_B - A_A) - s\frac{(N_B - N_A)}{N} \tag{14}$$

- Housing supply

$$r_c = z + k_c N_c \tag{15}$$

Effect of a labor demand shock on wages and prices

- Two periods
  - Period 1: cities are identical
  - Period 2: TFP increases in b: $X_{B2} = X_{B1} + \Delta$ where $\Delta > 0$
- Workers are more productive in B than A.

Effect of a labor demand shock on wages and prices

Change in nominal wages?

$$w_{B2} - w_{B1} = \Delta \tag{16}$$

$$w_{A2} - w_{A1} = 0 \tag{17}$$

Effect of a labor demand shock on wages and prices

Change in population?

$$w_{B2} = w_{A2} + (r_{B2} - r_{A2}) + s\frac{(N_{B2} - N_{A2})}{N} \tag{18}$$

$$w_{B1} = w_{A1} + (r_{B1} - r_{A1}) + s\frac{(N_{B1} - N_{A1})}{N} \tag{19}$$

$$\Delta = k_B N_{B2} - k_A N_{A2} - k_B N_{B1} + k_A N_{A1} + s\frac{(N_{B2} - N_{A2} - N_{B1} + N_{A1})}{N} \tag{20}$$

$$(N_{B2} - N_{B1}) = \frac{N}{N(k_B + k_A) + 2s}\Delta \geq 0 \tag{21}$$

Effect of a labor demand shock on wages and prices

Change in housing markets?

- In B

$$r_{B2} - r_{B1} = \frac{Nk_B}{N\left(k_B + k_A\right) + 2s}\Delta \geq 0 \tag{22}$$

- In A

$$r_{A2} - r_{A1} = \frac{-k_A N}{N\left(k_B + k_A\right) + 2s}\Delta \leq 0 \tag{23}$$

Effect of a labor demand shock on wages and prices

Change in Real wages?

- In B

$$(w_{B2} - w_{B1}) - (r_{B2} - r_{B1}) = \frac{Nk_A + 2s}{N(k_B + k_A) + 2s}\Delta \geq 0 \tag{24}$$

- In A

$$(w_{A2} - w_{A1}) - (r_{A2} - r_{A1}) = \frac{k_A N}{N(k_B + k_A) + 2s}\Delta \geq 0 \tag{25}$$

Effect of a labor demand shock on wages and prices

Marginal worker?

$$V_c^i = w_c - r_c + A_c + \epsilon_c^i \tag{26}$$

- The change in the relative preference for city a of the marginal worker is equal a

$$\left(\epsilon_{A2}^i - \epsilon_{B2}^i\right) - \left(\epsilon_{A1}^i - \epsilon_{B1}^i\right) = \frac{2s}{N\left(k_B + k_A\right) + 2s}\Delta \geq 0 \tag{27}$$

- The marginal worker in period 2 is different from the marginal worker in period 1.
- Since city b offers higher real wages in period 2, the new marginal worker in period 2 has stronger preferences for city a.

Effect of a labor supply shock on wages and prices

- Two periods
  - Period 1 both cities are identical
  - Period 2 amenity increases in B: $A_{B2} = A_{B1} + \Delta'$ where $\Delta' > 0$

Spatial Equilibrium with Agglomeration Economies

- Consider the case where there are agglomeration economies so that the productivity of firms in a locality is an endogenous function of the level of economic activity in that locality.
- This amounts to endogenizing the city-specific productivity shifter.
- Eg. productivity in a locality is a function of the number of workers in that locality

$$X_c = f(N_c) \tag{28}$$

- with $f' > 0$
- Decisions of workers generates a positive externality.

Spatial Equilibrium with Agglomeration Economies

- Assume

$$X_c = x_c + \gamma N_c \tag{29}$$

- The MPL

$$w_c = x_c + (\gamma - (1-h))\, N_c + (1-h)\, K_c + ln(h) \tag{30}$$

Spatial Equilibrium with Agglomeration Economies

- Two periods
  - Period 1 both cities are identical
  - Period 2 amenity increases in B: $X_{B2} = X_{B1} + \Delta$ where $\Delta > 0$