

UC Berkeley

Earlier Faculty Research

Title

Geographical Theories

Permalink

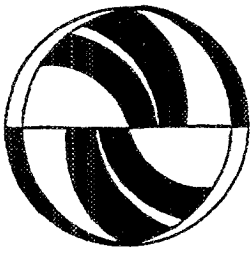
<https://escholarship.org/uc/item/1vr640bw>

Author

Golledge, Reginald G.

Publication Date

2001



Geographical Theories

Reginald G. Golledge

Reprint
UCTC No 471

The University of California
Transportation Center
University of California
Berkeley, CA 94720

The University of California
Transportation Center

The University of California Transportation Center (UCTC) is one of ten regional units mandated by Congress and established in Fall 1988 to support research, education, and training in surface transportation. The UC Center serves federal Region IX and is supported by matching grants from the U.S. Department of Transportation, the California Department of Transportation (Caltrans), and the University.

Based on the Berkeley Campus, UCTC draws upon existing capabilities and resources of the Institutes of Transportation Studies at Berkeley, Davis, Irvine, and Los Angeles, the Institute of Urban and Regional Development at Berkeley, and several academic departments at the Berkeley, Davis, Irvine, and Los Angeles campuses. Faculty and students on other University of California campuses may participate in

Center activities. Researchers at other universities within the region also have opportunities to collaborate with UC faculty on selected studies.

UCTC's educational and research programs are focused on strategic planning for improving metropolitan accessibility, with emphasis on the special conditions in Region IX. Particular attention is directed to strategies for using transportation as an instrument of economic development, while also accommodating to the region's persistent expansion and while maintaining and enhancing the quality of life there.

The Center distributes reports on its research in working papers, monographs, and in reprints of published articles. It also publishes *Access*, a magazine presenting summaries of selected studies. For a list of publications in print, write to the address below.



University of California
Transportation Center

108 Naval Architecture Building
Berkeley, California 94720
Tel 510/643-7378
FAX: 510/643-5456

DISCLAIMER

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the Department of Transportation, University Transportation Centers Program, in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof.

The contents of this report reflect the views of the author who is responsible for the facts and accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the State of California or the U.S. Department of Transportation. This report does not constitute a standard, specification, or regulation.

Geographical Theories

Reginald G. Golledge

Department of Geography
University of California
Santa Barbara, CA 93106

Reprinted from

International Social Science Journal, Geography State of the Art I – The Environmental Dimension
UNESCO, no 150, pp 461-476 (1996)

UCTC No. 471

The University of California Transportation Center
University of California at Berkeley

Geographical theories

Reginald G. Golledge

Why geographical theory?

The emergence of geographical theory was an inevitable product of the desire to systematize existing geographic knowledge and to use that systematized base to explore new areas of knowledge. Although the usefulness of theory and predictive models in geography is by now a matter of record, it was not always the case. The usefulness and need for theories was often disputed, despite the oft-repeated argument that theories of location explained the laws of spatial distributions, theories of interaction explain the laws of movement and spatial behaviour, theories of growth and development explain the nature of past, present, and future states of being, and theories of decision-making and choice explain observable regularities and repeatable trends in individual, group, institutional and governmental behaviours. Hudson (1969) argued that a theory represents a direct attempt to provide a logical system or nesting place for previously noted regularities – in his case concerning changes in rural settlement patterns. While Hudson's task was specific, the sentiment he expressed has widespread relevance for the emergence and adoption of geographical theories generally.

Within the realm of geographical theory, two singular paths were followed. One focused

Reginald G Golledge is Professor of Geography at the University of California Santa Barbara, Santa Barbara, California 93106, USA. He is interested in human spatial behaviour, knowledge acquisition, and cognition, as well as cognitive mapping, disaggregate transport modelling and geography and disability. His most recent co-authored book was *Behavior and Environment*. Forthcoming co-authored books include *Spatial Decision Making and Choice Behavior* and *Spatial and Temporal Reasoning in Geographic Information Systems*.

on 'form', or the patterns or peculiar spatial configurations that are distributed in a well defined area. As Amedeo and Golledge (1975) state, 'Reference to the configuration of a distribution would presumably give a *spatial* summary indication of the dispersion of events with respect to some designated area, the arrangement of events with respect to one another, the connection or links between and among events and the hierarchical relations inherent in the distribution of concern' (p. 176). On the other hand, the term 'process' was defined as being '... a collection of interrelated activities that operate on a set of events and, consequently, produce (or in some cases, prevent) changes in the characteristics of these events through time and over space. When we speak of processes, we imply, and normally expect, that the "behavior" of their activities generates relationships and interrelationships that are changing over time. Hence, the manifestations of these relationships and interrelationships at any moment in time may have an appearance that exists only temporarily, sometime later, these manifestations may take on an altogether different appearance' (p. 177). One of the more significant characteristics of the theoretical and quantitative revolution was a gradual shift from the early emphasis on deterministic theories of form and structure, to process theories of interactions,

interrelationships, behaviour and dynamics

Most geographers refer to the disciplinary upheaval of the late 1950s and 1960s as the 'quantitative revolution'. It was during this period that a significant part of human geography began to emphasize analysis and interpretation of data via mathematical and statistical models and methods. Such approaches had been used more extensively in physical geography where the rigours of scientific experimentation, careful experimental design, accurate measurement and explanatory model-building had been more widespread. Such was not the case in human geography. Much of human geography was descriptive rather than analytic, interpretive rather than logical, and dependent on the judgmental 'expertise' of persons claiming specific knowledge of systematic areas or of regional systems.

But the quantitative revolution was much more than the simple introduction of mathematics and statistics into human geography. Parallel to it, and in some senses preceding it, was a *theoretical revolution*, and mathematics and statistics were merely the necessary languages that allowed this theoretical revolution to emerge and develop. For more than a decade, however, the theoretical and quantitative revolutions were so intimately entwined that few were able to differentiate between them. In the eyes of many geographers, introducing theory and introducing measurement, mathematical modelling and statistical analysis were identical. To do this meant wholesale philosophical reorientation in the discipline, as positivist philosophy and scientific method were seen to be the primary mechanisms by which theory and quantitative methods could be introduced and spread.

As with any innovation there were innovators who went beyond the bounds of the discipline to search for the training that was required to encourage this new way of thinking. Such excursions were documented by Gould (1985), Johnston (1979), Gale and Olsson (1979), Amedeo and Golledge (1975) and others. It is not the purpose here to trace historically the introduction of theory and quantitative methods in the discipline but rather to try to understand why this happened, how it happened and what areas of the discipline were most impacted.

Descriptive and normative approaches

Providing an exact explanation or an exact prediction for an event or events that happened in the real world requires an extremely complex and complete system of reasoning. Rarely are such exact explanations or predictions provided. Instead we tend to make approximations. We have different degrees of confidence in these approximations. Using them we may be able to achieve some accurate predictions and indeed some acceptable explanations. In the geographic domain, little success has been achieved in formulating well defined theories or models of spatial structure that allow exact predictions and explanations of real world phenomena to take place. Both deterministic and stochastic inferential processes have been used in the attempt to develop geographic theories and models to provide appropriate levels of satisfaction.

Theories and the models that represent them are usually presented in a logical form. Throughout the early period of development of geographic theory, so little was known about the general influence of geographic space on human activity and so little information was formally structured concerning the patterns and distributions of phenomena over the Earth's surface and the interactions between them, that the idea of producing reliable and valid geographic theory with exact predictive or explanatory capabilities was a hopeless dream. However, geographers began the long and arduous task of moving in the direction of this distant goal. Like many other disciplines they initially embraced the idea that one can simplify the real world with an appropriately selected set of assumptions about the environment, about the people and about human-environment relations. Similarly, since relatively little was known about the stochastic nature of geographic events or about human behaviour in space, the initial step taken by many geographers was to embrace deterministic rather than stochastic inferential processes. In this way they were able to begin moving towards their distant explanatory goals by developing theory and models that described conditions as they might be or ought to be rather than as they were. Thus, while some geographers were content with descriptive models of change (such as detailed analysis of

the nature of growth and change within spatial systems), others adopted a more rigorous but, to many, less satisfying normative mode. The descriptive modellers made little attempt to abstract from reality and were content to analyse and interpret empirical data. As Amedeo and Golledge (1975) put it, '... they were in fact concretely tied to actual events in the real world' (p. 290). On the other hand, those choosing to work in normative worlds sought to describe in a more general way the variations in spatial structure that *could* conceivably exist (or perhaps even should exist) if specific constraints were adhered to. The value of this latter approach lay in the fact that it was possible to work with a well structured logical system, the method of inference was deduction and successive waves of information could be obtained by unravelling the effects of changing assumptions or by linking heretofore unlinked axioms. In a real sense any new output represented only the logical outcome of individual and joint consideration of the basic axioms and propositions that formed the base for a theory or its model. Given this type of an approach, the worth of the consequent theory and model must not be estimated by attempting to compare them with real world situations, unless of course the theory or model is so complete that the real world situation closely approximates to the conditions under which the theory or model was constructed. Rather, the efficacy of the theory or model built in a normative framework lay in the accuracy of its logical structure and the legitimacy and validity of the inferences made from that structure. The bulk of these normative theories were developed in the context of economic geography. It is not surprising, therefore, that these theories focus on the spatial manifestations of sets of economic premises such as least cost, maximum profit, level of competition, economic rationality, scale economies, and so on.

The influence of economics

Agricultural land use theory

Most geographers will agree that the greatest theoretical impact on human geography came as the result of an increasing fascination with the spatial implications of various economic

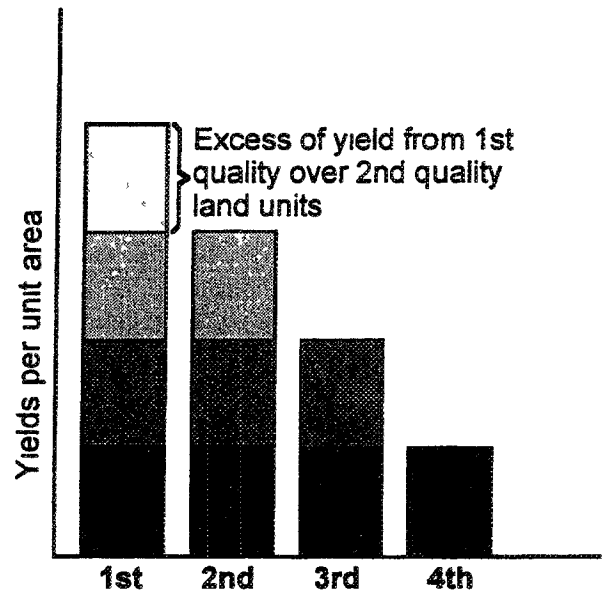


FIGURE 1 Land rent and land quality

theories. Perhaps the earliest influence derived from the work of German spatial economists. Certainly Heinrich von Thunen (translated 1966) and Alfred Weber (1909) have become household names to virtually every geography student who takes an introductory course in human geography or an advanced course in economic geography.

The pervasive influence of space was for the most part ignored in classical economic theory. For example, in formulating a theory of land rent, Ricardo (1817) argued that as soil fertility changed, productivity levels would also change, thus ensuring that land of the highest fertility or quality would provide more output per unit area than land of inferior quality. This difference was termed 'rent' and was interpreted as a return to the specific characteristics of the factor land (Figure 1). But Von Thunen wondered if there were other deep-seated and not so obvious factors influencing how land uses developed and were distributed in space. He experimented on his family holdings at Tellow near Rostock in Germany. He found that some land uses required intensive rather than extensive cultivation, that some products were required at market in a fresh and unspoiled manner, that other products could be transported longer distances without any deterioration in

their quality or usefulness, and that the longer distance a product had to be transported then the greater part transportation costs played in the total cost picture. Von Thunen thus established the importance of the tyranny of distance in the process of production and movement of agricultural and pastoral land products to market. He argued that even if the land was uniformly homogeneous with respect to climates, soil fertility and slope, then distance from market would still mitigate the type of land use that could economically be undertaken. Thus he developed his famous Concentric Zonal Theory of land uses and land rents.

One can assume that most land areas are suitable for a variety of uses and that for the most part uses have the ability to produce a positive net return on many different land units. It would simplify things if land owners used their resources optimally (i.e. for the purpose that gave the highest net return). This would allow a principle of profit maximization to be used as the basis for building a comprehensive theory of agricultural land uses. Such a principle could be termed the principle of 'best use' and the land use that gave that best return would be said to have the 'highest use capacity' of the land unit. However, for many land units to achieve a net return, differential quantities of the input factors such as labour, capital and technology must be used in the attempt to obtain a net return. Consequently it is a simple inferential matter to deduce that at some locations some land uses are likely to give a higher return because of the particular combination of environmental, economic and human factors that are available to be input at that site. Recognizing, therefore, that the tyranny of space is likely to encourage different land uses to locate in different places, one can assume that some type of normative theory which describes the best possible or ideal distribution of land uses given certain restraining conditions or assumptions, would be a valuable tool for comprehending the geographic nature of land use patterns. McCarty and Lindberg (1966) graphically summarized the idea of best use in an environmental sense in terms of an optima and limits diagram (Figure 2). Even just considering simple environmental constraints it is obvious that there will be optimal places which require minimal factor input in order to produce returns

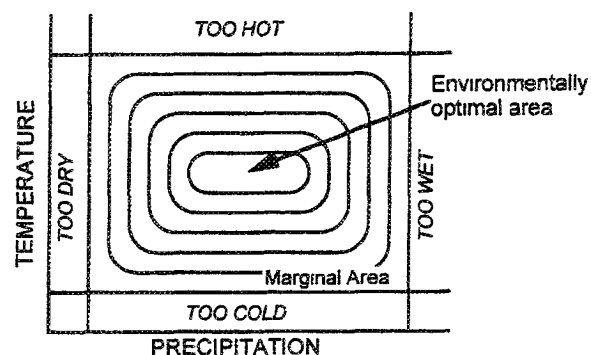


FIGURE 2 Optima and limits for crop-growing. Source: Modified from H. H. McCarty and J. B. Lindberg, 1966.

for certain land uses, and there will be limits beyond which it would be impractical to pursue those land uses unless the need was great and the resources to do so were readily available.

Next, considering that some land uses can be undertaken at a greater level of intensity than others (e.g., a skyscraper can be built on a half-acre of land or it may take that same half-acre to produce one unit of beef or ten sheep), one might naturally expect that land uses will enter into competition with each other to determine which use occupies a given location. If we assume that the use likely to occupy a land unit will be the one for whom the land has the greatest value, and in turn the one which can consequently bid the highest for rights to occupy the land unit, we could infer that land uses will generate families of bid rent curves for different land sites based on their value and need. One can further deduce that as one use achieved a capacity to outbid others, it would therefore occupy a land area. Thus using the notion of bid rents, a simple schema of land use gradient can be developed (Figure 3). It was on the basis of this idea that land use types would change as the marginal rent changed that Von Thunen built his model. It required the following normative (or variability reducing) assumptions:

- 1 A state isolated from other areas by an impenetrable barrier or wilderness,
- 2 A city acting as the primary market located at the centre of the state,
- 3 One agricultural hinterland surrounding that city,

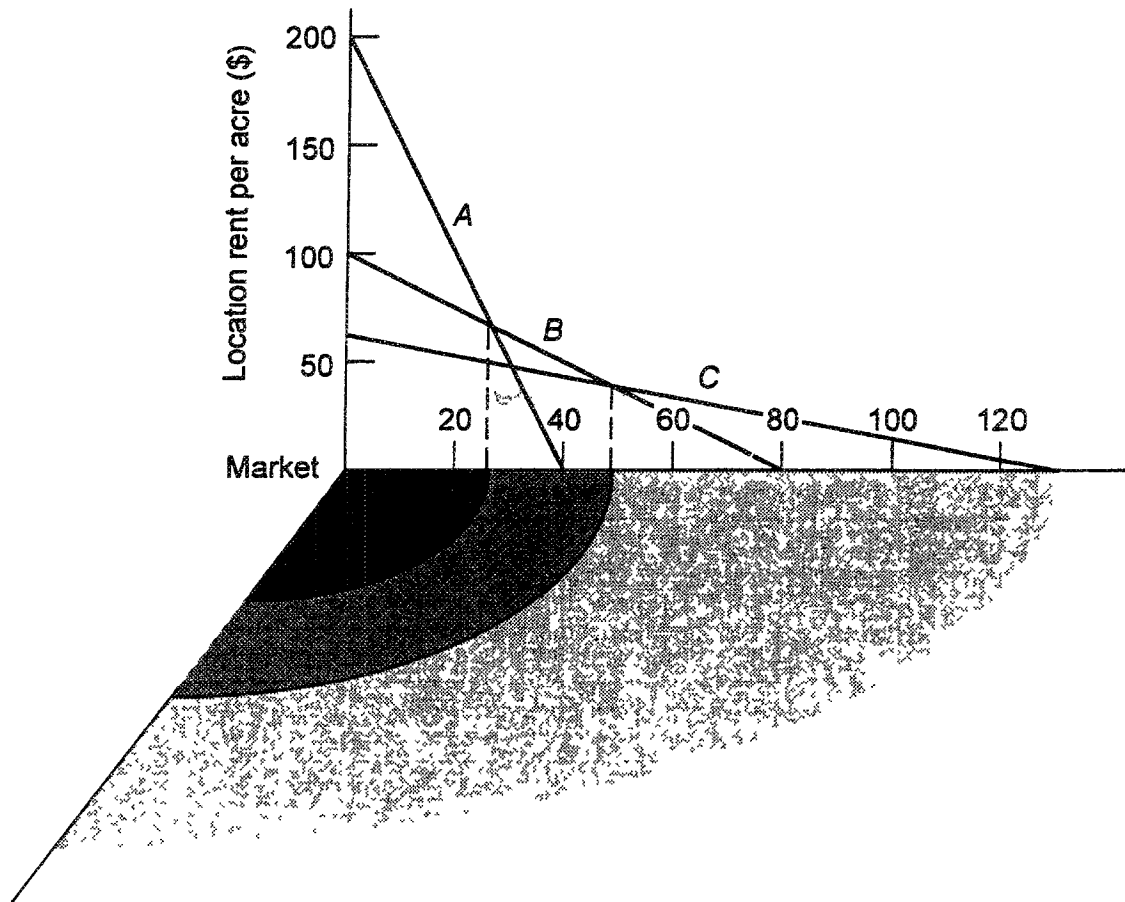


FIGURE 3 Land use zones from bid rent curves

- 4 All agricultural products are exchanged in that city,
- 5 No other area provides agricultural products for that city,
- 6 The city's hinterland sends no products to other cities;
- 7 Farmers aim to maximize profits.
- 8 Farmers are capable of adjusting agricultural operations to maximize their profits,
- 9 The hinterland in the isolated state consists of a uniform plane with respect to soil, topography, rainfall, temperature, sunny days and other environmental conditions;
- 10 The only means of transportation, horse-drawn wagon, is available equally in all directions, non navigable rivers or canals exist;
- 11 Transportation is available equally in all directions,
- 12 Transportation costs vary directly with distance, are the same in all directions for each commodity, but differ from commodity to commodity because of the inherent nature of the commodities themselves (e.g., fragility, durability, perishability).
- 13 Transportation costs are borne by the farmer;
- 14 The producer delivers his products fresh to the market;
- 15 Price at the market and production costs are the same for all farmers regardless of their location in the isolated state,
- 16 Market prices differ for different commodities

The question pursued by Von Thunen was how will agricultural production arrange itself spatially under these circumstances? Given these preliminary assumptions within this iso-

lated state, Von Thunen assumed that land use zones will emerge and that the types and quantities of products demanded in the city and the consequent prices that commodities generated would influence the total volume of production of each type of land use. Under these assumptions and using his marginalist theory (Leigh, 1946), Von Thunen deduced the following results:

- 1 Land rent for a particular product varies inversely with distance from the city, reaching zero where marginal costs and marginal revenue for a product are equal,
- 2 The product with the highest yield per acre is produced closest to the market,
- 3 Intensity of cultivation decreases with distance from the market,
- 4 Perishability of products decreases with increasing distance from market,
- 5 Price of land varies directly with its rent yielding capability,
- 6 Crops which have a considerable weight or bulk in proportion to their value or crops which require high transportation costs, will be produced close to the market,
- 7 The greater the distance from the city the more likely that land users occupying those locations will require lower transportation costs per unit per mile,
- 8 Recognizable zones of different types of land use will develop around the city,
- 9 Because of the different lifestyles associated with different forms of production, the nature of economic life will also vary distinctly with increasing distance from the city,
- 10 Because of the initial environmental uniformity assumptions in the isolated state, land users will be arranged in concentric circles around the central city (Figure 4)

This descriptive model was, in 1954, operationalized in the form of a mathematical model by E. S. Dunn, Jr. (1954). Dunn concurred with Von Thunen's basic ideas when considering product type and isolation. He formalized the Von Thunen model in a functional relationship expressed as follows.

$$R = E(p - a) - Efk$$

where R = rent per unit of land;

E = yield per unit of land,

p = market price per unit of commodity;

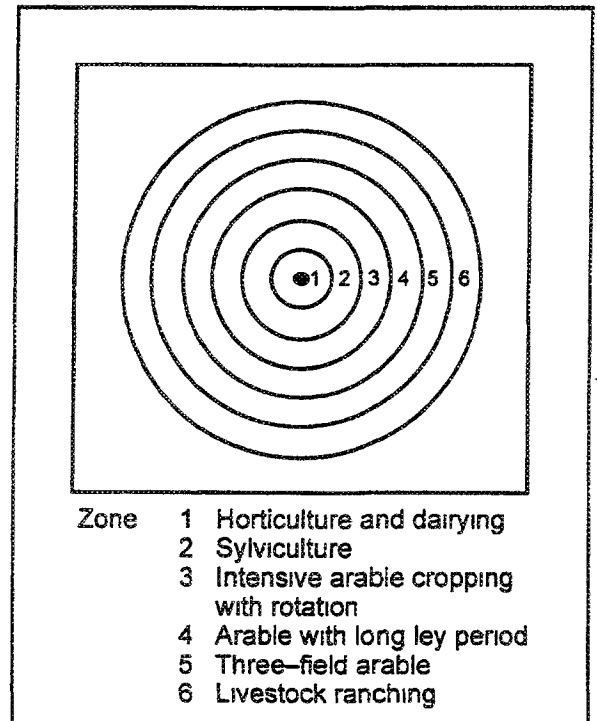


FIGURE 4 Von Thunen's land use zones

a = production cost per unit of commodity,
 f = transport rate per unit of distance for each commodity, and
 k = distance from market.

This summarizes a one-product linear model in which a given land use is capable of producing a net return out to that point where its marginal revenue and marginal costs are equal. Here receipts will be maximized at a site as close to the market as possible and will be minimized at that distance where marginal returns are zero.

However, when considering competitive products, Dunn showed that a variety of situations might occur. For example, given two products, where Product A always produced a higher rent than Product B (Figure 5), Product A will be grown exclusively and Product B will not be cultivated. However, where land uses have different intensities, the more usual case is that the bid rent curves will cross. At this point the returns from either land use are equal. Up to that point Use A will be cultivated and beyond that point Use B would give a higher rate of return (Figure 6). Thus, Dunn modified

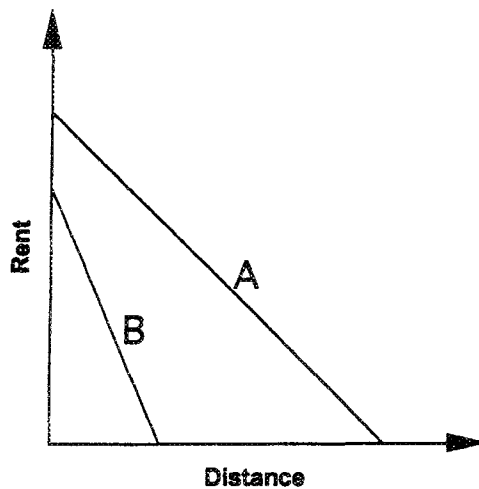


FIGURE 5 Land use competition effect on rent dominance

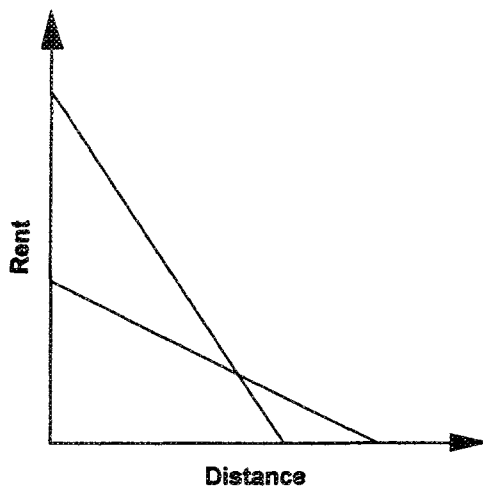


FIGURE 6 Rent lines for two products

the Von Thunen marginalist theory in the presence of competition to include the idea of land use transference not at the theoretical margin of equality between receipts and costs, but at that point when the returns from one use are equalled or exceeded by the returns from a different use. In a simplified environment similar to that constructed by Von Thunen, the end product is basically similar. Land use zones will develop as concentric circles around the single market centre and all the conclusions regarding intensity of cultivation, transportability and

value deduced by Von Thunen will similarly hold for the competitive case.

Geographers gradually became aware of this model, regarding it as a norm for interpreting land use patterns on a national scale (particularly when a country or a region was dominated by a single large market), they also investigated the potential of the theory for looking at the gradient of land uses around individual cities. As they expanded the frontiers of their knowledge base to include land resource economics and agricultural economics, the Von Thunen ideas were given more prominence. They were given more prominence because this theory was explicitly spatial. It was a theory that established without doubt the importance of geographic space even when many of the normal environmental features such as slope, soil fertility, climate, temperature and rivers were assumed away. The Von Thunen model as it was first introduced and used in geography was largely descriptive. It was not until the 'quantitative revolution' that researchers with the necessary mathematical and model-building background were able to examine the full range of geographic implications of this theory. The end result is a spatial theory, a theory that established a norm for capitalist societies and a point of departure for examining agricultural practices in developing economies or in non-capitalist economies.

Industrial location theory

As distinct from agricultural location theory (which was a topic dear to the heart of many geographers who were concerned with describing and explaining the pattern of human activities across the surface of the earth), a more specialized and systematic location theory evolved from the work of Alfred Weber (1909). Weber too was interested in the tyranny of distance and the role that distance played in location decision processes. As a spatial economist, however, distance to him had to be translated into a cost unit, and his choice was transportation costs. Weber then proceeded to build a series of models that helped to explain what particular combination of regional and local factors influenced the locational decision processes of different industries. Let us briefly examine this particular theory.

In developing theories of industrial location two dominant lines of thought have been perceived the first is that decision-makers choose locations at which costs are minimized, while the second argues that locations will be chosen where profits are maximized. We begin this summary of the original industrial location theory as it was imported into geography by focusing on the least cost approach.

By definition, manufacturing performs some transformation operation on a set of raw materials. These raw materials can be natural resources or the final product of another industry. For most industries more than one raw material is used. To geographers it was obvious that raw materials were not uniformly available at all locations. Likewise it was obvious that potential markets were distributed differentially in space. The questions raised by location theorist Alfred Weber (1909) included the following: To what extent is the attraction of raw materials influential in the location decision-making process? To what extent is the physical composition of the materials or the nature of the process using those materials important in the location decision-making process? How does the influence of raw materials change as technologies of use and distribution change? Do the raw materials used in the production process lose or gain weight, bulk, or volume during manufacture? Is the raw material perishable? How does the value per unit weight of raw material influence the location decision-making process? Are substitute raw materials readily available? In what proportion are different materials used in the manufacturing process? And finally, how will the answers to these questions help decide on a geographic location for a given industry?

Weber's reasons for asking some of these questions are straightforward. If a bulky or weighty material can be reduced in bulk or weight by a preliminary manufacturing process, then some savings of cost will be obtained by locating at the material site. This saves payment for the transfer of waste product. If the raw material is perishable but is made more durable during manufacture, economic losses may be avoided by doing this transformation at or near a raw material site (e.g., the production of cheese or butter or vegetable canning). If the value of a raw material is high relative to its

weight or volume it can be transported long distances and the need for a raw material orientation is reduced (e.g., wool). If the value of raw material is low in its initial state (e.g., copper ore) it cannot be transported too far before transport costs exceed the value of the materials.

Products gain value in the course of manufacturing. The higher this added value, the more likely it is that the manufactured product can absorb the transportation costs. However, when the sum total of raw materials used to make a product greatly increase the weight, volume or bulk of a product, there is a tendency to achieve cost minimization by locating at the market and bringing the raw materials to that site (e.g., beverage production). These principles are summarized by Weber into two major hypotheses: the Weight Loss Hypothesis and the Weight Gain Hypothesis. In simple cases where one could assume a linear system with a single raw material and a single market, then location was obviously either at the raw material, the market, or at some other location (usually between the two). If the raw material was weight-losing then it tended to attract the industry, if it was weight-gaining then the market site proved to be more attractive. If one added alternative sites of labour between the two initial locations, then industries having a very high labour cost component in their total cost structure might move to one of the intermediate labour points if labour costs were low enough.

One of the principal reasons for the development of a body of industrial location theory apart from classical economic theory was the failure of the latter to consider the effects of varying geographic factors (e.g., raw material locations), spatial interactions (e.g., as expressed through transportation costs), and behavioural variability (in spatial choice and decision-making). By adding the spatial or geographic domain to classical economic theory, it is possible to explain why some locations were chosen rather than others, why agglomeration at certain locations occurs and not at others, why occasional mass migrations of labour occur, why human settlement patterns adopt the form they do, and why some regions are comparatively richer, better off and faster developing than other regions. As it developed, location theory in its various forms had the goal of

helping to explain the impact of space on social and economic activity

As one moves to more complex cases with two raw materials and one market, or one raw material and two markets, or as other factors of production become more significant in the total process of manufacturing, then they exert an attractive influence in the locational decision-making process. Some manufacturing can be considered to be oriented to sites of cheap power (e.g., aluminium), some to cheap labour (e.g., textiles and fabrics), some to indentured labour (e.g., garment-making)

In his theory, Weber took account of geography and environmental variations by focusing on transportation costs. Differences in locational advantages were thus expressed in terms of the differences in the total transportation costs associated with assembling materials or distributing products. Usually the costs of processing materials were seen to be constant at all locations, while some factors were considered to be ubiquitous (i.e., things available in reasonable abundance at all locations at approximately the same cost, e.g., air, water, clay)

Geographic location theory

Weberian theory stimulated considerable geographic and regional science research into the industrial locational decision-making process. Within geography, an initial attempt was made by behavioural geographers to examine decision-making traces with respect to firms and industries, and to develop typical profiles of decision-making that helped explain why some locations are chosen more than others. An alternative analytical approach, however, developed in the mid- to late 1960s (Teitz and Bart, 1968), which established one of the most productive developments in geographic theory – the expansion of simple locational theory to location/allocation theory. Working initially with network models and maximizing access of network regions to selected network nodes, this location/allocation theory rapidly developed as an extremely powerful tool capable of handling a wide variety of locational problems. Typical examples of problems solved using location/allocation procedures included the location of emergency facilities (e.g., police/fire), location

of schools, location of health facilities (e.g., hospitals, medical centres), location of urban parks and recreational areas, and so on. The basic format for these location/allocation models was expressed by Church and ReVelle (1974), and sets of location-allocation computer algorithms were distributed by Rushton, Goodchild and Ostresh (1973). The essential form of a basic set covering location-allocation model is as follows

$$\begin{aligned} \text{Minimize} \quad & \sum_{i=1}^n \sum_{j=1}^m I_{ij} C_{ij} \\ \text{Subject to} \quad & \sum_{j=1}^m I_{ij} = S_i \\ & \sum_{i=1}^n I_{ij} = D_j \end{aligned}$$

where

I_{ij} = unknown allocation from source i to demand j

C_{ij} = known cost per unit allocated (i.e., distance surrogate)

S_i = known supply available at source i (capacity)

D_j = known demand at point j (population)

Given X_i, Y_i as *unknown source* locations and U_j, V_j as *known demand* locations then cost is equated with linear transportation

$$C_{ij} = \sqrt{(X_i - U_j)^2 + (Y_i - V_j)^2}$$

[i.e., given a destination of demand, the sources with known capacities, the aim of the model is to locate sources so that total cost of transportation required to satisfy the known demand is minimized.]

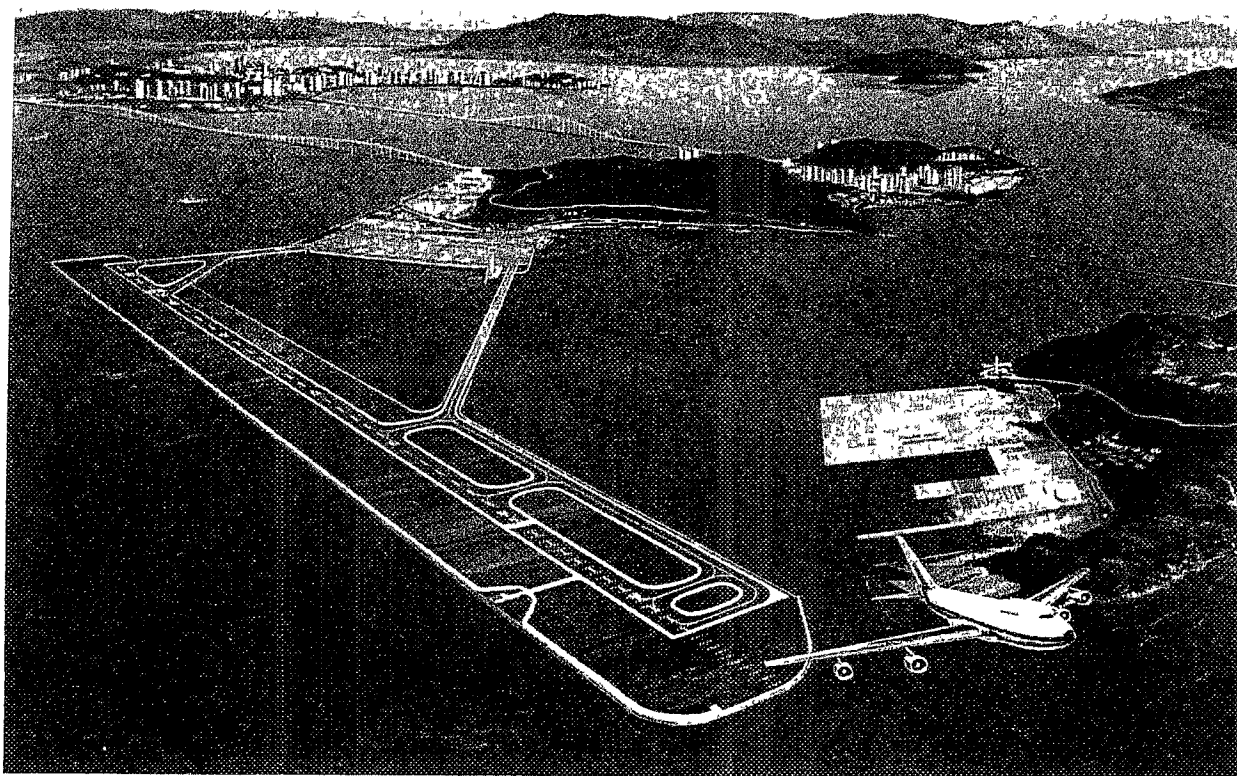
If the first constraint is dropped the problem is an 'unconstrained location-allocation problem'

If the locations of sources are known and fixed, this is the simple transport problem

If there is only one source, then it equals the Weber problem

In general, therefore, the location-allocation theory applies whenever centres have to be located so that each centre serves a prescribed number of people (e.g., school locations).

In general, the following steps are required:



Computer projection of a proposed new international airport onto a photograph of Macao Luc Olivier/Gamma

- 1 Initial sources are supplied by the user or generated by the programme,
- 2 Allocation is made to these locations;
- 3 Sources are relocated with respect to demand points,
- 4 Repeat the above (Cooper algorithm).

Each step consists of a solution to the transportation problem. Allocations are made from sources to demand points so as to minimize transport cost with sources at their existing temporary locations

An example of process theory: the marketing process

The marketing processes can be said to consist of three sets of interrelated activities which take place through time and space. The three activities identified in classical economic theory are production, consumption and exchange. Sets of actions associated with these activities include the following.

- (a) Consumers express demands for goods and services and support their demands with purchasing power,
- (b) The purchasing behaviour of consumers over time establishes demand schedules for different products;
- (c) Entrepreneurs service consumer demands by establishing firms at locations convenient to consumers,
- (d) The interaction of decision processes of entrepreneurs and consumers identify prices for different products or services,
- (e) As entrepreneurs establish their prices, consumers become aware of the possibility for different prices to exist at different geographic locations,
- (f) Consumers explore the distribution of price offerings and evaluate the utility or disutility associated with interacting with specific locations in order to purchase a good
- (g) Entrepreneurs begin competing with each other spatially and otherwise for consumer purchases,
- (h) Consumers and entrepreneurs act in their own best interests,

- (i) Different locations at which products are offered become market centres and establish trade areas or market areas of different spatial extent,
- (j) Apart from location, entrepreneurs begin to differentiate their products in the minds of consumers in order to obtain differential shares of the market

This brief history of events involved in the marketing process allows one to begin systematically exploring the nature of the distribution of market centres and trading areas that could eventuate as key factors in the marketing process are manipulated

Given the above simple history one can see that the marketing process can be discussed without any consideration of space as long as some assumption is made that everyone is at 'the market'. This allows one to deduce the occurrence of demand and supply relationships, sets of prices and market structures. But no spatial distribution of any kind is generated. If we introduce the geographic idea that in any economic system individuals do not all exist 'at the market', but are distributed geographically over space, then the marketing process becomes more geographic in nature. It can be seen immediately that there is potential for developing different spatial arrangements and distributions of market centres. If we explore the effects of the introduction of space into this simple type of economic system, then geographic theory concerning the distribution of the arrangement, size and frequency, functional complexity and distance apart of the various market centres that make up a system, begins to fall out. During this inferential process one can draw on significant steps that had been undertaken somewhat in isolation in economic theory but which prove necessary for the development of strong geographic theory about the marketing process.

For example, Losch (1954) spatialized the idea of a demand curve by inferring that the delivered to customer price of a good consisted of a base price and an increment proportional to the transportation cost that a dispersed individual would have to absorb in order to get to a centre to purchase a product. Consequently, by rotating the axis of a standard demand and supply curve and using that part of the price

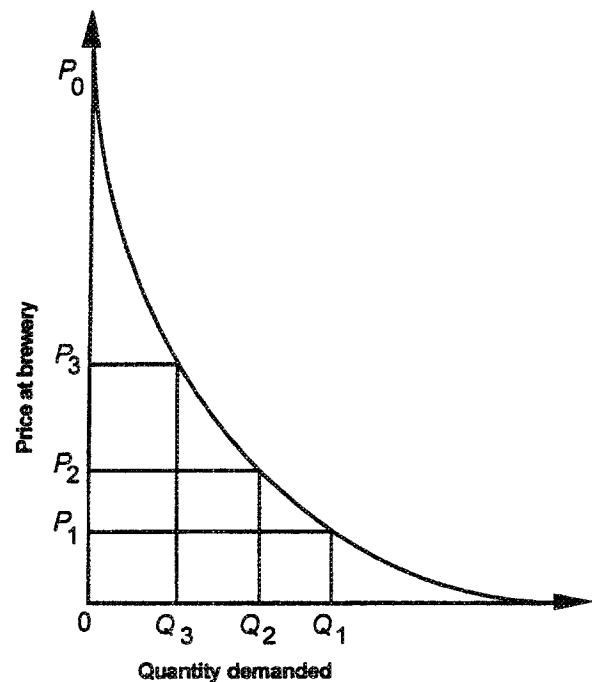


FIGURE 7 Quantity of beer demanded as a function of brewery price

axis above the base price as an indication of the impact that increased distance would have on the purchasing power (or quantity demanded) of potential consumers, Losch established a demand cone which identified the spatial extent of a potential market area at a given base price and given rate of transportation cost (Figure 7); this area would be circular in form. As prices change the size of this market is increased or decreased. Given that more than one entrepreneur existed in an area, a series of demand cones would emerge identifying the trade areas of viable entrepreneurs. Through the life and death process of economic survival which depended on covering costs and obtaining a normal profit, the circular market areas would fluctuate until a type of equilibrium or stable pattern emerged consisting of tightly packed tangent circles of equal size.

Let us now digress for a moment to yet another stream of market area analysis that focused more on competition. Hotelling (1929) had shown that in a given linear market, competing producers could change locations to give them an economic advantage over competitors. Under conditions of competition, an equilibrium

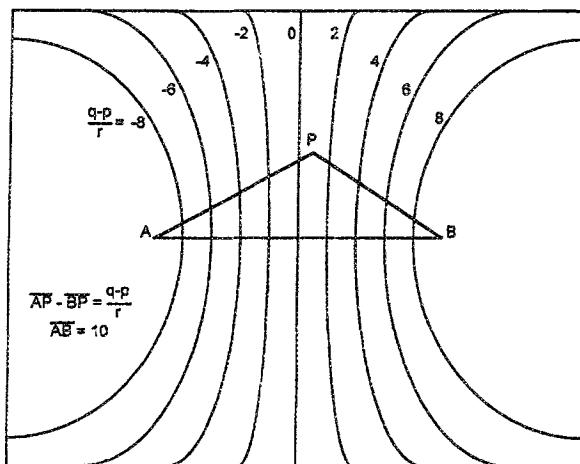


FIGURE 8 Changing market area boundaries *Source* After Hyson and Hyson, 1950

would be reached in a duopolistic market only when the competitors were located back to back in the centre of the market. Any deviation from this location would result in an advantage to one competitor. This, however, could inconvenience the more distant customers, who may have to travel excessive distances to purchase the product and would therefore be able only to purchase smaller amounts. Hotelling also showed, however, that acting in collusion, competitors could decentralize to the quartiles of the linear market and there maximize their profits. They would each gain access to exactly half the customers in the market and no consumer would have to travel more than one-quarter the length of the market in order to purchase a product. These ideas of an equilibrium point or point of indifference occurring between spatially distributed competitors in a market produced the notion of a market area boundary Fetter (1924) had already formulated an Economic Law of Market Areas for the simple case of duopolistic competition with uniform prices and transportation costs Hyson and Hyson (1950) showed that as entrepreneurs manipulated prices or made deals with transport companies for more favourable rates, the boundaries of market areas would shift They also showed such boundaries need not remain as linear segments but could degenerate to hyper-circles packed around the higher priced centre (Figure 8).

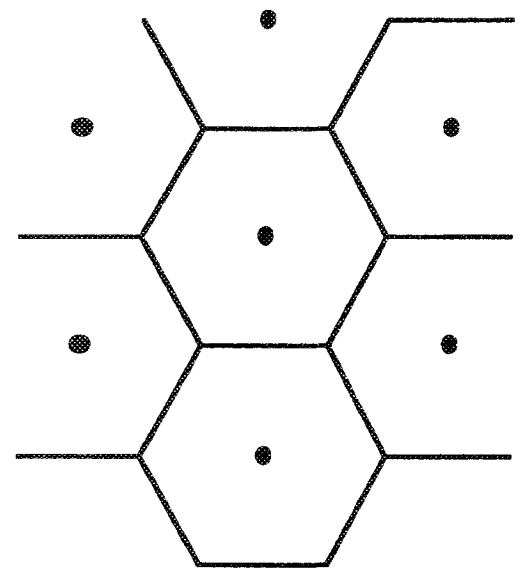


FIGURE 9 Non-overlapping hexagonal market areas

Given this parallel history, what happens to Losch's solidly packed circular trade areas? It must be obvious that tangent circles would produce areas of non-service, thus potentially encouraging instability in the system as bold new entrepreneurs established in these interstitial areas and attempted market penetration of surrounding areas. If it is assumed that all potential consumers in an area need to be served, then the circular packing principle produces an insufficient solution.

A German geographer, Walter Christaller (1933), had empirically developed a theory of the distribution of urban settlements which included a possible solution to these problems of instability and non-service. Working in a simplified environment with uniform topography and uniformly distributed populations of identical tastes and preferences, Christaller showed that Fetter type market area boundaries could be defined between each competing pair of centres As the distribution of entrepreneurs settled into an equilibrium, the irregular Thiessen polygons formed by the Fetter linear boundaries, settled into a uniformly shaped hexagonal market area (the famous Christaller $K = 3$ system) (Figure 9)

Thus, by combining form-related theories of the potential spatial distribution of market centres with the more process-oriented theories

derived from examining how economic systems work, a formal geographic theory emerged. 'Central Place Theory', as Christaller called it, has developed as perhaps the single most powerful and influential theory to have developed in geography. But, like Von Thunen's and Weber's initial theories, it is normative – it tells us what *should* be the case under certain types of constraining assumptions. And that is how many theories are presented. There is no reason to expect them to reflect the world exactly as it is at any particular time, for the world as it is rarely, if ever, matches the assumptions necessary to build the theory. However, as a normative principle, the theory and its spatial manifestation can be used to show departures from this norm. Identifying the reasons for such departures has continued to add significantly to our total understanding of the marketing process and its expression through a system of settlements.

Location of urban places

The results of pooling this literature on market area analysis and location under various types of competition together allowed first Christaller and then Losch to produce spatial theories of settlement systems. The properties of these systems are well described in many urban geography textbooks (e.g., King and Golledge, 1978; Cadwallader, 1985; Yeates and Garner, 1980; Berry, 1967). This settlement theory was hierarchically as well as spatially organized. Hierarchically one was able to deduce that as the functional nature of the system of settlement changed from primarily marketing to primarily transportation-oriented to primarily administration, then there were specific numerical relations between the number of centres at each level of the hierarchy and specific spatial relations concerning the distances apart of these centres. For example, in the simple marketing system ($K=3$), for every centre of one order there is a fixed number of centres at lower and higher orders. In the $K=3$ system, for every place at one level, there would be the equivalent of two *more* centres at the next lowest order. Remember that each centre incorporates functions of all lower levels and therefore acts as a centre at all those levels. Therefore, if there

was one largest centre, there would be the equivalent of two new lower order market areas plus the larger centre containing a second order market area, giving the equivalent of three market areas of the next order. Following the same reasoning, there would be nine market areas at the next order, twenty-seven at the next order, and so on. Remember, these are equivalent numbers. If at each level we subtract the number of centres already existing, then the sequence would be as follows: 1, 2, 6, 18, 54, etc. Spatially, the distances apart of centres at each successively lower order would increase by a factor equal to $\chi \sqrt{3}$ where χ is the distance apart of lower order centres. Thus, if the lowest order centres were located four miles apart, the next order centres would be $(4 \times \sqrt{3})$, the next highest order centres would be $(4 \times \sqrt{3} \times \sqrt{3})$, and so on.

Theories of urban structure

When geographers turned their interest to the analysis of cities as discrete units rather than elements of a settlement system, they found little theory within their own discipline. Studying cities had been the prerogative of historians, archeologists, sociologists and planners. So, it was to these disciplines that geography turned to search for organizing principles of urban structure, form and content. This search for general principles was undertaken as part of a search for a framework to bring together the disparate works primarily focusing on urban morphology and form that at that time was the dominant product of urban geographers.

In 1945, Harris and Ullmann brought to the attention of geographers the sociologically-based concentric zonal hypothesis of city form and the economics-based wedge and sector theories of urban form offered by Homer Hoyt (Harris and Ullmann, 1945). To this they added their own descriptive model of what they argued was a more realistic form for examining current cities whose growth had produced multiple nuclei rather than being dominated by a single centrally located high density business area. The Concentric Zonal Hypothesis argued that land uses would be arranged more or less concentrically about a single dominant central business district (CBD); immediately surrounding the

CBD would be an area of industry and wholesaling with intermittent slums. Surrounding that would be a zone of blue collar working-class homes, next would be a middle class suburbia, and finally there would be isolated pockets of high-income neighbourhoods scattered around the periphery

The wedge and sector hypothesis suggested that like land uses would tend to cluster around significant lines of transportation and that they would be stratified according to their rent-paying ability. Thus those urban functions that could afford to pay the highest rent would be concentrated along the major lines of transportation. This included industrial and wholesale activities along major arterials and some high quality professional and business areas extending as wedges out from the city centre. The land market would act in such a way as to segregate geographically professional and business sectors from industrial and wholesaling sectors. Lower-income housing would be concentrated between wedges dominated by commerce, manufacturing and wholesaling, and mid- to higher-income sectors would be arranged beyond these areas on different sides of the city. High-income areas may be located next to the professional and business sector close to the city centre and would again develop isolated pockets further out towards the periphery.

In the Harris and Ullmann model the single city centre was replaced by multiple nodes which included the central business district and suburban or regional shopping areas, or possibly a decentralized business area. Again, however, there would tend to be changes in the quality of residence as one moved from the inner city to the periphery.

These descriptive models achieved some theoretical credence beginning with the work of regional scientist William Alonso (1964). Alonso developed a model of how the urban land market would act, but initially confined his model to a city with a single CBD. He argued that centrality and location were critical factors in determining land values and in general land values would decrease from the 100 per cent spot in the CBD towards the periphery. At the centre, those users capable of paying the highest price per unit of land would compete most favourably for the land. Such uses included

multi-storey commercial establishments, department stores and specialty shops, financial and other professional and business activities, and some high-income/high-rent apartment buildings. Industry and wholesaling activities were capable of paying the next highest rents for land and would occupy the zones surrounding the central area. Residential activity would occupy the bulk of the land in the city, and isolated suburban commercial pockets would provide local peaks in the land value gradient. Alonso also pointed out the paradox of many of today's cities in which the poorer people tended to live closer to the centre, thus occupying some of the most valuable land in the city and consequently being able to consume only extremely small units of it. As one moved towards the periphery land units got larger and at the periphery lived many of the richer families who were thus able to consume large quantities of land per person. This theory of land rents accounted for the observed variation in population densities throughout cities that had intrigued many geographic researchers. Further research by geographers pinpointed the significant difference between North American and European cities on the one hand and many cities in developing economies and in Asia on the other, where the rich or high-income areas were closer to the city centre and the poor were often found in the barrios quite distant from the city centre. Thus regional variation and cultural differences were found to impact the population density and land value gradients in cities in different parts of the world, adding an essential geographic component to what was originally envisaged as an economic land-rent model.

Further developments in geographic theory relating to the internal structure of cities came after the quantitative revolution. Building on some work by sociologists Shevky and Bell (1955), Brian Berry and his students introduced geography to the idea of factorial ecology. Essentially this used a statistical data reduction tool called factor analysis, applying it to large quantities of social, economic, planning, ethnic, occupational and other characteristics of urban populations. The technique isolated a minimal number of dimensions or factors that underlay the database so collected. Segments of the city that loaded high on each factor were then mapped and the underlying social, economic,

demographic, ethnic, occupational and other areas of similarity within cities were thus exposed. Based as it was on the analysis of factual empirical data, factorial ecology provided a more realistic framework for comparing and analysing the internal structure of cities than did the more idealized and descriptive forms that had been borrowed from sociology and economics. The analytical methods built into factorial ecology proved useful for allowing comparisons to be made between different economic systems at different time periods and encompassing different social and cultural ideologies. Exposing the differences produced in different economic systems contributed to the development of views of city structure that went beyond the mere description of form and focused on the processes that produced those structures. In particular it opened the door to escape from the traditional capitalist models and to examine city growth and development from alternative points of view (e.g., Marxist).

Summary

In this article we have by no means dealt exhaustively with the variety and range of geographical theories. As is detailed elsewhere in this issue, geographers have developed spatial theories of innovation diffusion, economic growth and development, migration and mobility, social gravity models and theories of demographic transition, and have ventured into postmodern deconstruction and reconstruction theories of economy and society, they have covered scales from micro-behavioural to global and are exploring both qualitative and quantitative models. Instead, what we have tried to do is to show how geographical theory can evolve from theories not explicitly designed to account for spatial variation. The significance of place-to-place differences is not generally embodied in the theories of many other economic, social, and behavioural science disciplines. Making this spatial component explicit not only brings the essence of spatial diversity in real world existence to the fore, but also increases our understanding of why things are distributed in the

space where they are and not somewhere else. Geographical theories for the most part have evolved from theories in other disciplines. But it is not a simple matter to add geography into those pre-existing theories and models. The essence of geography is variation. Most normative models assume variation away. Descriptive models show variation but do not satisfactorily account for it. The emergence of geographical theories, therefore, is a step taken to increase our level of understanding and comprehension of the objective environment and of the way that humanity interacts with and changes those environments.

A fundamental assumption of geography is that there is not one single environment. Not only are there the various physical and biotic environments, but there are a host of social, cultural, political, economic, legal and other environments which daily impinge on our activities, beliefs and values. In addition to these many hidden environments there are those internal to humanity. These are the cognized environments or the ones transformed by personal experience, by information obtained through personal interaction or via the mass media, or environments that are the product of our imagination. The common element to all is that it is the spatial dimension that is a dominant one. Bringing this spatial dimension more to the fore has been a driving force in the various attempts to develop geographic theory. As shown numerous times elsewhere in this issue, there is an abundance of regularity that can be observed and explained in even the most complex of all these different environments in which we live. Geographers search for this regularity and try to explain why it occurs. But they also are conscious of the important notion that environmental elements and natural causes, human causes, or other events occur differentially throughout these environments, and that there are sets of underlying principles and laws that help us understand why this spatial variation occurs. This search for explanation and understanding of places and events in a spatial context is a dominant underlying theme throughout this article and others in this issue.

References

- ALONSO, W, 1964 'The historic and the structural theories of urban form Their implications for urban renewal', *Land Economics* 49, pp 227-31
- AMEDEO, D, GOLLEDGE, R G, 1975 *An Introduction to Scientific Reasoning in Geography* New York John Wiley & Sons
- BERRY, B J L, 1967 *Geography of Market Centers and Retail Distribution* Englewood Cliffs, NJ Prentice-Hall
- CADWALLADER, M T, 1985 *Analytical Urban Geography* Englewood Cliffs, NJ Prentice-Hall
- CHURCH, R L, RE VELLE, C S, 1974 'The maximal covering location problem', *Papers of the Regional Science Association* 32, pp 101-18
- CHRISTALLER, W, 1933 *Die Zentralen Orte in Suddeutschland* (trans Baskin, C W, 1963, as *Central Places in Southern Germany*) Englewood Cliffs, NJ Prentice-Hall
- DUNN, E S, Jr, 1954 *The Location of Agricultural Products* Gainesville University of Florida Press
- FETTER, F A, 1924 'The economic law of market areas', *Quarterly Journal of Economics* 38, p 525
- GALE, S, OLSSON, G (eds), 1979 *Philosophy in Geography* Dordrecht, Holland D Reidel Publishing Co
- GOULD, P, 1985 *The Geographer at Work* Boston Routledge & Kegan Paul
- HARRIS, C D, ULLMAN, E, 1945 'The nature of cities', *Annals of the American Academy of Political Science and Social Science* 242, pp 7-17
- HOTELLING, H, 1929 'Stability in competition', *The Economic Journal*, March, pp 41-57
- HUDSON, J C 1969 'A location theory for rural settlement', *Annals of the Association of American Geographers* 59, pp 365-81
- HYSON, C D, HYSON, W P, 1950 'The economic law of market areas', *Quarterly Journal of Economics*, pp 319-27
- JOHNSTON, R J, 1979 *Geography and Geographers* New York John Wiley & Sons
- KING, L J, GOLLEDGE, R G 1978 *Cities, Space, and Behavior The Elements of Urban Geography* Englewood Cliffs, NJ Prentice-Hall
- LEIGH, A, 1946 'Von Thunen's theory of rent and the advent of marginal analysis', *Journal of Political Economy*, p 54
- LOSCH, A, 1954 *The Economics of Location*. New Haven, CT Yale University Press
- MCCARTY, H H, LINDBERG, J, 1966 *A Preface to Economic Geography* Englewood Cliffs, NJ Prentice-Hall
- RICARDO, D, 1817 *The First Six Chapters of The Principles of Political Economy and Taxation of David Ricardo, 1817* New York Macmillan and Co, 1924
- RUSHTON, G, GOODCHILD, M F, OSTRESH, L M, Jr, 1973 *Computer Programs for Location-Allocation Problems* Monograph #6. Department of Geography, University of Iowa
- SHEVKY, E, BELL, W., 1955 *Social Area Analysis* Stanford, CA Stanford University Press
- TEITZ, M B, BART, P, 1968 'Heuristic methods for estimating the generalized vertex median of weighted graph', *Operations Research* 16, pp 955-61
- VON THUNEN, H, 1966 *Von Thunen's Isolated State* (trans C M Wartenberg, ed P Hall) London Pergamon Press Ltd
- WEBER, A, 1909 *Theory of The Location of Industries* (trans C J Friedrich, 1929) Chicago University of Chicago Press
- YEATES, M, GARNER, B, 1980 *The North American City* (3rd edn) San Francisco, CA Harper & Row, Publishers