Spatial Agglomeration Dynamics

By DANNY QUAH*

Income inequality across geography is as profound as it is across people. In spatial inequality, agglomeration and clustering constitute the observations to be formalized and explained. Understanding their evolution draws on ideas in economic development, growth, and economic geography.

When spatial inequality analyses are motivated by contrasting, say, New York City and Yuma, Arizona, the set of economic forces a researcher identifies distinguishes prototypes of the two locations one has in mind. New York might have economic activity showing high increasing returns; Yuma, by contrast, might produce goods with significant transport costs. The reasoning, canonical in economic geography, then addresses why New York enjoys an income level higher than Yuma's; that is, it explains income inequality.

That analysis is silent, however, on a number of interesting questions. Why are the clusters potentially only in Yuma and New York, not anywhere in between or beyond? How many clusters should endogenously emerge? If N locations are a priori possible, does Yuma/New York reasoning predict $N/2$ high-income agglomerations, or just 1? If $N/2$, are they interspersed in between low-income points, or do they collect all together at one end of the physical geography? (And what if N is not given?) Does it matter that from Yuma, New York is 2,100 miles northeast, or would the same reasoning work for comparing Yuma with San Francisco? How do spatial relations matter? Put differently, where is geography in this model of economic geography?

This paper describes research in spatial dynamics that address these and related questions (Paul Krugman and Anthony Venables, 1997; Quah, 2000, 2001).[1] I illustrate the ideas in a dynamic perfect-foresight equilibrium model that integrates growth, geography, and distribution in an explicit geographical space, namely a three-dimensional globe. The model determines the number of spatial modes (agglomerations, clusters) in economic activity on this globe, the locations of these agglomerations relative to one another, and their dynamics along convergent paths to balanced-growth steady state. (Multiple spatial modes can imply the kind of twin-peaked income distributions described in Quah [1997].)

The model is neoclassical with the driver of economic growth being technology, or knowledge accumulation.[2] In the model, new knowledge is generated exogenously. Transportation costs are zero so that knowledge potentially disseminates freely across space. However, in any given location and at any given time, the effective use of knowledge depends on past choices made there on the use of knowledge, and on current and past choices made in surrounding locations. The model displays imperfect productivity spillovers across space and time and determines spatial and dynamic fluctuations jointly.

Uniformity, an egalitarian spatial income distribution, is always an equilibrium and characterizes balanced-growth steady state. However, spatial agglomerations or clusters appear in perfect foresight Cass-Koopmans saddlepath transitions: such inequality dynamics are *necessary* for convergence to balanced-growth steady state.

Close in spirit to this paper (despite the differences in model, methods, and conclusions), Jan Eeckhout and Boyan Jovanovic (2001) study knowledge spillovers in production where permanent inequality resolves a tension between

* Department of Economics, London School of Economics, Houghton Street, London WC2A 2AE. I thank the ESRC (award R022250126) and the Andrew Mellon Foundation for supporting parts of the research reported here.

[1] The mathematical tools here might appear unfamiliar but are firmly classical: their core goes back at least to Ulf Grenander and Gabor Szegö (1958) and Alan Turing (1952).

[2] James Feyrer (2001) shows it is total factor productivity, rather than capital or labor, that accounts for the twin-peakedness of the cross-country income distribution given in Quah (1997).

catching-up and free-riding. In Dilip Mookherjee and Debraj Ray (2002), equality is unstable. Here, equality is stable, but spatial inequality is needed to achieve it. Kiminori Matsuyama (2002) studies the stability and general structure of discrete equilibria in complementarity games, in a way related to the concerns expressed above on the unsatisfactory nature of two-point (or, more generally, discrete) equilibrium outcomes.

I. The Model

Let z denote a representative location on a geography $\mathcal{G}$, and normalize $\int_{\mathcal{G}} dz$ to 1. (Section II-B below specializes $\mathcal{G}$ to the surface of a three-dimensional globe, but the discussion until then is general. If it helps intuition, the reader can, without loss, visualize what happens between now and then using that special case.)

The technology level in use at time t in location z is $A_t(z) \geq 0$. Write the spatial profile or distribution as $A_t = \{A_t(z) | z \in \mathcal{G}\}$. In this model, technology and (accumulated) knowledge are synonymous, so that the average state of knowledge worldwide is $\bar{A}_t = \int_{\mathcal{G}} A_t(z)\, dz$. Uniformity has $\bar{A}_t = A_t(z)\ \forall z$. Spatial agglomerations or clusters are modes in the spatial distribution A_t, inducing in turn modes in the spatial distribution of incomes (5) below.

As much A as demanded is supplied: Think of this as "D" in R&D, where R perpetually runs ahead of D (i.e., R grows by rate at least g given in equation (8) below), and is financed by nondistortionary taxation with output made costlessly available to everyone.

Each location is its own infinitesimally small nation. It discounts the future at constant rate $\rho > 0$ and produces gross output

$$(1) \qquad F(A_t(z)|A_t) = W_z(A_t) \times A_t(z).$$

with

$$(2) \qquad W_z(A_t) = \int_{\mathcal{G}} K(z, z')A_t(z')\, dz'$$

where for all z the weighting function $K(z, \cdot)$ is a probability density or probability kernel on $\mathcal{G}$, so that $K(z, z') \geq 0$ for all z' and $\int_{\mathcal{G}} K(z, z')\, dz' = 1$. By equations (1)–(2), output is linear in the state of knowledge $A_t(z)$, with coefficient (marginal product) equal to a weighted average of the current levels of A in the appropriate neighborhood of z.

Weighting function K is time-invariant; allowing it to evolve adds no additional insight. In section II-B I restrict K further, in line with the related specialization of $\mathcal{G}$. Until then, however, the discussion requires no further assumptions on the pair $(\mathcal{G}, K)$.

Robert Lucas (1988) described how, for economic growth, knowledge is necessarily at once global—not Chinese, or Korean, or American. This model maintains that; but how effective global knowledge A is at location z depends both on one's current state of knowledge and on one's neighbors'. Assume that training or retooling costs need to be expended before knowledge can be used in production. These training costs at z are quadratic in A's rate of change:

$$(3) \qquad C(\dot{A}_t(z)) = \frac{1}{2}\zeta \times \dot{A}_t(z)^2 \qquad \zeta > 0.$$

While A's effectiveness in equation (1) is history- and geography-dependent, from equation (3) the cost of changing it is neither, and it is the same everywhere. Transforming global knowledge to local use can also be interpreted as changing general, codifiable knowledge to specific, tacit knowledge.

Coefficient ζ parameterizes retooling costs: The larger is ζ, the more sluggish will be changes in A. Assume that

$$(4) \qquad 4\zeta^{-1} < \rho^2.$$

That is, relative to how much the future is discounted, retooling costs generate sufficiently high sluggishness.

Income, net of retooling costs, is then

$$(5) \qquad y_t(z) = F(A_s(z)|A_s) - C(\dot{A}_s(z)).$$

Economy z at time t solves

$$(6) \qquad \sup_{\{A_s(z): s \geq t\}} \int_{s \geq t} y_s(z) e^{-\rho s}\, ds$$

subject to conditions (1)–(5) and

$$\begin{cases} A_t(z) \\ \{A_s(z') : s \geq t,\ z' \neq z\}. \end{cases}$$

An equilibrium is a collection of mutually consistent time paths $\{A_s(z) : s \geq t\}$, one for each z, solving (6), or alternatively, a time path of profiles $\{A_s : s \geq t\}$ such that each z-section $A(z)$ solves (6) and follows what others expect of z. The equilibrium is rational-expectations Nash in the strategy space comprising time paths $\{A_s(z) : s \geq t\}$ since expectations are realized in equilibrium and since (6) requires that each location select a feasible time path taking as given the choices made in all other locations.

The model is one not only of a set of given locations choosing alternative patterns of development, but upon reinterpretation, also a model of location choice (i.e., of a planner deciding where to place resources, subject to feasibility constraints).

II. Results

At each z, program (6) has as a necessary first-order condition the Euler equation

$$W_z(A_t) + \zeta \frac{d\dot{A}_t(z)}{dt} - \rho\zeta\dot{A}_t(z) = 0$$

which implies the decision rule

$$\text{(7)} \quad \dot{A}_t(z) = \zeta^{-1} \times \int_0^\infty W_z(A_{t+s})e^{-\rho s}\, ds$$

where I have solved stable roots backward and unstable roots forward (see e.g., Thomas Sargent, 1987 Ch. 9).

A. *Balanced-Growth Steady State*

Uniformity then is an equilibrium with $\bar{A}$ growing at proportional rate

$$\text{(8)} \quad g = \frac{\rho}{2} - \frac{(\rho^2 - 4\zeta^{-1})^{1/2}}{2} \in (0, \rho).$$

If, as in Section II-B below, $\mathcal{G}$ is the surface of a three-dimensional globe, then the balanced-growth steady-state equilibrium given in (8) can be visualized easily: Across geography the level of knowledge is a globe concentric with $\mathcal{G}$, growing proportionally outward; so too then the spatial distribution of incomes. Nations everywhere have identical and growing incomes. Convergence, equality, and globalization are total.

Balanced-growth steady-state (8) is a uniform equilibrium, however, even without this restriction on $\mathcal{G}$. To see this, note that at uniformity with $\bar{A}$ growing at rate $g < \rho$, the right-hand side of equation (7) becomes

$$\zeta^{-1} \int_0^\infty \bar{A}_{t+s}e^{-\rho s}\, ds = \zeta^{-1}\bar{A}_t \int_0^\infty e^{-(\rho-g)s}\, ds$$

$$= [(\rho - g)\zeta]^{-1}\bar{A}_t.$$

Location z therefore has

$$\dot{A}_t(z)/\bar{A}_t = [(\rho - g)\zeta]^{-1}$$

so that in uniform equilibrium with $\dot{A}_t(z) = \bar{A}_t$, this becomes

$$g = [(\rho - g)\zeta]^{-1}.$$

This quadratic in g has one root given by (8); the other root exceeds $\rho/2$, implying infinite value to (6).

Notice that $dg/d\rho < 0$ and $dg/d\zeta < 0$. The intuition is straightforward. The more myopic are decision-makers, the lower is the steady-state growth rate; similarly, the higher are retooling costs, the slower the world grows.

The analysis thus far has been blind to any special structure (e.g., radial homogeneity) in $\mathcal{G}$ and K. Only in transition dynamics will that matter.

B. *Transition Dynamics*

Specialize $\mathcal{G}$ to the surface of a three-dimensional globe. Using polar coordinates, the representative location is $z = (z_1, z_2) = e^{-i\omega} = (e^{-i\omega_1}, e^{-i\omega_2})$, with $i = \sqrt{-1}$ and $\omega \in [-\pi, \pi] \times [-\pi, \pi]$.

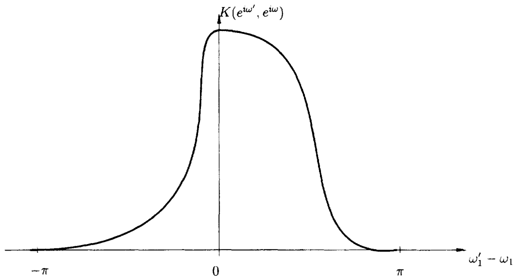

FIGURE 1. WEIGHTING KERNEL

Notes: Spillovers across geography can be asymmetric and unimodal. The section through K shown has $\omega' = (\omega_1', \omega_2')$ and $\omega = (\omega_1, \omega_2)$, with ω_2' and ω_2 fixed.

Next, require that K be nondegenerate [i.e., not constant, and that $K(z, \cdot)$ places positive weight outside a small open neighborhood of z], continuously differentiable, and radially homogeneous [i.e., $K(e^{i\omega'}, e^{i\omega})$ depends only on $\omega' - \omega$]. Radical homogeneity differs from symmetry, which would require instead dependence on $|\omega' - \omega|$.

Figure 1 shows a section through a possible K. By radial homogeneity, K is graphed as a function of only $\omega' - \omega$. Spillover weighting can increase in that separation and be asymmetric about the origin. From both these properties, it differs from the usual decay due to physical distance. A unimodal K is not ruled out and, indeed, will suffice to generate multiple modes in spatial outcomes below.

Outside of steady state, the transition behavior can be quite intricate. To rule out extraordinary but nonetheless uninteresting outcomes, assume that equilibrium is Markov; that is, at each location z there is a time-invariant mapping M_z such that equation (7) becomes $\dot{A}_t(z) = M_z(A_t)$. The present discounted value on the right-hand side of (7) can be calculated as depending only on the current knowledge profile A_t. Detrending both sides around the equilibrium growth path $\bar{A}_0 e^{gt}$ and then stacking into a spatial profile gives

$$\dot{\tilde{A}}_t = \tilde{M}(\tilde{A}_t) \tag{9}$$

which is an ordinary differential equation in the space of bounded positive functions on $\mathcal{G}$.

To obtain intuition for what follows, recall Cass/Koopmans-type dynamics when the transition equation (9) is finite and linear, with steady state $\tilde{A}_t = \bar{\tilde{A}}$; that is,

$$\dot{\tilde{A}}_t = \tilde{M} \times (\tilde{A}_t - \bar{\tilde{A}}). \tag{9'}$$

Suppose the finite matrix $\tilde{M}$ is diagonalizable:

$$\tilde{M} = \mathbf{\Phi V \Phi}^{-1} \qquad \mathbf{V} = \text{diag}\{\nu_1, \nu_2, \ldots\}.$$

Stable or unstable dynamics hinge on the sign of the real parts of eigenvalues ν_j by

$$(\mathbf{\Phi}^{-1}\dot{\tilde{A}}_t)_j = \nu_j \times ((\mathbf{\Phi}^{-1}\tilde{A}_t)_j - (\mathbf{\Phi}^{-1}\bar{\tilde{A}})_j). \tag{10}$$

The convergent manifold is the collection of initial states $\tilde{A}_0$ such that (9′) takes the system to steady state $\bar{\tilde{A}}$. Suppose $\mathbf{V}$ collects all ν_j with negative real parts in its leading entries. From equation (10), the convergent manifold has representation

$$\tilde{q}\colon \tilde{A}_0 - \bar{\tilde{A}} = \mathbf{\Phi} \times \begin{pmatrix} \tilde{q} \\ \underline{0} \end{pmatrix} \tag{11}$$

zeroing out components that multiply into unstable eigenvalues.

In equation (9) for the current model, the convergent manifold (a subset of the collection of spatial functions on $\mathcal{G}$ integrating to zero) is well defined. That manifold is the collection of initial states such that (9) converges to zero. However, $\tilde{M}$ is an infinite-dimensioned nonlinear mapping. Any linearization into an equation such as (9′) will have the eigenvector–eigenvalue decomposition evolving in time.

Recognize, however, that by its construction from a radially symmetric K the operator $\tilde{M}$ is Toeplitz (see e.g., Turing, 1952; Grenander and Szegö, 1958); if it were a matrix, its rows would be simply rotational shifts of one another. Alternatively, every fixed diagonal section comprises only identical entries. Then $\tilde{M}$ has a spectrum (counterpart to the set of eigenvalues) that is discrete and composed of Fourier transforms of horizontal sections of $\tilde{M}$, while its eigenfunctions (counterparts to eigenvectors) comprise only complex exponentials $e^{-i\omega j}$ for integer j.

In parallel with equation (10) the spectrum determines the dynamics of detrended profiles $\tilde{A}$

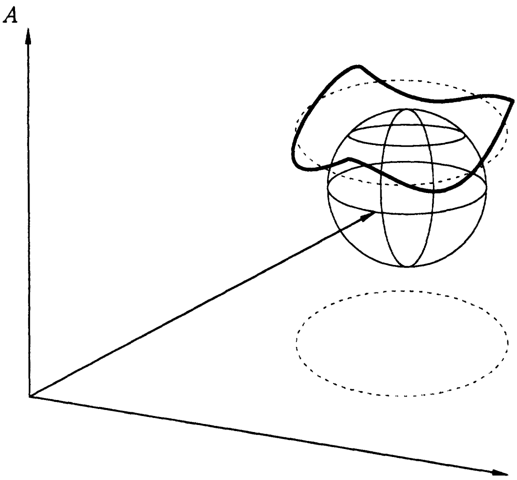

FIGURE 2. SPATIAL AGGLOMERATION ON A THREE-DIMENSIONAL GLOBE

Note: The figure shows a horizontal slice through the local perturbations that converge back to steady state only with clusters in the spatial distributions.

about the steady-state growth path. In parallel with (11) the eigenfunctions determine the convergent manifold. Here, every element of the convergent manifold is a linear combination of complex exponentials (i.e., two-dimensional waves of $\mathcal{G}$). Moreover, when retooling costs ζ are neither too large nor too small, the nullifying of spectral elements required in (11) makes the convergent manifold a strict subset of the full span of the eigenfunctions.[3] Multiple modes then necessarily appear in the spatial distributions that comprise the convergent manifold (see Fig. 2).

Along the equilibrium path, where the spatial agglomerations locate relative to one another in $\mathcal{G}$ (the bumps in Fig. 2) will be determined by the ω's that remain active in (11). These ω's are, in turn, functions of all the parameters of the model (K, ζ, ρ). Economic activity across space has a profile that depends on both those active ω's and the corresponding spectra. The dynamics of the spatial distribution of economic activity can, in turn, be read off equation (10).

[3] In Quah (2000, 2001), I provide explicit technical details for this reasoning.

III. Conclusions

When spatial inequality is studied in a model with discrete locations fixed, many interesting questions cannot be addressed. This paper develops a model of economic growth and activity that permits a richer analysis needed for that discussion.

The spatial neoclassical growth model in this paper has knowledge accumulation as the engine of growth. Equilibrium in the model is rational-expectations and Nash. In equilibrium, knowledge spillovers across geography and optimal knowledge-accumulation decisions determine the distribution of knowledge used across space and over time. The resulting pattern of economic activity is not concentrated on discrete isolated points, but is instead a dynamically fluctuating, smooth spatial distribution. Spatial inequality is a Cass/Koopmans saddlepath in the space of spatial distributions, and the global distribution of economic activity converges toward egalitarian growth. Equality is stable, but spatial inequality is needed to attain it.

REFERENCES

Eeckhout, Jan and Jovanovic, Boyan. "Knowledge Spillovers and Inequality." Working paper, University of Pennsylvania, February 2001.

Feyrer, James. "Convergence by Parts." Working paper, Dartmouth College, December 2001.

Grenander, Ulf and Szegö, Gabor. *Toeplitz forms and their applications,* 2nd ed. New York: Chelsea, 1958.

Krugman, Paul and Venables, Anthony J. "The Seamless World: A Spatial Model of International Specialization." Working paper, London School of Economics, London, U.K., April 1997.

Lucas, Robert E. "On the Mechanics of Economic Development." *Journal of Monetary Economics,* July 1988, *22*(1), pp. 3–42.

Matsuyama, Kiminori. "Explaining Diversity: Symmetry-Breaking in Complementarity

Games." *American Economic Review,* May 2002 (*Papers and Proceedings*), *92*(2), pp. 241–46.

Mookherjee, Dilip and Ray, Debraj. "Is Equality Stable?" *American Economic Review,* May 2002 (*Papers and Proceedings*), *92*(2), pp. 253–59.

Quah, Danny. "Empirics for Growth and Distribution: Polarization, Stratification, and Convergence Clubs." *Journal of Economic Growth,* March 1997, *2*(1), pp. 27–59.

______. "Internet Cluster Emergence." *European Economic Review,* May 2000, *44*(4–6), pp. 1032–44.

______. "Demand-Driven Knowledge Clusters in a Weightless Economy." Working paper, London School of Economics, London, U.K., April 2001.

Sargent, Thomas J. *Macroeconomic theory,* 2nd Ed. New York: Academic Press, 1987.

Turing, Alan M. "The Chemical Basis of Morphogenesis." *Philosophical Transactions of the Royal Society of London Series B,* August 1952, *237*, pp. 37–72.