B: *Planning and Design* 2012, volume 39, pages 262 – 286

doi:10.1068/b37078

Slime mold cities

David Barker
Tippie College of Business, University of Iowa, Iowa City, IA 52242-1994, USA;
e-mail: david-barker@uiowa.edu
Received 2 June 2010; in revised form 28 March 2011

Abstract. Inspired by the behavior of slime mold cells, Paul Krugman developed a simple one-dimensional model in which moving firms self-organize into cities. In this paper I show that extending the model into two dimensions significantly improves its realism. Cities in the two-dimensional model are similar in several respects to real cities: they grow and decline, they cluster near rivers and coasts, and, given certain parameters, their distribution follows Zipf's law. A calibration exercise, however, suggests that observed levels of agglomeration must be due to factors beyond those included in the model.

Keywords: urban, slime mold, self-organization, Zipf, Krugman

1 Introduction

The humble slime mold has inspired revolutionary changes in theories of the emergence of order from chaos. Biologists have learned that individual slime mold cells following simple rules, without central direction, form structures and act as a community. This discovery has influenced research in many fields, including urban economics. An example can be found in the work of Paul Krugman (1998), who adapted models of slime mold behavior to the growth of cities and found that simulated firms following simple economic rules formed cities.

Well-fed slime mold cells behave very simply. They move randomly, reproduce by dividing, and operate independently of other slime mold cells. Hungry slime mold cells are very different. They clump together and form structures designed to increase their chances of surviving and reproducing. In particular, the newly formed group moves in search of food and shapes itself into something resembling a plant, with a stalk holding up a body that eventually releases spores, some of which drift to more favorable environments.

Until the late 1970s it was believed that this change was directed by a hierarchy of slime mold cells. 'Pacemaker' cells were thought to coordinate the process of agglomeration and transformation. It is now generally believed that there is no such hierarchy. Individual cells react to a lack of food by emitting a chemical to which other cells are attracted. Cells detect the direction in which the chemical is most abundant, and tend to move in that direction (Resnick, 1997). Agglomeration of cells is a form of self-organization, accomplished without any central direction. Similar discoveries have been made about the behavior of flocks of birds, which also organize themselves without leaders.

Simple mathematical models of slime mold behavior demonstrate that the behavior of these systems is sensitive to several parameters, such as the rate of evaporation of the chemical, the randomness of cell motion, and the ability of cells to detect the chemical. Some parameter values produce large agglomerations quickly; others produce smaller, more numerous agglomerations; and others produce no agglomerations at all. One illustration of the value of the model is that the effects of changes in many of the parameters are counterintuitive (Resnick, 1997). Without the mathematical model, it is difficult to predict the outcome of a system with a given set of parameter values. The most important aspect of these models is the tension between forces leading to aggregation and forces leading to disaggregation. For example, attraction to the chemical leads to aggregation, while random motion causes many cells to break away from clusters.

Cities are good candidates for modeling as self-organizing systems. In most countries people and firms are free to move between cities, or to rural areas, as they choose in response to prices and other information. Without central direction, people and firms organize themselves into complex systems of cities of different sizes. Urban systems, like biological systems, are characterized by tension between aggregating and disaggregating forces.

In this paper a model is developed in which cities of different sizes emerge out of an initially random distribution of firms. The model extends the one-dimensional 'racetrack' geometry of Krugman (1998) into two dimensions, significantly adding to the ability of the model to make predictions that can be compared with real cities.

A brief review of relevant literature is contained in section 2. Krugman's original model based on slime mold behavior is presented in section 3. In section 4 the geometry of the model is extended from a circle to the surface of a dodecahedron. In section 5 the geometry is changed to a flat two-dimensional surface. Section 6 concludes.

2 Literature review

Urban economic models, like biological models of organism clusters, are based on tension between aggregating and disaggregating forces. Models of the internal structure of cities, from von Thünen (1966 [1826]) to the modern urban models of Alonzo (1964), Mills (1967), and Muth (1969), have viewed city structure as the result of the competing forces of land and transportation costs. These models, however, have little to say about systems of cities, or relative sizes of cities.

Systems of cities have been modeled since the work of Christaller (1933), Lösch (1954 [1940]), and Isard (1956). This work is often referred to as central place theory, and it describes an urban hierarchy in which large cities supply a broad range of products to surrounding cities, and smaller cities supply a smaller range of products to their surrounding cities. Under certain assumptions, a structure with hexagonal market areas for cities was shown to be the most efficient organization. Others have viewed systems of cities as resulting from cities with different functions growing to their optimal sizes. Tolley (1974) and Henderson (1974) both saw optimal city size as determined by tensions between economies of scale and diseconomies such as congestion and commuting costs.

Central place theory described final equilibrium urban structures. Influenced by the concept of 'self-organization' in biology and physics (Glansdorff and Prigogine, 1971; Nicolis and Prigogine, 1977), Allen and Sanglier (1978; 1979; 1981) developed a dynamic version of central place theory. From a set of differential equations governing population, transportation, and competition between urban centers, they were able to simulate the emergence, growth, and decline of cities in a spatial system. One of their concerns was to replicate Zipf's law of city-size distribution. The authors found that Zipf's law approximated the hierarchy of city size that resulted from their simulation, and commented:

> "Any model purporting to describe the evolution of a system containing several interacting urban centres must be able to take into account the effects of historical circumstances as well as the operation of 'universal' laws" (Allen, 1996, page 30).

Another goal of this literature was replication of real geographic conditions. Sanglier and Allen (1989) used a dynamic model based on central place theory to compare simulations with the actual geography of European cities.

Another approach to the study of urban systems is to look for causes of the growth of individual cities. Explanations of the fate of individual cities are easy to manufacture, but a convincing general theory of city growth is much more difficult. Careful empirical papers, such as those by Beeson et al (2001) and Carlino and Mills (1987), have regressed growth rates of US counties on their initial characteristics. They find that climate and access to inexpensive transportation are important factors in city growth. Beeson et al (2001) argued that these initial characteristics led to city formation and the development of educational infrastructure that, in turn, led to additional growth.

The 'new economic geography' began with Krugman (1991) and was declared to have 'come of age' by Neary (2001). New economic geography is characterized by models that incorporate rational decision making and general equilibrium into the study of aggregation and disaggregation of economic activity. In these models cities develop naturally as a result of the interactions of individual agents. Krugman's early models continue to be refined, as in the paper by Mossay (2006), where the conditions needed for equilibrium in the model are derived.

Davis and Weinstein (2002) divided new economic geography theories of urban development and growth into three categories: increasing returns, random growth, and location fundamentals. They examined data on Japanese city populations dating from prehistoric times to the present for evidence for and against these theories. Zipf's law holds throughout Japanese history and this was taken as evidence against increasing returns and for the remaining two theories. The rise in variation of city sizes after the industrial revolution was taken as evidence for increasing returns and against the remaining two theories. Mean reversion of the populations of Hiroshima and Nagasaki after the atomic bombings of 1945 was taken as evidence in favor of location fundamentals.[(1)] In short, Davis and Weinstein (2002) found some evidence in favor of each of the three theories.

As was pointed out by Fujita et al (1999), most of the theoretical work that has been done on urban systems does not present these systems as the product of self-organization, and there is no explicit mechanism in the models to create cities of certain sizes. Krugman (1998) and Fujita et al (1999) do this and they incorporate elements of two of the three theories described in Davis and Weinstein (2002). Differentiated products by the firms in the model are a form of economies of scale, and movement of firms from city to city has a random component. The racetrack model of Krugman (1998) continues to be refined mathematically, as in the work of Akamatsu et al (2009) and Picard and Tabuchi (2010).

The one-dimensional racetrack universe, however, does not allow any locations to have advantages over other locations, which means that the third of the three theories in Davis and Weinstein (2002)—location fundamentals—cannot be incorporated into the model. This limitation has certain benefits, since other factors can be examined in isolation from location fundamentals, but it also limits the realism of the models.

Several writers have discussed the need to extend new economic geography models into two dimensions. Krugman (1998) wrote:

[(1)] One problem with this analysis is that a great deal of city infrastructure, including roads and rail beds leading to the cities, remained after the bombings. Because of this infrastructure, rebuilding the cities at their previous locations made more economic sense than abandoning them, just as individual buildings that burn are usually rebuilt at their previous locations.

> "To make this a truly convincing model, of course, considerable work is needed. Perhaps the most obvious task is to extend the analysis to two dimensions, something that is not at all easy, even if one is interested only in simulation results" (page 31).

Neary (2001) noted the lack of two-dimensional models in the new economic geography, adding that the difficulties involved would likely result in models that are "long on trigonometry and short on elegance." Stelder (2005) wrote:

> "In all the NEG literature of the past ten years, Neary (2001) is just about the only one who has actually discussed the introduction of a two-dimensional plane in a Krugman model" (page 658).

Stelder (2005) went on to extend the model of Krugman (1998) to a grid of points superimposed on a map of Europe. He found that cities that developed in model simulations showed a rough similarity to actual city locations, and that his model did a better job of predicting the sizes of European cities than assuming a simple random growth process. Stelder (2005) also found some evidence that adjusting the distance matrix of the grid points to take account of mountains as transportation barriers and of low cost coastal sea transport improved the performance of the model.

Two-dimensional new economic geography models remain rare, however. Bosker et al (2010) recently wrote:

> "For reasons of analytical tractability, new economic geography (NEG) models treat geography in a very simple way, focusing on stylized 'unidimensional geography structures [e.g. an equidistant or line economy]. All the well-known NEG results are based on these simple geography structures" (page 793).

Bosker et al (2010) also conclude that the difficulties of developing an analytically solvable two-dimensional model are probably impossible to overcome, so they resort to simulation. Their model consists of 194 regions, the same as the number of regions with available statistical data in the European Union. The distance matrix follows actual distances between these regions. Bosker et al (2010) replicate some results from one-dimensional models, but find that in two dimensions very different agglomeration patterns can result in the same summary measures of agglomeration.

In this paper the model of Krugman (1998) is extended into two dimensions. As in the papers described above, simulation is employed because of the difficulty of obtaining analytical results. The two-dimensional model can be configured without locational differences, so that factors other than location in city evolution can be examined, and it can also be configured to have central places, coasts and rivers. Other additions also help to enhance the realism of the model. In the model, systems of cities resembling the real world in many ways develop as products of self organization. With these additions, all three of the theories described by Davis and Weinstein (2002) can be incorporated into the same model.

3 Krugman's model

Krugman's slime mold model was originally published as a chapter in the book, *Topics in Public Economics* (Krugman, 1998). Because of its limited exposure, the chapter has not received the attention of related work, such as Krugman (1991) and Fujita et al (1999). Because many readers will be unfamiliar with the model, I present it here in detail. My presentation of the model follows Krugman (1998) closely.

The model begins with a simple economy composed of farmers and firms. Farmers cannot move and are evenly distributed across space, but firms move in response to profit opportunities. There are L possible locations, all evenly spaced around a circle. In this one-dimensional 'clock-face' world, each location is symmetric; no location is central or peripheral. For example, with twelve locations, just as on a clock, the distance between locations 1 and 12 is 1; the distance between locations 5 and 8 is 3, etc.

Farmers produce a single product with constant returns to scale which has zero transportation costs. Each firm, by contrast, produces a different product, and each firm produces one unit of this product per period. Transportation costs for the products manufactured by the firms are like Samuelson's iceberg; $\exp(-\tau D)$ units remain from an original shipment of one unit after being transported over a distance D. Farmers and firms share a utility function of the agricultural good, A, and a composite of the various manufactured goods, M, shown in equation (1).

$$U = C_{\mathrm{M}}^{\mu} C_{\mathrm{A}}^{1-\mu} \ . \tag{1}$$

Choosing units such that the number of farmers is equal to $1-\mu$ is a trick which simplifies the algebra of the model. The total population of the economy is then divided into a fraction, $1-\mu$, of farmers and a fraction, μ, of firms. The first-order conditions for utility maximization reveal that a fraction, μ, of total income will be spent on manufactured goods, and a fraction, $1-\mu$, will be spent on the agricultural good.

Firms use the products of other firms as inputs according to a constant elasticity of substitution production function shown in equation (2).

$$C_{\mathrm{M}} = \left(\sum_i C_i^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}} \ . \tag{2}$$

With this production function, using a wide variety of inputs allows a firm to produce more. In other words, each firm has an incentive to purchase the products of all of the other firms. Since all products enter the production function in the same way, the price (free on board: fob) of each good will be identical. Transportation costs will be the only source of difference between the prices of products by type and location. Firms must balance their need for a variety of products with the high transportation costs of products from distant locations.

The nominal, gross income of farmers will not depend on their own location or the location of firms, since transportation costs for agricultural goods are zero. Farmers will be affected by location, however, since farmers that are more distant from aggregations of firms will pay more for their consumption of manufactured goods. Since they are spread evenly by location, the number of farmers at any location will be $1-\mu$ divided by the number of locations, J. If the income of a farmer is used as a numeraire, then this will also equal the total income of farmers at any location. Firms are not necessarily spread evenly, so there will be n_j firms at each location j. The total income, including transportation charges, earned by farmers and firms at any location j will therefore be equal to:

$$Y_j = \frac{1-\mu}{J} + \mu n_j p_j \ . \tag{3}$$

Since a fraction, μ, of total income will be spent on manufactured goods, then μY_j will be equal to the sum of expenditures on the manufactured products at location j. Farmers facing high transportation costs for manufactured goods will substitute to agricultural goods, but the expenditure share (not consumption share) of manufactured goods will always be equal to μ. If $\hat{p}_{ij}$ is the delivered price of good i at location j, and consumption of good i at location j is equal to c_{ij}, then

$$\mu Y_j = \sum_i c_{ij} \hat{p}_{ij} \ . \tag{4}$$

In other words, expenditure on manufactured goods at a location is equal to the sum of expenditures on each good at that location. A first-order condition from maximizing equation (1) subject to equation (2) is, for any product variety k:

$$\frac{c_{ij}}{c_{kj}} = \left(\frac{\hat{p}_{ij}}{\hat{p}_{kj}}\right)^{-\sigma} . \tag{5}$$

This is obtained simply by substituting equation (2) into equation (1), maximizing subject to a budget constraint for products i and k, and dividing the two resulting equations. This first-order condition implies:

$$c_{ij} = \frac{c_{kj}\hat{p}_{kj}^{\sigma}}{\hat{p}_{ij}^{\sigma}} . \tag{6}$$

Substituting this into equation (4) yields:

$$\mu Y_j = \sum_i \frac{c_{kj}\hat{p}_{kj}^{\sigma}}{\hat{p}_{ij}^{\sigma}} \hat{p}_{ij} , \tag{7}$$

which simplifies to:

$$\mu Y_j = c_{kj}\hat{p}_{kj}^{\sigma} \sum_i \hat{p}_{ij}^{1-\sigma} . \tag{8}$$

This can be rewritten as:

$$c_{kj}\hat{p}_{kj}^{\sigma} = \frac{\mu Y_j}{\sum_i \hat{p}_{ij}^{1-\sigma}} , \tag{9}$$

or

$$c_{kj}\hat{p}_{kj} = \mu Y_j \frac{\hat{p}_{kj}^{1-\sigma}}{\sum_i \hat{p}_{ij}^{1-\sigma}} . \tag{10}$$

This is the total expenditure on good k at location j.

The next step is very important, for it not only simplifies the algebra, but later it will be used to define real income at different locations. Firms will use this information to decide where to move. A true price index reflects the change in income needed to maintain some utility level after prices of goods have changed (Klein and Rubin, 1948). For the constant elasticity of substitution production function of equation (2), the derivation of the true price index is explained in Lloyd (1975). The true price index is as follows:

$$T_j = \left(\sum_i \hat{p}_{ij}^{1-\sigma}\right)^{\frac{1}{1-\sigma}} . \tag{11}$$

In other words, if a firm moves from one location to another, input prices will change because transportation costs for various goods will be different. If firms remained evenly distributed, the price index would not change, since the goods of different firms are symmetric. The distribution of firms is not even, however, and, as the model evolves, firms sometimes cluster into cities. This means that a move by a firm can have a significant effect on its overall costs. The ideal price index is a way of measuring the overall change in cost from such a move. The expression for the ideal price index can be substituted into equation (10) to produce:

$$c_{kj}\hat{p}_{kj} = \mu Y_j \left(\frac{\hat{p}_{kj}}{T_j}\right)^{1-\sigma} . \tag{12}$$

Since fob prices differ from delivered prices by transportation costs, an explicit formulation of transportation is needed to express T_j in terms of fob prices. D_{ij} is the distance between firm i and firm j, and the transportation cost per unit of distance is given by the parameter τ:

$$\hat{p}_{ij} = p_{ij} \exp(\tau D_{ij}) \ . \tag{13}$$

Substituting equation (13) into equation (11) produces another expression for T_j:

$$T_j = \left\{ \sum_i \left[p_{ij} \exp(\tau D_{ij}) \right]^{1-\sigma} \right\}^{\frac{1}{1-\sigma}} \ . \tag{14}$$

But since each firm is symmetric to each other firm except for location, this can be expressed in terms of a single good, k:

$$T_j = \left\{ \sum_k n_j \left[p_{kj} \exp(\tau D_{jk}) \right]^{1-\sigma} \right\}^{\frac{1}{1-\sigma}} \ , \tag{15}$$

where n_j is the number of firms at location j.

Finally, the system can be solved for the fob prices. A firm at location j produces one unit of its good. Total expenditure by all consumers in the economy on this good must be equal to the fob price of the good. Equation (12) tells us the expenditure on this good at some location j. Summing over all locations also gives us total expenditure on the good, so we have the following equation:

$$p_j = \mu \sum_j Y_j \left[\frac{p_j \exp(\tau D_{jk})}{T_j} \right]^{1-\sigma} \ . \tag{16}$$

Solving for p_j yields:

$$p_j = \left\{ \mu \sum_j Y_j T_j^{\sigma-1} \exp\left[-\tau(\sigma - 1) D_{jk} \right] \right\}^{\frac{1}{\sigma}} \ . \tag{17}$$

For a given distribution of firms, the preceding equations can be solved for equilibrium prices in this economy. In a numerical simulation, they can be solved iteratively.

The next task is to describe the dynamics of firm movement. Firms move in response to real income they can earn at their present location and at neighboring locations. Real income at a location is obtained by deflating nominal income by the true price index discussed earlier. Nominal income is simply the price a firm obtains for its good, p_j. The exponent of $-\mu$ on the ideal price index is from the utility function in equation (1). Intuitively, if μ were equal to 1, then agricultural goods would not be consumed at all, so the appropriate deflator would simply be the ideal price index for manufactured goods. If μ were equal to 0, then only agricultural goods would be consumed, so the deflator for manufactured goods would not be relevant at all. For μ between 0 and 1, μ is the share of manufactured goods in total consumption, and so the true price index of manufactured goods is appropriately weighted:

$$\omega_j = p_j T_j^{-\mu} \ , \tag{18}$$

where ω_{j+1} is equal to real income at the location clockwise to the right of the firm's current location, and ω_{j-1} is real income in the location clockwise to the left. If real income is higher in one direction than current income, the firm will be more likely to move in that direction. Noisy information and other factors cause randomness in the firm's responses, but it will tend to move in the direction of higher real income.

The probability that the firm will move to the right is given by:

$$\mathrm{pr}(j, j+1) \;=\; \frac{\exp(\gamma\omega_{j+1})}{\exp(\gamma\omega_{j+1}) + S\exp(\gamma\omega_j) + \exp(\gamma\omega_{j-1})} \;. \tag{19}$$

And the probability that the firm will move to the left is given by:

$$\mathrm{pr}(j, j-1) \;=\; \frac{\exp(\gamma\omega_{j-1})}{\exp(\gamma\omega_{j+1}) + S\exp(\gamma\omega_j) + \exp(\gamma\omega_{j-1})} \;. \tag{20}$$

Note that these two equations correct mistakes in equations (1-3) and (1-4) in Krugman (1998). A high value of γ reduces the randomness of firm movements; firms will be likely to move toward areas that will allow them to earn higher income. A low value of γ means that firms are much less influenced by the income potential of neighboring locations, and move more randomly. A high value of S means that the firm is less likely to move at all. For example, suppose that ω_{t+1}, the potential income if a firm moves to the right, is 0.9 then ω_{t-1}, the potential income if the firm moves to the left, is 1.1. Income at the firm's current location is equal to 1. If S is equal to 1 and γ is equal to 1, then there is a 30% chance of moving right, towards lower income, a 37% chance of moving left, toward higher income, and a 33% chance of staying put. If γ is equal to 10, however, there will be only a 9% chance of moving to the right, a 67% chance of moving to the left, and a 24% chance of staying put. If S is also increased to 10, then the chance of staying put increases to 76%. Thus γ is a key parameter of the model. Changes in γ can cause dramatic changes in the simulation results.

This completes the description of the economy. Equilibrium prices are determined for an initial location of firms, and firms then move in response to observed neighboring real incomes. New equilibrium prices are then determined, and firms move again in response to these new prices. The firms continue to move about in this manner, perhaps settling down to a stable distribution, or perhaps moving around randomly forever. The kind of outcome obtained will be determined by the parameters μ, γ, S, τ, and σ.

Simulation of this model can be done by assigning random locations to a number of firms, solving the pricing equations, generating random numbers and making

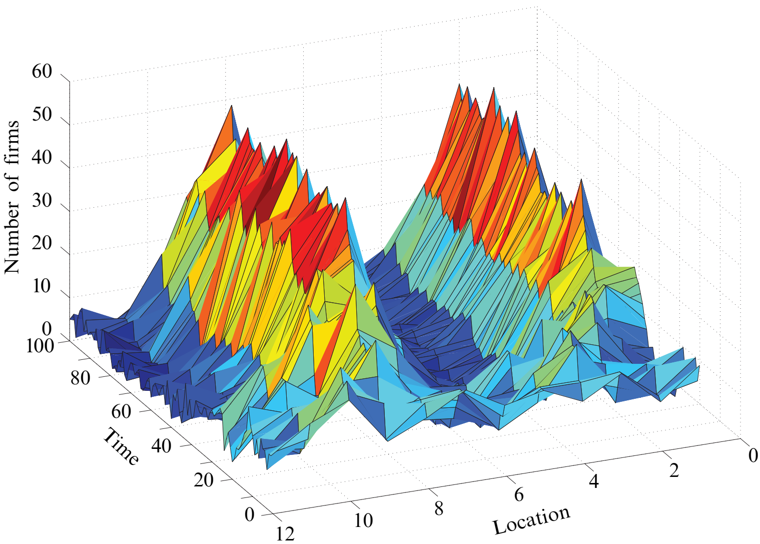

Figure 1. [In color online, see http://dx.doi.org/10.1068/b37078] Evolution of cities in a one-dimensional model.

decisions about firm movement on the basis of these random numbers and calculated probabilities of movement, and then repeating the process. These simulations were programmed by the author in Matlab, and the code is available upon request.

Figure 1 shows the evolution of the system over time with parameter values $\mu = 0.57$, $\sigma = 4$, $\tau = 0.3$, $\gamma = 80$, and $s = 2$. In figure 2, γ is equal to 40 and τ is equal to 0.2. These figures closely match figures 1 and 2 of Krugman (1998). In the first the system self-organizes into cities, but in the second the distribution of firms remains random. Other parameter values result in different patterns of cities.

Figure 3 presents a different perspective on the evolution of the system which is perhaps more appropriate for the geometry of the model. A cross-section of the cylinder represents the locations of firms at a particular time. The system starts at the front of the cylinder, where firms are randomly scattered across the twelve locations. Over time two cities emerge on opposite sides of the cylinder. In this figure μ is

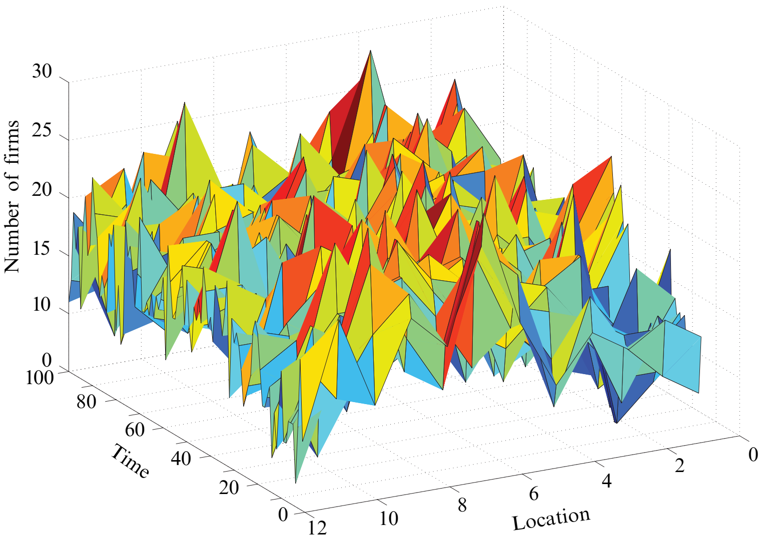

Figure 2. [In color online.] Failure of cities to evolve in a one-dimensional model.

Figure 3. [In color online.] Evolution of two cities along a circle.

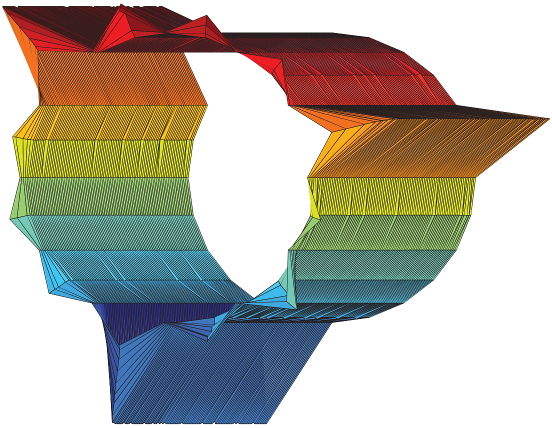

Figure 4. [In color online.] Evolution of three cities along a circle.

equal to 0.65, τ is equal to 0.3 and γ is equal to 40. In figure 4, γ is increased to 120, and three cities quickly emerge. Small towns occasionally form, but their firms are quickly absorbed by the dominant cities. It is interesting to note that it is difficult to predict where the dominant cities will emerge from the initial distribution of firms. A location that happens to start with the most firms does not always end up as a city at all.

4 Extending the model to two dimensions: the surface of a dodecahedron

In this section the model of Krugman (1998) is extended into two dimensions, as was suggested by Krugman (1998), Neary (2001), Stelder (2005), and Bosker et al (2010).

An advantage of Krugman's geometry is that there is no central place, so the evolution of cities can be seen as a result of self-organization instead of the natural advantages of certain sites. An obvious extension of Krugman's circular space is a sphere, the surface of which also has no central point. The vertices of the platonic solids define equidistant points on the surface of a sphere, which can be the locations of firms. One of the platonic solids, the dodecahedron, has twenty vertices, each connected by an edge to three other points. Three changes must be made to Krugman's model to change its geometry from points on a circle to points on a dodecahedron.[2] First, a matrix of distances must be calculated. This is more complicated for a dodecahedron than for points on a circle. In figure 5, for example, the distance between points 10 and 9 is 2, while the distance between 10 and 12 is 5. No two points are separated by a distance of more than 5. Second, a matrix of movement between points must be constructed. On a circle, movement between points must be constructed. On a circle, movement to the right simply means moving from point j to point $j+1$ until point 12 is reached, then moving to the right takes a firm to point 1. At a point on the dodecahedron, there are three possible moves, which could be called left, right, and up. All twenty points can be numbered, and aligning each point in turn to the front, the points that are to the left, right, and above can be noted. For example, starting from point 10, moving up leads to point 5, moving left leads to point 15, and moving right leads to point 14. Since all points are symmetric, the dodecahedron can be rotated to place any point where point 10 is located in figure 5,

[2] Note that the pentagons of the surface of the dodecahedron and the tessellated hexagons used later in the paper play a different role in this model than the hexagons in Christaller's central place theory. Christaller's cities were located in the center of the hexagon, and the boundaries of the hexagon were the boundaries of the city's market area. In the model of this paper, cities locate at the vertices of the polygons and move between vertices along the edges.

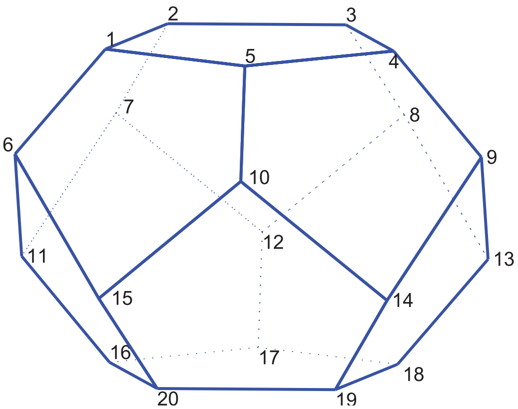

Figure 5. [In color online.] A dodecahedron.

and the points left (l), right (r), and up (u) from any starting point can be obtained.[3] Third, equations (19) and (20) must be modified as follows:

$$\mathrm{pr(r)} = \frac{\exp(\gamma\omega_\mathrm{r})}{\exp(\gamma\omega_\mathrm{r}) + \exp(\gamma\omega_\mathrm{l}) + S\exp(\gamma\omega_j) + \exp(\gamma\omega_\mathrm{u})}, \qquad (21)$$

$$\mathrm{pr(l)} = \frac{\exp(\gamma\omega_\mathrm{l})}{\exp(\gamma\omega_\mathrm{r}) + \exp(\gamma\omega_\mathrm{l}) + S\exp(\gamma\omega_j) + \exp(\gamma\omega_\mathrm{u})}, \qquad (22)$$

$$\mathrm{pr(u)} = \frac{\exp(\gamma\omega_\mathrm{u})}{\exp(\gamma\omega_\mathrm{r}) + \exp(\gamma\omega_\mathrm{l}) + S\exp(\gamma\omega_j) + \exp(\gamma\omega_\mathrm{u})}. \qquad (23)$$

Figure 6 shows an evolution of this model with $\mu = 0.6$, $\sigma = 4$, $\tau = 0.2$, $\gamma = 80$, and $s = 2$. Unlike the earlier simulations in one-dimensional space, it is common for a single city to emerge—not double cities on either side of the globe. At first there are three cities, about equal in size. The city at location 6 dies out, retaining only four firms. The city at location 12 survives with sixty-two firms and small suburbs next to it, but does not grow nearly as large as the city at location 17.

The final outcome can be better visualized in figure 7, which shows the cities on a globe. This figure makes it clear that the cities are clustered in one hemisphere, and much of the globe has no cities at all.

An interesting feature of new urban economics models is their tendency to exhibit multiple equilibria (Krugman, 1998; Neary, 2001). The empirical relevance of multiple equilibria has been debated, however (Bosker et al, 2007; Davis and Weinstein, 2008). In the model developed in this paper, multiple equilibria can be identified by adjusting parameter values to a point close to where the model 'switches' from one equilibrium to another. For example, lowering S increases the randomness of firm movement. If γ remains high enough, firms move to where income is highest, and a single large city develops. If, from there, S is raised to a certain point, twin cities develop instead of a single large city. With S equal to -1.6 and γ equal to 50, these two outcomes—twin cities and a single city—are approximately equally likely. Random movements of firms are enough to tip the system into either a single large city or two cities of equal size adjacent to each other. Parameter values that allow the creation of large cities, however, tend to produce stable outcomes. For these parameter values, the same initial conditions, or even a different initial distribution of cities, is always followed by a similar pattern of city sizes, although where on the dodecahedron the cities form is random.

[3] Some rule for orienting the dodecahedron is needed for consistency. For example, the number of the 'up' point could always be the highest of the three adjacent points.

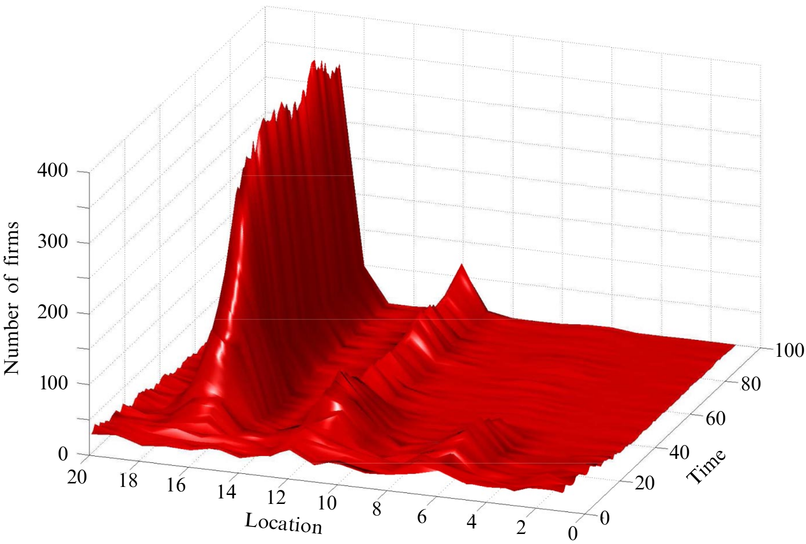

Figure 6. [In color online.] Evolution of cities on the surface of a dodecahedron.

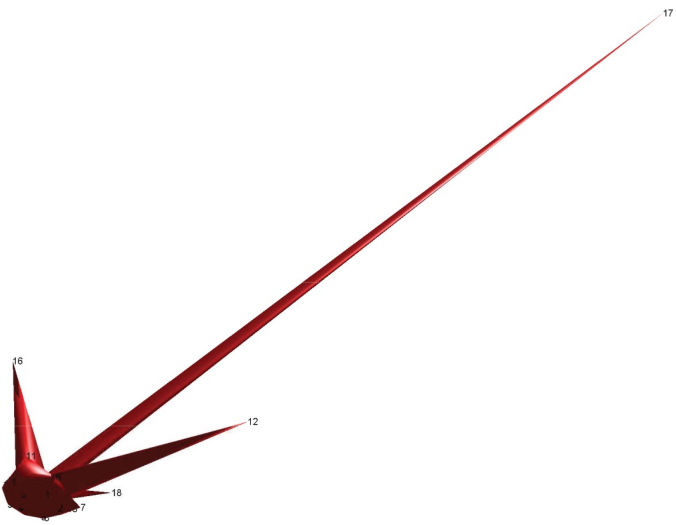

Figure 7. [In color online.] Cities on the surface of a dodecahedron.

A test of the model's realism is its ability to adhere to Zipf's law. Zipf's law says that the size distribution of cities follows a power law. More specifically, many studies have found that regressing the log of the rank of a list of cities on the log of the population of these cities produces a slope coefficient close to one with a very high R^2.[(4)] Gabaix (1999a, page 739) wrote that Zipf's law "constitutes a minimum

[(4)] Nitsch (2005) produced a meta-analysis of previous studies of Zipf's law, which finds strong evidence of the relevance of the law, although coefficient estimates often differ from 1. Eeckhout (2004) disputes the law's validity. Using data from all US Census Department "Places" he finds that the law breaks down for small towns. These data, however, are very incomplete. Many small settlements with fewer than 3000 people are not included in the census list of places, which could explain their result. In this paper, a database of world cities and towns is used, but towns smaller than 10 000 people are not used due to the current impossibility of enumerating every small hamlet in the world.

criterion of admissibility for any model of local growth, or any model of cities", a comment similar to that of Allen and Sanglier (1978) quoted earlier. Other papers that attempt to match simulations of urban systems with Zipf's law include those by Duranton (2007) and Rossi-Hansberg and Wright (2007).

It turns out that there are many combinations of parameter values that produce a close approximation to the distribution of cities implied by Zipf's law. Figure 8 shows an example using all of the parameter values in Krugman (1998), except that γ has been increased to 55.

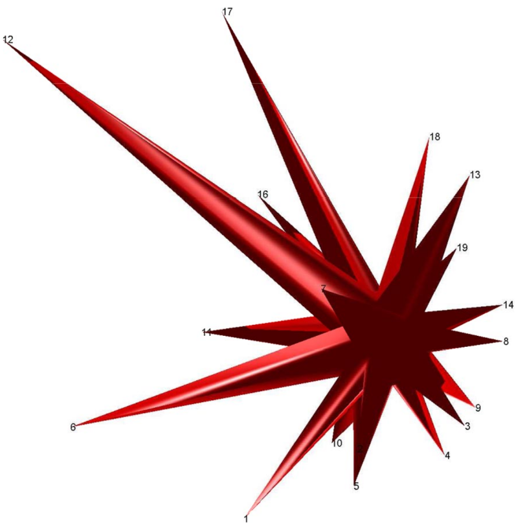

Figure 8. [In color online.] Cities on the surface of a dodecahedron, parameters chosen so the city distribution approximates Zipf's law.

A stricter test of the model's realism is to see whether parameters calibrated to match the economy of the United States produce a matching distribution of city size. The model contains twenty population centers, and the remaining population is assumed to be rural, so it seems reasonable to assume that these cities represent the largest twenty cities in the US, and the rest of the population is rural. Approximately 38% of the US population lives in these cities, so μ will be set to 0.38. The transportation sector of the US economy represents approximately 3% of GDP,[(5)] so τ will be set to 0.03. This means that to transport a product from one city to a neighboring city costs 3% of the product's value. Note that this is a significant overestimate of τ since many products are transported further than the distance to a neighboring city. Parameter σ represents the elasticity of substitution between products produced in different cities. Many papers have estimated the elasticity of substitution between products produced in a home country and imported products. Estimates vary widely, but some researchers have chosen 1.25 as a reasonable value (Alfaro et al, 2010), so σ will be set to 1.25.[(6)]

From equations (21) – (23), if income in a city is equal to that in neighboring cities, the probability of a firm moving in a given period is shown in equation (24).

$$\mathrm{pr(move)} = \frac{3}{3+S} . \tag{24}$$

(5) See http://www.bea.gov/industry/gdpbyind_data.htm

(6) For a discussion of the practical usefulness of aggregate production functions, see Miller (2008).

The US Census Bureau reports[7] that 39.5% of Americans changed residences between the years 2000 and 2005. It also reports that only 13% of moves were motivated by higher income in another region, so 34% of Americans can be expected to move for noneconomic reasons over five years. Movement for noneconomic reasons is equally likely in all three directions, so the probability of moving to the right for noneconomic reasons is 11.3%. Setting the probability of a move equal to 0.113 in equation (24) yields an estimate of S of 23.55.

Kennan and Walker (2011) show a roughly 5 percentage point increase in the chance of moving over five years when income is 10% higher in a different region. To find a value of γ that is consistent with the results of Kennan and Walker (2011) we can set the average slope of pr(r) in equation (21) equal to 0.05 as ω_r goes from 1.0 to 1.1, with ω_l, ω_u, and ω_j equal to 1.0, and S equal to 23.55. This calculation results in a value of 33.92 for γ.

Alternatively, we can set pr(r) in equation (21) equal to 0.163, the sum of the probabilities of moving to the right for economic and noneconomic reasons, S, equal to 23.55, ω_r equal to 1.1, and ω_l, ω_u, and ω_j equal to 1.0, and solve for γ. This calculation results in a value of 16 for γ.

Either of these parameter values, however, produces much less agglomeration than what is seen in the US. Figure 9 shows the populations of the largest twenty US cities normalized so that the population of the largest city is equal to 1. These city sizes are compared with the average distribution of city size from 100 runs of the model. Apparently, some factor not captured by the model is causing additional agglomeration. Glaeser (2008) discusses the productivity gains that result from proximity in cities. The model parameter best able to capture this effect may be τ. If, for example, personal communication is a factor of production, and it works better when individuals are in close proximity, it is as if this factor is expensive

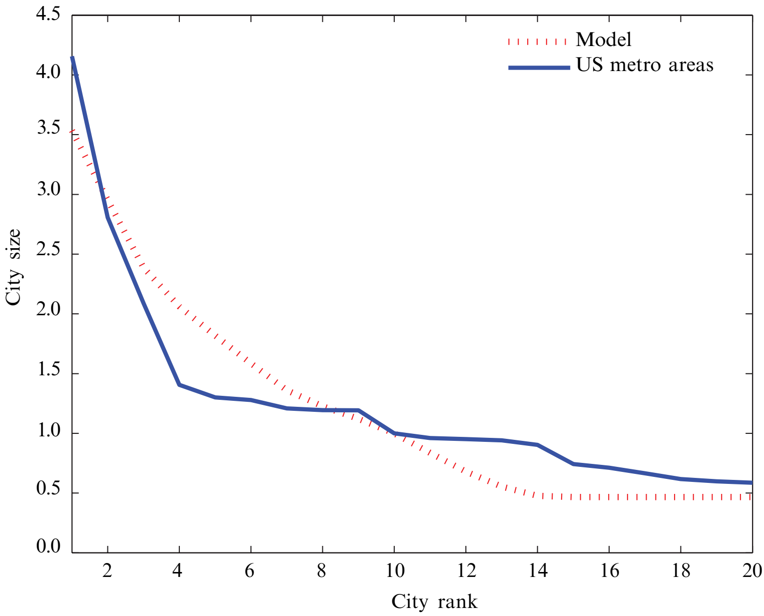

Figure 9. [In color online.] City size and rank, low transportation costs.

[7] See http://www.census.gov/population/www/socdemo/migrate/cps2005-5yr.html

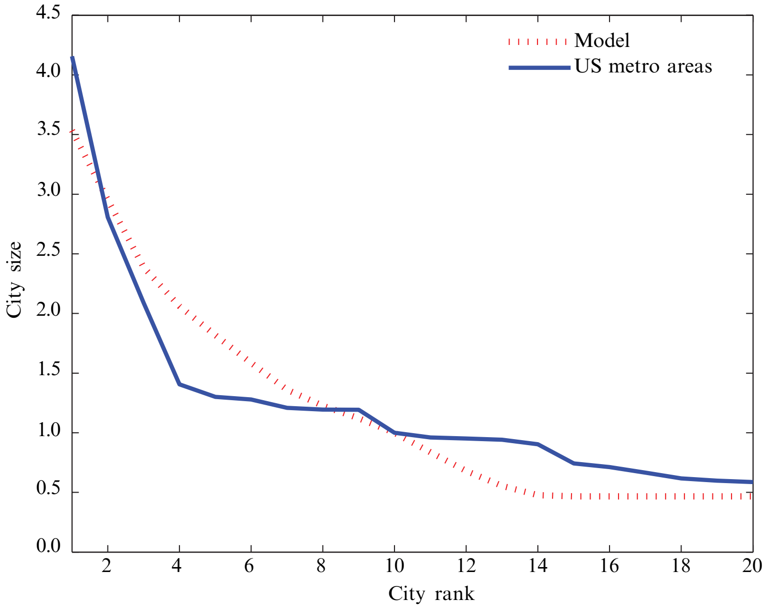

Figure 10. [In color online.] City size and rank, high transportation costs.

to transport. In this case τ, the average transportation cost of factors of production, would be higher. In figure 10, τ is increased to 1.2, meaning that moving from one city to a neighboring city would reduce the value of the factor being transported by 70%. This value of τ produces a city distribution that is fairly close to the actual distribution of US cities.

The model is able to meet the minimum criteria specified by Gabaix (1999a) and Allen and Sanglier (1978) by producing distributions of cities that are consistent with Zipf's law. In addition, it provides insight into the economics of cities by showing that plausible parameter values do not match the distribution of US cities unless they are adjusted to take account of very strong agglomeration economies.

Another aspect of the simulations that fits reality is the dynamics of city growth. Cities that are the largest in early iterations do not always maintain their dominant position. New cities spring up, and often surpass older, established cities. For example, city 12 starts out in 9th place, falls in rank, then rises to first, then falls to 18th, then rises and competes with city 6 for the top ranking, and eventually becomes the largest city again. This time, once it becomes the largest city it retains the position indefinitely. City 6, after losing the top spot to city 12, falls to the rank of 5. Real cities have displayed this kind of behavior throughout history. Herodotus wrote in 450 BC that

> "Those [cities] which in old times were great have for the most part become small, while those that were in my own time great used to be small" (2004, pages 4–5).

In the United States, Saint Louis fell from 4th place among US cities to 53rd between 1900 and 2003. Indianapolis rose from 21st to 12th and Memphis rose from 27th to 17th, but Boston fell from 5th to 23rd. Figure 11 shows rankings of a selection of US cities over time. There is considerable variability over time but, once cities attain a certain size, they often remain dominant over time.

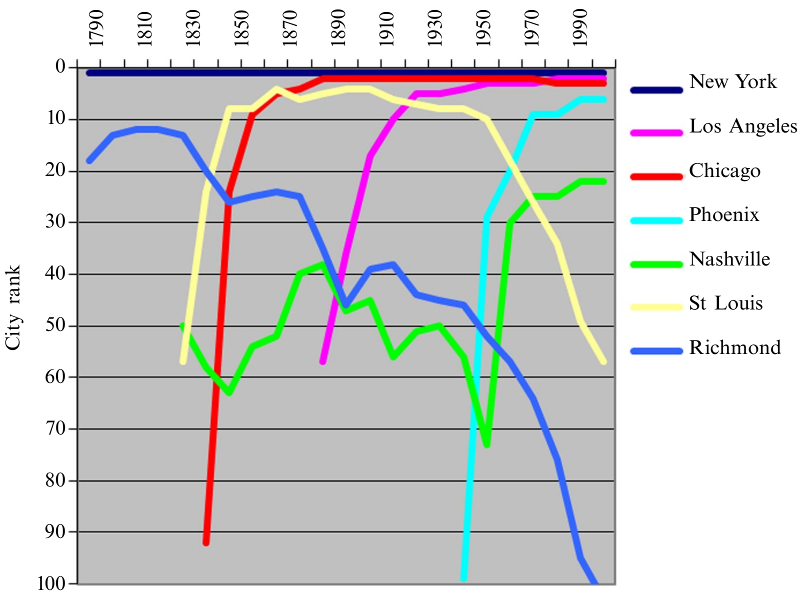

Figure 11. [In color online.] City rank over time.

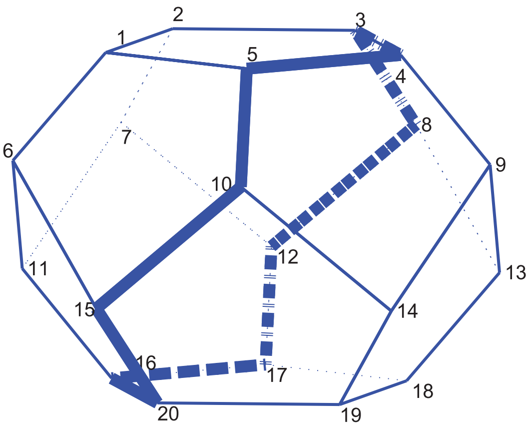

Figure 12. [In color online.] Dodecahedron with a river.

The realism of the model can be enhanced in various ways. One is to introduce a river that lowers transportation costs between some points. Figure 12 shows such a river traveling all the way around a dodecahedron. To travel from, say, point 6 to point 5, the fastest route without the river is in two steps, through point 1. If river transportation is less than half the cost of land transportation, however, it would be cheaper to go first to point 15, then travel two steps on the river to point 5 for a total of 3 steps. Using the same parameter values as in figure 6, except that transportation on the river is assumed to be half the cost of ground transportation, figure 13 shows that all of the cities develop along the river. It is also interesting that no single city dominates—the cities along the river are roughly equal in size around the entire globe.

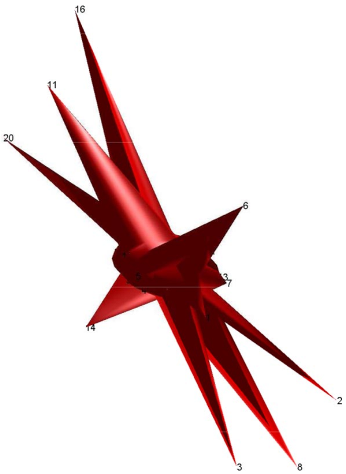

Figure 13. [In color online.] Cities along a river.

5 Flattening the dodecahedron

Although the earth itself is round, economic activity takes place on continents, which have coasts and central areas.[8] A geometry similar to the dodecahedron can be created with a flat tessellation of hexagons. Such a geometry can be seen in figure 14. Firms can locate at the vertices of lines, but cannot travel beyond the coastline. If transportation is possible only by land, central firms have an advantage over coastal firms, since they can transport goods to and from more cities more cheaply. Figure 15 shows the city distribution that evolves under these conditions. The parameter values are the same as for figure 6. A large central city surrounded by suburbs is formed, with no firms located on the coast. If we assume that firms can transport by sea, then the balance will shift toward coastal cities. Figure 16 shows the result of assuming that ocean shipping costs are the same between any coastal locations and are equal to the cost of land transportation over 2 units of distance. All of the firms locate on the coast, and all of the cities are approximately equal in size. Interestingly, as ocean shipping costs rise, even though ocean shipping costs do not depend on distance shipped, firms eventually concentrate into a single coastal city. If shipping costs increase further, the firms suddenly shift to a large central city. A larger grid size, however, increases the opportunities for central cities to develop, since there are more neighboring areas

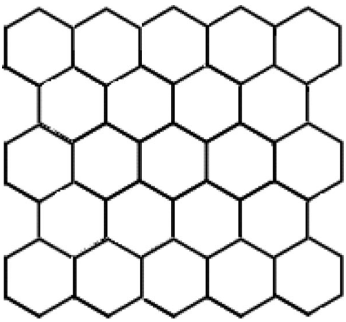

Figure 14. A tessellation of hexagons.

[8] The importance of coastal location for cities is discussed by Rappaport and Sachs (2003) and Serow and O'Cain (1992).

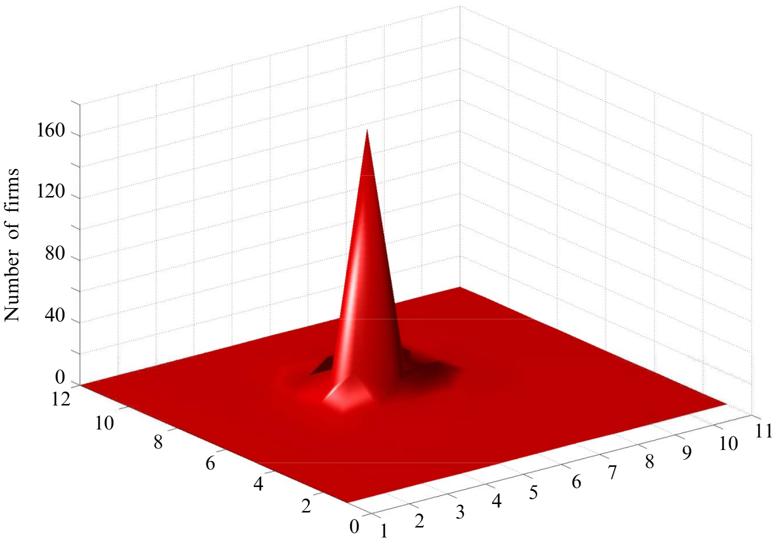

Figure 15. [In color online.] A large interior city with suburbs.

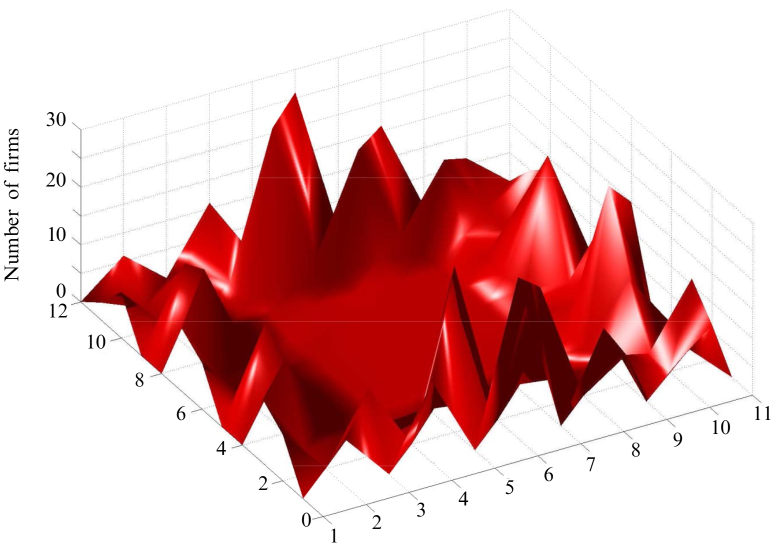

Figure 16. [In color online.] Coastal cities.

to trade with that cannot easily access coastal shipping. Figure 17 shows a grid of 496 points after 1000 iterations. Most firms locate on the coast, but interior cities survive.

The mix of ocean and land shipping can also be calculated. Equation (10) can be rewritten to obtain an expression for the consumption of good k at location j:

$$c_{kj} = \mu Y_j \frac{\hat{p}_{kj}^{-\sigma}}{\sum_i \hat{p}_{ij}^{1-\sigma}} . \tag{25}$$

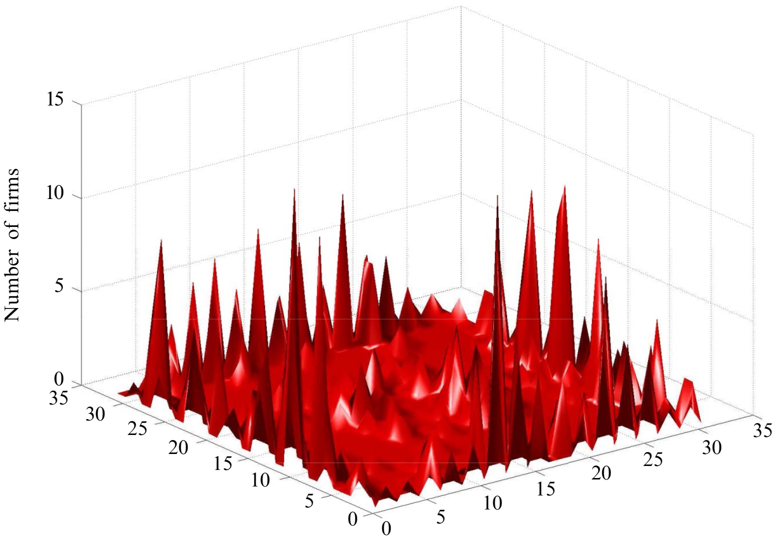

Figure 17. [In color online.] A mixture of coastal and interior cities.

Using the fact that the delivered price is equal to $p_j \exp(\tau D_{jk})$, and substituting in equation (11), consumption can also be expressed in as terms for which solutions have already been found:

$$c_{kj} = \mu Y_j \frac{[p_j \exp(\tau D_{jk})]^{-\sigma}}{T_j^{1-\sigma}} . \tag{26}$$

Consumption of any good at any location can be determined from equation (26).

By calculating the lowest cost transportation method between the location of firms i and j, the amount that goes by sea and land can be calculated. In figure 17, for example, 62% of good/miles are shipped by sea, while in figure 18, a city distribution obtained by increasing the cost of sea transport to 6 units of land transport, the fraction shipped by sea is 35%. For comparison, in Australia, an island nation with a geometry that resembles the model, 30% of the ton/miles of freight transported within Australia is sent by sea. The correct statistic to compare, however, would be the value of products transported by sea and land. Since products with low value per unit weight are often shipped by sea, the value ratio for Australia would be lower.

Another advantage of coastal cities is their ability to ship products overseas, not just to other cities along their own coastline. Overseas markets can be easily added to the model by including an additional point containing a fixed number of firms. Movement of firms to or from this point is not allowed, but goods can be moved back and forth. By changing the size of the overseas market, the relative advantage of coastal versus interior cities can be changed in a similar manner to the previous changing of the cost of ocean shipping. A small overseas market favors a large interior city, while a large overseas market favors many cities along the coast.

The twenty-one countries shown in table 1 are those with geographies that are the most appropriate for comparison with the model. All of them have seacoasts on at least two sides, have less than 25% of their exports going to countries sharing land boundaries, and are at least 30 000 square miles in area. Countries that are much longer than they are wide, such as Panama and Italy, are excluded, as are countries made up of many islands. The model predicts that countries with a high ratio of exports to GDP will have more cities near the coast. Figure 19 shows a scatterplot

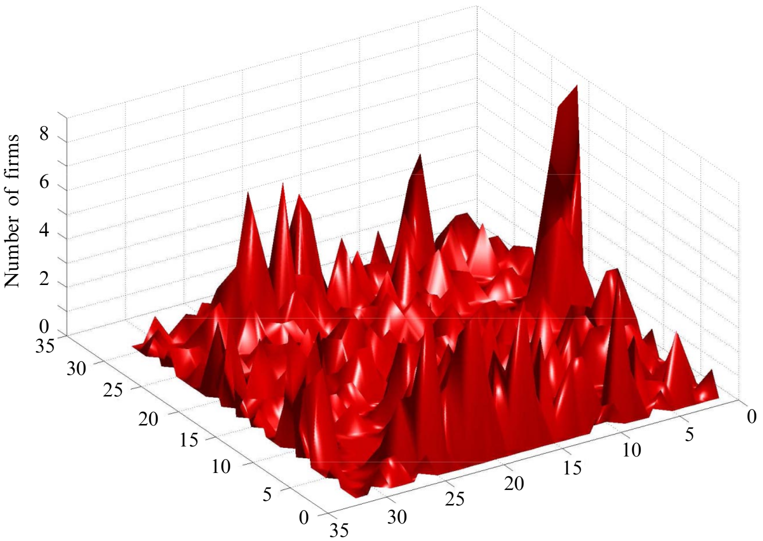

Figure 18. [In color online.] A mixture of coastal and interior cities.

Table 1. Data for selected countries.

	Percentage near coast	Percentage exports to GDP	Population per square mile	Rivers (miles)	Coastline (miles)	Area (square mile)
Australia	88.6	17.6	6.7	0	16 007	2 967 908
Brazil	32.6	16.4	56.4	5 157	4 655	3 265 075
Egypt	25.1	21.6	196.9	2 175	1 522	386 662
Iceland	100.0	36.2	7.4	0	3 099	39 769
India	17.0	14.8	839.1	3 532	4 350	1 269 345
Ireland	89.5	83.7	174.1	0	900	32 587
Korea, South	48.1	37.9	1 278.1	448	1 499	38 023
Madagascar	33.3	23.1	77.2	373	3 000	226 657
Morocco	52.4	32.5	186.8	0	1 140	172 414
Nicaragua	23.8	24.5	107.2	1 379	565	49 998
Norway	89.9	41.3	36.5	980	15 626	125 182
Oman	75.3	56.2	35.4	0	1 300	82 031
Saudi Arabia	31.8	46.1	34.1	0	1 640	756 985
South Africa	32.5	27.9	90.7	0	1 739	471 010
Spain	53.8	26.3	206.7	621	3 084	194 897
Sweden	76.8	43.8	51.7	1 275	2 000	173 732
Tunisia	70.7	43.8	157.9	0	713	63 170
Turkey	47.5	27.4	228.6	0	4 474	301 383
UK	39.1	25.4	657.5	385	7 723	89 073
US	38.7	9.6	82.8	12 000	12 380	3 537 437
Yemen	32.7	30.0	98.2	0	1 184	203 850

of the percentage of the population (in towns with at least 10 000 inhabitants) living within 20 miles of the coast along with exports as a percentage of GDP for the twenty-one countries. The figure also shows the predicted values from the model. There are three significant outliers: Australia, Iceland, and Saudi Arabia. Saudi Arabia has higher exports than other countries with the same fraction of their population living near the coast, probably due to its high level of oil exports. Australia and Iceland are the two countries with by far the lowest population density. Desert and glaciers in the

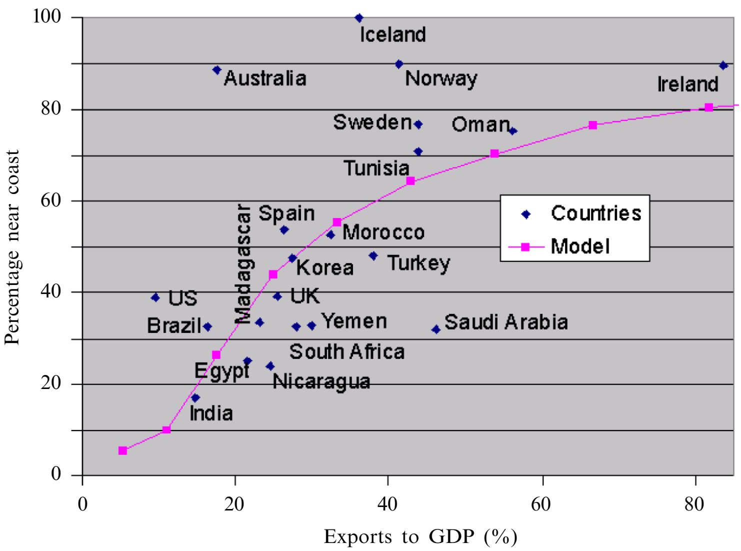

Figure 19. [In color online.] Population near coast and exports to GDP.

interiors of these two countries cause a higher percentage of the population to live near the coast than would be predicted by their trading activity. Other than these three countries, the fit between the model and the country data appears to be reasonably good, even without taking account of factors other than the level of exports.

With only twenty-one observations in table 1, it is difficult to come to definitive conclusions using statistical analysis. Regression analysis of these data should be interpreted with caution, since the small number of observations could cause one unusual country to have a large influence on the results. A variety of specifications and data subsets, however, produced similar results. The regression results in table 2 are not intended as a rigorous empirical investigation, but more as an aid in interpreting the data in table 1.

Table 2. Regression results, dependent variable is percentage of population near coast.

Variable	Parameter estimate	*T*-ratio	*P*-value
Intercept	1.414	4.573	0.000
Exports/GDP	0.425	2.680	0.017
Population density	−0.186	−3.977	0.001
Rivers/area	−13.572	−2.201	0.044
Coastline/area	0.163	2.601	0.020
Oil exporter	−0.417	−2.082	0.055
R^2	0.889		

The shape shown in figure 19 produced by the points predicted by the model suggests a linear relationship between the logs of the percentage of the population near the coast and exports. It seems reasonable to expect that the percentage of the population near the coast would fall with increased population density and rise with an increase in coastline. More navigable rivers would reduce the population near the coast, but this effect might diminish with river mileage, since rivers narrow with distance from the coast. For this reason, the variable for rivers is the navigable mileage, not the log of mileage. The variables for population density and coastline are in logs.

The variable for oil exporters is equal to 1 for countries with oil exports greater than 75% of total exports (Saudi Arabia, Oman, and Yemen) and 0 otherwise. The theoretical relationship is as follows, with P equal to the percentage of population living near the coast, X equal to exports as a percentage of GDP, D equal to population density, R equal to the navigable length of rivers, C equal to the length of the coastline, and O equal to 1 if oil exports are greater than 75% of total exports, and equal to 0 otherwise.

$$P = \alpha X^{\beta_1} D^{\beta_2} \exp(\beta_3 R) C^{\beta_4} \beta_5^O \ . \tag{27}$$

Taking logarithm of both sides gives an equation that can be estimated using ordinary least squares regression.

$$\ln P = \alpha + \beta_1 \ln X + \beta_2 \ln D + \beta_3 R + \beta_4 \ln C + \ln(\beta_5) O \ . \tag{28}$$

The regression results in table 2 confirm the impression given by figure 19 that higher exports as a percentage of GDP increase the percentage of the population living near the coast. Low population density and longer coastlines lead to more people on the coasts. More rivers lead to fewer people on the coasts, and oil-exporting countries have fewer people near the coast than would be expected given only their total level of exports. The fit of the mode is remarkably good, and each of the variables is statistically significant with the expected signs.

Comparing the world as a whole with the model presents some difficulties. The fraction of the world's population that lives near coasts depends on the shapes of land masses as well as on trade patterns. For example, if the world contained a single, perfectly round continent of radius equal to 1 and the world's population were randomly distributed, then the fraction of the world's population living within d miles of the coast would be equal to $2d - d^2$. For example, 19% of the population would live within a distance of one tenth of the radius of the continent from the coast. If the world's continents were long and narrow, then a much larger fraction of the population would live near the coast. To compare the world with the model, the fraction of the world's population near the coast should be compared with the fraction we would expect if people were randomly distributed on the continental shapes of the real world and of the model. The map of the world's land masses in figure 20 shows the distance from a coast for all inhabited areas. Land with low population density is assumed to be uninhabitable due to climate, topography, or other factors. Uninhabited land is shown in figure 20 as white. The color of inhabited areas changes with distance from a coast.

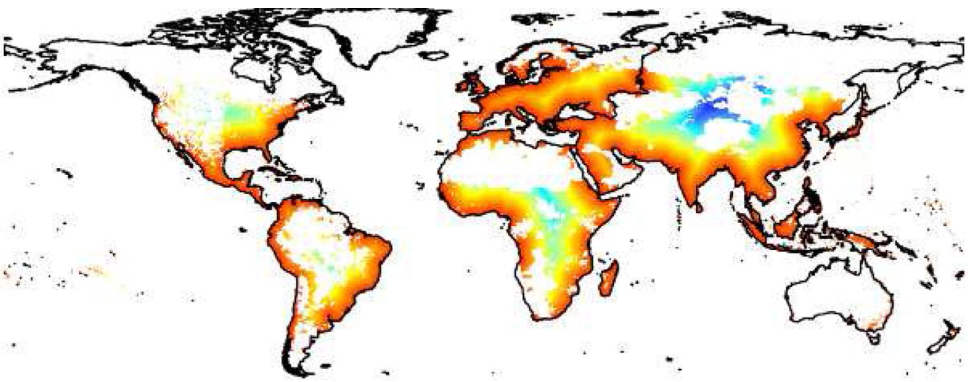

Figure 20. [In color online.] Distance from coast in inhabited areas.

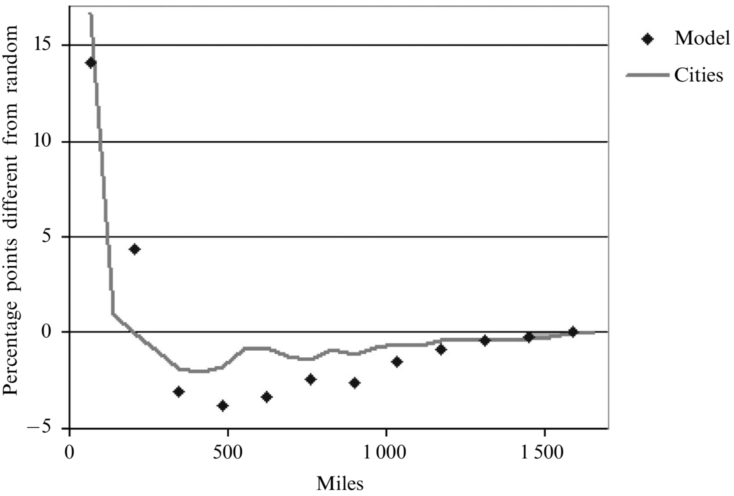

Figure 21. Distribution of cities by distance from coast.

Of inhabited land 31% is within 70 miles of a seacoast. Of all the world's population living in towns of at least 10 000 people, 47% live within 70 miles of the coast. Figure 21 shows the difference between the fraction of the land area and population that is within different distances from a coast. The dots show a comparable value for the model with 25% of the firms located overseas. This figure is roughly equal to the percentage of world GDP that travels by sea.[9] The basic pattern in figure 21 is similar for the model and for real cities. Real cities and simulated cites are roughly 15% more likely to locate near coasts than would be the case if cities located randomly.

6 Conclusion

The model of city evolution in this paper appears to be able to replicate several aspects of the actual distribution of cities in the world. Given the right parameter values, it follows Zipf's law: cities grow, decline, and cluster near rivers and coasts. The mixtures of sea and land transport of goods and of coastal and interior cities can also be similar to those of the real world if the parameter values are right. Krugman's (1998) presentation of a version of this model demonstrated that firms agglomerate into cities, but the geometry was very simple, and two large cities had a tendency to form on opposite sides of the world. The realism that can be obtained by expanding the model into two dimensions should produce more confidence in the potential of the model.

A modest claim for the model would be that it could help to develop better understanding of the factors that influence city distribution. For example, the effect of lower transportation costs and the development of new transportation corridors on the distribution of cities could be investigated, and the predictions of the model could be tested. A more ambitious project would be the development of a model of the world's geography, complete with islands, peninsulas, bays, and harbors. Mountain ranges would affect transportation costs between certain cities, and deserts would affect the density of the agricultural population. The model could be used to predict the effect on the world's urban system of changes in population, transportation innovations, and other factors. The computational demands of such a model would be great,

[9] The Institute for National Strategic Studies reports in their Strategic Assessment 1995 that seaborne commerce accounts for 80% of international trade, and the World Trade Organization reports in their World Trade Report 2004 that international trade accounts for 30% of world GDP. Multiplying these figures together produces an estimate that 24% of world GDP travels by sea.

but would be within the reach of modern computers. A model of this kind would be similar to the large models of world climate that have been developed and that now help predict hurricanes as well as large-scale climate change. Historical data could be used to calibrate the model to the actual evolution of cities over human history. There are many ways to further develop the assumptions of the model. For example, the location decisions of firms might depend on potential income over a larger geographic scale. Products with different characteristics, such as substitutability and different modal transportation costs could be included. Durable structures could be included in the model, which might slow urban decline (Glaeser and Gyourko, 2004).

The most important modification, however, would be to add agglomeration economies. In the current model, unless transportation costs are raised to unrealistic levels, city distribution is far more level than is actually observed. Perhaps the production function could be modified to make firms more productive when they are in close proximity to other firms. With these and other modifications, the model could be a useful tool for understanding the economic geography of the world.

Acknowledgements. I thank Richard Arnott, Rakesh Mohan, and Meg Richardson for many helpful ideas and suggestions. I also thank seminar participants at the 2006 American Real Estate and Urban Economics Association international meetings.

References

Akamatsu T, Takayama Y, Ikeda K, 2009, "Spatial discounting, Fourier, and racetrack economy: a recipe for the analysis of spatial agglomeration models", Munich Personal RePEc Archive Paper 21738, http://mpra.ub.uni-muenchen.de/21738/

Alfaro L, Chandac A, Kalemli-Ozcan S, Sayek S, 2010, "Does foreign direct investment promote growth? Exploring the role of financial markets on linkages" *Journal of Development Economics* **91** 242 – 256

Allen P M, 1996 *Cities and Regions as Self-organizing Systems: Models of Complexity* (Gordon and Breach, New York)

Allen P, Sanglier M, 1978, "Dynamic models of urban growth" *Journal of Social and Biological Structures* **1** 265 – 280

Allen P, Sanglier M, 1979, "A dynamic model of urban growth: II" *Journal of Social and Biological Structures* **2** 269 – 278

Allen P, Sanglier M, 1981, "Urban evolution, self-organization, and decisionmaking" *Environment and Planning A* **13** 167 – 183

Alonso W A, 1964 *Location and Land Use: Toward a General Theory of Land Rent* (Harvard University Press, Cambridge, MA)

Beeson P, DeJong D, Toesken W, 2001, "Population growth in US counties, 1840 – 1990" *Regional Science and Urban Economics* **31** 669 – 699

Bosker M, Brakman S, Garretsen H, 2007, "Looking for multiple equilibria when geography matters: German city growth and the WWII shock" *Journal of Urban Economics* **61** 152 – 169

Bosker M, Brakman S, Garretsen H, Schramm M, 2010, "Adding geography to the new economic geography: bridging the gap between theory and empirics" *Journal of Economic Geography* **10** 793 – 823

Carlino G, Mills E, 1987, "The determinants of county growth" *Journal of Regional Science* **27** 39 – 54

Christaller W, 1933 *Die Zentralen Orte in Suddeutschland* (Gustav Fischer, Jena); translated by C W Baskin in 1966 as *Central Places in Southern Germany* (Prentice-Hall, Englewood Cliffs, NJ)

Davis D, Weinstein D, 2002, "Bones, bombs, and break points: the geography of economic activity" *American Economic Review* **92** 1269 – 1289

Davis D, Weinstein D, 2008, "A search for multiple equilibria in urban industrial structure" *Journal of Regional Science* **48** 29 – 65

Duranton G, 2007, "Urban evolutions: the fast, the slow, and the still" *American Economic Review* **97** 197 – 221

Eeckhout J, 2004, "Gilbrat's law for (all) cities" *American Economic Review* **94** 1429 – 1451

Fujita M, Krugman P, Mori T, 1999, "On the evolution of hierarchical urban systems" *European Economic Review* **43** 209 – 251

Gabaix X, 1999a, "Zipf's law for cities: an explanation" *Quarterly Journal of Economics* **114** 739 – 767

Gabaix X, 1999b, "Zipf's law and the growth of cities" *American Economic Review* **89** 129 – 132
Glaeser E, 2008 *Cities, Agglomeration, and Spatial Equilibrium* (Oxford University Press, New York)
Glaeser E, Gyourko J, 2005, "Urban decline and durable housing" *Journal of Political Economy* **113** 345 – 375
Glansdorff P, Prigogine I, 1971 *Thermodynamic Theory of Structure, Stability, and Fluctuations* (John Wiley, New York)
Henderson J V, 1974, "The sizes and types of cities" *American Economic Review* **64** 640 – 656
Herodotus, 2004 *The Histories* translator G C Mccaulay (Banes and Noble, New York)
Institute for National Strategic Studies, 1995 *Strategic Assessment, Engaging Power for Peace* (National Defense University, Washing, DC)
Isard W, 1956 *Location and the Space Economy* (John Wiley, New York)
Kennan J, Walker J, 2011, "The effect of expected income on individual migration decisions" *Econometrica* **79** 211 – 251
Klein L, Rubin H, 1948, "A constant-utility index of the cost of living" *Review of Economic Studies* **15** 84 – 87
Krugman P, 1991, "Increasing returns and economic geography" *Journal of Political Economy* **99** 483 – 499
Krugman P, 1998, "A slime mold model of city formation" *Topics in Public Economics* Eds D Pines, E Sadka, I Zilcha (Cambridge University Press, Cambridge) pp 15 – 32
Lloyd P, 1975, "Substitution effects and biases in nontrue price indices" *American Economic Review* **65** 301 – 313
Lösch A, 1954 [1940] *The Economics of Location* translation (Yale University Press, New Haven, CT)
Miller E, 2008, "An assessment of CES and Cobb – Douglas production", WP 2008-05, Congressional Budget Office, Washington, DC
Mills E S, 1967, "An aggregative model of resource allocation in a metropolitan area" *American Economic Review* **57** 197 – 210
Mossay P, 2006, "The core – periphery model: a note on the existence and uniqueness of short-run equilibrium" *Journal of Urban Economics* **59** 389 – 393
Muth R F, 1969 *Cities and Housing* (University of Chicago Press, Chicago, IL)
Neary J P, 2001, "Of hype and hyperbolas: introducing the new economic geography" *Journal of Economic Literature* **39** 536 – 561
Nicolis G, Prigogine I, 1977 *Self-organization in Non-equilibrium Systems* (John Wiley, New York)
Nitsch V, 2005, "Zipf zipped" *Journal of Urban Economics* **57** 86 – 100
Picard P, Tabuchi T, 2010, "Self-organized agglomerations and transport costs" *Economic Theory* **42** 565 – 589
Rappaport J, Sachs J, 2003, "The United States as a coastal nation" *Journal of Economic Growth* **8** 5 – 46
Resnick M, 1997 *Turtles, Termites, and Traffic Jams* (MIT Press, Cambridge, MA)
Rossi-Hansberg E, Wright M, 2007, "Urban structure and growth" *Review of Economic Studies* **74** 597 – 624
Sanglier M, Allen P, 1989, "Evolutionary models of urban systems: an application to the Belgian provinces" *Environment and Planning A* **21** 477 – 498
Serow W, O'Cain S, 1992, "The role of region and coastal location in explaining metropolitan population growth differentials during the 1980s" *Review of Regional Science* **22** 217 – 225
Stelder D, 2005, "Where do cities form? A geographical agglomeration model for Europe" *Journal of Regional Science* **45** 657 – 679
Tolley G, 1974, "The welfare economics of city bigness" *Journal of Urban Economics* **1** 324 – 345
von Thünen J, 1966 [1826] *Von Thünen's Isolated State* (Pergamon, New York)
World Trade Organization, 2004 *World Trade Report 2004: Exploring the Linkage Between the Domestic Policy Environment and International Trade* (World Trade Organization, Geneva)