A Unified Theory of Cities

Jacques-François Thisse

UCLouvain and HSE

Matthew Turner

Brown and NBER

Philip Ushchev

HSE

March 2022

People can choose where to work and where to live, it is costly to commute, there may be increasing returns to scale in production. What happens?

We address this question in a model with heterogenous households, discrete space and local returns to scale.

Idea: Apply new QSM toolbox to old stylized geography to understand IRS and QSM.

- This is a defining question for urban economics and is central to the study of economic growth.
- It will help us distinguish theorems from facts in the QSM literature, and may suggest testable implications of these models.

Results

We characterize equilibrium throughout the parameter space.

- Returns to scale can be a centralizing or decentralizing; income effects matter. Commuting costs can be centralizing or decentralizing.
- The discrete choice framework with commute costs implies preferences for central work and residence. This is not like classical urban economics.
- Corner equilibria occur when there are increasing returns to scale. Multiple interior equilibria may exist and are discontinuous at empirically relevant thresholds of returns to scale.
- Corner equilibria are unstable. Stable equilibria need not exist. Stability is subtle.
- As returns to scale increases, model behavior resembles a simplified history of cities from the medieval period to the present.

Literature

- *Urban Economics*; continuous space, homogenous agents, simple geographies, analytic/numerical solutions: Ogawa and Fujita (1980), Fujita and Ogawa (1982), Lucas (2001), Lucas and Rossi-Hansberg (2002).
- *QSM*; discrete space, realistic geographies, numerical solutions: Ahlfeldt et al. (2015), and: Severen (2018), Balboni (2019), Dingel and Tintelnot (2020), Allen, Arkolakis and Li (2015), Tsivanidis (2019), Heblich, Redding and Sturm (2020), Herzog (2020).

Model I

A static city with heterogenous households, discrete space, and local returns to scale

- A city is a finite set of locations $\mathcal{I} = \{-1, 0, 1\}$. One unit of land per location.
- Population is a continuum $[0, 1]$. Household ν chooses residence $i \in \mathcal{I}$, workplace $j \in \mathcal{I}$, tradable numeraire good and land.
- Usually, a city is three identical locations uniformly spaced along a line, the simplest geography with a 'center' (cf. 'linear city').
- Households commute between workplace and residence. Commuting incurs an iceberg cost $\tau_{ij} \geq 1$. Commute cost is constant for all households and psychic.

Model II

- W_j is wages, R_i is land rent. Households have indirect utility

$$V_{ij}(\nu) = z_{ij}(\nu)\frac{W_j}{\tau_{ij}R_i^{\beta}}.$$

- Household $\nu \in [0,1]$ has a type $\mathbf{z}(\nu) \equiv (z_{ij}(\nu)) \in \mathbb{R}_+^{3\times 3}$. $\mathbf{z}(\nu) : [0,1] \to \mathbb{R}_+^{3\times 3}$. Distribution of types is the product measure of 9 identical Fréchet distributions. ε is (inverse) taste dispersion.
- The set of types $\mathbf{z}$ such that ij is (weakly) preferred to all other location pairs rs is

$$S_{ij} = \left\{\mathbf{z} \in \mathbb{R}_+^{3\times 3}\,;\; V_{ij}(\mathbf{z}) = \max_{r,s\in\mathcal{I}} V_{rs}(\mathbf{z})\right\}$$

- Share of households at ij

$$s_{ij} = \mu\left(\mathbf{z}^{-1}(S_{ij})\right) = \frac{\left[W_j/(\tau_{ij}R_i^{\beta})\right]^{\varepsilon}}{\sum_{r\in\mathcal{I}}\sum_{s\in\mathcal{I}}\left[W_s/(\tau_{rs}R_r^{\beta})\right]^{\varepsilon}}.$$

Model III

- Costlessly traded numeraire produced at j

$$y_j = A_j L_j^{\alpha} N_j^{1-\alpha},$$

A_j depends only on the mass of employment at j,

$$A_i = L_i^{\gamma} \text{ for } \gamma \geq 0$$

- *Say that increasing returns to scale (IRS) are:*
  i. *weak* $\Leftrightarrow 0 < \gamma < \gamma_m$,
  ii. *moderate* $\Leftrightarrow \gamma_m < \gamma < \gamma_s$,
  iii. *strong* $\Leftrightarrow \gamma > \gamma_s$,

  for $\gamma_m = \alpha/\varepsilon$ *and* $\gamma_s \in (\gamma_m, \frac{1+\varepsilon}{(1-\beta)\varepsilon} - \alpha]$.
- (1) If population is homogeneous (ε infinite), weak IRS cannot occur. (2) If commute costs are high or land unproductive, moderate IRS cannot occur. (3) We experimented with 'spillovers'.

Model IV

- A *spatial pattern* is an element of R_+^3: residence, (M_{-1}, M_0, M_1); employment, (L_{-1}, L_0, L_1), housing, (H_{-1}, H_0, H_1); commercial land, (N_{-1}, N_0, N_1); wages, (W_{-1}, W_0, W_1); rent, (R_{-1}, R_0, R_1). *Symmetric* if peripheral values are equal.
- *Centrality ratios*:

$$m \equiv \frac{M_0}{M_1}; \quad \ell \equiv \frac{L_0}{L_1}; \quad h \equiv \frac{H_0}{H_1};$$
$$n \equiv \frac{N_0}{N_1}; \quad w \equiv \frac{W_0}{W_1}; \quad r \equiv \frac{R_0}{R_1}.$$

- Note that, e.g., symmetry and ℓ and $L_0 + 2L_1 = 1$ identifies employment pattern.

Model V

A spatial equilibrium is a set of patterns, $\{\mathbf{M}^*, \mathbf{L}^*, \mathbf{H}^*, \mathbf{N}^*, \mathbf{W}^*, \mathbf{R}^*\}$, such that

i. All households make utility-maximizing choices of workplace, residence, housing and consumption,
ii. The tradable good is produced by a competitive sector. First order conditions are satisfied wherever production occurs.
iii. All households live and work somewhere, and
iv. Land markets clear.

Linear City with three Locations I

- Iceberg commuting cost matrix

$$\begin{pmatrix} \tau_{-1,-1} & \tau_{-1,0} & \tau_{-1,1} \\ \tau_{0,-1} & \tau_{0,0} & \tau_{0,1} \\ \tau_{1,-1} & \tau_{1,0} & \tau_{1,1} \end{pmatrix} = \begin{pmatrix} 1 & \tau & \tau^2 \\ \tau & 1 & \tau \\ \tau^2 & \tau & 1 \end{pmatrix}.$$

- To ease notation, we define

$$\begin{aligned} a &\equiv \frac{\beta\alpha}{1-\alpha} = \frac{\text{housing share} \times \text{labor share}}{\text{commercial land share}} > 0, \\ \phi &\equiv \tau^{-\varepsilon}. \end{aligned}$$

and

$$\omega \equiv w^{\varepsilon} \quad \text{and} \quad \rho \equiv r^{-\beta\varepsilon},$$

Linear City with three Locations II

- Suppose $W_j / R_i^{\beta} = V$, then a household faces the discrete choice

$$\max_{ij} \left\{ \begin{array}{ccc} z_{-1,-1} V, & \frac{z_{-1,0}}{\tau} V, & \frac{z_{-1,1}}{\tau^2} V \\ \frac{z_{0,-1}}{\tau} V, & z_{0,0} V, & \frac{z_{0,1}}{\tau} V \\ \frac{z_{1,-1}}{\tau^2} V, & \frac{z_{1,0}}{\tau} V, & z_{1,1} V \end{array} \right\}.$$

- The distribution of tastes is identical for each of the pairwise choices $\Rightarrow$ average utility for central residence

$$E\left(\max\left\{\frac{z_{0,-1}}{\tau} V,\ z_{0,0} V,\ \frac{z_{0,1}}{\tau} V\right\}\right) = \Gamma\left(\frac{\varepsilon - 1}{\varepsilon}\right)\left[(1 + \frac{2}{\tau^{\varepsilon}})\right]^{1/\varepsilon} V,$$

$\Rightarrow$ average utility for peripheral residence,

$$E\left(\max\left\{z_{1,1} V,\ \frac{z_{1,0}}{\tau} V,\ \frac{z_{1,-1}}{\tau^2} V\right\}\right) = \Gamma\left(\frac{\varepsilon - 1}{\varepsilon}\right)\left[(1 + \frac{1}{\tau^{\varepsilon}} + \frac{1}{\tau^{2\varepsilon}})\right]^{1/\varepsilon} V.$$

- $1 + \frac{1}{\tau^{\varepsilon}} + \frac{1}{\tau^{2\varepsilon}} < 1 + \frac{2}{\tau^{\varepsilon}}$, so the payoff for peripheral residence is less than central. There is an *average preference for central residence*. Choice of workplace is symmetric.

Linear City with three Locations: Equilibrium I

Summing up, we have:

- Two centralizing forces: average preference for central work, average preference for central residence. These are not conventional ‘agglomeration forces’; they attach to a place, not to density.
- Two centrifugal forces: congestion in residential land market and congestion in commercial land market. Both push activity toward land abundant periphery.
- One conventional ‘agglomeration force’, increasing returns to scale.

What should happen?

- Central land rent and wage should capitalize preferences for central work and residence, $R_0 / R_1 \uparrow$, $W_0 / W_1 \downarrow$.
- Central/peripheral specialization in residence/employment should depend on land demand. The activity with the greatest demand for land should be relatively peripheral.
- Returns to scale is complicated. It increases the returns to concentrated employment and incomes.

Linear City with three Locations: Equilibrium II

Proposition 1: *The equilibrium demand for housing and commercial space and the equilibrium supply of workers and residents are:*

$$M_0 = \frac{\rho(\omega + 2\phi)}{\rho(\omega + 2\phi) + 2(\phi\omega + 1 + \phi^2)} \qquad M_1 = \frac{1 - M_0}{2}$$

$$L_0 = \frac{\omega(\rho + 2\phi)}{\omega(\rho + 2\phi) + 2(\phi\rho + 1 + \phi^2)} \qquad L_1 = \frac{1 - L_0}{2}$$

$$H_0 = \frac{a\rho(1 + 2\phi\omega^{-\frac{1+\varepsilon}{\varepsilon}})}{a\rho(1 + 2\phi\omega^{-\frac{1+\varepsilon}{\varepsilon}}) + \rho + 2\phi} \qquad N_0 = 1 - H_0$$

$$H_1 = \frac{a(\phi\omega^{\frac{1+\varepsilon}{\varepsilon}} + 1 + \phi^2)}{a(\phi\omega^{\frac{1+\varepsilon}{\varepsilon}} + 1 + \phi^2) + \phi\rho + 1 + \phi^2} \qquad N_1 = 1 - H_1.$$

Linear City with three Locations: Equilibrium III

Proposition 2: *Assume $\gamma \neq \alpha/\varepsilon$. Then, a pair (ρ^*, ω^*) is an interior spatial equilibrium if and only if it solves the following two equations:*

$$\omega^{\frac{1+\varepsilon}{\varepsilon}} = f(\rho) \equiv \frac{\phi\rho - 2a\phi\rho^{1+\frac{1}{\beta\varepsilon}} + (1+\phi^2)(1+a)}{(1+a)\rho^{1+\frac{1}{\beta\varepsilon}} + 2\phi\rho^{\frac{1}{\beta\varepsilon}} - a\phi},$$

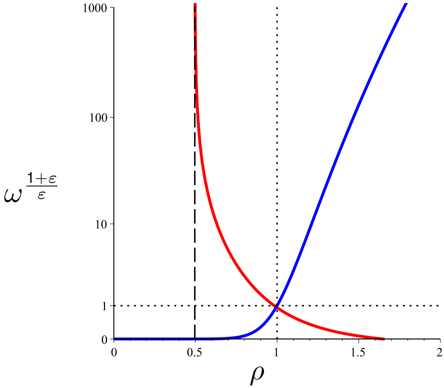
$$\omega^{\frac{1+\varepsilon}{\varepsilon}} = g(\rho;\gamma) \equiv \rho^{\frac{b}{1-\gamma\varepsilon/\alpha}} \left(\frac{\rho + 2\phi}{\phi\rho + 1 + \phi^2}\right)^{\frac{\gamma\varepsilon/\alpha}{1-\gamma\varepsilon/\alpha}\frac{1+\varepsilon}{\varepsilon}}$$

- existence $\Longleftrightarrow$ f and g cross.
- uniqueness/multiplicity $\Longleftrightarrow$ count crossings.
- (1) f monotone decreasing with asymptote ρ_0, the zero of the denominator. (2) $g' > 0 \Longleftrightarrow \gamma\varepsilon/\alpha < 1$ and $g \to$ step function as $\gamma \to \alpha/\varepsilon = \gamma_m$ with step at ρ_L,.
- $\rho_0 < \rho_L \Longleftrightarrow$ Commute cost or land productivity high.

Existence of Equilibrium

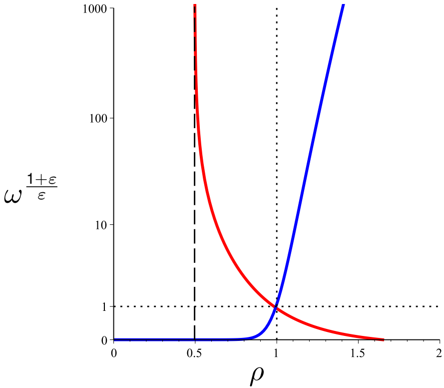
Proposition 3: If $\gamma \geq 0$, an interior equilibrium always exists. Furthermore, there exist corner equilibria if and only if $\gamma > 0$. These corner equilibria are given by $L_0^* = 1$ and $L_1^* = 1/2$.

(1) Commute cost or land productivity high, $\gamma = 0$, $f = g$



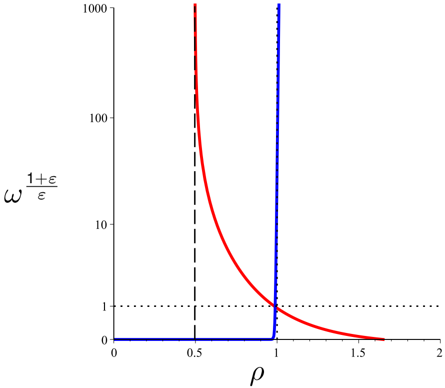
Constant returns to scale. Unique interior equilibrium. ρ^*, $\omega^* < 1$.

(2) Commute cost or land productivity high, $\gamma < \gamma_m$, $f = g$



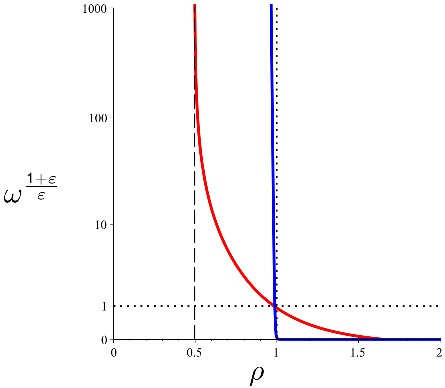
Increasing returns to scale are *weak* ($\gamma = 0.16$). Unique interior equilibrium. $\rho^* < 1$.

(3) Commute cost or land productivity high, $\gamma < \gamma_m$, $f = g$



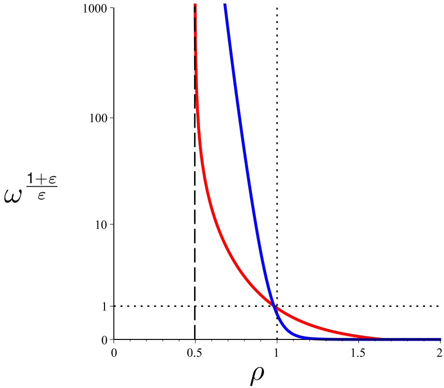
Increasing returns to scale are *weak* ($\gamma \nearrow \gamma_m$, $\gamma = 0.38$). Unique interior equilibrium. $\rho^* < 1$.

(4) Commute cost or land productivity high, $\gamma > \gamma_m$, $f = g$



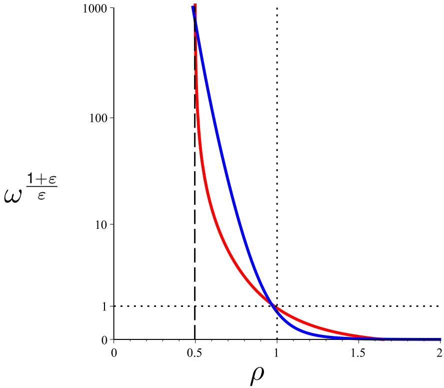
Increasing returns to scale are *moderate* ($\gamma \searrow \gamma_m$, $\gamma = 0.42$).
Three interior equilibria.

(5) Commute cost or land productivity high, $\gamma \in (\gamma_m, \gamma_s)$, $f = g$



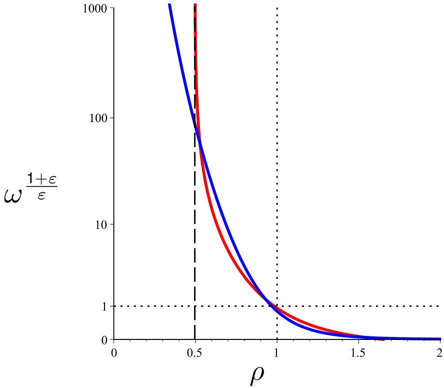
Increasing returns to scale are *moderate* ($\gamma = 0.66$). Three interior equilibria.

(6) Commute cost or land productivity high, $\gamma \in (\gamma_m, \gamma_s)$, $f = g$



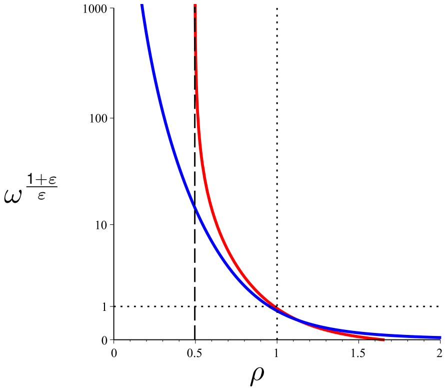
Increasing returns to scale are *moderate* ($\gamma = 0.90$). Three interior equilibria.

(7) Commute cost or land productivity high, $\gamma \in (\gamma_m, \gamma_s)$, $f = g$



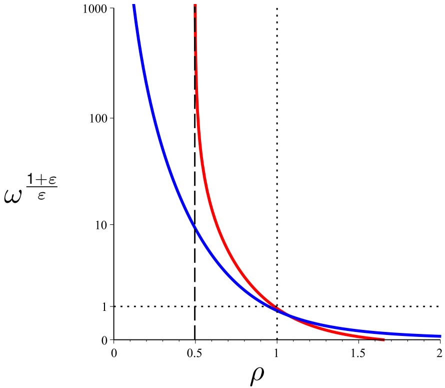
Increasing returns to scale are *moderate* ($\gamma = 1.14$). Three interior equilibria.

(8) Commute cost or land productivity high, $\gamma > \gamma_s$, $f = g$



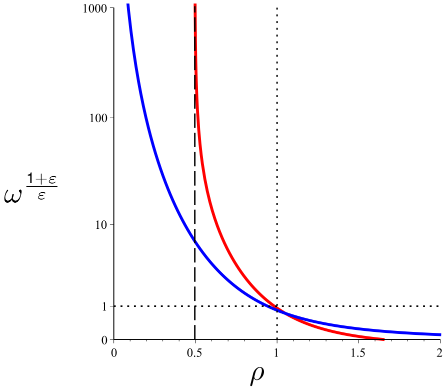
Increasing returns to scale are *strong* ($\gamma = 1.62$). Unique interior equilibrium. $\rho^* \to 1$.

(9) Commute cost or land productivity high, $\gamma > \gamma_s$, $f = g$



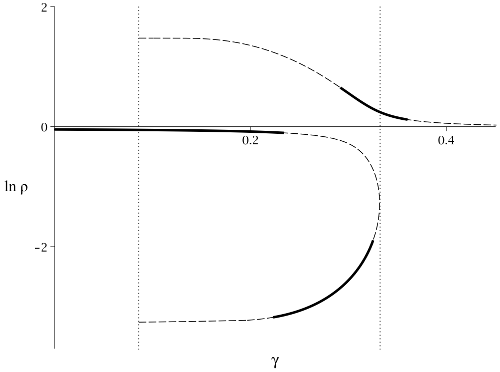
Increasing returns to scale are *strong* ($\gamma = 1.86$). Unique interior equilibrium. $\rho^* \to 1$.

(10) Commute cost or land productivity high, $\gamma > \gamma_s$, $f = g$



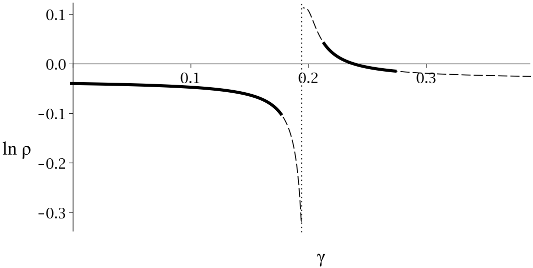
Increasing returns to scale are *strong* ($\gamma = 2.10$). Unique interior equilibrium. $\rho^* \to 1$.

Commute cost or land productivity high, summary



Interior equilibria only. Corner equilibria occur if $\gamma > 0$.

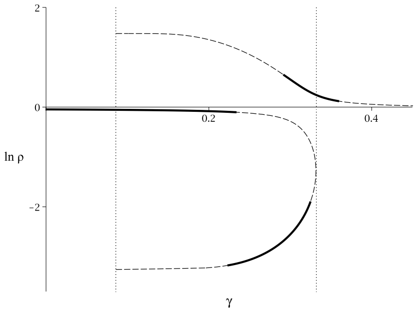
Commute cost and land productivity low, summary



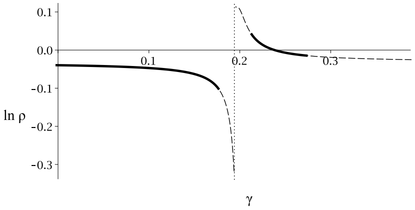
Interior equilibria only. Corner equilibria occur if $\gamma > 0$.

Comparative statics for γ

τ or β high



τ and β low



- Equilibrium is always discontinuous at γ_m.
- Economic activity always concentrates and then disperses as γ increases.
- The figures do not show corner equilibria.
- These figures are a complete characterization of equilibrium. Our propositions are slightly less complete.

Stability I

Stability is subtle. Three candidate definitions,

- Define a system of differential equations in L_0, M_0 and s_{00} with the same steady states as $f(\rho) = g(\rho, \gamma)$. This requires ad hoc assumptions about the adjustment process.
- 'An iterative process will find it'. If $f(\rho) = g(\rho, \gamma)$ defines equilibrium, then $h(\rho) = \rho$ for$h = f^{-1}(g)$ does too, and this is 'stable' if $|h'| < 1$. But...

$$h(\rho) = \rho \Longrightarrow h(\rho) = \theta\rho + (1-\theta)h(\rho)$$

and this means that

$$\widetilde{h}(\rho) = \frac{(h(\rho) + (1-\theta)\rho)}{\theta} = \rho.$$

So $\widetilde{h}(\rho)$ also defines equilibrium and is not stable for θ small. That is, iterative stability is not well defined.

Stability II

- Something ‘like’ trembling hand perfection. This is a static notion, like our model, and does not require us to specify an adjustment process.

Stability III

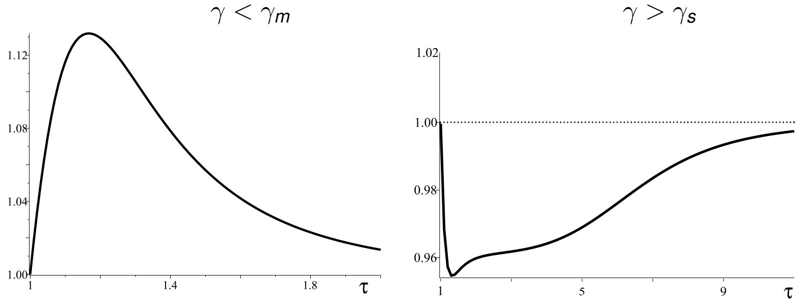
- Say an equilibrium is unstable if for all $\Delta > 0$ we can find a measure $\Delta > 0$ of people and work-residence pairs *ij* and *kl* such that if we move Δ from *ij* to *kl* they will not want to move back. Stable otherwise.
- We can always find an agent indifferent between any pair *ij* and *kl* (support of z_{ij}'s is unbounded).
- Check stability by checking if this indifferent agent can be made better off by moving, along with $\Delta > 0$ almost indifferent agents.
- Analytic results: Corner and 'near corner' equilibria are unstable.
- We have a rule that allows us to evaluate the stability of any equilibrium numerically.
- In every case we have checked, stability is similar to the last figure.

Comparative statics for τ, ε

Recall that $\phi = \tau^{-\varepsilon}$ is the 'spatial discount factor' or iceberg commute cost to travel one unit distance.

- $\phi \to 1$ as $\tau \to 1$ or $\varepsilon \to 0$
- $\phi \to 0$ as $\tau \to \infty$ or $\varepsilon \to \infty$.

Comparative statics for τ, ε



Notes: x-axis is τ, y-axis is $\ell = L_0 / L_1$.

- Why?
- $\tau \to \infty$ and $\varepsilon \to 0$ are easy.
- $\tau \to 1$ is 'like' $\varepsilon \to 0$. $\varepsilon \to \infty$ is 'like' $\tau \to \infty$.
- Decreasing commute costs can concentrate or disperse employment.

A 500 year history of returns to scale

- Agglomeration effect are estimated to be around 5% using contemporary data.
- We don't have corresponding estimates historically. But
  - The number and size of cities increases slowly from 1500 to the present (Bairoch, 1988),(de Vries, 1984).
  - The urban share of population increases slowly from 1500 to the present.

  this must be partly because cities are getting more productive.
- The spread of printing (Dittmar, 2011), universities (Cantoni and Yuchtman, 2014), and apprenticeship (de la Croix et al., 2018) gradually increased the productivity of pre-industrial cities.
- Pre-industrial manufacturing was often organized by 'letting out' (de Vries, 1984).
- This suggests that returns to scale in cities have been increasing for a long time.

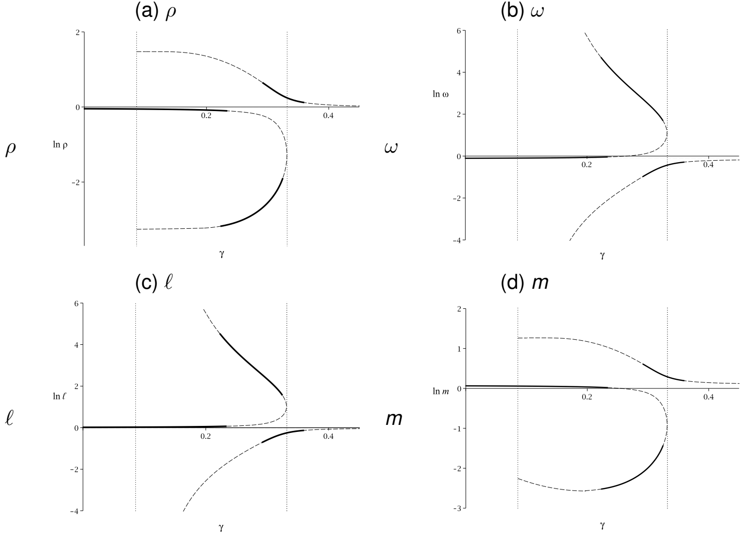
A 500 year history of cities

- Pre-industrial walled cities were for protection not production. Many lived inside and worked outside. Poverty was pervasive and extreme (de Vries, 1984). Farmers were taller/healthier than city dwellers (Costa, 2015).
- Mill towns were very dense, were organized around production, and saw wages increase for urban residents (Clark, 1951).
- Pedestrian based mill towns give way to less dense cities with extensive suburbs.
- Complicated multicentric cities arise with edge cities and complicated commuting patterns (Glaeser and Kahn (2004), Garreau (1992), McMillen and McDonald (1998)).
- The late 20th century started to see a resurgence of central cities (Couture and Handbury (2020), Couture et al. (2019)).

Simplified history versus model

- Think of history as a gradual increase in γ.
- Ignore other changes in parameters, we don't really need them.
- Focus attention on the stable $\rho_0 < \rho_L$ equilibria.

Evolution of cities and returns to scale



(1) $\rho_0 < \rho_L$ equilibria. (2) The *x*-axis is γ. (3) Note concentration pre-industrial revolution, dispersion post (4) Define the 'industrial revolution' as the first discontinuity. 20th century suburbanization is second discontinuity/multiple equilibrium.

Spillovers I

Suppose that productivity depends on nearby employment, not just own employment (like Fujita and Ogawa (1982)),

$$A_0 = (L_0 + 2\delta L_1)^\gamma, \qquad A_1 = (\delta L_0 + (1+\delta^2)L_1)^\gamma.$$

- $\delta \in [0, 1)$ distance decay rate.
- $\gamma \geq 0$ local IRS (as earlier).
- So we have $A_i L_i^\alpha N_i^{1-\alpha} \to (L_i^\gamma) L_i^\alpha N_i^{1-\alpha}$ as $\delta \to 0$ and spillovers stop.
- This technology nests the one used everywhere else.
- Note that spillovers and local returns to scale are different.

Spillovers II

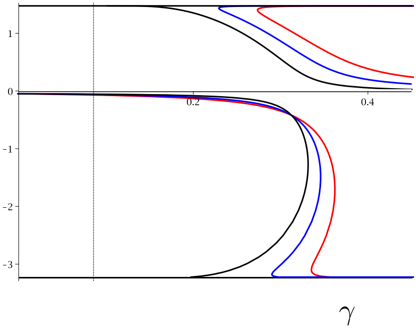
- We can solve for

$$f(\rho) = g(\rho, \gamma, \delta)$$

  just as we did with purely local spillovers.
- We investigate equilibrium numerically for $\rho_0 < \rho_L$.
- Interior equilibrium appear to change continuously with δ
- Corner equilibria disappear when $\delta > 0$

Spillovers III

Figure: Equilibrium with productivity spillovers.



Notes: Equilibrium correspondences between $\ln(\rho)$ *and* γ. *Black* $\delta = 0$, *interior and corner equilibria. Blue,* $\delta = 0.02$. *Red,* $\delta = 0.04$.

Conclusion I

We characterize equilibrium in a discrete city where people choose where to work and live, and production is subject to returns to scale. This is a central question for urban economics.

- Returns to scale can centralize or decentralize. Income effects are important. Commuting costs centralize or decentralize.
- Corner equilibria are pervasive. Multiple interior equilibria may exist. Corners are unstable. All interior equilibria can be unstable.
- The discrete choice framework with commute costs implies preferences for central work and residence, not like the standard literature.
- Interior equilibria are discontinuous at an empirically relevant threshold value of returns to scale.
- As returns to scale increases, model behavior resembles a simplified history of cities from the medieval period to the present.

Conclusion II

- With the introduction of spillovers, corner equilibria disappear. Interior equilibria appear to change continuously.

Bairoch, P. (1988). *Cities and economic development: from the dawn of history to the present*. University of Chicago Press.

Cantoni, D. and Yuchtman, N. (2014). Medieval universities, legal institutions, and the commercial revolution. *The Quarterly Journal of Economics*, 129(2):823–887.

Clark, C. (1951). Urban population densities. *Journal of the Royal Statistical Society*, 114(4):490–496.

Costa, D. L. (2015). Health and the economy in the united states from 1750 to the present. *Journal of Economic Literature*, 53(3):503–70.

Couture, V. and Handbury, J. (2020). Urban revival in america. *Journal of Urban Economics*, 119:103267.

de Vries, J. (1984). *European Urbanization, 1500-1800*. Routledge.

Dingel, J. I. and Tintelnot, F. (2020). Spatial economics for granular settings. Technical report, National Bureau of Economic Research.

Dittmar, J. E. (2011). Information technology and economic change: the impact of the printing press. *The Quarterly Journal of Economics*, 126(3):1133–1172.

Fujita, M. and Ogawa, H. (1982). Multiple equilibria and structural transition of non-monocentric urban configurations. *Regional Science and Urban Economics*, 12(2):161–196.

Garreau, J. (1992). *Edge city: Life on the new frontier*. Anchor.

Glaeser, E. L. and Kahn, M. E. (2004). Sprawl and urban growth. In *Handbook of Regional and Urban Economics*, volume 4, pages 2481–2527. Elsevier.

Herzog, I. (2020). The city-wide effects of tolling downtown drivers: Evidence from london's congestion charge. Technical report, University of Toronto.

Lucas, R. E. (2001). Externalities and cities. *Review of Economic Dynamics*, 4(2):245–274.

Lucas, R. E. and Rossi-Hansberg, E. (2002). On the internal structure of cities. *Econometrica*, 70(4):1445–1476.

McMillen, D. P. and McDonald, J. F. (1998). Suburban subcenters and employment density in metropolitan chicago. *Journal of Urban Economics*, 43(2):157–180.

Ogawa, H. and Fujita, M. (1980). Equilibrium land use patterns in a nonmonocentric city. *Journal of Regional Science*, 20(4):455–475.

Severen, C. (2018). Commuting, labor, and housing market effects of mass transportation: Welfare and identification. Technical report, Federal Reserve Bank of Philadelphia.

Tsivanidis, N. (2019). The aggregate and distributional effects of urban transit infrastructure: Evidence from Bogotá's Transmilenio. *Unpublished manuscript*.