Journal of Economic Literature 2014, 52(3), 805–819
http://dx.doi.org/10.1257/jel.52.3.805

The New Science of Cities by Michael Batty: The Opinion of an Economist†

Jacques-François Thisse*

Cities are the cradle of a wide range of cultural, social, and technological innovations that are at the heart of modern economic growth and development. Half of humanity today lives in cities but, until the last two decades, economists have paid much less attention to cities than have other social scientists. By contrast, geographers have long studied the role of cities in human affairs. Michael Batty, a distinguished scholar in the field of human geography, has recently written The New Science of Cities, *a synthesis of his work and of some other prominent urban geographers. A review of his book is the first objective of this essay. The second is to discuss and compare the tools and concepts developed by urban economists with those of urban geographers in the hope of triggering a fruitful debate between those two groups of social scientists.* (*JEL* R10, R23, R30, R40, R58)

1. *Introduction*

Michael Batty is one of the most distinguished scholars in the field of human geography, as well as a scientist with broad interests. He was awarded a CBE in June 2004 for services to geography and made a Fellow of the Royal Society in 2009. A few years ago, Batty summarized his view of cities as being "complex systems that mainly grow from the bottom up, their size and shape following well-defined scaling laws that result from intense competition for space" (Batty 2008, p. 769). His latest book, *The New Science of Cities*, may be viewed as a synthesis of his work in urban geography that aims at developing, as the title suggests, a new science of cities.

The book is divided into three main parts. In part I, Batty discusses the main concepts he uses in his book. Part II focuses on a positive analysis of cities and urban systems, whereas part III, which addresses the issue of urban design, has a normative flavor. Providing a thorough discussion of a book as rich as Batty's cannot be done in a review without being superficial. I will therefore focus on the ideas and concepts underpinning the book. I will also discuss, from an economist's point of view, what I see as the book's main contributions, as well as the topics where there is a real potential for interaction between urban geographers and urban economists. I apologize to Batty for not covering issues that are too far from my domain of expertise.

* CORE-UCLouvain (Belgium)

† Go to http://dx.doi.org/10.1257/jel.52.3.805 to visit the article page and view author disclosure statement(s).

Batty, as an urban geographer, aims to build a *geographic* theory of cities where aggregates such as populations and various flows are central to the argument, whereas urban economics focuses on the spatial behavior of individual agents—consumers and firms. Yet, in my view, such a difference in approaches is not necessarily an issue. Indeed, it would be futile to expect an integrated theory of cities that would appeal equally to economists, regional scientists, geographers, and urban planners. In a word, cities are simply too complex. As suggested by Batty, we must learn how to live with "a series of approaches that adopt . . . various methods and tools" (p. 118). There will always be something to learn from alternative approaches, and not just in urban analysis.

The first three chapters (part I) describe the theoretical framework underlying the rest of the book; therefore, I find it appropriate to discuss in detail the ideas and concepts developed in these chapters.

2. *The Proximity/Market Crowding Trade-Off*

For Batty, locations are important, but only as places that anchor interactions (p. 8). To put it differently, to understand cities, we must view them not simply as places in space but as systems of networks and flows. Thus, we are far from the attitude of some geographers who avoid abstract theories and focus on real communities.[1] In contrast, we are close to urban economics, where first nature geography (e.g., natural resources and climate differences) is considered weak as a way of explaining second nature geography, which involves large cities and big trade flows.

[1] The tribe of geographers is more heterogeneous than that of economists. In this review, by geographers I mean quantitative human geographers.

For centuries, the main distinctive feature of cities was their separation from the countryside. This separation no longer exists. Urban activities have gradually extended beyond the city's physical boundaries to create suburbs and city-edges. These ones are now very much part of the city, which is considered as an all-encompassing economic agglomeration. The main reason for the existence of cities is to *connect people*. This vision is now widely embraced by urban economists (Glaeser 2011).

Batty's analysis rests on the key idea that the form of a city reflects a fundamental tension between the desire to be as close as possible to everyone else and the desire to consume as much space as possible (pp. 19 and 253). In such a setting, the need to interact with others plays the role of a centripetal/agglomeration force, whereas competition for land acts as a centrifugal/dispersion force. Urban economists agree with this view because similar trade-offs are at the heart of their field and the new economic geography. As shown by Beckmann (1976), the mere need to interact turns out to be sufficient to generate a single-peaked distribution of individuals who compete for space.

Ogawa and Fujita (1980) go one step further and combine consumers and firms within a full-fledged model of general equilibrium. The centripetal force—the exchange of information among firms—is subject to distance-decay effects, so the value of a firm's location is based on its proximity to other firms. Since workers are keen to minimize commuting costs, the centrifugal force goes through the land and labor markets. The clustering of many firms in a single area increases the average commuting distance for their workers, which in turn increases both the land rent workers pay and the wage rate paid by firms. As locations are endogenous and interdependent, the equilibrium distribution of firms and households is the balance between these opposing forces. Ogawa and Fujita show that the equilibrium city may display different configurations. For example, when commuting costs are high in relation to the distance-decay effect, the equilibrium involves a full integration of business and residential activities. By contrast, when interactions among firms are strong enough to dominate the dispersion force generated by commuting costs, the equilibrium outcome involves a monocentric city with a central business district (CBD) flanked by two residential areas.

Until the 1990s, these and related contributions have attracted little attention in economics for two reasons. On the one hand, the economics profession has paid little attention to the organization of the space-economy because the constant returns–perfect competition paradigm can hardly cope with the emergence of large economic agglomerations and sizable regional disparities. On the other hand, urban economists probably spent too much time streamlining the monocentric city model, which has the following two advantages: it is compatible with the competitive paradigm and it treats the CBD as a given. What could bring geographers and economists together is the clear recognition that nonmarket interactions and externalities are fundamental to the process that gives rise to the city.

3. *Power Laws*

Power (or scaling) laws have been used successfully for studying the relationships between body size and shape, so it is easy to see why urban geographers are interested in such tools when trying to understand how the size and morphology of cities are related. In chapter 1, Batty considers various power functions of city population $P(t)$ at time t:

$$Y(t) = \alpha P(t)^{\alpha}, \tag{1}$$

where $Y(t)$ can stand for very different magnitudes, such as total income, but also for the number of new patents or road surfaces, while α is the elasticity of Y with respect to P.

The meaning of an elasticity larger or smaller than one depends on what Y stands for. If this variable represents the total income, an exponent larger (smaller) than one implies the existence of agglomeration economies (diseconomies). If Y stands for the stock of a particular type of infrastructure, an exponent smaller (larger) than one indicates that the provision of this infrastructure displays increasing (decreasing) returns. Power functions yield a very good fit when tested against a great number of city variables (see, e.g., Bettencourt et al. 2007).

As argued by Batty, power functions may generate the celebrated rank–size distribution: if we rank cities from the largest to the smallest, the population (or other measures) of a city at any rank in the hierarchy is the size of the largest city divided by its rank. In other words, as cities get larger, there are fewer of them. Evidently, the law $P(r) \propto 1/r$ is a special case of (1) where α is close to one.

However, (1) assumes implicitly that the population P is *the* explanatory variable, thereby neglecting the well-known endogeneity problem, which has two main sources. First, population size is likely to capture the impact of omitted variables, such as the supply of public services and the presence of natural amenities, which influence the population size. Second, city size is itself an endogenous variable because people are drawn to areas with higher incomes and better infrastructures. Estimating (1) shows only the extent of correlation between Y and P. Therefore, one should refrain from building too much on such simple relationships.

There is another reason to be unhappy about power functions. Indeed, these functions generate scaling effects that have an undesirable consequence: *cities that are apparently different are scaled versions of one another*. In other words, using power functions amounts to assuming that the city industrial structure doesn't matter, as if cities were more or less the same at various spatial scales. This seems to be implausible. The sector in which cities are specialized is critical in explaining their difference in size and activities (Henderson 1988). Large cities are often diversified, while the development of new information technologies has fostered a transformation of their industrial structure, which has shifted from sectoral specialization to functional specialization. Owing to cities' industrial composition and jobs, large cities are not scaled-up versions of small cities, in a manner similar to matryoshka dolls.

Batty states seven "loose laws" (p. 40) that may be given microeconomic underpinnings:

1) As cities get bigger, the number of potential connections increases as the square of the population size.
2) As cities get bigger, their average income increases more than proportionally.
3) As cities get bigger, there are fewer of them.
4) As a city grows around its CBD, various densities, including the land rent and the population density, decline nonlinearly with distance from the CBD.
5) As cities get bigger, interactions between them scale with their size.
6) As cities get bigger, their population densities flatten.
7) As cities get bigger, they become greener.

Each of the first six laws can be shown to hold in the monocentric city model of urban economics (Fujita 1989) or the economic theory of urban systems (Henderson 1988). The last law is a more recent finding that needs confirmation.

Using power functions, Batty then discusses, in a very ingenious way, some implications of his seven laws. For example, the potential interaction within a population of size P is $P(P-1)/2$. Since distance is an impediment to interaction, the potential interaction is proportional to $P(P-1)/2d^2$, where d is the radius of the city area. We thus fall back on the Newtonian gravitational model. But, unfortunately, Batty has chosen to leave this theory "hanging for future work" (p. 43). I am not convinced he made the right choice. Furthermore, I didn't find any detailed discussion of how competition for land is related to the existence of power laws. In a way, this is not surprising because Batty does not consider a land market explicitly. It would have been interesting and informative to read a more detailed discussion of the socioeconomic foundations of power laws.

4. *Spatial Interaction Theory*

In chapter 2, Batty surveys one of the main theories of human geography, the *spatial interaction theory*.[2] The aim is to study the formation of different types of flows, i.e., goods, people—customers or workers—and information, between a given set of origins O and a given set of destinations D. The units of flow, $T_{ij} \geq 0$, from origin $i \in O$ to destination $j \in D$, are designated as *trips*. In operations research, T_{ij} represents the quantity of a certain commodity shipped from zone i to zone j. In urban geography, O often stands for consumers' residences, while D represents firms' locations (shops, plants, or offices) set up in the city. In this case, a trip describes the number of customers (commuters) residing at i and patronizing shops (working in firms) established at j. The key issue is to determine how the volumes available in the various origins are *allocated* among the destinations.

[2] Some of the material presented here is discussed in chapter 9 of the book. However, Batty stresses the relationships between spatial interaction theory and physics, whereas I stress those with economics.

4.1 *The Transportation Problem*

The basic model, which lies at the root of operations research, is the *transportation problem* embedded into the framework of economic analysis by Dorfman, Samuelson, and Solow (1958). It derives its name from the many applications to problems involving transporting commodities from several sources to several destinations. More precisely, the supply of a certain commodity at origin i (the number of consumers) is perfectly inelastic and given by $s_i \geq 0$, while the demand for this good at destination j (the number of shops) is also perfectly inelastic and given by $d_j \geq 0$. Thus, the total amount of trips

$$(2) \quad T = \sum_i \sum_j T_{ij} = \sum_i s_i = \sum_j d_j$$

is constant. With the unit cost of movement from i to j denoted by $c_{ij} > 0$, the purpose of the model is to find the commodity flows $T_{ij} \geq 0$ that minimize the total transportation cost

$$(3) \qquad C = \sum_i \sum_j c_{ij} T_{ij}$$

subject to the constraints

$$(4) \quad \sum_j T_{ij} = s_i \qquad \sum_i T_{ij} = d_j,$$

where all numbers are nonnegative integers.

Since both the objective function and the constraints are linear, this optimization problem belongs to the field of linear programming. The optimal graph of flows is sparse: a source supplies a small number of destinations, while a destination is supplied by a small number of origins. To be precise, the matrix $[T_{ij}]$ includes many zeroes: there are $|O||D|$ potential flows, but the number of positive flows at the equilibrium outcome is equal to $|O| + |D| - 1$.

Then the focus shifted rapidly from optimality to equilibrium (Samuelson 1952). When supply and demand functions are elastic, the competitive equilibrium is reached when the demand price equals the supply price plus the transportation cost for all positive flows. If the demand price is less than the supply price plus the transportation cost, then no trade flow occurs. The total number of trips T is now endogenous, but the matrix $[T_{ij}]$ still includes many zeroes. Samuelson's model could have been used to develop a trade theory with transportation costs, which came into being only with the new trade theories.

4.2 *Gravity Models*

Geographers have followed a different path from the economists. To be precise, spatial interaction theory builds on the Newtonian gravitation law. The canonical model assumes that flows between spatially separated populations are given by the following gravitational law (p. 62):

$$(5) \qquad T_{ij} = KP_i P_j d_{ij}^{-\theta}.$$

In this expression, K is a constant, P_i is the (possibly unknown) population size at size i, d_{ij} is some measure of distance between i and j with $d_{ij} = d_{ji}$, and θ a parameter measuring the intensity of the friction that distance imposes on interaction. The matrix $[T_{ij}]$ contains no zeroes because the functional form in (5) implies that there is a positive flow, however small it is, for every origin–destination pair, provided that P_i and P_j are both positive. Three families of models have been developed.

1) If the total number of trips T and the populations P_i and P_j are given a priori, the parameter θ in (5) may be

estimated provided that the constant K takes on the following form:

$$(6) \qquad K = \frac{T}{\sum_i \sum_j P_i P_j d_{ij}^{-\theta}}.$$

Observe that the denominator of K accounts for multilateral, instead of bilateral, resistance terms. It was not until Anderson and van Wincoop (2004) that trade economists became aware of the importance of such effects in gravity equations.

2) The above model can be relaxed by taking only the constraint on the sum of the origin-populations P_i (or the destination-populations P_j) into account. In the former (latter) case, the constant K is replaced with KA_i (KA_j) in T_{ij}, where

$$(7) \qquad A_i = \sum_j P_j \, d_{ij}^{-\theta} \qquad \left(A_j = \sum_i P_i \, d_{ji}^{-\theta}\right).$$

Notice that A_i (A_j) captures the idea of the *accessibility* of origin i (destination j) to (from) the various destinations (origins) through the sum of all flows emanating (converging) from i (to j). The access to opportunities is a basic issue in economic theory. However, apart from the spatial mismatch model in labor economics and very few others, economists have paid little attention to the spatial accessibility to opportunities, that is, *where* these opportunities are available. Here lies a whole research domain where economists and geographers could fruitfully learn from each other, the aim being to better understand *why* opportunities are made available in some places, but not in others.

In the origin-constrained model, firms' sizes are endogenous and determined by the spatial behavior of consumers. In the destination-constrained model, the housing stock used in each origin is endogenous. Thus, these two models allow one to determine both the location of firms and consumers.

3) Lastly, when trips are observed, the quantities P_i and P_j are endogenous and determined from

$$(8) \qquad P_i = \sum_j T_{ij} \qquad P_j = \sum_i T_{ij}.$$

In this way, we start from the flows across places to get the distribution of activities over the location space, while in spatial economics, flows are obtained when the economic agents have chosen their locations.

But where does (5) come from? And how to relate T_{ij} to the optimizing behavior of agents? To the best of my knowledge, few geographers have explored the micro foundations of the gravitational forces they use. Batty makes a brief reference (p. 69) that these models can be generated by using random utility theory. This attitude is surprising because gravity-like models have recently attracted a lot of attention in the trade literature and beyond, and we now have a better understanding of the microeconomic underpinnings of such models (Head and Mayer 2014). In a book on a new science of cities, I would have expected Batty to pay more attention to the relationships between aggregate flows and individual spatial behavior. Here we encounter what I view as a distinctive feature of the approach of many quantitative geographers: they focus on semiaggregated approaches that provide a good fit of spatial magnitudes and are content with intuitive explanations of their fundamentals.

Given the importance of gravity-like equations in both economics and geography, it is important to look at how geographers and transportation analysts have studied this family of models before economists started their investigation. From the historical point of view, the first sound approach was pioneered by Wilson (1970), who used the concept of entropy to explain the formation of flows. The argument goes as follows: The analysts often have access to some information about certain aggregates such as the populations P_i and P_j, the total number of trips T, and the total (time or energy) cost C spent in commuting or shopping. However, they seldom know how individual agents make their trip decisions.

4.3 *Entropy-Like Models*

The entropy within the transportation system is defined (up to an additional constant) by the following expression:

$$\mathbf{E} = \sum_i \sum_j T_{ij} \ln (P_i P_j) - \sum_i \sum_j T_{ij} \ln T_{ij}.$$

The maximization of $\mathbf{E}$ yields the trips T_{ij} that satisfy the following two conditions: (i) they are consistent with given macroscopic constraints, such as (8), which are known by the analyst, and (ii) they minimize the analyst's bias about the spatial behavior of individual agents, which is unknown to the analyst. This approach is thus reminiscent of that adopted by the econometrician who uses discrete choice models in a situation of incomplete information.

In most entropy models of trip formation, (3) is added as an additional constraint. It is then straightforward to show that maximizing $\mathbf{E}$ under the constraints (3) and (8) yields

$$T_{ij} = P_i P_j \exp(\lambda_i + \mu_j - \beta c_{ij}), \tag{9}$$

where λ_i and μ_j are the Lagrange multipliers associated with the population constraints (8), while β, which is the multiplier associated with the constraint (3), reflects individual spatial behavior within the transportation system.

If $c_{ij} = \ln d_{ij}$, we fall back to the negative power relationship (5). According to the value of β, (9) is consistent with a variety of individual spatial choice behaviors. In the limit, when β is arbitrarily large, the entropy maximization boils down to the transportation problem discussed above.

Relaxing the number of trips to be made to each destination allows consumers located at i to choose how to distribute their trips within D. In this case, the constrained maximization of $\mathbf{E}$ leads to

$$T_{ij} = \frac{\exp(-\beta c_{ij})}{\sum_k \exp(-\beta c_{kj})} P_i, \tag{10}$$

where the ratio is the multinomial logit probability that a consumer residing in i visits zone j. Evidently, when β tends to infinity, every consumer located in i patronizes only the shops established in the lowest-travel-cost destination. It should be stressed that the most popular justification of the gravity equation used in trade presupposes that consumers are endowed with preferences where the elasticity of substitution between any two goods is constant (CES), so that consumers have a taste for variety. This, in turn, implies that they divide their purchases among different countries. Although seldom noticed, there is a close relation between the approaches followed by geographers, transportation analysts, and economists to study (trade) flows.

Geographers have thus developed, before the 1990s, tools and concepts that would have been very useful to economists. But, in the huge number of publications coping with the gravity equation in international trade, I have never found a reference to spatial interaction theory. So, if geographers' approach to power laws disregards the various econometric problems pointed out by economists, it can also be said that economists have ignored spatial interaction theory.

4.4 *Random Utility Models*

Despite the appeal of the entropy modeling strategy, I find it fair to say that the easiest way to generalize the transportation problem is probably to use the framework of *random utility* maximization developed by McFadden (1974). This is done by adding a random term ε_j to the unit transportation costs c_{ij}. If the variables ε_j are identically and independently Gumbel-distributed with a variance proportional to $1/\beta^2$, (10) is equal to the probability that destination j maximizes the utility $-c_{ij} + \varepsilon_j$ of a consumer located at i. As a consequence, using entropy amounts to assuming that each origin accommodates a representative consumer endowed with an entropy-like subutility nested in quasi-linear preferences (Anderson, de Palma, and Thisse 1992).

More generally, if the individual utility is given by

$$\tilde{u}_{ij} = v_{ij} + \varepsilon_j,$$

where v_{ij} is the deterministic/measured part of the utility, while ε_j accounts for the analyst's uncertainty about a consumer's spatial behavior, the probability that a consumer visits j is equal to the probability of the various occurrences for which j is the consumer's most preferred destination:

$$\mathbf{P}_j^i = E(\tilde{u}_{ij} = \max_k \tilde{u}_{ik}),$$

where E is the expected value operator. The random utility and entropy approaches are thus formally equivalent, as shown by Anas (1983). They are built on the same idea of incomplete information on the analyst' side, but the levels of aggregation differ. Economists are likely to prefer the random utility approach because of its deeper microeconomic underpinnings and because it permits using a wider range of econometric techniques. Yet, the entropy approach remains a valuable tool for studying specific topics, especially the impact of different levels of spatial aggregation.

It is worth stressing that the random utility model may be reinterpreted by assuming that a particular realization of ε_j is the matching value between a specific consumer residing at i and the shopping or job opportunities available at j. In this event, consumers residing at i are heterogeneous, in the sense that they have different tastes, since they make different choices within the same set of opportunities. Thus, contrary to general belief, geographic models of entropy are consistent with rational choice behavior. This observation should be sufficient to convince economists to pay more attention to what geographers have accomplished with spatial interaction models. The entropy, multinomial logit and CES approaches are very close relatives.

That said, whereas prices and many other economic variables are absent in spatial interaction theory, they can be grafted onto random utility models to study markets equilibria and related issues (Anderson, de Palma, and Thisse 1992). For example, one of the main differences between the standard gravity equation used in trade and its microeconomic rationalizations is precisely the presence of prices in the generalized gravity equations. This highlights differences between those variables on which economists and geographers focus.

5. *Graphs and Networks*

The third chapter of the book is devoted to graph/network theory. Batty observes that flows and networks are the opposite sides of the same coin (p. 79) and, therefore, deserve equal attention. In my opinion, chapter 3 could have benefited from a fuller discussion of the following three topics: 1) The chapter does not really say how graph theory can be used to cope with urban problems. However, more detailed and stimulating material on networks is provided in chapters 6 and 7 of this book. 2) More importantly, Batty does not develop, or sketch out, any theory that might allow us to better understand how physical and social networks interact to shape cities. Admittedly, the subject is a difficult one and, to the best of my knowledge, the situation is not better in urban economics where most contributions ignore the importance of networks. (Helsley and Zenou 2014 is a valuable exception.) 3) The location of (private or public) facilities on networks in not addressed at all, despite a vast and well-established literature developed in operations research and to which geographers have contributed (Labbé, Peeters, and Thisse 1995). This body of literature both highlights how networks affect the location of activities and provides efficient algorithms to solve numerically large-scale problems. Facility location models could be grafted onto the operational models proposed by Batty in chapter 9 and part III.

In sum, the first three chapters, which serve as a foundation for the rest of the book, contain a wealth of interesting material with which economists (not just urban) should be familiar. In particular, I agree with Batty, who says that "cities are sets of actions, interactions, and transactions" (p. 115). But this only begs the question of who are the relevant decisionmakers and how do they interact? More specifically, how are their transactions conducted? In this respect, these chapters offer little that is new to urban economists. For example, Batty recognizes that agglomeration economies are at work in cities. But to understand how cities function, don't we need a better understanding of the forces shaping agglomeration economies? Are these forces present in small and large cities alike? Probably not, but why not? Perhaps Batty has chosen to leave these issues to economists who are making substantial progress in this area. If so, the trade seems fair. But then I would have expected these three chapters to present a more detailed discussion of network and complexity theories, which Batty sees as the main ingredients of the new science of cities.

In part II, Batty turns his attention to processes that generate cities' physical forms. The focus is mainly on city growth, systems of cities, accessibility within cities, and bottom–up hierarchical processes within and between cities.

6. *City Size Distributions*

Chapters 4 and 5 are closely related. The former addresses the issues of city size distribution and urban hierarchy in a nonspatial context, while the latter deals with the same issues when cities are embedded in a grid-like network. In chapter 4, Batty aims to build models that are stripped to the essence of the issue in question and, as in chapter 1, city population is the main driver. To achieve his goal, Batty starts with the simplest model in which the population of city i changes according to

$$(11) \quad P_i(t+1) = [1 + \lambda_i(t+1)]P_i(t),$$

where the growth (decay) rates $\lambda_i(t)$ are independent random variables with a finite support centered at 0. In other words, cities grow randomly, but proportionally to their size. In addition, city size cannot be smaller than a certain threshold. Since growth shocks are independent random variables, it is not surprising that the probability of a series of positive (negative) shocks in any given place is low. Therefore, few cities grow very big while few disappear; most places are clustered around the middle. More surprisingly, this model generates a distribution that bears some strong resemblance to the lognormal distribution, which is known to provide a very good approximation of the real-world city distribution. This is a remarkable result because it would be hard to develop a simpler model for this result to hold. Indeed, there is no interaction or competition across cities in the random proportionate growth model. However, as acknowledged by Batty (p. 129), this model includes nothing specific to cities, since the aforementioned result applies equally to different objects, such as the top one hundred U.S. firms from 1955 to 1994 or the skyscrapers in New York City.

Furthermore, when a population conservation constraint is accounted for, the only way for a city population to change is through migration. If the size of the migration flow between any two cities is given while the two cities i and j involved in swapping population at time t are chosen randomly, the steady-state distribution of the process is the negative exponential that, like power functions, gives rise to a log-linear relationship.

Although the big picture seems promising, there is too much volatility within the top one hundred places, whereas the population size of large cities displays a fair amount of inertia. In addition, nothing is said about where cities are located. So, richer models are required.

In chapter 5, Batty explores different alternatives that all rely on some form of interactions between cities. One of the most interesting processes is the Gibrat interaction model, where links are randomly added to cities. First, a city i is randomly chosen in proportion to its size. The probability of adding a link from i to city j depends positively on the size of city i and negatively on the distance between i and j. Hence, in accordance with gravity-like models, larger and closer places get preferential treatment.

Starting with an undifferentiated set of locations, the Gibrat interaction model generates a pattern of cities that is in accordance with the hierarchical principle of central place theory, i.e., the number of cities decreases as we go up the hierarchy. Unfortunately, despite its stronger intuitive appeal, this model still lacks any substantial socioeconomic foundations. The difference is striking when it is compared to with what Fujita, Krugman, and Mori (1999) have accomplished. These authors show that a regular hierarchical central place system emerges as the equilibrium outcome of a setting where manufacturing firms and farmers compete for workers who choose their residential locations and types of job while shipping goods is costly. Manufacturing clusters form cities with different sizes and industrial structures. Because manufactured goods are differentiated, trade between cities of the same rank may be larger than trade with smaller cities. As a result, different types of hierarchies interlock within the urban system.

Batty and Fujita et al. thus concur in viewing the urban hierarchy as the unintentional consequence of a decentralized process. Furthermore, both approaches are able to replicate real-world regularities from the bottom up. However, Batty's models remain, at least to me, black boxes. By contrast, in Fujita, Krugman, and Mori (1999), the urban system arises as the aggregate outcome of interactions among firms and households competing on the product, land, and labor markets.

7. *Urban Morphology*

In chapter 6, but also throughout his book, Batty stresses the importance of the physical form (or spatial organization) of cities, from which "we infer processes that create the structures we see in cities" (p. 179). But there is a lack of any compelling evidence supporting such a strong statement. Obviously, I agree that the design of a city may foster or dampen particular social relations, but it does not determine many socioeconomic relations. In addition, various facets of the spatial organization of a city, such as the number of subcenters, are endogenous and determined by the interplay between different socioeconomic forces.

For sure, studying the physical form of a city, its built spaces, and their evolution are important issues for urban planners and architects, and probably for real estate economists too. By contrast, the physical form of cities is not a top priority in urban economic theory, where many papers happily focus on linear cities. Only a handful of contributions to the literature pay attention to the fact that cities are two- or three-dimensional. There are at least two reasons for this neglect. First, many fundamental economic issues (e.g., the shape of the land rent profile and population density, the role of spatial externalities in explaining the emergence of a CBD, or the mono- or polycentric nature of the metropolitan area) may be studied by using linear models. Second, working with two (or three) dimensions requires the choice of a particular metric expressing the distance between any two locations. One implication, which is often ignored by urban economists, is that different metrics need not yield the same equilibrium outcome. To Batty, this should not come as a surprise: when urban transportation networks have different shapes, economic activities are not distributed in the same way within cities. This issue is not a trivial one. As transportation and land use are intimately related, transportation policies are one of the main instruments used by governments to control and promote urban growth. So, Batty is right when he reminds us to pay more attention to the configuration of spatial networks. To the best of my knowledge, very little is known about the possible impact of transportation network configurations on the economic performance of cities. This can be accomplished only in a genuinely two-dimensional network.

Chapters 6 and 7 are linked. Many concepts used in these chapters come from the literature on urban planning and design, as well as graph and network theories. In particular, Batty shows how difficult it is to measure a site's accessibility and the distance between two sites in a city, where different types of networks are mixed and superimposed. Although the focus is different, the content of these two chapters is more in tune with network economics. The discussion provided by Batty also suggests that studying coupled networks is a hard task; this probably explains why the subject is not addressed in any depth in chapter 3.

8. *Cellular Automata Models*

Large cities are not clusters organized around a CBD, but are formed by a hierarchy of clusters whose distribution seems to obey certain rules. Analysis of how the structure of a city evolves in response to some shock has been pioneered by Fujita and Ogawa (1982). These authors showed that the interplay between commuting costs and spatial externalities among firms may generate a variety of land developments and urban forms nesting different types of centers. Thus far, despite quite a few valuable empirical studies and a handful of theoretical papers, urban economic theory doesn't have much to say about the transition from a monocentric to a polycentric urban structure and the organization of subcenters within the metropolitan area. Batty seems to have more to offer in chapters 8 and 9, but his models, like this of Fujita and Ogawa, are also solved numerically. Still, his approach is different in many respects from that followed by those authors. It epitomizes how geographers tackle this difficult problem.

Chapter 8 shows the author at his best. The story goes like this. Assume a trader located at the crossing of two rivers. For quite a while, this trader remains alone. However, other traders also seek a place to settle down. In doing so, they move randomly north or south or west or east across the locational plane. Given enough time and enough traders, a second trader will locate in the vicinity of the first one because of the advantages generated by the division of labor. As time goes by, owing to the benefits generated by a large size, more and more traders become attracted by this area. Since traders consume land, new traders are more likely to find the available locations at the fringe of the settlement. Does the cluster look like von Thünen's circles? Unexpectedly, the answer is no. *Traders distribute themselves across space according to a tree-like pattern*. In other words, this model—which relies on the tension between agglomeration effects and competition for land—generates a network structure. It then becomes clear why networks and graphs are so central to Batty's analysis of cities. Moreover, the resultant structure is *fractal*, that is, it is the same at different spatial scales. All of this differs from the few attempts made in urban economics to move away from the canonical monocentric city model.

Batty goes on to impose more control over the process of land development by appealing to the concept of cellular automata (CA), one the main tools used in contemporary human geography. Recall that an automata is a virtual machine, consisting of a finite number of states driven by inputs that change the states of the machine into outputs, e.g., from undeveloped to developed. At the next period, these outputs are used as inputs to generate new changes. In a CA, changes are initiated only from neighboring cells located on a rectangular grid. To illustrate, consider the Moore neighborhood N_i of cell i, which is formed by the eight cells in the N–S and W–E axes, plus NW–SE and NE–SW diagonals. The working of a CA may be described by the following principle (Batty 2005): *If* something happens in the Moore neighborhood of a cell, *then* something may happen to the cell in question.

In its simplest form, the state A_{it} of cell i (a land plot) at time t is given by $A_{it} = 1$ (0) if i is developed (undeveloped). For the "then" to apply, we have to specify a transition rule describing how cell i changes from t to $t + 1$. One such rule is given by

$$F_{it} = \sum_{k \in N_i} A_{kt} \tag{12}$$

and

$$\text{if } F_{it} > \Psi \text{ (which is given),} \quad \text{then } A_{it+1} = 1. \tag{13}$$

In other words, if cell i has more than Ψ developed cells in its Moore neighborhood, then it is developed at time $t + 1$ if it is undeveloped at time t and remains (un)developed otherwise. This is in accordance with the idea underlined in new economic geography—large places are more attractive than small places. In particular, when $\Psi = 1$ all cells belonging to the Moore neighborhood of a developed cell at time t are developed at time $t + 1$, and so on. Evidently, more sophisticated transition rules can be used.

One of the most natural extensions is to consider probabilistic transition rules. For example, (12) and (13) are replaced by

$$p_{it} = \sum_{k \in N_i} A_{kt} \,/\, 8$$

and

$$\text{if } \overline{\Psi} < p_{it}, \text{ then } A_{it+1} = 1,$$

where p_{it} is the probability of development and $\overline{\Psi}$ is a positive random variable whose realization must be smaller than the percentage of developed sites in the neighborhood of i.

In another extension, which is in the spirit of the gravity models discussed above, it is assumed that

$$F_{it} = \sum_{k \in N_i} A_{kt}/d_{ik}^2,$$

which allows us to capture more elaborated neighboring effects than (13). In a third extension, one distinguishes between different types of activities and makes the development of a cell dependent on the activities located in its vicinity.

To an extent, the transition rule specifies the potential of a particular piece of land to be developed in a certain environment. By applying the chosen rule recursively, one generates a pattern of land use over time, that is, the way a city grows. By tweaking rules over time, it is even possible to adjust to unexpected evolutions. Unfortunately, sophisticated rules seem to yield land development patterns that are hard to trace. In addition, the number of possibilities is gigantic, so that choosing the right transition rules is very difficult. This leads Batty to write that the problem is "to select those rules that are observable in the way real cities develop, while keeping such models as simple as possible" (p. 269).

As shown by the work of urban geographers, especially Batty's, CA-based models have proven to be useful "to replicate many different generative phenomena that characterize many different forms of cities" (p. 267). However, by using transition rules that are often ad hoc or simple rules of thumb, these models fail to explain why and how cities actually grow and develop. CA-based models cannot have a real predictive power as long as they do not recognize that cities are the result of choices made by agents with their own motivations. Most important, how can these models help to design public policies if they ignore know how agents react to the new incentive systems generated by these policies.

I am unable to see how CA-based models allow for a more disaggregated approach than the spatial interaction models discussed above. To what extent does a CA really differ, say, from the Gibrat interaction model? In addition, I don't understand why Batty has chosen, in his new science of cities, to ignore agent-based models, which allow for more sophisticated behaviors than CA-based models (Batty 2005). In this alternative framework, agents—not to be confused with the agents of microeconomic theory—are entities that act by themselves according to some protocol. For example, Caruso et al. (2007) studied local interactions among residents who also value green externalities, stemming from the proximity to farmers. These authors combine an urban economics model with a CA to simulate neighborhood interactions over time in a two-dimensional space. Their model generates a path-dependent equilibrium that, according to the parameter values, can display a wide range of urban morphologies, which need not be star-shaped as in the monocentric city model. By studying both the equilibrium outcomes and their evolution, Caruso et al. (2007) offer a promising potential for interaction between economists and geographers. Note also the existence of a connection between the agent-based models used by geographers and evolutionary game theory, which seems worth exploring.

In chapter 9, Batty argues that "fast, simple, visual, and accessible models" are needed (p. 276). And indeed, if these models are to be used to help decisionmakers and stakeholders, we had better tools these people understand. That is a sensible objective but not necessarily what I see as the main objective of the new science of cities. Moreover, I found it hard to figure out what kind of objective the decisionmakers have in such models, apart from seeking a compromise. And, for a compromise to arise, it must be that distinct parties have diverging but specific goals.

In this chapter, Batty also discusses how to use spatial interaction models in applications to modal choice. His settings differ greatly from, say, Regional Economy, Land Use and Transportation (RELU-TRAN), that is, a computable general equilibrium model of a metropolitan economy developed by Anas and coauthors (Anas and Liu 2007) that studies modal choices but accounts for product, land, and labor markets. Anas and Hiramatsu (2012) have used the model RELU-TRAN to simulate the impact of a gas price increase on trips and transport modes in the short run, and location patterns in the long run.

In the foregoing, I've stressed the similarities between the approaches followed by geographers and economists, and thus the resulting potential for interaction. On the other hand, the operational models above epitomize the differences between urban geography and urban economics: the role of markets is meager in urban geography and the concept of equilibrium is not in the picture. Note, however, that Batty concludes chapter 9 with an overview of some ongoing extensions of his models. Interestingly, they seem more in line with what Anas does.

In part III, the focus shifts "rather dramatically from understanding cities to their design" (p. 301). A priori, this bears some resemblance to what is known in urban economics as the optimal city. However, the aim here is different: Batty wants to take into account the ways local governments intervene in cities. To achieve his goal, he uses to the same tools as in parts I and II, which makes the whole book very coherent from the methodological viewpoint. But it might have been more interesting for the reader if what we have learned from urban public finance (Glaeser 2012) had been incorporated. The material presented in part III might be more useful to urban planners than, in my opinion, to urban economists. This is why I have chosen not to discuss this part but to concentrate in detail on parts I and II, which address issues related to what we, as economists, do.

9. *Concluding Remarks*

By focusing on urban form, networks, and flows, Batty has made a substantial contribution to the science of city design and planning. His book also contains many concepts, ideas, and results that should be of interest to urban and trade economists. Although urban economists, very much like urban geographers such as Batty himself, aim to develop a bottom–up approach to cities, the two approaches differ. The main source of disagreement appears to lie in the methodological foundations of these two sciences of cities. As a microeconomist, I believe that macroscopic phenomena can best be understood by examining how they arise from the motivated actions and interactions of individual agents in a world where various types of uncertainty prevail. Batty's approach treats cities as entities formed by purposeless agents whose aggregate behavior is governed by stochastic laws. By contrast, the approach of urban economics and the new economic geography treats cities as emerging from the decisions of many rational agents—firms, households, developers, and local governments—whose actions are often coordinated by markets. However, as seen above, I do not see this difference as an insurmountable obstacle to interactions between urban geographers and urban economists.

References

Anas, Alex. 1983. "Discrete Choice Theory, Information Theory and the Multinomial Logit and Gravity Models." *Transportation Research Part B: Methodological* 17 (1): 13–23.

Anas, Alex, and Tomoru Hiramatsu. 2012. "The Effect of the Price of Gasoline on the Urban Economy: From Route Choice to General Equilibrium." *Transportation*

Research Part A: Policy and Practice 46 (6): 855–73.

Anas, Alex, and Yu Liu. 2007. "A Regional Economy, Land Use, and Transportation Model (RELUTRAN©): Formulation, Algorithm Design, and Testing." *Journal of Regional Science* 47 (3): 415–55.

Anderson, James E., and Eric van Wincoop. 2004. "Trade Costs." *Journal of Economic Literature* 42 (3): 691–751.

Anderson, Simon P., André de Palma, and Jacques-François Thisse. 1992. *Discrete Choice Theory of Product Differentiation*. Cambridge, Mass. and London: MIT Press.

Batty, Michael. 2005. *Cities and Complexity: Understanding Cities with Cellular Automata, Agent-Based Models, and Fractals*. Cambridge, Mass. and London: MIT Press.

Batty, Michael. 2008. "The Size, Scale, and Shape of Cities." *Science* 319 (5864): 769–71.

Beckmann, Martin J. 1976. "Spatial Equilibrium in the Dispersed City." In *Mathematical Land Use Theory*, edited by Yorgo Papgeorgiou, 117–25. Lexington: Lexington Books.

Bettencourt, Luís M. A., José Lobo, Dirk Helbing, Christian Kühnert, and Geoffrey B. West. 2007. "Growth, Innovation, Scaling, and the Pace of Life in Cities." *Proceedings of the National Academy of Sciences* 104 (17): 7301–06.

Caruso, Geoffrey, Dominique Peeters, Jean Cavailhès, and Mark Rounsevell. 2007. "Spatial Configurations in a Periurban City: A Cellular Automata-Based Microeconomic Model." *Regional Science and Urban Economics* 37 (5): 542–67.

Dorfman, Robert, Paul A. Samuelson, and Robert M. Solow. 1958. *Linear Programming and Economic Analysis*. New York: McGraw-Hill.

Fujita, Masahisa. 1989. *Urban Economic Theory: Land Use and City Size*. Cambridge and New York: Cambridge University Press.

Fujita, Masahisa, Paul Krugman, and Tomoya Mori. 1999. "On the Evolution of Hierarchical Urban Systems." *European Economic Review* 43 (2): 209–51.

Fujita, Masahisa, and Hideaki Ogawa. 1982. "Multiple Equilibria and Structural Transition of Non-monocentric Urban Configurations." *Regional Science and Urban Economics* 12 (2): 161–96.

Glaeser, Edward L. 2011. *Triumph of the City*. London: Macmillan.

Glaeser, Edward L. 2012. "Urban Public Finance." National Bureau of Economic Research Working Paper 18244.

Head, Keith, and Thierry Mayer. 2014. "Gravity Equations: Workhorse, Toolkit, and Cookbook." In *Handbook of International Economics, Volume 4*, edited by Gita Gopinath, Elhanan Helpman, and Kenneth Rogoff, 131–195. Oxford and Amsterdam: Elsevier, North-Holland.

Helsley, Robert W., and Yves Zenou. 2014. "Social Networks and Interactions in Cities." *Journal of Economic Theory* 150: 426–66.

Henderson, J. Vernon. 1988. *Urban Development: Theory, Fact and Illusion*. Oxford and New York: Oxford University Press.

Labbé, Martine, Dominique Peeters, and Jacques-François Thisse. 1995. "Location on Networks." In *Handbooks in Operations Research and Management Science, Volume 8: Network Routing*, edited by M. O. Ball, T. L. Magnanti, C. L. Monma, and G. L. Nemhauser, 551–624. Amsterdam: North-Holland.

McFadden, Daniel. 1974. "Conditional Logit Analysis of Qualitative Choice Behavior." In *Frontiers in Econometrics*, edited by Paul Zarembka, 105–42. New York: Academic Press.

Ogawa, Hideaki, and Masahisa Fujita. 1980. "Equilibrium Land Use Patterns in a Nonmonocentric City." *Journal of Regional Science* 20 (4): 455–75.

Samuelson, Paul A. 1952. "Spatial Price Equilibrium and Linear Programming." *American Economic Review* 42 (3): 283–303.

Wilson, A. G. 1970. *Entropy in Urban and Regional Modelling*. London: Pion.