65

Urban Economics

James D. Burnell

College of Wooster

The field of urban economics was developed based on the observation that population and economic activity are concentrated in geographic space. Thus, one can define the field of urban economics as the study of the spatial relationships between individuals, households, and firms from an economic perspective. Much of urban economic analysis extends the maximizing behavior of individuals and firms from microeconomics to include how location affects this behavior. A focal point of urban economic analysis is how distance affects maximizing behavior. The fact that there is increased population and employment density at certain points in geographic space indicates that there are advantages of clustering of activity at certain locations. The concentration of activity also has consequences, both positive (i.e., enhanced productivity) and negative (i.e., increased congestion), and urban economists also consider these consequences.

Urban economics had its origins in the area of location theory and regional economics. The foundation was the role transport costs played in location decisions (Richardson, 1979), the hierarchy of trading areas, the system of cities (Christaller, 1933; Lösch, 1940), and the theory of land rent (von Thunen, 1966). The inclusion of geography in the decisions made by economic agents set the stage for urban economics to become a separate field in economics in the 1960s. A major reason was the important contributions of William Alonso (1964), Richard Muth (1969), and Edwin Mills (1967, 1972) in presenting theoretical and empirical analysis on the organization of urban space. The 1960s were also a period in which there was a significant focus on problems associated with the current spatial organization. Problems of poverty and social unrest experienced by large central cities made cities the focal point in the media, in public policy, and in academia.

The application of urban economic analysis has extended to subareas of the field, such as urban transportation, housing and real estate, and urban public finance. The field has evolved as the structure of urban areas has evolved. The focus of transport costs on location decisions has diminished and been replaced with the analysis of the impact of agglomeration economies on the size and composition of urban areas. The Alonso-Muth-Mills focus on monocentric urban areas has evolved to incorporate the growing multicentric nature of urban areas and the relatively rapid rate of urban sprawl. The empirical approach to analyzing urban economic phenomena has also evolved with the development of spatial econometric techniques that allow empirical modeling to account for the influences of the contiguity of urban space.

The remainder of this chapter is organized as follows. First, the concept of the urban area is considered from the economist's viewpoint. Next, a discussion of the elements that affect urban areas in an interurban context is presented. This section considers dominant explanations for the productivity of urban areas. The following section considers the intraurban relationships that exist and how they explain the urban spatial structure that exists in metropolitan areas. In particular, this section considers the monocentric model of urban land use as well as growing suburban and multicentric land use patterns. Finally, some outcomes and issues related to modern urban spatial structure are considered. The focus of this section is how urban economists address the spatial pattern of households.

The Concept of an Urban Area From an Economist's View

Undergraduate urban economics textbooks typically provide different definitions of urban areas that reflect a hierarchy of urban areas, usually incorporating the definitions used by the Bureau of the Census. Currently, the Census defines an urban area as a community with a population of 2500 or more. A micropolitan area is an area with an urban core population of between 10,000 and 50,000 people, while a metropolitan area has an urban core of at least 50,000. The Census defines the spatial reach of these areas in terms of one or more counties that have a high degree of economic and social interaction with the urban core.

The urban core refers to an important city that is the focal point of the interactions with the areas outside the city. Cities are defined by political boundaries: the legal boundaries that define the political authority of these areas. Edwin Mills and Bruce Hamilton (1989) point out that to economists, political boundaries are less important than the market forces that contributed to the increased density of individuals and firms. The market forces that define the economic concept of an urban area can be considered from two perspectives: interurban and intraurban. These perspectives can be examined in terms of the location decisions made by households and firms.

One may consider interurban analysis as the study of competition across urban areas. Interurban analysis considers location decisions and their consequences for different urban areas. Households and firms evaluate the locational advantages of different urban locations. For example, a manufacturing firm may be assessing the productivity of the labor force for the type of workers it needs between the St. Louis and Minneapolis metropolitan areas. A household may evaluate which urban area to locate in based on the availability of employment, the cost of living, or the existence of desirable amenities. The economic concept of urban area in this case transcends the existence of the political boundaries of cities. The consequences of these decisions affect the growth and income-creation ability of the area.

Intraurban analysis addresses the location decisions and their consequences within urban areas. Once an urban area is chosen, both firms and households will then decide where within the area to locate. The market allocation of land among households and firms within an urban area was the focus of the Alonso-Muth-Mills monocentric model. Although the existence of political boundaries does not necessarily define the overall urban area in an economic sense, decentralization of households and employment within urban areas suggested that political boundaries do have some influence. The increasing fragmentation of urban areas since the middle of the twentieth century gave rise to a large number of suburban jurisdictions that could compete for households and firms based on their tax and service packages and their ability to use zoning to influence land market outcomes. The consequences of intraurban location decisions by households and firms and the existence of interjurisdictional competition affect the fiscal viability of large central cities, the distribution of employment opportunity, and the distribution of income and minority groups throughout the urban area.

Interurban Analysis: The Urban Hierarchy

The focus of interurban analysis is on the process of urbanization, in which economic activity concentrates at particular locations, and the factors that contribute to the extent of this concentration. The starting point of the urbanization process is location theory, whereby the profit-maximizing decisions of firms specifically considered how location affected profitability. Although location theory could explain the location decision of individual firms, urbanization also meant that there were other factors that would influence the size and growth of urban areas. Early explanations were based on central place theory, which described a hierarchy of urban places in a system of urban areas based on the market area for goods and services (Christaller, 1933; Lösch, 1940). The most recent emphasis on explaining the process of urbanization is on the importance of agglomeration economies, factors external to the firms that provide advantages of clustering economics activity (Fujita, Krugman, & Venables, 1999; Rosenthal & Strange, 2004).

Location theory includes transportation costs as well as the costs of inputs in the cost functions of firms. To illustrate the importance of transport costs in location decisions, the basic model assumed labor and capital costs were equal across space and the firm used a raw material input that was available at one location while the market for the firm's product occurred at a different location. Thus, the firm would have to choose whether to transport the raw material to the market location to produce its product or to locate at the raw material site and transport the product. The firm's profit-maximizing location would be that which minimized transport costs. The usefulness of this simple location decision was that it introduced the concept of an economic location weight that was determined not only by the physical weight but also by the unit transport cost per mile (O'Sullivan, 2009). Traditional examples include weight-gaining and weight-losing production processes. Weight-gaining processes include products that gain physical weight, such as water added in the beverage industries, or products whose transport costs are high because of their fragile nature. Weight-losing processes include mining and lumber.

Relaxing the assumption that the costs of other inputs used by the firm are equal across space allowed the spatial variation in cost and productivity of inputs to influence the location of the firm. Inputs prices that are generally considered to vary across space include labor, capital, land,

energy, and raw materials. Another factor considered important by firms is the impact of the tax and service package that can be offered by these governments as they compete among themselves for firms to locate within their boundaries.

Location theory could explain why firms choose among different locations, but it could not effectively address the fact that urban areas varied in size and in their ability to grow. Central place theory developed by Walter Christaller (1933) and August Lösch (1940) provided explanations for the location of market-oriented firms and the resulting hierarchy of urban areas that results.

The basic assumptions of the theory were that consumers were evenly distributed across space and transportation costs were equal in all directions. A firm existed at a particular location to serve the population. Consumers would travel to the firm to obtain the good, and the effective price paid by the consumer would be the price established by the firm plus transportation costs. The firm's market area would stretch to the point that the good's effective price was such that consumers were no longer willing to purchase the product from the firm. If the firm earned an economic profit, other firms producing the product would enter at another location.

As firms continued to enter, the market area of existing firms would decrease as consumers encountered lower effective prices due to closer proximity to firms. Firms would continue to enter only until normal profits were made and spatial equilibrium was reached. Each firm would have a spatial monopoly over a particular area, creating a spatial network for that product. Networks of different sizes for different goods would exist based on the size of the market area for the product. Smaller order goods are goods with small geographic market areas, and higher order goods have larger geographic market areas. A central place would be a location where one or more networks locate. The size of the urban area would be determined by the market area of the product with the largest geographic reach—the highest order good present at that location—and would contain all successively lower order goods. The market reach of its highest order good would define an urban area's position in the hierarchy. As an example, consider medical services. Small places would be expected to have a number of general practitioners and basic medical testing services. More specialized medical services would be found in a larger place. This establishes a hierarchical spatial link whereby residents of smaller places are linked to large places to obtain higher ordered services.

Central place theory's value in urban economics is primarily in its explanation of the pattern of retail activity. Other approaches to explaining the size and growth of urban areas also developed. Economic base theory posited that an area's economy was composed of two sectors. The basic sector was composed of firms that exported their product beyond the area's boundaries. The local sector was composed of firms that provided products to the area's residents. The growth of the area was determined by the export demand for the area's products. As income flowed into the area from exporting basic sector goods, a fraction of this income would be spent in the local sector through successive rounds of spending. The successive rounds of spending defined the multiplier effect—a dollar's worth of income would generate a larger amount of total income, depending on the size of the multiplier. The size of the multiplier was dependent on how much of the basic sector income was spent locally (McDonald & McMillen, 2007).

Economic base theory explains that the growth of an urban area is dependent on the change in the export demand for the area's basic sector products and how well developed the local sector is in providing goods and services for the area's population. This simple theory of urban growth has found wide practical application in the area of economic impact analysis, in which predictions are made regarding the impact of the attraction or departure of a key economic activity on the local area's economy.

The most recent emphasis in research on the process of urbanization has been on the advantages of the clustering of economic activity. This recent emphasis was stimulated by the emergence of the new economic geography based on the work of Paul Krugman (1991) in the area of international trade, but it has found much theoretical application in explaining the clustering of economic activity in urban areas (Fujita et al., 1999). Masahisa Fujita and Tomoya Mori (2005) describe the new economic geography as being a general equilibrium approach that specifically considers agglomeration in explaining the pattern and structure of urban areas. A primary consideration of the new economic geography is the specific incorporation of economies of scale and imperfect competition.

Scale economies are a fundamental requirement for increases in employment density in geographic space, since larger production facilities allow efficiencies that reduce average costs as output produced rises, and larger facilities lead to more workers locating in proximity to their jobs. The advantages of clustering are attributed to the existence of agglomeration economies, which are external benefits to firms that reduce their average costs at all levels of output.

Alfred Marshall (1920) originally identified three sources of agglomeration economies. One source is input sharing, in which final product firms purchase intermediate inputs from a specialized provider who is able to use economies of scale because of increased demand resulting from the clustering of the final product firms. Stuart Rosenthal and William Strange (2006) refer to input sharing as a form of local outsourcing. For example, computer hardware firms may locate near a computer chip manufacturer. The concentration of hardware firms allows the chip manufacturer to realize economies of scale that result in lower chip prices for the hardware firms.

A second source of agglomeration economies is labor pooling, which occurs when firms are able to draw from a large pool of specialized labor. The concentration of high-tech computer firms in Silicon Valley and biotech firms in the Boston metropolitan area reflect a large pool of highly educated individuals that offer small startup computer or biotech firms a labor force with requisite skills. This benefits not only firms who have a large demand for specialized labor, but also the workers who have other job opportunities if some of the small startups fail.

The third source is attributed to knowledge spillovers, whereby the presence and interaction of those with specialized knowledge about their products and production processes will stimulate a higher rate of innovation. The interaction of highly educated workers in Silicon Valley and Boston and the existence of (well-known) research universities increase the likelihood of innovation.

Agglomeration economies are classified according to the type of firm receiving the external benefit. Localization economies are external to the firm but internal to the industry, since they occur based on the extent of the presence of firms in the same industry. Urbanization economies are external to both the firm and industry and are attributed to the size of the area. Larger areas have a diverse set of firms and labor, which allows firms across industries to benefit.

The existence of agglomeration economies contributes to the size and growth of urban areas. Agglomeration economies increase the productivity of the area, which results in higher rates of growth. Increased demand for the area's products leads to higher labor demand. In areas realizing agglomeration economies, the existence of this greater productivity also provides a self-reinforcing effect. As production increases to meet the increased demand, there is an additional pull of firms who would benefit (O'Sullivan, 2009).

The consideration of the self-reinforcing effects of agglomeration suggests that agglomeration economies can be dynamic as well as static. Static agglomeration economies relate to the industrial and geographic dimensions of the effect of agglomeration economies on firms and the urban area and help explain why some urban areas are larger than others. The industrial dimension relates to the industries experiencing the benefits of localization or urbanization economies. The geographic dimension relates to the proximity of establishments in industries experiencing agglomeration economies and the diminishing effect of the agglomeration advantage as proximity between establishments decreases (Rosenthal & Strange, 2004). Static agglomeration economies refer to a one-time cost reduction associated with the clustering of firms (McDonald & McMillen, 2007).

Dynamic agglomeration economies relate to a time dimension of the impact and help explain not only the growth of particular urban areas but also the rate of growth. Dynamic localization economies cause continual reductions in the firm's average costs as the size of the industry in the area increases. Dynamic urbanization economies yield continual reduction in a firm's average costs as the size of the area increases compared to static agglomeration economies that are considered a one time reduction in cost (McDonald & McMillan, 2007). Rosenthal and Strange (2004) attribute dynamic agglomeration economies to knowledge spillovers whereby the acquisition and transfer of knowledge between firms occurs over time, resulting in cost advantages realized in the future.

Much empirical work has been done to determine the impact of agglomeration economies on the urbanization process. Rosenthal and Strange (2004) provide a comprehensive review of the empirical literature regarding agglomeration economies. The empirical literature follows a number of approaches to test various aspects of the impact of agglomeration economies on urban productivity. These include the testing for the importance of localization and urbanization economies, identifying the appropriate geographic level, and using a microbased versus an aggregate approach to estimating productivity effects. A brief representation of the empirical literature is considered here.

Edward Glaeser, Hedi Kallal, Jose Scheinkman, and Andrei Shleifer (1992) consider the role of knowledge spillovers on the growth of industry employment in the largest 170 cities. They test for three potential impacts of knowledge spillovers on industry growth. The first two relate to localization economies whereby within-industry knowledge spillovers lead to higher rates of growth, and the difference between them centers on the degree of local competition within the industries. On one hand, less competition allows innovating firms to realize the gains of the innovation internally and provides an incentive to innovate. On the other hand, greater local competition within the industry stimulates innovation as firms try to stay ahead of their competitors. The third potential impact of knowledge spillovers reflects urbanization economies. The more diverse the representation of industries, the greater the interchange of ideas across firms, which leads to faster growth.

The empirical model of Glaeser et al. (1992) considers the growth rate of employment in the six largest industries in 170 of the largest cities as a function of variables that measure the three potential impacts of knowledge spillovers, controlling for regional and natural characteristics that might affect local growth. Their findings show that growth is faster in cities where local industry competition is greater and where industry diversity is greater.

Rosenthal and Strange (2001) consider the three sources of localization economies to determine the level of geography at which the sources are important in explaining the concentration of industry employment. The authors hypothesize that the impact of the agglomeration advantages may be influenced by the spatial concentration of industry activity.

Some sources of agglomeration may be more important at close proximity, while others may be important over a larger area. They regress an index of spatial concentration for an industry on variables that measure input sharing, labor market pooling, and knowledge spillovers, controlling for transport cost and natural advantage. Versions of the model are estimated for the zip code, county, and state levels. The results provide evidence that labor market pooling is an important agglomerative source at all three levels of geography, input sharing was important at the state level but not lower levels of geography, and knowledge spillovers were important only at the zip code level.

J. Vernon Henderson (1986) addresses the importance of localization versus urbanization economies. From an industry location view, the difference is whether areas that specialize in particular industries have greater advantages or whether it is the size of the area that provides the greatest impact on productivity. Henderson employs a production function approach, where industry output is a function of inputs. He uses local industry employment as a measure of localization economies and urban size as a measure of urbanization economies. From his empirical results, Henderson concludes that there is strong evidence of localization economies for almost all industries considered and that the localization effects were large. He finds almost no evidence of urbanization economies. He also finds evidence that the agglomeration effects diminish for larger urban areas.

One of the empirical issues regarding estimating the effect of agglomeration economies has been the level of aggregation of the industry data. For example, Henderson (1986) used the two-digit level to define the industry. A more recent work by Henderson (2003) estimates plant-level production functions using panel data to provide a microlevel assessment of the effect of localization and urbanization economies and the ability to capture dynamic localization economies. Plant-level data is desirable since it is the plant that realizes the agglomeration economies at its particular location. Henderson considers the effect of localization economies and urbanization economies on what he identifies as machinery industries and high-tech industries. He also considers whether the plants are single plant or multiplant; single plants are more reliant on the local economic environment than plants that are affiliated with a corporation since the operation of these plants reflects the internal linkages determined by corporate decisions.

Henderson (2003) estimates plant-level production functions as a function of plant-level inputs plus measures to account for localization and urbanization economies. The number of plants in the same industry in the same county is used to measure localization economies and a measure of diversity of manufacturing employment for urbanization economies. Dynamic localization economies are measured by lagging the localization economies measure. The results indicate that localization economies are important for plants in high-tech industries but not for machine industries. The impact of localization economies

is stronger for single plant versus multiplant affiliates. High-tech single plant firms also benefit from dynamic localization economies. There is no evidence that urbanization economies had an influence for either high-tech or machinery industries.

Rosenthal and Strange (2003) also use a microlevel approach with a focus on the birth of new establishments for six industries. They hypothesize that the presence of agglomeration economies results in new establishments clustering around existing establishments, while a dispersed pattern of new establishment locations would occur if there were no agglomeration economies. The sample used is composed of data measured at the zip code level, which is a departure from earlier analysis that typically used metropolitan areas as the geographic reference point. This allowed the authors to test for whether the effects of agglomeration economies diminish as distance from existing establishments increases. The authors focus on six industries that represent both innovative (e.g., software) and traditional (e.g., machinery) industries. The number of new establishments or the amount of new establishment employment in a zip code is regressed against variables that include the number of establishments per worker in the industry and outside the industry as a measure of competitiveness and diversity of economic activity.

Measures of urbanization and localization economies are also included. Urbanization economies are measured as employment outside the industry, while localization economies are measured as employment within the industry. To test whether the impact of the agglomeration economies diminishes with distance, Rosenthal and Strange (2003) include within- and outside-industry employment in a series of concentric rings around the initial zip code. The results are that localization economies are important for five of the six industries, while there is little evidence that urbanization economies matter. The results also indicate that the impact of localization economies diminishes with distance.

The review of interurban analysis indicates that many factors affect the location of economic activity across geographic space. The recent focus on the role of agglomeration economies is important given the changing nature of the economic base in many urban areas. As urban governments try to influence the location of firms to generate growth within their boundaries, the understanding of the type of agglomeration economies and the geographic extent of these economies become important information for policy makers.

Intraurban Analysis: Urban Spatial Structure

Intraurban analysis considers how space is organized within an urban area. Historically, the development of urban areas was such that the intensity of land use was much greater at

the urban core, generally referred to as the central business district (CBD) and diminished as distance from the CBD increased. The urban land market is the mechanism by which land is allocated to the competing residential and business users. Residential and business users establish their willingness to bid for a certain location based on the location's value in terms of utility for residents or profit for business. Competition in the land market would allocate space to the highest bidder.

The origin of the land allocation process was the theory of land use developed by Johann von Thunen (1966). His contribution was to consider the competition for land use among users who valued proximity to a central location, and he developed the concept of the land rent function, whereby agricultural land users who desired to minimize the costs of producing and transporting their outputs to the market would bid for locations in close proximity to the market. Bids would decline for sites farther from the market center to account for the increase in transportation costs. A land rent existed based on the value that users place on the scarce locations surrounding the desired central location.

The modern version of the land allocation process is generally labeled the Alonso-Muth-Mills approach. Alonso (1964), Muth (1969), and Mills (1967) formally developed the theory of spatial equilibrium in the land market. The Alonso-Muth-Mills approach considers a competitive urban land market in which the demand for urban land is based on the utility-maximizing decisions of households and the profit-maximizing decisions of firms in a monocentric city. Firms value CBD locations since the CBD contains the transport node where firms export their products. A CBD location would reduce the cost of transporting products to the export node. For households, the CBD is the location of employment, and households value proximity to the CBD to reduce commuting costs. The value that firms and households place on proximity to the CBD defines their bid-rent function, which is the willingness of a user to pay for a location at a particular location from the CBD.

For firms, the bid-rent function is based on the profits of the firm:

$$\text{Profit} = pq - cq - txq - Rs, \tag{1}$$

where p is product price, c is unit cost, q is quantity sold, t is the transport cost per unit of distance, s is the size of the site, and R is the rent bid per unit of site size. R is the amount the firm is willing to pay at distance x, holding the level of profit constant. Assuming competitive equilibrium where profits are zero and solving for R yields the firm's rent bid function:

$$R = \frac{pq - cq - txq}{s}. \tag{2}$$

As distance from the CBD increases, the firm will reduce its bid by the increase in transport cost incurred for shipping its product to the CBD. The slope of the bid rent function,

$$-\frac{tq}{s}, \tag{3}$$

represents the reduction in land costs necessary to compensate for increased transport costs as distance from the CBD increases, and it measures the value of accessibility to the CBD. As compared with firms that have smoother bid-rent curves, firms with steeper bid-rent functions assign greater value to accessibility since the reduction in rent bid will be greater in order to compensate for the higher level of transport costs.

For households, the utility-maximizing decision defines the concept of the bid-rent function. Household utility depends on the consumption of housing services and other goods, and households face a budget constraint where income is spent on other goods, housing services, and commuting costs. Holding income and expenditures on other goods constant, consider the expenditures on housing services and commuting costs, which depend on distance to the CBD:

$$Rh + tx, \tag{4}$$

where R is the rent bid per unit of housing services, h is the amount of housing service, t is the commuting cost per unit of distance, and x is distance from the CBD. Spatial equilibrium requires that utility is constant at different distances. This means that expenditures on housing and commuting must remain constant for a constant level of utility. As distance from the CBD increases, households will change their bids to compensate for the change in commuting costs:

$$\Delta Rh + t\Delta x = 0. \tag{5}$$

The slope of the household's bid-rent function is

$$\frac{\Delta R}{\Delta x} = -\frac{t}{h}, \tag{6}$$

and it represents the amount rent bid will go down as distance increases.

Land market equilibrium requires that land goes to the highest bidder and that neither firms nor households can gain profits or utility by moving to a different location. The graphical representation of the urban land market is depicted in Figure 65.1. The slopes of the bid-rent curves reflect the value of accessibility to the three users. Commercial firms have the greatest value of accessibility to the CBD and outbid resident and manufacturing firms.

The Alonso-Mills-Muth model has been useful in explaining patterns of urban land use evident in American cities. The declining importance of the CBD and the increased presence of population and employment in the suburbs can be explained by a reduction in transport costs,
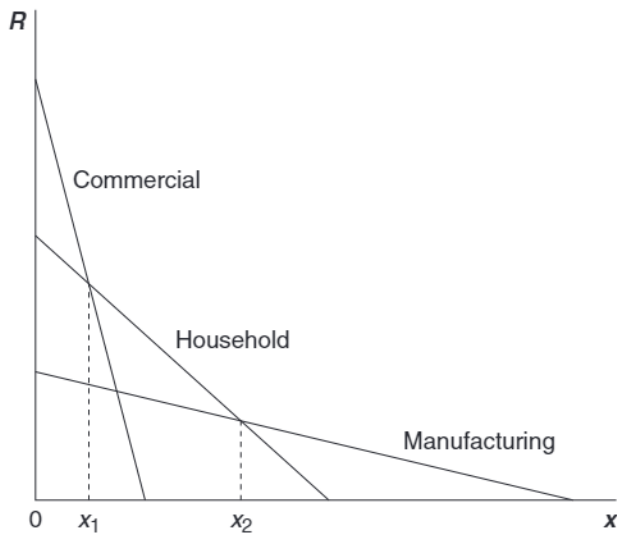
Figure 65.1 The Urban Land Market

Figure 65.2 Disamenities and the Household Bid-Rent Curve

which leads to flatter sloped bid-rent curves. The model also provided an explanation for the fact that lower income groups lived in the central city and higher income groups lived in the suburbs. Both Alonso (1964) and Muth (1969) consider this in their formulations of the monocentric model. The explanation is based on the income elasticity of demand for housing relative to the income elasticity of commuting. If the income elasticity of demand for housing were greater than the income elasticity of commuting, then high-income populations would locate in the suburbs where housing prices are lower and the amount of land available for housing is greater.

The Alonso-Mills-Muth model can be extended to consider other factors related to the characteristics of urban areas. Douglas Diamond and George Tolley (1982) incorporate the concept of amenities to determine their impact on households' bid-rent functions. Amenities are location-specific characteristics that yield utility to households but are not purchased directly. Diamond and Tolley incorporate amenities as an additional component of the household's utility function. The resulting utility-maximizing outcomes determine the values households place on amenities and yield bid-rent functions that can be positively as well as negatively sloped functions. As depicted in Figure 65.2, the positive-sloped portion of the function would represent proximity to a negative amenity, such as pollution, in which households would need to be compensated for locating at or near the source of the negative amenity. As the distance from the source increases, the effect of the negative amenity on the household diminishes, resulting in higher bids as distance from the negative amenity increases. This occurs up to x_1. Beyond x_1, the effect of pollution has dissipated, and compensating for commuting costs now dominates the bids made.

The Alonso-Mills-Muth model and its variations provided a theoretical structure for empirical work on urban spatial structure. Empirical approaches have included the estimation
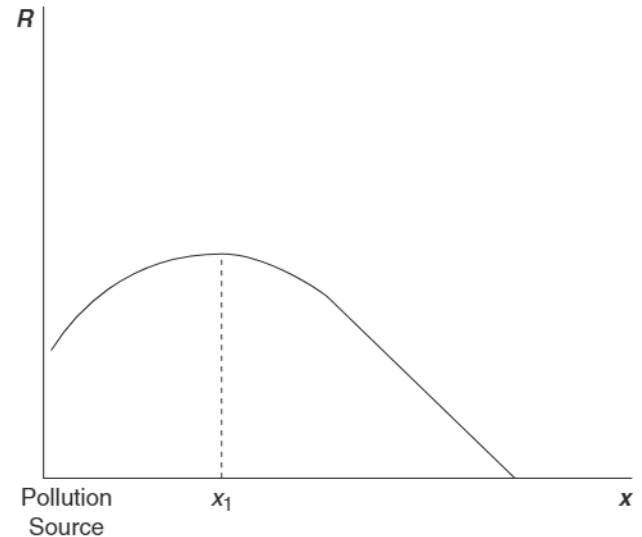
of population and employment density functions. Here, population and employment density are regressed against the distance from the CBD to determine the intensity of land use, with the monocentric model predicting that density decreases as distance from the CBD increases. Estimation of a hedonic housing price function is another empirical tool used to measure the value that households place on characteristics that comprise housing services. These characteristics typically include structural characteristics such as size of the dwelling and the number of rooms, quality characteristics such as the age of the dwelling, neighborhood characteristics such as crime rate or school quality, and proximity characteristics such as distance to the CBD or to particular amenities (Sirmans, MacPherson, & Zietz, 2005).

Early empirical work estimated population and employment density functions to test the predictions of the monocentric model of decreasing population and employment density as distance from the CBD increases. The early evidence provided by Muth (1969) and Mills (1972) supported these predictions. Even recent estimation of population density functions indicates population density is negatively related to distance from the CBD (McDonald & McMillen, 2007). Glaeser (2007) estimates the relationship between housing prices and distance from the Boston CBD and also finds support for the monocentric model.

Although the monocentric focus of the Alonso-Muth-Mills model remains the basis of much of the analysis of urban spatial structure, there has been recognition that the decentralization of population and particularly employment represents a transformation of the urban landscape (Glaeser & Kahn, 2003). Although the recent estimation of population functions and housing price gradients suggest that there is still the negative relationship between density and distance from the CBD, the empirical results suggest that the explanatory power of the monocentric model is relatively low.

Population decentralization has been attributed to factors such as poorer quality housing and the concentration

of minority groups and low-income groups in the central city that act to divert middle- and upper-income groups to the suburbs (Anas, Arnott, & Small, 1998). The politically fragmented nature of urban areas has also contributed to this decentralization through a Tiebout (1956) process, whereby households have the ability to match their preferences for local public services to particular communities.

Alternative explanations of the decentralization of employment have also been offered. In considering the intraurban location decisions of firms, Glaeser and Kahn (2003) offer three potential explanations of the draw to suburban locations. The first reflects the development of transportation infrastructure in the suburbs that offers savings in transportation costs compared to CBD locations. The second is that there may be differences in productivity between central locations and the suburbs. Knowledge and information spillovers are more likely to occur in the dense CBD areas, and this creates a productivity advantage for what Glaeser and Kahn refer to as idea firms—those that require high human capital and computers. The third explanation is that firms that do not benefit from these spillovers may find other cost advantages, such as proximity to workers, at suburban locations. They postulate that firms may follow the population to the suburbs. Based on their empirical research, Glaeser and Kahn conclude that the extent of decentralization in an urban area is greater when the presence of manufacturing firms in the area's industry mix is greater, that knowledge spillovers are important for idea-intensive industries, and that the labor force location has a strong influence on firm location.

The increasing focus on decentralization in urban areas suggests that there has been an evolution in the spatial structure such that many urban areas are considered polycentric. This relates to the observation that in many metropolitan areas, there exist edge cities (Garreau, 1991) that contain a concentration of office and commercial activity compared with other locations throughout the urban area. The existence of subcenters reflects agglomeration advantages away from the traditional CBD. In part, the technological advances of communication and the existence of the transportation infrastructure in suburban areas have contributed to this concentration of activity.

A major contributor to subcenter formation is also a contributor to the importance of monocentric city: agglomeration economies. Both localization and urbanization economies provide clustering advantages at locations away from the central business district. Robert Helsley and Arthur Sullivan (1991) and Denise DiPasquale and William Wheaton (1996) suggest that the advantages of these dispersed locations are enhanced by the increase in commuting costs associated with an increasingly dense monocentric city and by the existence of public infrastructure. Increased commuting costs necessitate higher wage costs for central business district firms to attract workers. Dispersed locations offer lower wages and land costs. The existence of public

infrastructure at these dispersed locations provides an additional locational advantage for firms to cluster at particular locations. All of these factors enhance the external economies that occur as firms cluster at these sites.

Much work has been done to determine what empirical evidence exists for the existence of the polycentric form of urban areas. Three examples will be considered here. Kenneth Small and Shunfeng Song (1994) estimate polycentric employment and population density functions for the Los Angeles metropolitan area for the years 1970 and 1980. They find that the polycentric functions have greater explanatory power for the density patterns compared with the monocentric estimates. Daniel McMillen and John McDonald (1998) estimate employment density functions for Chicago for 1980 and 1990 to determine the effect of agglomeration economies on the existence of subcenters. They hypothesize that the presence of highway interchanges and commuter rail contributes to increased density at the subcenters. They also include proximity to subcenters as a measure of agglomeration economies. They find evidence that the agglomeration effects of shared transportation infrastructure and information and shopping externalities associated with subcenters contribute to higher employment density at these locations. Last, McMillen and Stefani Smith (2003) consider the number of subcenters in a metropolitan area. They consider the characteristics of 62 metropolitan areas and find that the number of subcenters increases as the population and commuting costs of an area increase.

The Spatial Patterns of Residents

An important aspect of research in the urban economics field is the spatial pattern of the location of particular groups in urban areas and the consequences of this pattern. The spatial pattern that exists in most metropolitan areas is the centralization of low-income and minority groups in the central city or inner-ring suburbs. This spatial pattern evolves from household decisions to rent or own and from institutional factors such as lending and insurance practices and community land use policies.

Housing is important from an urban economics perspective because it is fixed in location and is a major component of housing services in the residential utility functions in the Alonso-Muth-Mills model. Along with a residential location decision, households must also decide whether to purchase or rent, which is referred to as *tenure choice*. Housing is also a durable good. The durability affects homeowners in two ways. First, housing deteriorates over time, and homeowners must decide on maintenance and repair. Second, homes generally appreciate in value over the length of ownership; thus, housing can represent an important asset to the owner. In the process of deciding to own a home, the household will require access to mortgage institutions. Housing market

outcomes have important implications for the distribution of population groups throughout the urban area. Different income and racial groups are affected by the availability of affordable housing, the preferences individuals have for neighborhood racial composition, and the existence of discriminatory behavior by institutions that facilitate home ownership.

The tenure choice decision of the household is based on the costs of homeownership versus renting. One element of the costs of home ownership is the mortgage payment. Mortgage payments are based on ability to obtain a loan and the resulting terms of the loan. The ability to obtain a loan is affected by the income and wealth of the household, the household's credit risk, and the value of the property. Lack of wealth has been a factor for low-income households generally; it has also affected the home ownership rates of African Americans (Charles & Hurst, 2002; Gyourko, Linneman, & Wachter, 1999; Wachter & Megbolugbe, 1992).

Racial segregation is a feature of the residential living patterns in metropolitan areas. The concern for racially segregated living patterns is the impact that segregation has on the outcome of minorities who live in more centralized neighborhoods. These outcomes include living in neighborhoods with poorer housing quality, lower quality educational opportunities, and restricted job opportunities (Cutler & Glaeser, 1997; Ross, 1998).

The preferences of racial subgroups are incorporated in residential location choice models by considering composition of a neighborhood as one of the neighborhood characteristics that affect the household's utility. In the case where a racial subgroup has an aversion to living with other racial subgroups, this will lead to segregated outcomes in housing markets. Analysis of racial preferences indicates that whites have the greatest aversion to living with nonwhites, particularly African Americans (Charles, 2005).

Another area that has been the focus of segregated living patterns has been the role of racial discrimination in housing and mortgage markets. In the case of housing markets, the analysis of how discrimination occurs centers on the behavior of real estate agents who provide information on housing and mortgage availability that may steer clients to neighborhoods of a particular racial composition (Yinger, 1995). The most recent empirical analysis to determine the existence and extent of racial discrimination in housing markets was sponsored by the Department of Housing and Urban Development in 2000 (Turner, Ross, Galster, & Yinger, 2002). Researchers used paired testing methodology, whereby equally qualified white and minority home seekers and renters interact with real estate and rental agents to determine whether they receive the same treatment in their housing search. Compared to the results of an identical study conducted in 1989, the evidence showed that African American and Hispanic home seekers and renters suffered from discrimination, but the incidence of discrimination was lower for African American home seekers and renters and Hispanic home seekers compared to 1989. Hispanic renters faced the same incidence of discrimination as they did in 1989.

Access to mortgage credit also plays an important role in the tenure decision. There is growing research interest in the question of how decisions by financial institutions affect urban residents. Much of the interest has centered on whether financial institutions provide access to mortgage credit based on the creditworthiness of the household rather than the household's membership in a particular income or racial group. Since the 1990s, a dual mortgage market has developed. The prime market is composed of low-risk borrowers who obtain loans with lower interest rates compared to borrowers in the subprime market. The subprime market is composed of those considered to be high-risk borrowers, and as a result, they pay higher mortgage interest rates than prime borrowers (Apgar & Calder, 2005). Mortgage discrimination occurs when borrowers are denied credit or are channeled to the higher rates in the subprime market or when the financial institution bases its decision by incorporating group membership as an additional determinant of the loan decision. The consequences are not restricted to the influence on the tenure decision of a minority or low-income household, but also include the increased risk of foreclosure.

The empirical research in mortgage lending discrimination typically shows that there are racial differences in mortgage lending outcomes, but it is more difficult to conclude that these differences are attributed to discrimination on the part of lenders. Alicia Munnell, Geoffrey Tootell, Lynn Browne, and James McEneaney (1996) provided a detailed analysis of factors that contribute to mortgage loan denials in Boston. Their empirical model includes variables that reflect the mortgage applicant's risk and cost of default, loan characteristics, and personal characteristics, with race as one of the personal characteristics. The results showed that African American and Hispanic applicants had a greater likelihood of being denied a mortgage loan than whites. Stephen Ross and John Yinger (2002) and Anthony Yezer (2006) discuss the criticisms of this approach, which range from exclusion of relevant variables that reflect creditworthiness to inability to account for the interaction of the applicant with the loan originator. The conclusion from these criticisms is that the results on the race variables will be biased.

A more recent approach applies the paired-testing methodology to the mortgage market. Ross, Margery Austin Turner, Erin Godfrey, and Robin Smith (2008) apply the paired-testing methodology to the pre-application process, whereby potential applicants interact with loan originators to obtain information regarding various types of loan products. The sample is composed of 250 paired tests in the Los Angeles and Chicago metropolitan areas. The results showed that African American and Hispanic testers in Chicago were less likely to be provided information, or they

received information on fewer products than white testers. The results for Los Angeles did not show a statistical difference in the treatment between African Americans, Hispanics, and whites. The implications for their results in Chicago are that minorities may have limited access to prime sources of mortgage credit, and this may lead qualified applicants to rely more heavily on the subprime market.

Conclusion

As the presentation of this chapter demonstrates, the field of urban economics concerns itself with the location decisions of firms and households and how those decisions affect the population who live in densely populated urban areas. Urban economics addresses the circumstances and consequences of living in metropolitan areas. Its emergence as a separate field in the 1960s coincided with the changing urban spatial structure—the decentralization of households and employment and the centralization of poverty and minority households—and the consequences of these changes on central city and suburban residents.

Urban areas have evolved from being largely centers of manufacturing activity to having more service- and information-oriented activity as important components of their economic bases. The focus on the role of agglomeration economies has important policy implications for areas undergoing the restructuring of their economic base. Policy makers trying to retain or attract new firms need to be aware of the important contributions of agglomeration economies that exist in their areas.

Research in the field has also addressed the change in urban spatial structure that has occurred within metropolitan areas over the past 40 years. The theory of monocentric urban areas remains the foundation of understanding how the role of distance affects the location decisions of firms and households within urban areas. Extensions of the theory have been used to explain the development of multicentric urban areas and urban sprawl. Urban economics has also developed a number of subfields. Real estate economics and finance extend theoretical and empirical approaches to real estate markets. Governmental issues such as jurisdictional fragmentation and competition, land use policy, and urban fiscal issues are widely addressed. Policy makers interested in the future outcomes of their local areas need to be aware of the interaction that occurs between the pattern of increasing suburbanization and sprawl and its consequences not only for newly formed suburban communities but also for the central city, which may bear the burden of being home to an immobile population.

There are a number of areas that will be pursued in future research. As the national economy continues to undergo structural change, the impact of these changes on the productivity of urban areas will be explored. Much of the literature on agglomeration economies has focused on manufacturing. Future research should extend to the growing service sectors to determine the consequences of this restructuring on the size and viability of urban areas. The development of geographic information data by local governments and the development of spatial econometric software will become an increasingly important component of future research. These developments allow inclusion of spatial spillover effects and their influence on firm and household decisions. Finally, future research will also explore the consequences of urban sprawl and political fragmentation for metropolitan areas. An increasingly important area in urban policy is the concept of regionalism, where metropolitan governments may supersede particular decisions of local governments. Future research should develop a theoretical approach to consider cooperative decision making in metropolitan areas.

References and Further Readings

Alonso, W. (1964). *Location and land use.* Cambridge, MA: Harvard University Press.

Anas, A., Arnott, R., & Small, K. A. (1998). Urban spatial structure. *Journal of Economic Literature, 36,* 1426–1464.

Apgar, W., & Calder, A. (2005). The dual mortgage market: The persistence of discrimination in mortgage lending. In X. Briggs (Ed.), *The geography of opportunity: Race and housing choice in metropolitan America* (pp. 101–126). Washington, DC: Brookings Institution.

Arnott, R. J., & McMillen, D. P. (Eds.). (2006). *A companion to urban economics.* Malden, MA: Blackwell.

Brueckner, J. (2001). Urban sprawl: Lessons from urban economics. In W. G. Gale & J. Rothenberg Pack (Eds.), *Brookings-Wharton papers on urban affairs 2001* (pp. 65–98). Washington, DC: Brookings Institution.

Brueckner, J. (2006). Strategic interaction among governments. In R. J. Arnott & D. P. McMillen (Eds.), *A companion to urban economics* (pp. 332–347). Malden, MA: Blackwell.

Carlino, G. A., Chatterjee, S., & Hunt, R. M. (2007). Urban density and the rate of invention. *Journal of Urban Economics, 61,* 389–419.

Charles, C. Z. (2005). Can we live together? Racial preferences and neighborhood outcomes. In X. Briggs (Ed.), *The geography of opportunity: Race and housing choice in metropolitan America* (pp. 45–80). Washington, DC: Brookings Institution.

Charles, K. K., & Hurst, E. (2002). The transition to homeownership and the black-white wealth gap. *Review of Economic Statistics, 84,* 281–297.

Christaller, W. (1933). *Central places in southern Germany.* London: Prentice Hall.

Cutler, D. M., & Glaeser, E. L. (1997). Are ghettos good or bad? *Quarterly Journal of Economics, 112,* 827–872.

Diamond, D. B., & Tolley, G. S. (1982). The economic roles of urban amenities. In D. B. Diamond & G. S. Tolley (Eds.), *The economics of urban amenities* (pp. 3–54). New York: Academic Press.

DiPasquale, D., & Wheaton, W. C. (1996). *Urban economics and real estate markets.* Englewood Cliffs, NJ: Prentice Hall.

Fischel, W. (2001). *The homevoter hypothesis: How home values influence local government taxation, school finance, and land-use policies.* Cambridge, MA: Harvard University Press.

Fujita, M., Krugman, P., & Venables, A. J. (1999). *The spatial economy: Cities, regions, and international trade.* Cambridge: MIT Press.

Fujita, M., & Mori, T. (2005). Frontiers of the new economic geography. *Papers in Regional Science, 84,* 377–405.

Garreau, J. (1991). *Edge city: Life on the new frontier.* New York: Doubleday.

Glaeser, E. L. (2007). *The economic approach to cities* (NBER Working Paper No. 13696). Cambridge, MA: National Bureau of Economic Research.

Glaeser, E. L., Hanushek, E. A., & Quigley, J. M. (2004). Opportunities, race, and urban location: The influence of John Kain. *Journal of Urban Economics, 56,* 70–79.

Glaeser, E. L., & Kahn, M. E. (2003). *Sprawl and urban growth* (NBER Working Paper No. 9733). Cambridge, MA: National Bureau of Economic Research.

Glaeser, E. L., Kallal, H. D., Scheinkman, J. A., & Shleifer, A. (1992). Growth in cities. *Journal of Political Economy, 100,* 1126–1152.

Gyourko, J., Linneman, P., & Wachter, S. (1999). Analyzing the relationships among race, wealth, and home ownership in America. *Journal of Housing Economics, 8,* 63–89.

Helsley, R., & Sullivan, A. (1991). Urban subcenter formation. *Regional Science and Urban Economics, 21*(2), 255–275.

Henderson, J. V. (1974). The sizes and types of cities. *American Economic Review, 64,* 640–656.

Henderson, J. V. (1986). Efficiency of resource usage and city size. *Journal of Urban Economics, 19,* 47–70.

Henderson, J. V. (2003). Marshall's scale economies. *Journal of Urban Economics, 53,* 1–28.

Kain, J. (1968). Housing segregation, negro employment and metropolitan decentralization. *Quarterly Journal of Economics, 82,* 175–197.

Krugman, P. R. (1991). *Geography and trade.* Cambridge: MIT Press.

Lösch, A. (1940). *The economics of location.* New Haven, CT: Yale University Press.

Marshall, A. (1920). *Principles of economics.* London: Macmillan.

McDonald, J. F., & McMillen, D. P. (2007). *Urban economics and real estate: Theory and policy.* Malden, MA: Blackwell.

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

McMillen, D. P., & Smith, S. C. (2003). The number of subcenters in large urban areas. *Journal of Urban Economics, 53,* 321–338.

Mills, E. S. (1967). An aggregative model of resource allocation in a metropolitan area. *American Economic Review, 57,* 197–210.

Mills, E. S. (1972). *Studies in the structure of the urban economy.* Baltimore: Johns Hopkins University Press.

Mills, E. S., & Hamilton, B. W. (1989). *Urban economics* (4th ed.). Glenview, IL: Scott, Foresman.

Mills, E. S., & Lubuele, L. S. (1997). Inner cities. *Journal of Economic Literature, 35,* 727–756.

Munnell, A. H., Tootell, G. E. B., Browne, L. E., & McEneaney, J. (1996). Mortgage lending in Boston: Interpreting HMDA data. *American Economic Review, 86,* 25–53.

Muth, R. (1969). *Cities and housing.* Chicago: University of Chicago Press.

O'Sullivan, A. (2009). *Urban economics.* Boston: McGraw-Hill.

Richardson, H. W. (1979). *Regional economics.* Urbana: University of Illinois Press.

Rosenthal, S. S., & Strange, W. C. (2001). The determinants of agglomeration. *Journal of Urban Economics, 50,* 191–229.

Rosenthal, S. S., & Strange, W. C. (2003). Geography, industrial organization and agglomeration. *Review of Economics and Statistics, 85,* 377–393.

Rosenthal, S. S., & Strange, W. C. (2004). Evidence on the nature and sources of agglomeration economies. In J. V. Henderson & J. F. Thisse (Eds.), *Handbook of urban and regional economics* (Vol. 4, pp. 2119–2171). Amsterdam: Elsevier North-Holland.

Rosenthal, S. S., & Strange, W. C. (2006). The micro-empirics of agglomeration economies. In R. J. Arnott & D. P. McMillen (Eds.), *A companion to urban economics* (pp. 7–23). Malden, MA: Blackwell.

Ross, S. L. (1998). Racial differences in residential and job mobility: Evidence concerning the spatial mismatch hypothesis. *Journal of Urban Economics, 43,* 112–135.

Ross, S. L., Turner, M. A., Godfrey, E., & Smith, R. R. (2008). Mortgage lending in Chicago and Los Angeles: A paired testing study of the pre-application process. *Journal of Urban Economics, 63,* 902–919.

Ross, S. L., & Yinger, J. (2002). *The color of credit: Mortgage discrimination, research methodology, and fair-lending enforcement.* Cambridge: MIT Press.

Sirmans, G., MacPherson, D. A., & Zietz, E. (2005). The composition of hedonic pricing models. *Journal of Real Estate Literature, 13,* 3–43.

Small, K. A., & Song, S. (1994). Population and employment densities: Structure and change. *Journal of Urban Economics, 36,* 292–313.

Tiebout, C. (1956). A pure theory of local expenditures. *Journal of Political Economy, 64,* 416–424.

Turner, M. A., & Ross, S. L. (2005). How racial discrimination affects the search for housing. In X. Briggs (Ed.), *The geography of opportunity: Race and housing choice in metropolitan America* (pp. 81–100). Washington, DC: Brookings Institution.

Turner, M. A., Ross, S. L., Galster, G. C., & Yinger, J. (2002). *Discrimination in metropolitan housing markets: National results from phase 1 of HDS 2000.* Washington, DC: Department of Housing and Urban Development.

von Thunen, J. H. (1966). *The isolated state* (P. Hall, Ed., & C. M. Wartenberg, Trans.). Oxford, NY: Pergamon Press. (Original work published 1826)

Wachter, S. M., & Megbolugbe, I. F. (1992). Racial and ethnic disparities in homeownership. *Housing Policy Debate, 3,* 333–370.

Yezer, A. M. (2006). Discrimination in mortgage lending. In R. J. Arnott & D. P. McMillen (Eds.), *A companion to urban economics* (pp. 197–210). Malden, MA: Blackwell.

Yinger, J. (1995). *Closed doors, opportunities lost: The continuing costs of housing discrimination.* New York: Russell Sage.