We are very grateful to Ken Caplan who provided excellent research assistance, extensive comments on a variety of chapters, and an undaunted spirit in conquering the gremlins in mapping software.

Our greatest debt and gratitude is to Jean L. Cummings, who worked on this project from beginning to end. She compiled most of the empirical evidence, read and commented on every chapter through what must have seemed like an endless succession of drafts, and suggested ways of improving the presentation of both models and data. Throughout this endeavor, Jean was our colleague, friend, sounding board, and, at times, referee. To the extent that the reader finds any glimmers of clarity in this book, it is safe to assume that Jean was involved.

DENISE DIPASQUALE
WILLIAM C. WHEATON

CHAPTER 1

THE PROPERTY AND CAPITAL MARKETS

What is real estate? How big is the real estate sector? How does the market for the use of real estate differ from the market for real estate assets? In this chapter, our objectives are (1) to define *real estate,* (2) to describe how the real estate sector operates within the national economy, and (3) to distinguish between the market for real estate use (where space is rented or purchased for occupancy) and the market for real estate assets (where buildings are bought and sold as investments).

The most common definition of *real estate* is the national stock of buildings, the land on which they are built, and all vacant land. These buildings are used either by firms, government, nonprofit organizations, and so on, as workplaces, or by households as places of residence. When defined this way, the value of all real estate makes up the largest single component of national wealth.

The yearly value of new buildings put in place represents an annual investment in the nation's stock of capital. The dollar value of these new buildings also has been the largest single category of national investment in recent years. Investment in new buildings accounts for roughly 7 percent of gross domestic product (GDP). Of this 7 percent, about 60 percent is payments to the construction sector (for labor, construction equipment, etc.), with the remaining 40 percent going to the producers of building materials. It is important to note that land is not counted as part of investment or GDP since it is not a produced commodity. On the other hand, land is a national asset, and so the value of land beneath buildings should be counted as stock or wealth.

The distinction between real estate as space and real estate as an asset is most clear when buildings are not occupied by their owners. The needs of tenants and the type and quality of buildings available determine the rent for space in the market for property use. At the same time, buildings may be bought, sold, or exchanged between investors. These transactions occur in the capital market and determine the *asset* price of space. In this chapter, we present a simple analytic framework which illustrates the connections between the market for real estate space (the property market) and the market for real estate assets (the asset or capital market). As we will show later in this chapter, this same approach can be used for owner-occupied real estate where the user of space is also the investor in the asset.

Our view of the real estate market as actually two markets helps to clarify how different forces influence this important sector. If, for example, there is a sudden demand by foreign investors to purchase U.S. office buildings, the impact on rents is very different than if firms suddenly decide that they wish to rent more office space for their use. A reduction in long-term mortgage rates has just the opposite effect on house prices from that caused by a reduction in short-term rates for construction financing. Distinguishing between the property and capital markets provides a clearer understanding of how such forces operate in the real estate sector as a whole.

In addition, there is a methodological objective in this first chapter—to reacquaint the reader with simple supply and demand analysis. By considering how the markets for both real estate property and assets operate within a global economy, we recall the distinction between endogenous economic variables and exogenous economic forces.[1] Within the markets for real estate, economic variables like prices, sales, and output are all determined endogenously. The outcome of the market, however, may be strongly influenced by exogenous forces such as world trade, interest rates, or even climate. Studying how changes in exogenous forces affect endogenous variables is one of the most important and fundamental pursuits of economics. How this methodology is applied to the real estate markets is a major focus of this book.

THE SIZE AND CHARACTER OF U.S. REAL ESTATE

Private real estate is composed of all types of buildings as well as the land they sit on. Houses, office buildings, warehouses, and shopping centers all are clear examples of real estate. Other, less obvious examples include privately owned ice-skating rinks and aircraft hangars. These are the buildings that various service firms (e.g., recreation, lodging, and travel) need for their operations. Finally, petroleum refineries, steel mills, and utility power plants are partially real estate—some portion of these structures is considered industrial "plant," while the remainder is "equipment." Public real estate is composed of all government-owned office buildings, schools, firehouses, military barracks, and the like.

[1] By *exogenous forces* we mean factors that influence real estate market outcomes but are not influenced by the real estate market. For example, interest rates have profound impacts on real estate market outcomes, but the real estate market has very little if any impact on interest rates. *Endogenous variables* are measures of real estate market outcomes. Prices and rents are endogenous, which means that they are determined by the real estate market.

Like any economic variable, real estate can be measured both as a flow and as a stock. The *flow* of real estate is the value of new buildings put in place each year, less losses from the stock through depreciation or demolition. New completions represent an important component of national investment, which is also a flow variable. On the other hand, the total value of all existing buildings and the value of all land are *stock* variables, and are part of national wealth—also a stock variable. Since land is nonreproducible, it is always a stock variable and never a flow variable.

Table 1.1 breaks down the value of new construction put in place in 1990. Virtually all private construction was in the form of buildings and represented $301 billion (5.5

TABLE 1.1 Value of New Construction Put in Place, 1990

	$ (in billions)	% of GDP
Private Construction	338	6.1
Buildings	301	5.5
Residential Buildings	183	3.3
Nonresidential Buildings	118	2.1
Industrial	24	0.4
Office	29	0.5
Hotels/Motels	10	0.2
Other Commercial	34	0.6
All Other Nonresidential	21	0.4
Nonbuilding Construction	37	0.7
Public Utilities	31	0.6
All Other	6	0.1
Public Construction	109	2.0
Buildings	46	0.8
Housing and Development	4	0.1
Industrial	1	0.0
Other	41	0.7
Nonbuilding Construction	63	1.1
Infrastructure	55	1.0
All Other	8	0.1
Total New Construction	446	8.1
Total GDP:	5,514	100.0

Source: Current Construction Reports, Series C30, U.S. Bureau of the Census; gross domestic product, *Economic Report of the President 1992*.

percent of GDP). Residential buildings accounted for about 61 percent of the dollar value of private buildings constructed, with office, industrial and other commercial structures representing the remainder. In the public sector, buildings represented only about 42 percent of the $109 billion in public construction. The largest component of public construction was investment in infrastructure (roads, bridges, airports, etc.).

Total gross private domestic investment, which measures purchases of fixed capital goods including structures and equipment, was $803 billion in the U.S. in 1990. (Board of Governors 1991). In these figures, structures include purchases of new residential and nonresidential structures, as well as net purchases of existing structures. The structure component accounted for just over half of the nation's gross investment ($415 billion), while the remaining $388 billion was investment in machinery or equipment.

Over the years, government statistics have tracked the flow of gross investment into real estate (i.e., new building construction) with a high degree of accuracy. Valuing the total real estate stock at any point in time, however, is far more difficult. The Bureau of Economic Analysis (BEA) provides estimates of "fixed reproducible tangible wealth," which attempt to value the stock of both residential and nonresidential structures. Estimates of gross stock are calculated by summing investment flows and deducting the cumulated investment in structures that have been discarded. The investment flow data are from the national income and product accounts. Net capital stock is calculated by using a depreciation formula to write off the value over the life of the stock. The Federal Reserve uses the BEA estimates of net stock and adds their estimates of the value of land to provide balance sheet estimates of the value of U.S. real estate (Miles 1990).

An important problem with this approach to value is that the cumulated investment flow less demolition and depreciation may bear little resemblance to the actual market value of real estate. While the investment flow data may be a fairly accurate measure of the value of additions to the stock in the current year, changes in the value of those additions to the stock over time are a function of market forces that are unlikely to be captured even by the most sophisticated estimates of depreciation. In addition, it is difficult to estimate the value of the land on which those structures are built because most observed transactions provide a single purchase price for the existing building and the land, with no breakdown by the land and structure components.

A 1991 study made a gallant attempt at estimating the value of all U.S. real estate according to type of real estate (e.g., residential, office, and retail) and type of owner (e.g., individuals, corporations) by piecing together data from a variety of sources (IREM 1991). The study employed standard government statistics, such as the American Housing Survey (AHS) for residential units, and trade association data such as those collected by the International Council of Shopping Centers. In addition, it used state and county property tax records as a basis for regression models that estimate real estate value. Finally, the study relied on interviews with industry experts to augment their data.

In this study, total real estate in the U.S. for 1990 was estimated to be worth $8.8 trillion. As shown in Table 1.2, almost 70 percent of all U.S. real estate was residential, and almost 90 percent of the value of residential real estate was in the nation's stock of single-family homes. The 30 percent of U.S. real estate that was nonresidential was dominated by office and retail space—at least in dollar value. While estimates of total national

TABLE 1.2 Value of U.S. Real Estate, 1990

	$ (in billions)	% of Total
Residential	6,122	69.8
Single-Family Homes	5,419	61.7
Multifamily	552	6.3
Condominiums/Coops	96	1.1
Mobile Homes	55	0.6
Nonresidential	2,655	30.2
Retail	1,115	12.7
Office	1,009	11.5
Manufacturing	308	3.5
Warehouse	223	2.5
Total U.S. Real Estate	8,777	100.0

Source: Managing the Future: Real Estate in the 1990s (Chicago IREM Foundation and Arthur Anderson, 1991), pp. 28, 31.

wealth often vary, the Federal Reserve estimates that national net worth was at $15.6 trillion in 1990. Thus, the figures in Table 1.2 suggest that real estate, valued at $8.8 trillion, constituted roughly 56 percent of the nation's wealth in 1990.[2]

Who owns American real estate? The legal ownership status of this wealth is shown in Table 1.3. In 1990, 83 percent of residential real estate was owned by individuals. It is important to remember that this figure covers not only individual ownership of personal residences but also sole-proprietor ownership of apartment buildings. Similarly, the 62 percent of nonresidential real estate that was owned by corporations consists of investment property as well as buildings occupied by their corporate owners. Partnerships own almost equal shares of residential and nonresidential property. Finally, the ownership of U.S. real estate by foreign entities is very small, despite considerable concern about this during the last few years.

In summary, U.S. real estate is the largest single component of national wealth and the largest component of annual net private investment. This huge base of assets, however, has been accumulated by devoting only about 5 to 7 percent of each year's GDP to the construction and renovation of that base. It is, of course, the durability of real estate that allows us to devote such a small fraction of GDP to the accumulation and maintenance of such a large share of our assets.

[2]The national net worth estimate is from *Balance Sheets for the U.S. Economy 1945–1990* (Washington, D.C.: Board of Governors of the Federal Reserve System, 1991). It should be noted that the Federal Reserve estimates the value of real estate at $10.7 trillion. Miles (1991:74) argues that the BEA/Federal Reserve estimates of the value of nonresidential real estate may be high because the data used for the stock estimates include special fixtures in manufacturing plants that are certainly part of the capital stock, but should not be included when measuring the value of real estate.

TABLE 1.3 Who Owns U.S. Real Estate, 1990

	All Real Estate		Residential Only		Nonresidential Only	
	$ (in billions)	%	$ (in billions)	%	$ (in billions)	%
Individuals	5,088	58.0	5,071	82.8	17	0.6
Corporations	1,699	19.4	66	1.1	1,633	61.5
Partnerships	1,011	11.5	673	11.0	338	12.7
Nonprofits	411	4.7	104	1.7	307	11.6
Government	234	2.6	173	2.8	61	2.3
Institutional Investors	128	1.5	14	0.2	114	4.3
Financial Institutions	114	1.3	13	0.2	101	3.8
Other (Includes Foreign)	92	1.0	8	0.1	84	3.2
Total:	8,777	100.0	6,122	100.0	2,655	100.0
% of All Real Estate:		100.0		69.8		30.2

Source: Managing the Future: Real Estate in the 1990s (Chicago: IREM Foundation and Arthur Anderson, 1991), pp. 29–33.

THE MARKETS FOR REAL ESTATE ASSETS AND REAL ESTATE USE

Since real estate is a durable capital good, its production and price are determined in an asset, or capital, market. In this market, the demand to own real estate assets must equal their supply. Thus, the price of houses in the U.S. largely depends on how many households wish to *own* units and how many units are available for ownership. Likewise, the value or price of shopping center space depends on how many investors wish to own such space and how many centers there are available to invest in. In both cases, all else being equal, an increase in the demand to own these assets will raise prices, while a greater supply of space will depress prices.

The supply of new real estate assets comes from the construction sector and depends on the price of those assets relative to the cost of replacing or constructing them. In the long run, the asset market should equate market prices with replacement costs that include the cost of land. In the short run, however, the two may diverge significantly because of the lags and delays that are inherent in the construction process. For example, if demand for the ownership of space suddenly rises, then, with a fixed supply of assets, prices will rise as well. With prices now above construction and land costs, new development takes place. As this space arrives on the market, demand is satisfied and prices begin to fall back towards the cost of replacement. A question to be addressed in future chapters is whether the cost of asset replacement is constant, varies with the level of development or depends on the total stock of assets.

What would cause the demand for owning real estate assets to suddenly increase? More generally, are there other determinants of asset demand besides simply the price of these assets? The answer is yes, and the most important of these determinants is the rental income that real estate assets earn. To understand rent, it is necessary to consider the market for the *use* of real estate.

In the market for real estate use or space (referred to here as the property market), demand comes from the *occupiers* of space, whether they be tenants or owners, firms or households. For firms, space is one of many factors of production, and, like any other factor, its use will depend on firm output levels and the relative cost of space. Households likewise divide their income into the consumption of many commodities, only one of which is space. The household demand for space depends on income and the cost of occupying that space relative to the cost of other commodities, such as food, clothing, or entertainment. For firms or households, the cost of occupying space is the annual outlay necessary to obtain the use of real estate—its *rent.* For tenants, rent is simply specified in a lease agreement. For owners, rent is defined as the annualized cost associated with the ownership of property.

Rent is determined in the property market for space use, not in the asset market for ownership. In the property market, the supply of space is given (from the asset market). The demand for space depends on rent and other exogenous economic factors such as firm production levels, income levels, or the number of households. The task of the property market is to determine a rent level at which the demand for space use equals the supply of space. All else being equal, when the number of households increases or firms expand production, the demand for space *use* rises. With fixed supply, rents rise as well.

The link between the asset market and the property market occurs at two junctions. First, the rent levels determined in the property market are central in determining the demand for real estate assets. After all, in acquiring an asset, investors are really purchasing a current or future income stream. Thus, changes in rent occurring in the property market immediately affect the demand for ownership in the asset market. The second link between the two markets occurs through the construction or development sector. If construction increases and the supply of assets grows, not only are prices driven down in the asset market, but rents decline in the property market as well. These connections between the two markets are illustrated in the four-quadrant diagram in Figure 1.1.

In explaining Figure 1.1, it is useful to refer to quadrants by their compass designation. The two right-hand quadrants (northeast and southeast) represent the property market for the use of space, while the two left-hand quadrants (northwest and southwest) deal with the asset market for the ownership of real estate. Let's begin with the northeast quadrant, where rents are determined in the short run.

The northeast quadrant has two axes: rent (per unit of space) and the stock of space (also measured in units of space, such as square feet). The curve represents how the demand for space depends on rents, given the state of the economy. Movement along that curve depicts how much space would be demanded given a particular rent level on the vertical axis. If households or firms tend to demand the same amount of space regardless of rent levels (inelastic demand), then the curve is nearly vertical. If space usage is quite sensitive to rents (elastic demand), then the curve is more horizontal. If the economy changes, then the entire curve shifts. An upward shift occurs with an increase in firms or households (economic growth) and signifies that more space is demanded for the same rent. Economic decline causes a downward shift in the line, with the reverse implications.

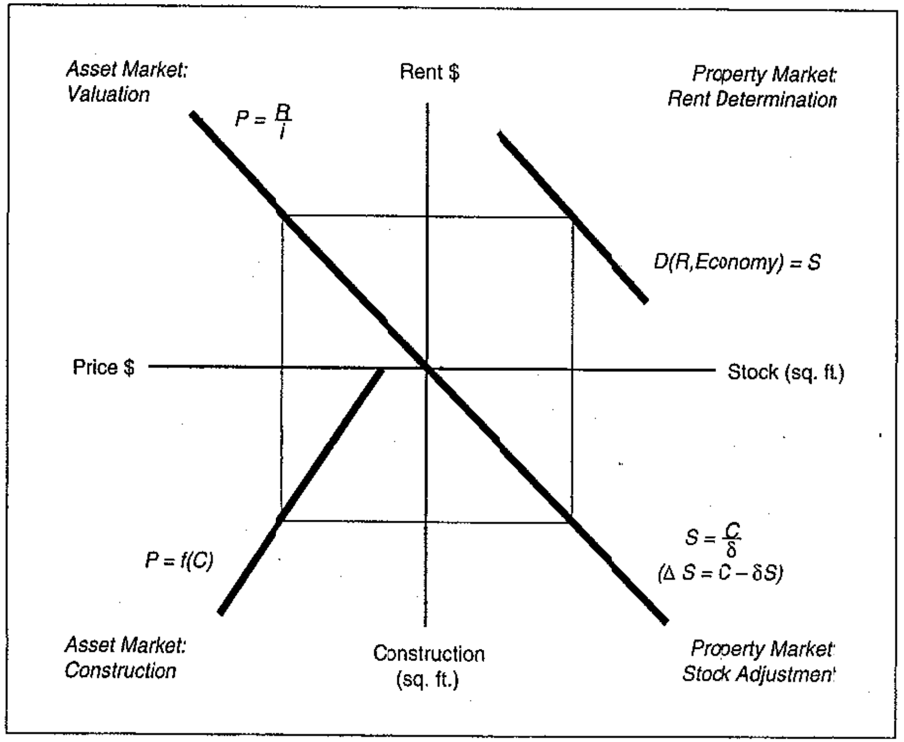

FIGURE 1.1 Real estate: the property and asset markets.

In equilibrium the demand for space, D, is equal to the stock of space, S. Thus, rent, R, must be determined so that demand is exactly equal to stock. Demand is a function of rent and conditions in the economy:

$$D(R, \text{Economy}) = S \tag{1.1}$$

Recall that in the property market the supply of stock is given from the asset market. In Figure 1.1, then, rent is determined by taking a level of stock of space on the horizontal axis, drawing a line up to the demand curve, and then over to the vertical axis. With this rent for the use of space, we then move to the asset market in the northwest quadrant.[3]

The northwest quadrant represents the first part of the asset market and has two axes: rent and price (per unit of space). The ray emanating from the origin represents the capitalization rate for real estate assets: the ratio of rent-to-price. This is the current yield that investors demand in order to hold real estate assets. Generally, four considerations make up this capitalization rate: the long-term interest rate in the economy, the expected growth in rents, the risks associated with that rental income stream, and the treatment of real estate in the U.S. federal tax code. A higher capitalization rate is represented by a clockwise rotation in the ray, while a decline in the cap rate is represented by a counterclockwise rotation. In this quadrant, the capitalization rate is taken as exogenous, based on interest rates and returns in the broader capital market for all assets (stocks, bonds, short-term deposits). Thus, the purpose of the northwest quadrant is to take the rent level, R, from the northeast quadrant and determine a price for real estate assets, P, using a capitalization rate, i:

$$P = \frac{R}{i} \tag{1.2}$$

This is done by moving from the rent level on the vertical axis in the northeast quadrant over to the ray in the northwest quadrant, and then down to the horizontal axis where asset price is given.

The next (southwest) quadrant is that portion of the asset market in which the creation of new assets is determined. Here, the curve $f(C)$ represents the replacement cost of real estate. In this version of the diagram, the cost of replacement through new construction is assumed to increase with greater building activity (C), and so the curve moves in a southwesterly direction. It intersects the price axis at that minimum dollar value (per unit of space) required to get some level of new development underway. If this construction can be supplied at any level with almost the same costs, then the ray will be close to vertical. Construction bottlenecks, scarce land, and other impediments to development lead to inelastic supply and a ray that is more horizontal. Given the price of real estate assets from the northwest quadrant, a line down to the replacement cost curve and then over to the vertical axis determines the level of new construction where replacement costs equal asset prices.[4] Lower levels of construction would lead to excess profits, whereas higher levels would be unprofitable. Hence, new construction occurs at that level, C, at which asset price, P, is equal to replacement costs, $f(C)$:

$$P = f(C) \tag{1.3}$$

In the southeast quadrant, the annual flow of new construction, C, is converted into a long-run stock of real estate space. The change in the stock, ΔS, in a given period is equal to new construction minus losses from the stock measured by the depreciation (removal) rate, δ:

$$\Delta S = C - \delta S \tag{1.4}$$

[3]If we take the office market as an example, it would be reasonable to assume that the demand for office space use (in square feet), D, is equal to: $E(400 - 10R)$, where E is the number of office workers in the economy (in millions), and R is the annual rent per square foot. In the northeast quadrant, we equate this demand to the existing stock, S, solving for rent. Equation (1.1) can be rewritten as:

$$R = 40 - \frac{S}{10E}$$

[4]For the office market example, a reasonable cost function (dollar per square foot) for the development of new office space would be: $200 + 5C$. The annual level of new construction, C, is in millions of square feet. If these costs are equated to the asset price (per square foot) of office space, Equation (1.3) can be rewritten to solve for the level of construction where costs equal asset prices:

$$C = \frac{P - 200}{5}$$

The ray emanating from the origin represents that level of stock (on the horizontal axis) that requires an annual level of construction for replacement just equal to that value on the vertical axis. At that level of stock and corresponding level of construction, the stock of space will be constant over time, since depreciation will equal new completions. Hence, ΔS is equal to 0 and $S = C/\delta$.[5] Future chapters will discuss this relationship in more detail; for now, it is important only to remember that the southeast quadrant assumes a certain level of construction and determines the level of stock that would result if that construction continued forever.

This completes a 360-degree rotation around the four-quadrant diagram. Starting with a level of stock, the property market determines rents, which are then translated into property prices by the asset market. These asset prices, in turn, generate new construction, which, back in the property market, eventually yields a new level of stock. The combined property and asset markets are in equilibrium when the starting and ending levels of stock are the same. If the ending stock differs from the starting stock, then the values of the four variables in the diagram (rent, prices, construction, and stock) are not in complete equilibrium. If the starting value exceeds the finishing value, then rents, prices, and construction must all rise to be in equilibrium. If the initial stock is less than the finishing stock, then rents, prices, and construction must be decreased to be in equilibrium. This journey around the four-quadrant diagram provides a simple, intuitive illustration of the solution to the simultaneous system of Equations (1.1)–(1.4).[6]

OWNER-OCCUPIED REAL ESTATE

A reasonable question is how all of this works in the case of real estate that is mainly occupied by its owner. In this case, the four quadrants still hold, but asset prices and rents are determined by the same market participants—the owner occupants. Consider for the moment the market for owner-occupied housing. The demand for single-family homes depends on the number of households, their incomes, and the annual costs of owning a home. This annual cost is equivalent to rent. A rise in the number of households shifts the demand curve out. With greater demand and a fixed stock of housing units, the annual payment to occupy a house must rise. The northwest quadrant then translates this payment into the actual price that households are willing to pay for the home.

Lower interest rates, for example, imply that with the same annual payment (rent), households can afford to pay a higher purchase (asset) price. Hence, with owner-occupied real estate, decisions by the user-owner determine both rent (the annual payment) and price. These decisions, however, are influenced by the same economic and capital market conditions that influence rental properties. Thus, owner-occupiers have the same investment motives as the owners of rental property. Once the purchase price is determined, then new housing development and eventually a new equilibrium stock of space follow in the other two quadrants (southwest, southeast).

This same analysis can be applied to the case of corporate-owned and -occupied industrial or office space. The demand for this space is determined by the annual cost of owning it (rent) as well as the number and size of firms in the market. In equilibrium, the annual cost of owning (rent) equates demand with the fixed stock of office or industrial space. The cost of capital for the corporation (capitalization rate) converts this annual cost into the asset price that corporations are willing to pay for the space.

[5] The increase each year in the stock of office space is the difference between construction and depreciation. Assuming depreciation, δ, is 1 percent annually, then $\Delta S = C - 0.01S$. Given a stable level of construction, the stock that will eventually emerge is: $S = C/0.01$. This is the steady-state version of Equation (1.4) and is the ray emanating out of the southeast quadrant.

[6] If the equations in footnotes 3–5 plus Equation (1.2) are successively substituted into each other, the result is:

$$S = \frac{800 - 2\frac{S}{E}}{i} - 4{,}000$$

Solving so that S is only on the left-hand side of the equation, we get the equilibrium stock level:

$$S = E\frac{[800 - 4{,}000i]}{iE + 2}$$

Thus, if there are 10 million office workers in the economy ($E = 10$) and the interest or capitalization rate is 5% ($i = 0.05$), the long-run stock of office space will be 240 square feet per worker, or 2,400 million square feet. This stock will require that 24 million square feet be constructed each year, and this will require that asset prices equal a replacement cost of $320 per square foot. Rents of $16 annually will sustain such asset prices with a 5% capitalization rate.

REAL ESTATE AND THE NATIONAL ECONOMY

Using Figure 1.1, we can trace the various impacts of the broader economy on the real estate market. The economy can grow or contract. Long-term interest rates or other factors can shift the demand for real estate assets. Changes in short-term credit availability or local regulations can alter the cost of supplying new space. Each has different repercussions, and these are easily determined by examining alternative solutions within the four-quadrant diagram. In each case, we can identify which quadrant is initially affected, trace the impacts through the other quadrants, and arrive at a new long-run equilibrium. This comparison of different long-run solutions (market equilibria) in a model is called "comparative static" analysis.

Economic Growth and the Demand for Real Estate Use

As the economy expands, the curve in the northeast quadrant shifts out to the northeast. This reflects a greater demand to use space at current (or other) rent levels, such as would occur with increases in production, household income, or the number of households. For a given level of real estate space, rents must therefore rise if the demand to use space is to equal available space. These higher rents then lead to greater asset prices in the northwest quadrant, which, in turn generate a higher level of new construction in the southwest quadrant. Eventually, this leads to a greater stock of space (southeast quadrant). As shown

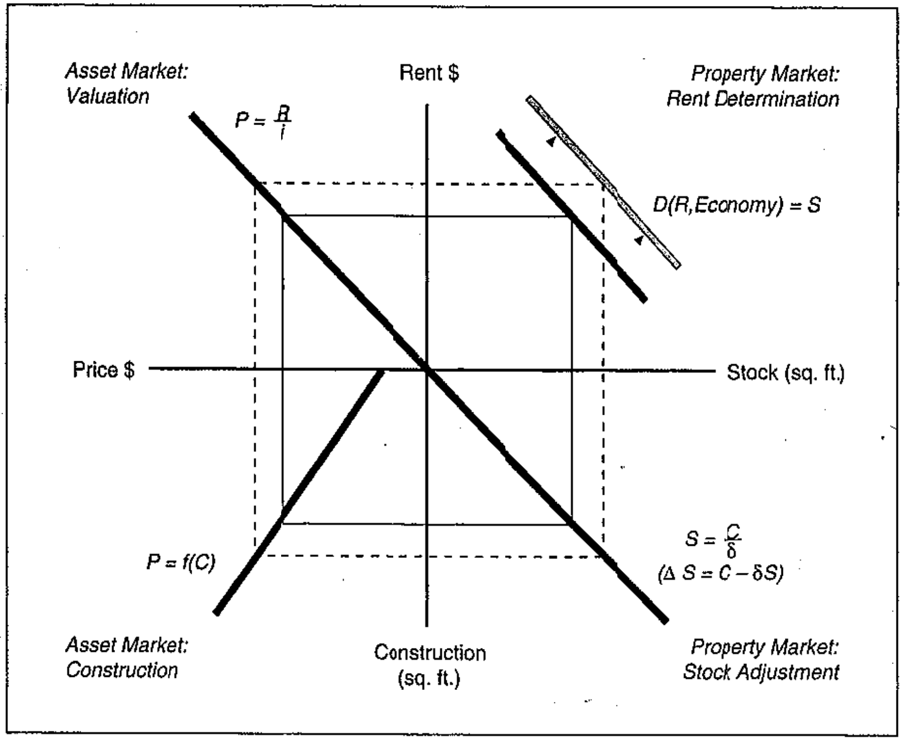

FIGURE 1.2 The property and asset markets: property demand shifts.

in Figure 1.2, the new market equilibrium involves a dashed box that in every direction lies outside of the box that connected the four curves in the original equilibrium.[7]

It should be clear that the equilibrium solution with an expanded economy must generate a box that lies outside of the original. Neither rents, prices, construction, nor stock can be less than in the initial equilibrium. The new solution, however, need not at all involve a proportional expansion of the original box. The shape of the new box depends on the slopes of the various curves. For example, if construction were very elastic with respect to asset prices (a nearly vertical curve in the southwest quadrant), then the new levels of prices and rents would be only slightly greater than before, whereas construction and stock would expand considerably.

[7]We can use growth in the number of office workers as an example of economic expansion. Returning to the equation in footnote 6, if the number of office workers increased from 10 million to 20 million, the long-run stock of space would increase to 4,000 million square feet, or just 200 square feet per worker. Rents would increase to $20 per square foot annually. Such rents would generate asset prices of $400 per square foot, which would just equal the replacement costs necessary to sustain the higher level of annual construction (40 million sq. ft.).

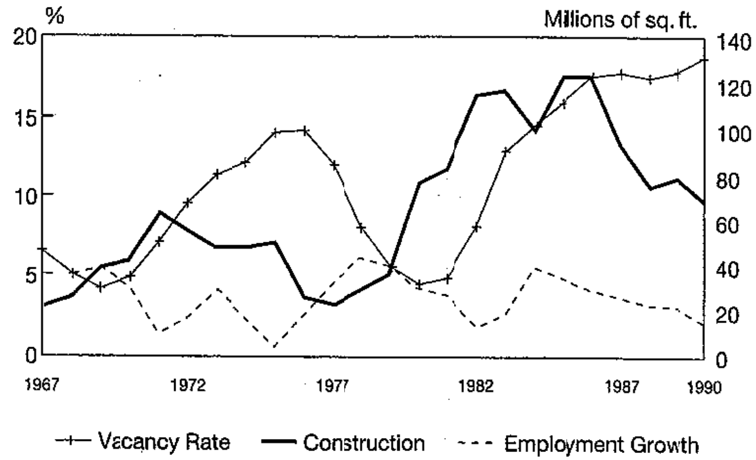

FIGURE 1.3 Office employment growth, vacancy rate, and construction, 1967–1990.

These data are aggregated from 30 metropolitan areas.

Source: Employment, adjusted U.S. government figures courtesy of Regional Financial Associates, Bala-Cynwyd, PA; vacancy and construction, CB Commercial.

Economic growth, then, increases all equilibrium variables in the real estate market(s), while economic contraction leads to decreases in all variables. Figure 1.3 compares the growth of total employment in the U.S. with construction of office space and the overall office vacancy rate. It is clear that the national office market moves with the economy: during recessions, vacancies tend to rise and construction falls, whereas the opposite occurs during recoveries.

Long-Term Interest Rates and the Demand for Real Estate Assets

If the demand to *own* real estate shifts, the impact on the combined markets is quite different than if the demand to *use* real estate changes. A number of factors can cause shifts in the demand to own real estate assets. If interest rates in the rest of the economy rise (fall), then the existing yield from real estate becomes low (high) relative to fixed income securities and investors will wish to shift their funds from (into) the real estate sector. Similarly, if the risk characteristics of real estate are perceived to have worsened (improved), then the existing yield from real estate may also become insufficient (more than necessary) to get investors to purchase real estate assets relative to other assets. Finally, changes in how real estate income is treated in the U.S. federal tax code can also greatly impact the demand to invest in real estate. As will be discussed in Chapter 8, if the depreciation allowance for real estate is increased, the same income stream generates a higher after-tax yield. This will increase the demand to hold real estate assets.

This book assumes that the capital market efficiently adjusts the prices of particular assets—so that each investment earns a common risk-adjusted, after-tax total rate of return. Thus, shifts in asset demand, such as described above, will alter the capitalization rate at which investors are willing to hold real estate. Reductions in long-term interest rates, decreases in the perceived risk of real estate, and generous depreciation or other favorable changes in the tax treatment of real estate all will reduce the yield that investors require from real estate. As shown in Figure 1.4 in the northwest quadrant, this has the effect of a counterclockwise rotation (the ray always goes through the origin) in the capitalization rate ray emanating from the origin, raising asset prices. Higher interest rates, greater perceived risk, and adverse tax changes rotate the ray in a clockwise manner, lowering asset prices.

Given a level of rent from the property market, a reduction in the current yield or capitalization rate for real estate raises asset prices, and in the southwest quadrant, this begins to expand construction. Eventually this increases the stock of space (in the southeast quadrant), which then lowers rents in the property market for space (northeast quadrant). A new equilibrium requires that the initial and finishing rent levels be equal. In

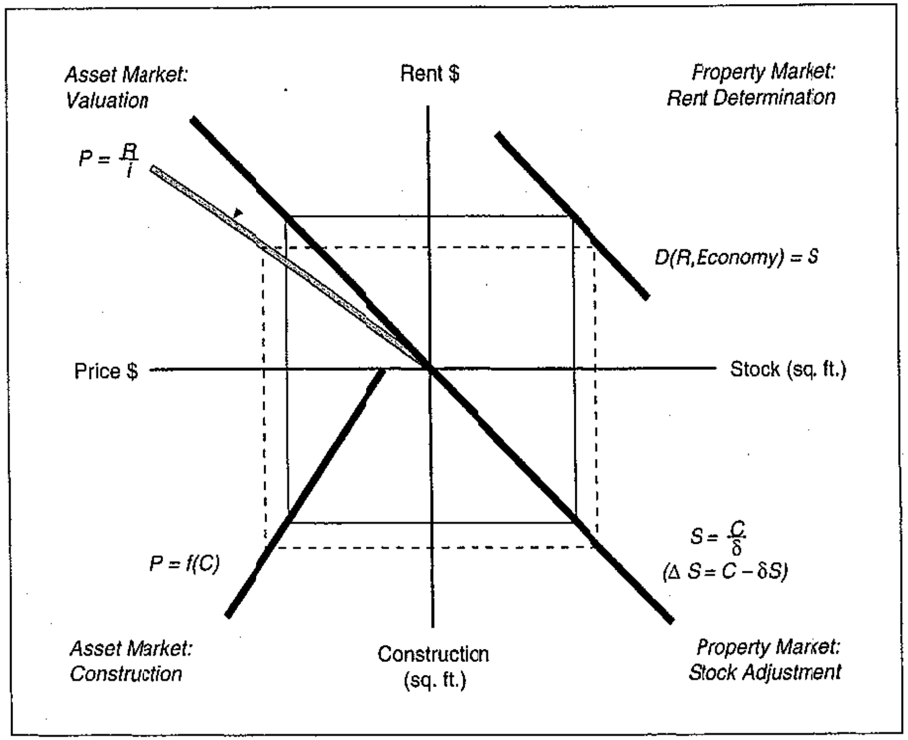

FIGURE 1.4 The property and asset markets: asset demand shifts.

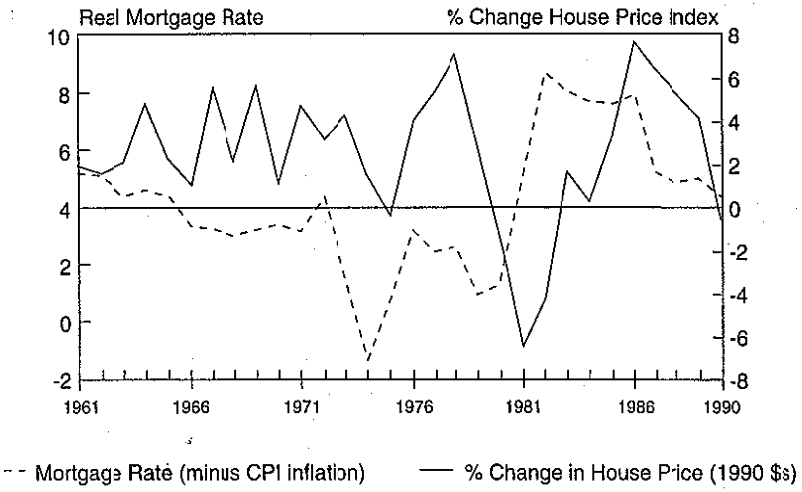

FIGURE 1.5 Change in house price versus mortgage rates (real), 1961–1990.

Source: Price index: 1960–69: Federal Housing Finance Board, *1990 Rates & Terms on Conventional Home Mortgages, Annual Summary.* 1970–90: Federal Home Loan Mortgage Corp., *Quality-Controlled Existing Home Price Index*, unpublished report. Mortgage rates: 1960–62: FHA insured rate plus .44 percentage points. 1963–90: Federal Housing Finance Board, *1990 Rates & Terms on Conventional Home Mortgages, Annual Summary.*

Figure 1.4, this new equilibrium results in a new solution box that is lower and more rectangular than the original.

It is important to be convinced that the new solution box is as portrayed in Figure 1.4. Asset prices must be higher and rents lower, while the long-term stock and its supporting level of construction must be greater. If rents were not lower, the stock would have to be the same (or lower), and this would be inconsistent with higher asset prices and greater construction. If asset prices were not higher, rents would be lower, which is inconsistent with the reduced stock (and less construction) generated by lower asset prices. A positive shift in asset demand, like a positive shift in space demand, will raise prices, construction, and the stock. It will, however, eventually lower—rather than raise—the level of rents. This inverse relationship between asset prices and long-term interest rates can be seen in Figure 1.5, which examines the historic movements in house prices (in 1990 dollars) and real mortgage rates (adjusted for inflation).

Credit, Construction Costs, and the Supply of New Space

The final exogenous change likely to impact the real estate market is a shift in the supply schedule for new construction. This can come about through several channels. Higher short-term interest rates and a general scarcity of construction financing will increase the costs of providing a given amount of new space, leading to less construction. Likewise, stricter local zoning or other building regulations will also add to development costs and (for a fixed level of asset prices) reduce the profitability of new construction. These kinds of negative supply changes have the effect of causing a westerly shift in the cost schedule of the southwest quadrant: for the same level of asset prices, construction will be less. Positive changes in the supply environment, such as the easy availability of construction financing or a relaxation of development regulations, move the curve in an easterly direction, and for the same asset price expand construction. Figure 1.6 shows the powerful inverse relationship that has existed historically between short-term real interest rates (inflation-adjusted) and new home construction.

Finally, Figure 1.7 traces the long-run implications of a negative supply change, such as would occur with higher short-term interest rates.[3] For a given level of asset prices, a negative shift in the new space supply schedule (southwest quadrant) will lower the level of construction and eventually lower the stock of space (southeast quadrant). With less space in the northeast quadrant, rent levels will have to rise, which will generate higher asset prices in the northwest quadrant. When starting and finishing asset prices

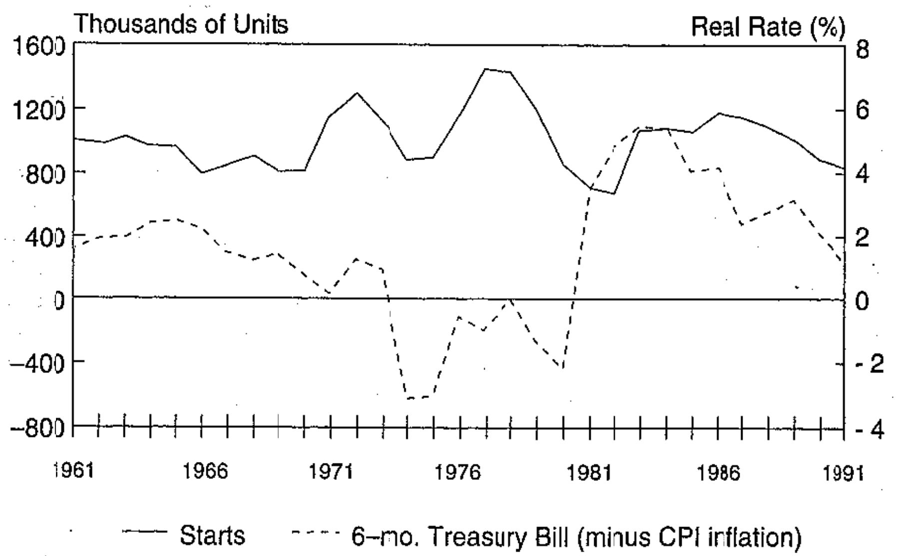

FIGURE 1.6 Single-family construction and real interest rates, 1961–1991.

Source: Construction, *Economic Report of the President 1992*; interest rates, The Federal Reserve.

[3]In Figure 1.7, this negative supply change is drawn as a shift in the cost curve, which means that the minimum level of asset price necessary to get construction going has increased. There may also be supply changes that leave this minimum asset price the same, but lower or raise the amount of construction resulting from increases in asset prices. Such changes would result in a pivot in the schedule, changing its slope (i.e., if the slope becomes flatter there is less construction with increases in asset prices; if the slope becomes steeper, there is more construction with increases in asset prices).

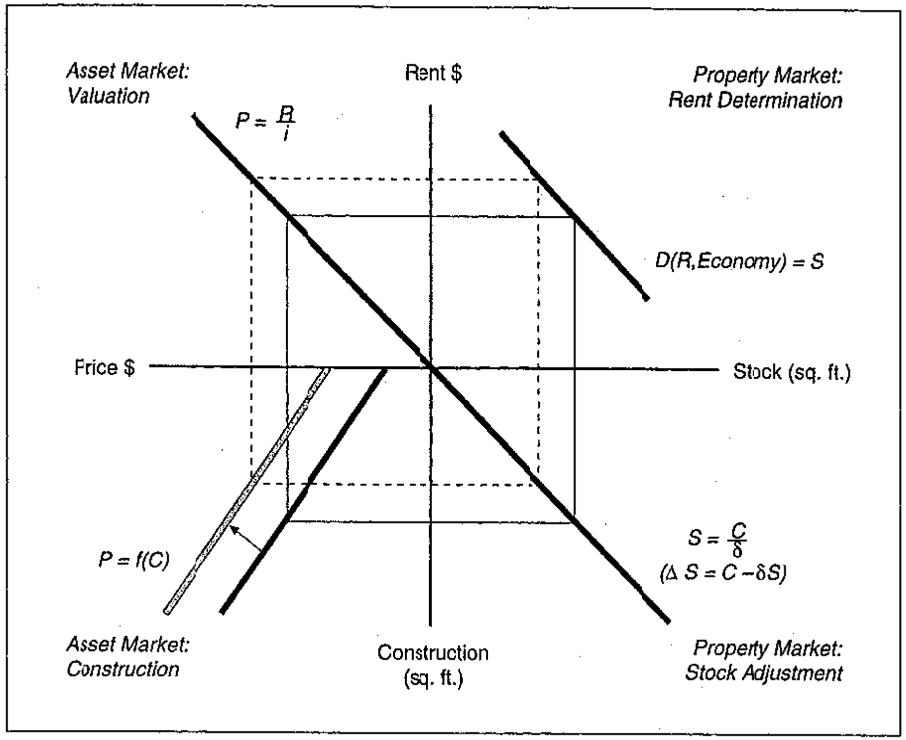

FIGURE 1.7 The property and asset markets: asset cost shifts.

are equal, the new solution box will lie strictly to the northwest of the original solution. Rents and asset prices will increase, whereas construction and stock levels will be less. The magnitude of these changes, of course, will depend on the slopes (or elasticities) of the various curves. For example, if the demand for space is very rent elastic (a nearly horizontal curve in the northeast quadrant), then the increase in rents will be slight. An inelastic (nearly vertical) demand curve will generate a larger increase in rents. It should be clear that if any of these variables moved in a different direction, the solution would be inconsistent—that is, not an equilibrium.

Occasionally, single, individual shifts, such as are portrayed in Figures 1.2, 1.4, and 1.7, do occur by themselves. For example, the Economic Recovery Tax Act of 1981 (ERTA), greatly shortened the depreciation period for rental housing. This generated a sharp reduction in the capitalization rate for this asset, and the ensuing construction boom and fall in rents are quite consistent with Figure 1.4. Some researchers are also convinced that much of the commercial building boom and eventual fall in office rents and asset prices during the 1980s was the result of deregulating the nation's thrift (savings and loan, or S&L) industry. Deregulation is argued to have resulted in a dramatic increase in the availability of cheap construction credit for new commercial development. As such, it would have caused a singular (easterly) shift in the construction cost schedule.

It is more likely the case that economic events cause several shifts to occur simultaneously. This is particularly true of movements in the nation's macroeconomy. As the national economy enters a slowdown, not only is there a contraction in output and employment (northeast quadrant), but there are usually increases in short-term interest rates as well (southwest quadrant). An economic expansion leads to the opposite combination. This combination of shifts can generate any pattern of new box solutions that lies between the two shown in Figures 1.2 and 1.7. Although the analysis gets more complicated in the case of multiple shifts, the net outcome is always some combination of the impacts from each individual change.

The simple framework represented by the four-quadrant diagram works well in illustrating the new equilibria that result as the exogenous environment changes. An important drawback of this framework is that it is not easy to trace the intermediate steps as the market moves to its new equilibrium. A dynamic system of equations is needed to depict the intermediate adjustments of the market, which significantly complicates our analysis. More complicated dynamic models will be developed in Chapters 10 and 12.

REAL ESTATE MARKETS AND PUBLIC POLICY

The real estate sector is affected by public policy changes at the federal, state, and local levels. However, the real estate sector may or may not be the primary target of a specific policy. For example, national monetary policy has several goals and implications. As a prime determinant of interest rate movements, monetary policy clearly affects the demand for real estate assets as well as the level of new construction. The four-quadrant diagrams presented in the previous section illustrate the impacts of such policy changes on the real estate markets. The public sector also creates policies that are aimed specifically at the real estate sector. In developing such policies, government decision makers must be able to anticipate the impacts of their actions on the broader real estate market. The framework represented by the four-quadrant diagram permits a useful analysis of the impacts of these policies. Consider the following examples of major public policies designed to impact the real estate markets directly.

Publicly Assisted Housing

Both federal and state governments provide a range of assistance programs to encourage the development of low- and moderate-income housing. Some programs directly build units for targeted groups, while other programs assist households in making rental payments. The construction of publicly owned units generally decreases the demand for privately owned rental units. This decrease in demand assumes that public units succeed in attracting tenants. This seems a safe assumption since public units generally carry subsidies, which mean that the units are offered at below market rents. In fact, in many urban areas in the U.S., there are long waiting lists for public housing units. This decrease in demand for private units results in an inward shift in the demand curve in the northeast quadrant which will, in turn, decrease rents, asset prices, construction and ultimately the stock of private units. Thus, building public housing produces a new equilibrium that is exactly the opposite of that illustrated in Figure 1.2, where demand is increasing. The reduction of private building activity due to public construction programs is sometimes referred to as public displacement of private construction.

Rental assistance programs, on the other hand, act to simulate housing demand in much the same way as an expansion in the economy. These programs will shift out the demand curve in the northeast quadrant, resulting in increases in rents, prices, construction, and the stock, as occurred in Figure 1.2. Proponents of rental assistance programs argue that such programs will encourage construction and have only a modest impact on rents. Hence, they argue that the net effect of such programs will be to expand the housing opportunities of low-income households. Opponents of such policies suggest that rental assistance only serves to increase rents in the market and has very little impact on construction. Hence, they argue that the main beneficiaries of the program are landlords. Clearly, the size of the impact on construction and rents depends on the relative elasticities of the curves presented in Figure 1.2.

Local Government Development Regulations

In the U.S., local governments exercise great control over the amount and type of development permitted on privately owned land. Such regulations often are in the public interest, but they do impose two additional costs on private development. First, they frequently extend the time necessary for completion of a project, since local governments require developers to apply for various permits in order to proceed. Second, they sometimes act to create a scarcity of sites. This can drive up land prices and add to site-acquisition costs. The more binding or restrictive such regulations become, the more they increase development costs. This increase in costs will shift the supply schedule in the southwest quadrant in a westerly direction, as occurred in Figure 1.7.

The Taxation of Real Estate

The federal tax code generally treats real estate favorably in several ways. Interest payments on real estate debt are fully tax deductible for both firms and households. Homeowners also are effectively exempt from the taxation of housing capital gains realized with the sale of their homes. For investors, generous depreciation deductions are allowed each year far in excess of actual economic obsolescence. Favorable provisions like these act to reduce the current yield necessary from real estate, as the tax advantage supplements property income. This will rotate the capitalization schedule of the northwest quadrant in a counterclockwise manner and lead to the higher asset prices and other impacts described in Figure 1.4.

At the local level, real estate is treated quite unfavorably, largely through the widespread use of property taxation. Most local governments finance their services with a flat tax rate on commercial, industrial, and residential property value. The effective rate of such taxation generally is in the 1 to 2 percent range, and this directly raises the capitalization rate necessary for real estate. Increased property taxation will rotate the capitalization ray clockwise, reducing asset prices, reducing construction, and raising rents.

Real Estate Financial Institutions and Regulations

In the U.S., the federal government has created a number of institutions whose purpose is to facilitate the financing of residential real estate. The system of S&L banks was established to channel local savings into mortgage lending, and the secondary market together with national mortgage insurance ensure that mortgages have almost instant liquidity. By facilitating the flow of investment funds into mortgage lending, such institutions have effectively reduced the cost of mortgage borrowing. This again will rotate the capitalization rate ray of the northwest quadrant in a counterclockwise manner, increasing asset prices and residential construction.

The federal government also regulates a number of financial institutions in ways that can increase or decrease the flows both of long-term investment funds and short-term construction lending to nonresidential real estate. For example, when the Employee Retirement Income Security Act of 1974 (ERISA) was passed by Congress, pension funds were required to increase diversification of their assets—a policy which many believe greatly increased the supply of funds for long-term commercial mortgages. In 1989, Congress enacted the Financial Institutions Reform, Recovery, and Enforcement Act (FIRREA) which was designed to deal with the S&L crisis. FIRREA increased the capital reserve requirements for loans that were viewed as risky, such as commercial mortgages and short-term construction loans. These requirements substantially decreased the willingness of lenders to originate and hold these loans. The history of the U.S. financial industry is full of many such examples, each of which has influenced the flow of funds into real estate. In our four-quadrant diagram, a decrease in the availability of long-term financing will shift the capitalization rate in the northwest quadrant; a decrease in the availability of short-term construction financing will shift the cost schedule in the southwest quadrant.

SUMMARY

In this chapter, we defined real estate and provided a simple analytic framework to examine the operation of this important market.

- Real estate is defined as the national stock of buildings, the land on which those buildings sit, and all vacant land. Real estate is a very important component of our nation's wealth; the total value of real estate was estimated at $8.8 trillion in 1990, representing 56 percent of the nation's net worth. Additions to the stock represent roughly 5 to 7 percent of annual GDP.
- The four-quadrant model divides the real estate market into the property market where space use is determined and the asset market where buildings are bought and sold. There are two critical links between these markets. First, rents determined in the property market are translated into asset prices in the asset market. Second, asset prices determine the level of new construction, which determines the amount of stock available in the property market.

Exogenous shocks to the property market can have very different impacts on the operation on the real estate market than exogenous shocks to the asset market. In this chapter, we showed that:

- An increase in the demand for space in the property market shifts out the demand curve in the northeast quadrant, increasing rents, which, in turn, increases asset prices, construction, and the stock of space.
- A decrease in the capitalization rate in the asset market increases the demand for real estate assets, which increases asset prices. Increased asset prices in turn bring forth more construction, increasing the stock of space and decreasing rents.
- An increase in construction costs decreases construction levels, which, in turn, decreases the stock of space, driving up both rents and asset prices.

REFERENCES AND ADDITIONAL READINGS

Balance Sheets of the U.S. Economy: 1945–1990. Washington, D.C.: Board of Governors of the Federal Reserve System, September 1991.

BRUEGGEMAN, WILLIAM, J. FISHER, AND LEO STONE, *Real Estate Finance,* 8th ed. Homewood, Ill: Irwin, 1989.

Managing the Future: Real Estate in the 1990s. Chicago: IREM Foundation and Arthur Anderson, 1991.

MILES, MIKE E., "What Is the Value of All U.S. Real Estate?" *Real Estate Review* (1990), 20(2): 69–77.

CHAPTER 2

The Operation of Property Markets: A Micro and Macro Approach

What is the fair market value of a specific house? How much rent can a property manager charge for his office space? Is now the best time to develop a site? What should the density and use be? Real estate decisions such as these must be based on an understanding of the economic environment of each parcel or property. This economic environment is constantly being changed by forces and events at two levels: the micro level and the macro level. Micro forces are those location-specific factors that influence the value or use of one particular site. Macro forces, on the other hand, are those broad economic factors that affect market timing and influence the profitability or use of all properties.

As a discipline, the study of economics is divided into two broad fields with a similar distinction: microeconomics and macroeconomics. *Microeconomics* refers to the study of how individual economic agents (such as households and firms) operate, whereas *macroeconomics* investigates the behavior of the overall economy. Real estate economics is divided in a similar way. In this book, the study of the use, development, or pricing of individual properties or parcels of land involves a microeconomic approach. In contrast, the study of the behavior of the overall market aggregating across individual properties involves using a more macroeconomic approach.

In real estate economics, location plays a crucial role in distinguishing between micro and macro approaches. Real estate microeconomics borrows heavily from the traditional urban economics literature, treating the operation of land and property markets with the explicit recognition of space or location. What is the demand for land at one site? What is that site's highest and best use? How do house prices vary across sites, and

why? What factors determine office rents or the location of firms and their plants? Real estate microeconomics investigates these questions by studying the operation of urban land markets and developing theories and explanations for the spatial structure of cities.

Real estate macroeconomics abstracts from the spatial dimension and considers the overall market for housing, land, or office space. The simplification that results from such aggregation can be justified on two grounds. First, as with any type of inquiry, insight often is achieved, perhaps at the expense of some realism, by abstracting or making simplifying assumptions. Without aggregating across locations and dealing with a market as a whole, the detailed time series models in Chapters 10 and 12 of this book simply could not be developed. Second, many factors that affect real estate are largely independent of location. Consider changes in interest rates or variations in the growth rate of an area's economy. These are forces that move over time and impact real estate at all locations. In these cases, it makes sense to study a market as an aggregate entity.

The distinction between real estate micro- and macroeconomics clearly hinges on the notion of an aggregate market. Within this context, then, how should a market be defined? The answer is that a market should represent a group of properties that react similarly with respect to macro factors (such as interest rates and economic growth). If the behavior of properties within a market is similar, then the macro approach will work. If properties react very differently to macro effects, then such modeling will be largely unsuccessful.

DEFINING MARKETS: PROPERTY TYPES

In addition to the distinction between microeconomic and macroeconomic approaches to real estate, throughout this book we distinguish between residential and nonresidential property markets. This delineation has both advantages and disadvantages, but we believe there are net benefits to the distinction. At the macroeconomic level, housing markets clearly behave differently from those of nonresidential property. Movements in housing prices and residential construction do not relate closely to movements in rents or construction for office, industrial, or retail properties. The institutions that guide each of these markets also are quite different. Residential contractors rarely build commercial space. Industrial or office space brokerage firms have no connection to firms undertaking residential brokerage. The financing of residential properties also takes place through a distinct mortgage origination process, and the residential mortgage market has a very active secondary market (the market in which mortgages are purchased and sold). Commercial financing takes place largely through private placements, but there is only a small formal secondary market. Thus, at the macro level, there are good reasons to consider different property types as different markets.

At the micro level, the distinction between residential and nonresidential property is not as clear, largely due to the fact that both types of property use and compete for a common resource: land. The price of commercial property, for example, bears a direct relationship to the price of residential property, since both uses compete for the same fixed supply of land. The locations of commercial and residential property also are closely linked through the commuting of workers and travel of shoppers. At the same time, the behavior of participants in the residential and nonresidential property markets is based on different economic theories and motives. Finally, the extensive government regulation of land in the two uses (e.g., through zoning) has important impacts on both residential and nonresidential property markets.

DEFINING MARKETS: AREAS

The second issue in defining a property market involves the question of spatial aggregation: what is the appropriate geographic definition of a real estate market? Should markets be defined at the neighborhood, town, metropolitan, state, or national level? There is no definitive answer to this question, but there are both conceptual and pragmatic criteria for making the choice.

Conceptually, the geographic definition of a real estate market should encompass real estate parcels that are influenced by the same economic conditions. Clearly, national economic conditions, such as interest rate levels, influence real estate markets. But we also know that real estate markets are profoundly influenced by economic conditions such as employment and income, which vary widely across regions of the country. How should these regional or local economies be defined? Urban economists have long struggled with this question.

As a practical matter, the geographic definition of a local economy is generally settled by the way in which data on urban economies are collected. The U.S. Census Bureau defines metropolitan statistical areas (MSAs) on the basis of economic behavior rather than the legal, political, or historic precedent that delineates towns or states. An *MSA* is generally defined as a county with a central city with a population of 50,000 or more and adjacent counties that are metropolitan in nature. The decision to include a county in an MSA is based on population density, the percentage of the population that is urban, the nonagricultural employment level, and the commuting patterns among counties and with the central city. As a result, MSAs are drawn to reflect a single labor market and the mobility of workers; in addition, MSA boundaries often change over time.[1] Commuting within a metropolitan area should not be so burdensome that it prevents most households from living and working at any two locations. In other words, a worker at a particular location within a metropolitan area should not be severely constrained from taking any job within that area. Conversely, commuting between metropolitan areas should present enough potential obstacles to make it a rare occurrence.

This labor-mobility definition of a local economy has broad ramifications for how real estate markets operate. If a worker within a metropolitan area can be reasonably expected to live anywhere in that area, then all houses in that area are, in some sense,

[1]In recent years, the Census Bureau has regrouped some adjoining and economically linked MSAs into common Consolidated Metropolitan Statistical Areas (CMSAs).

competitive with each other. For example, by commuting more or less, any worker could, in principle, bid on any house. Locations for firms are similarly competitive, in the sense that a business could choose any site and, in principle, still find workers willing to commute there. Between metropolitan areas, this degree of competitiveness normally does not exist. Households in San Francisco cannot reasonably be expected to work in Los Angeles, which means that the real estate markets of the two areas are in some fundamental sense disconnected. The competitiveness of properties within a metropolitan area comes from the ability of workers to change houses without switching jobs, and change jobs without moving their place of residence. The absence of such mobility between metropolitan areas separates their real estate markets.

Conceptually, if we define our unit of analysis at the macro level as the metropolitan area, what is the appropriate unit of analysis for the micro level? Within a market definition, there are distinct factors that affect the behavior of individual properties that are different from those factors that affect the overall market. This effectively means that there is something to be gained from the whole micro-macro delineation; that micro analysis is different from macro analysis, and that the two complement each other. If there were little difference between the types of theories and arguments used at the micro and macro levels, then there would not be any clear advantage to the twofold approach. In real estate economics, the role of location helps to create the distinction between the two levels, particularly when markets are defined at the metropolitan level.

The mobility of households and firms across locations within metropolitan areas is fundamental to the microeconomic study of real estate. Since locations within a market differ (by commuting for example), mobile households will quickly generate higher prices for more desirable (lower commuting) sites. The theory of urban land markets provides a methodological approach based on this high degree of spatial mobility. The mobility of firms and workers within metropolitan areas also means that an economic shock occurring at one site (e.g., loss or gain of jobs) will have impacts on prices at all sites, as mobile workers adjust their demands. Thus, sites within metropolitan areas should react similarly to changes affecting the overall market.

Throughout this book, our macro analysis is generally focused at the metropolitan level, although occasionally we deal with a national housing or commercial real estate market—particularly when there appears to be systematic patterns of behavior across many metropolitan areas. For the micro level analysis, we generally look within a metropolitan area, often at cities and towns within an MSA.

REAL ESTATE MICROECONOMICS: URBAN LAND AND LOCATION

When defined at the metropolitan level, real estate property markets contain thousands, and even millions, of individual parcels (sites, houses, buildings). Each property occupies a location that is technically unique: adjacent sites may have similar, but not exactly identical, locations. A market in which each good is somewhat unique is referred to as a *product-differentiated market.* In addition to urban land, labor is another major market that is fully product-differentiated. This type of market may be contrasted with a *commodity market*, such as that for corn, in which goods are largely identical.

The observed behavior of individual parcels within the urban land market tends to follow several patterns that are quite characteristic of product-differentiated markets:

1. The prices of individual properties or land parcels vary widely and systematically with the physical or location characteristics of the property. The valuation that households or firms place on these characteristics determines the property's overall value.
2. The *relative* prices of different properties remain very stable over time and change little as the overall market undergoes either cyclic fluctuations or long-term growth. Overall market movements tend to raise and lower all prices by proportionate amounts.
3. The relative price of a parcel tends to change mainly when the characteristics of that parcel are altered. Such changes include physical alteration of the structure as well as changes in the characteristics associated with the parcel's location or neighborhood.

The theory of urban location and land markets has been developed essentially to explain these observed patterns. Since some readers may be troubled, or not completely convinced by the assertions above, let's consider their conceptual and empirical foundations.

The fact that house prices vary according to the size, quality, and character of a unit's structure probably comes as no surprise. It is also true, however, that units with similar physical characteristics will vary enormously in price by location. Location characteristics that have been shown to affect house prices include commuting time or access to jobs, public services and neighborhood quality (e.g., school performance, crime rates, etc.), and natural or environmental features (water frontage, terrain with views, air or noise quality). These characteristics of a location can account for more than half of the overall value of a house. As an example of how important public services are, Table 2.1 compares the average price of a single-family house in a range of towns in 1990 metropolitan Boston to the average score of high school seniors in the town's school system on the test administered by the Massachusetts Educational Assessment Program. Figure 2.1 is a scatter plot of the house values and test scores and shows the regression line between the two variables. The relationship is striking and the statistical correlation between the two is quite strong.[2]

[2]In Table 2.1, a simple bivariate regression between test score (*TEST*) and house price (*P*) yields the following relationship:

$$P = -280811 + 369.4\,TEST \qquad R^2 = 0.38$$
$$(-6.1) \qquad (4.3) \qquad N = 32$$

The t-statistics are in parentheses below the coefficients. Thus, across the sample in the table, towns with average test scores 100 points higher have houses that are worth almost $37,000 more. Of course, we must be careful in drawing strong conclusions from this simple regression. Certainly, some of this large effect is caused by the fact that towns with higher test scores may also have lower crime, a better park system, and so on.

TABLE 2.1 Boston-Area House Values and Student Test Scores, 1990

	House Values (1990 $)	12th-Grade Achievement Scores
Brookline	377,800	1430
Wellesley	349,500	1437.5
Belmont	307,800	1455
Newton	293,400	1435
Lexington	282,800	1525
Cambridge	263,800	1222.5
Marblehead	257,200	1395
Needham	256,500	1470
Rockport	227,500	1445
Milton	219,600	1315
Arlington	209,200	1375
Reading	204,100	1347.5
Watertown	196,700	1280
Stoneham	194,900	1292.5
Waltham	191,100	1267.5
Burlington	191,100	1350
Wakefield	190,600	1282.5
Framingham	184,700	1415
Medford	182,400	1207.5
Sharon	182,100	1427.5
Dedham	177,500	1327.5
Peabody	177,100	1270
Ipswich	174,000	1415
Woburn	172,600	1247.5
Braintree	168,700	1387.5
Somerville	165,800	1155
Hopkinton	163,200	1265
Malden	162,900	1207.5
Boston	161,400	1180
Quincy	161,100	1295
Revere	160,500	1212.5
Randolph	155,500	1297.5

Scores represent averages of reading, math, science, and social studies scores.
Source: House Values, 1990 Census of Population and Housing, *Summary Tape File 1A*; Scores, *State Results of Massachusetts Testing Program*, Massachusetts Education Assessment Program, 1990.

Strong statistical relationships, such as that in Figure 2.1, exist throughout the housing market between prices and many structural or location attributes. They also characterize other property markets as well. Retail space rents vary systematically with the expected pedestrian traffic on downtown streets, and office space rents are higher around mass transit lines. In each of these cases, well-established price premiums or discounts exist for locational or structural features of a property. Location theory holds that such

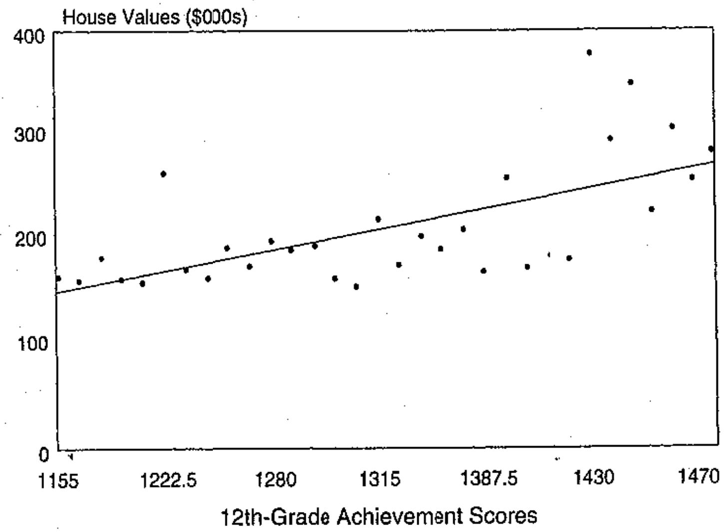

FIGURE 2.1 Relationship between house values and test scores of Boston-area towns, 1990.

Scores represent averages of reading, math, science, and social studies scores.

Source: House Values, 1990 Census of Population and Housing, *Summary Tape File 1A*; Scores, *State Results of Massachusetts Testing Program*, Massachusetts Educational Assessment Program, 1990.

price effects represent the long-term valuation of these attributes by households or firms. The theory also suggests that such valuations are relatively stable as the overall market undergoes growth or decline. In other words, rents or prices for all locations rise and fall with a market's fortune, but the relative price of more desirable versus less desirable locations changes very little.

The general stability of *relative* property prices or of property price premiums can be shown empirically. In Figure 2.2, the rents for office space during two very different periods are compared across several locations in the Boston metropolitan area. In 1980, the Boston office market was coming out of a long inactive period: construction was occurring everywhere and rents were rising rapidly. The economic growth of the defense, computer, and financial sectors in Boston created a real estate boom in the early 1980s. In contrast, by 1990, the economy had cooled and unemployment was rising. A changing computer technology market and the declining defense budget hit the area hard. In 1990, the real estate market had peaked and was declining rapidly; construction had almost ceased and rents were falling. Figure 2.2 shows that the relative rents for office space across locations changed very little between the beginning and end of this decade, despite considerable price inflation and very different overall market conditions.

The stability of relative property prices within metropolitan areas results from two features of the urban land market. The first is the high degree of household and firm mobility within metropolitan markets, as discussed in the previous section. Mobility acts to create a form of price "arbitrage." Locations within a market can rarely stay over- or underpriced with respect to other locations because of the mobility of potential property buyers or users. Another way of expressing this same condition is to say that the demand for any site or location is very *price elastic* with respect to its competitor sites. In this situation, small adjustments to prices (rents) should be sufficient to attract many buyers (tenants). Competition, demand elasticity, and arbitrage all imply that prices at one location cannot move independently of prices at other locations. Locations within a market are closely connected and rarely have independent price movements or cyclic behavior.

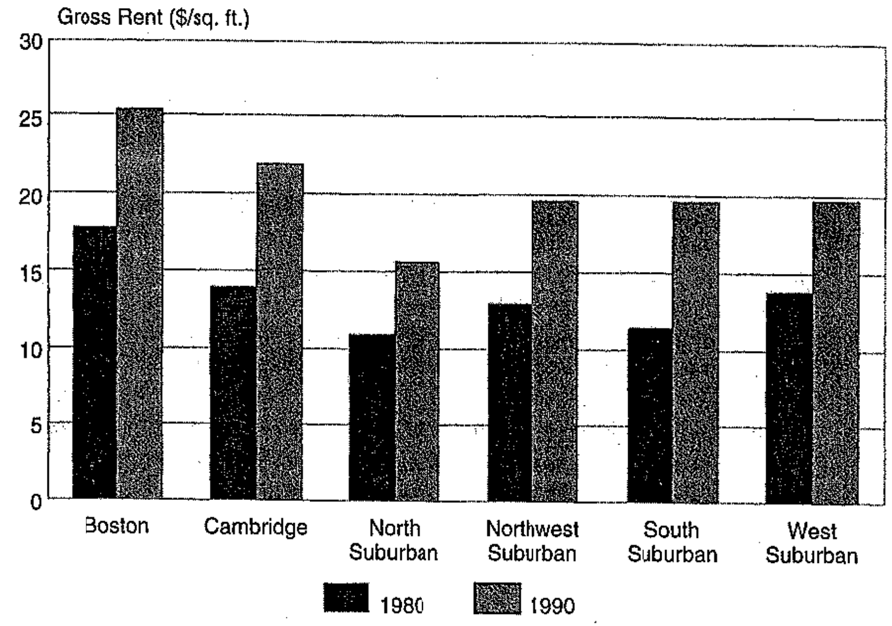

FIGURE 2.2 Comparison of Boston-area office rents, 1980 and 1990.

Source: 1980 Rents, Leggat and McCall and Spaulding and Slye, Inc.; 1990 Rents, CB Commercial.

The stability of relative property prices also results from the manner in which such prices are determined within metropolitan land markets. The premium for property at one location should reflect the present discounted value of the consumer's *utility*, or cost savings, from that location relative to others. Thus, for example, the premium for sites with an easy commute should be the discounted value of the time and costs savings that results from such commutes. If the physical and locational attributes of properties change only slowly, and if consumer valuations of those attributes are reasonably fixed, then the relative prices of different parcels should also change very little.

In fact, urban location theory suggests that the relative prices of different properties should change only in either of two situations. The first situation in which relative prices within a metropolitan market might change is if consumer valuations of particular physical or locational attributes change. For example, a sudden increase in gasoline prices would lead consumers to value sites with shorter commutes more highly. The gradual demographic change of an area from having predominantly households with children to those without might be expected to alter the valuation of houses with many bedrooms versus smaller units. Similarly, long-term shifts in the spatial distribution of jobs might also change the relative valuation for certain kinds of office or industrial space. Generally, such price changes occur only very slowly, since they are based on fundamental shifts in the makeup of an area's population or economy.

The second source of shifts in relative property prices is change in the attributes of a property. Obvious examples include physical change in the property's structure such as rehabilitation, renovation, or expansion. Changes in locational or neighborhood characteristics may also dramatically impact property prices. A common example is the construction of a new highway or transit facility, which alters the pattern of commuting costs for many locations. In metropolitan areas with new transit systems or extensions of older systems, it is common to find new developments emerging around the transit system. It is often difficult to measure directly the impact of such changes in the transit system on land values and development because the impacts are too localized to demonstrate with publicly available data. Changes in neighborhood crime or town school quality can also be expected to affect relative property prices in one area versus another.

Real estate microeconomics involves more than just the study of how property rents or prices are determined across locations. It also focuses on how the land market uses these prices to determine both the density of development and the location of different land uses. One of the important lessons that microeconomics teaches us is that the price of land and its density, or use, are determined simultaneously: denser uses generate higher land prices, whereas more expensive land encourages denser use. Another precept of real estate microeconomics involves the separation of uses. The land market works much as an auction, with each site being developed or occupied by that use offering the most for it. In this process, it is a natural market outcome to have each use occupying separate areas or distinct locations. The development of land proceeds under a number of such microeconomic principles.

A thorough understanding of these principles of microeconomics allows us to understand the evolution of cities over time and the spatial patterns of land prices, land uses, and density. Why is residential density normally higher in downtown areas of a city, or along beaches and other natural amenities? How did Central Business Districts (CBDs) come into existence, and why has there been a recent explosive growth of suburban employment districts? What explains the district hierarchical pattern of retail establishments—a large number of smaller stores or centers and a small number of larger centers? Why do property prices and density eventually rise as a city grows? The study of metropolitan land markets provides answers to these questions by emphasizing the long-term equilibrium outcomes of spatial competition. This microeconomic approach, however, does not examine the economic determinants of long-term metropolitan growth or the short-run cyclic behavior of metropolitan economies. Real estate macroeconomics focuses on these topics.

REAL ESTATE MACROECONOMICS: MARKET GROWTH AND DYNAMICS

While microeconomics studies prices and land use across space within a particular market, macroeconomics examines the overall movement of prices and real estate development for the metropolitan market as a whole. By abstracting from the spatial dimension, macroeconomics is able to focus more specifically on the time dimension that emphasizes short-run movements and temporary disequilibrium. As a result, macroeconomics deals with aggregate variables that are averages or aggregations of data measured at each location within the market. Given the theory of real estate microeconomics, the presumption in macroeconomics is that such aggregate measures depict the behavior of variables at most locations within the market.

Some macroeconomic variables are most commonly measured as averages. These variables include market prices, rents, or vacancy. For example, samples of apartments are repeatedly surveyed by the U.S. Commerce Department to produce estimates of average rents in metropolitan areas. Brokerage companies periodically survey the inventory of office buildings to produce estimates of average market vacancy rates. The Federal Housing Finance Board maintains data on single-family house transactions through their survey of mortgage lenders and reports the average price of existing house sales. As an example of such data, Figure 2.3 traces house price series for the Dallas and Boston metropolitan areas in constant (1990) dollars.

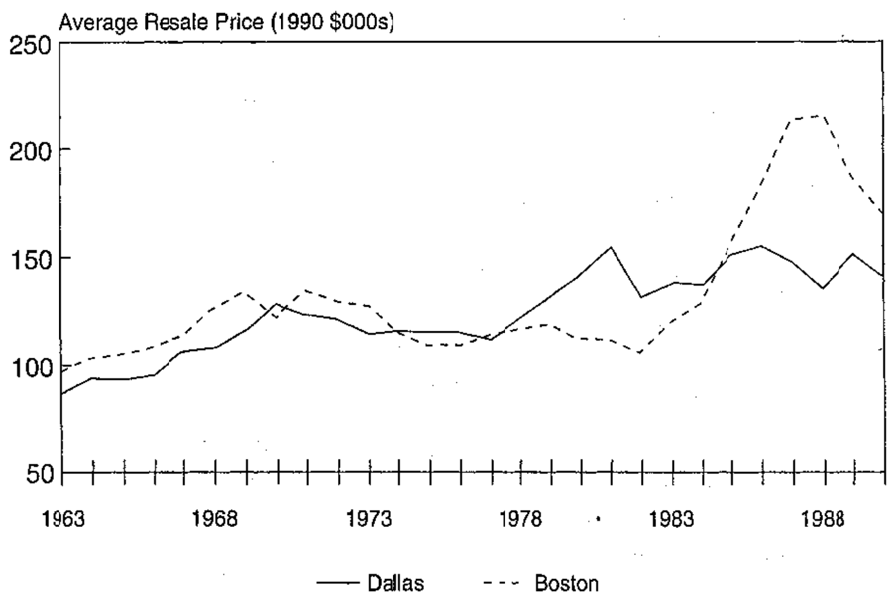

FIGURE 2.3 Boston and Dallas house prices, 1963–1990.

Source: 1963–72, average of monthly average purchase price for existing homes, Federal Home Loan Bank Board. 1973–90, annual average for existing and new homes, Federal Housing Finance Board.

One of the main objectives of real estate macroeconomics is to explain the long-run trends and short-run movements of price or rent data such as is shown in Figure 2.3. Sometimes house prices in many different areas are affected similarly by overall U.S. economic conditions. In both Boston and Dallas, for example, house prices dipped during the economic recessions of 1974 and 1982. At the same time, local market factors also are important. Dallas prices rose during the oil boom of the late 1970s, while Boston prices were flat. During the 1980s, Boston prices soared with the boom in technology industries, while the collapse in oil led to price declines in Dallas.

Often, the demand side of the market seems to account for much of the movement in prices. At other times, however, it is fluctuations in supply that explain price fluctuations. Consider the history of office-space vacancy rates in these same two markets in Figure 2.4.

In the market for office space, vacancy rates sometimes do move with a market's economic growth. For example, the vacancy rate soared in Dallas just after the recession of 1970 and during the Texas economic crash of the early 1980s. Similarly, Boston's vacancy rate peaked just after the recession of 1975. But why did Boston's vacancy rate rise in the mid-1980s, just when economic growth was strongest? Why did the vacancy rate in Dallas rise during the early 1970s, when its growth was strong? The answers lie with participants in the supply side of the market who built excessive amounts of new space, at least relative to market demand.

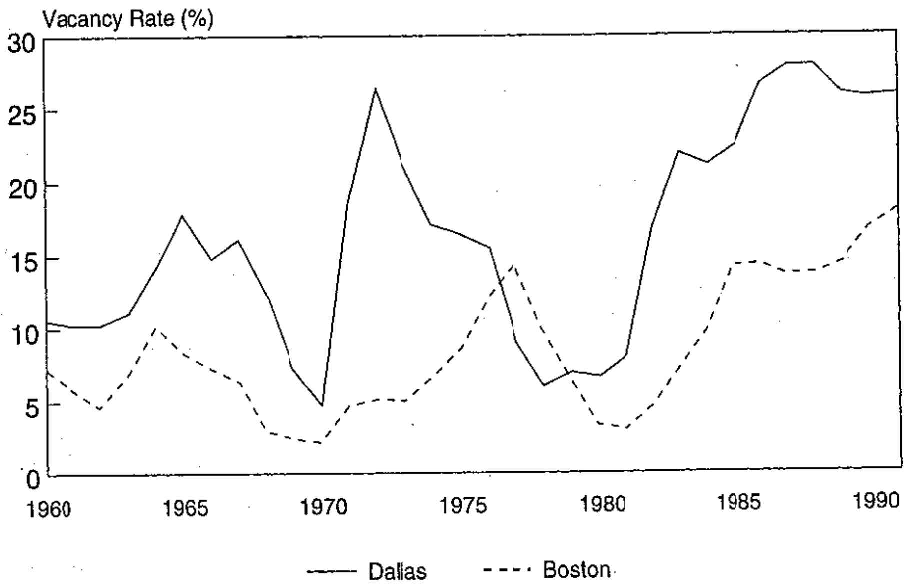

FIGURE 2.4 Boston and Dallas office vacancy rates, 1960–1991.
Source: CB Commercial.

Figure 2.5 portrays the growth in Dallas of office employment and office construction as a percentage of the stock of office space. Sometimes the two seem to move together, as during the period of Dallas's growing economy between 1972 and 1974, or during the decline in its economy in the mid-1980s. At other times, such as the periods between 1968 and 1970, 1974 and 1976, and 1988 and 1989, construction moved quite differently from employment growth. Clearly, the supply of real estate is not always in tune with demand, particularly over shorter intervals of time. Understanding the movements in supply and their determinants is as important as understanding the long- and short-term factors that affect property demand.

These examples illustrate some of the common patterns of behavior that real estate macroeconomics seeks to explain. These patterns give rise to three general principles that govern the aggregate behavior of metropolitan real estate markets:

1. The economic growth of each metropolitan area is determined in the short run by movements in the overall U.S. economy, together with the area's industrial mix and competitiveness. In the longer run, demographic changes and lifestyle preferences also play a role. Unlike different locations within a market, different metropolitan areas can simultaneously experience widely varying economic conditions. In the short run, there is no "arbitrage" between metropolitan areas.
2. The real estate market of each metropolitan area moves closely with the area's economic growth. In some situations, the supply or price of real estate can actually exert an influence on the area's overall economic development.

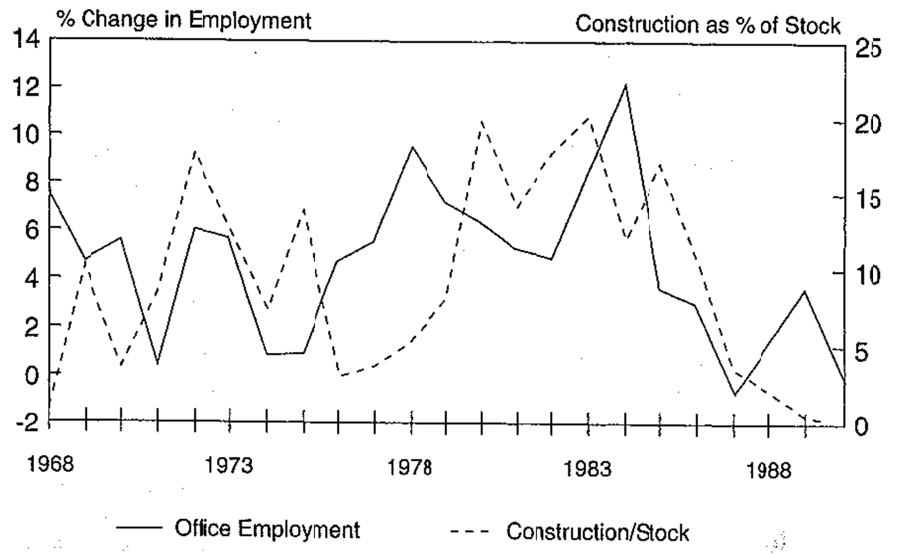

FIGURE 2.5 Dallas office employment growth and office construction, 1968–1990.
Source: Employment, adjusted U.S. government figures courtesy of Regional Financial Associates, Bala-Cynwyd, PA; construction/stock, CB Commercial.

3. Regions or metropolitan areas adjust slowly to economic change because resources are relatively immobile between markets. In response to changes in demand, the supply of factors into an area (labor, capital, structures) often occurs very gradually. Such slow adjustments give rise to temporary imbalances and help to generate cyclical patterns.

This macroeconomic behavior stands in noticeable contrast to the microeconomic principles discussed earlier. Within markets, individual parcels of real estate are intensely competitive with each other and have prices that adjust rapidly to maintain a degree of parity. While metropolitan areas also compete with one another over the very long run, the movement of economic resources between such areas is slow enough in the short run to largely uncouple the economies of different regions. Real estate microeconomics thus focuses on the equilibrium relationships between land or property parcels across locations within markets, while macroeconomics studies the disequilibrium that occurs as the markets grow and adjust to economic change.

SUMMARY

The micro and macro perspectives outlined in this chapter are the fundamental basis of analysis used throughout much of this book.

- The micro approach focuses on the importance of structural and locational characteristics on the prices and rents for a particular property or development. We often treat metropolitan areas as a single real estate market. While prices for real estate and land within a single market can vary enormously across locations, the relative price of real estate across locations is generally stable over time.
- The micro approach is examined in detail in the next four chapters. In Chapters 3 and 4 we explore the determinants of household location within a metropolitan area and examine the willingness to pay for structural characteristics, including the density of the development. In Chapters 5 and 6, we focus on the location decisions of firms within a metropolitan area.
- The macro approach examines how broad economic forces such as growth or decline in a metropolitan area's economy influence the area's real estate market. Metropolitan growth is determined by growth in the national economy as well as the area's industrial mix and competitiveness.
- The macro approach is examined in detail in Chapters 7 through 12. In Chapter 7 we examine the determinants of metropolitan growth. In Chapters 8 through 10 we provide a detailed examination of metropolitan housing markets, and in Chapters 11 and 12 we investigate metropolitan nonresidential real estate markets.

CHAPTER 3

THE URBAN LAND MARKET: RENTS AND PRICES

How is it that Manhattan has a dense urban skyline, while Los Angeles does not? How is it possible to find farmland within the city of Phoenix? Why is land in downtown Tokyo a thousand times more valuable than its suburban counterpart? This chapter begins a section of this book devoted to the study of residential development and the urban land market.

In economics, the markets for land and housing are often referred to as completely product-differentiated because each product sold in the market (every house or location) is unique. They stand in stark contrast to commodity markets, such as that for corn, oil, or minerals, in which a uniform good is traded in bulk quantities. The markets for most manufactured durable goods are considered partially differentiated, falling between these extremes. In the automobile market, for example, a large number of different models are regarded as being close, although not perfect, substitutes for each other. In the case of land or housing, no two parcels are exactly alike.

The fact that urban land is a completely differentiated product makes it difficult to speak about the supply or demand for sites at any particular location. By definition, the supply of land at each location is fixed; hence, it is quite price *inelastic*. The demand for a particular site, on the other hand, is likely to be quite sensitive, or *elastic*, with respect to its price. This results from the fact that numerous competitive sites, or substitutes, exist at adjoining locations. For almost two centuries, economists have recognized these distinctive features of the land market and have developed a simple approach for determining land or housing prices. The approach argues that land must be priced at each site so that its occupant is charged for the value of whatever locational advantages exist at that site. Understanding these advantages and how consumers evaluate them, therefore, becomes the key to understanding the spatial pattern of land or housing prices. This theory of *compensating differentials* assumes that only demand considerations determine the *relative* value of land or housing at different locations. The supply of land does play a role, but only in setting the overall level of prices.

We begin this chapter by examining a simple model in which the rent for housing and land is determined according to the compensation principle. This rent is referred to as *Ricardian rent,* since the approach was developed by Ricardo (1817). We then embellish this model to make it more realistic and examine what the model implies about how rents change over time and across cities. Ricardian theory also explains why land uses and different types of households tend to be separated spatially. Since sites go to those offering the highest rent, spatial separation occurs naturally in these markets. The chapter finishes by examining how Ricardian rent is converted into land or house prices. The capitalization of rent into prices is shown to vary by location within cities, particularly when urban areas are growing and such growth is anticipated to continue in the future. Finally, we examine data concerning house rents and prices, demonstrating that a century-old theory holds up fairly well against modern reality.

URBAN COMMUTING AND RICARDIAN RENT

The first fundamental characteristic of urban housing and land markets is that housing and land are more expensive at better locations and cheaper at less advantageous sites. This holds whether we consider natural locational amenities, such as lakes or an ocean, or man-made locational advantages, such as distance to employment or cultural centers. To illustrate how rent and locational advantage interact, we begin with a very simple city. In this city, commuting or access to a place of employment is the only locational advantage that is considered. Following a long literature, our city will be *monocentric,* meaning that it has only one employment center (Alonso 1964, Mills 1972, Muth 1969). Commuting to this center gives rise to what is called Ricardian rent. In the Ricardian definition, *rent* refers either to the payments that a tenant would offer for housing, or, alternatively, to the *annual* amount that an owner would be willing to pay for the right of occupancy or use. Later in this chapter we will examine how these rents get capitalized into prices. The model also assumes that the density of development is fixed, which means that structure capital cannot be substituted for land. This absence of any factor substitution may seem unrealistic, but we will deal with determining density in Chapter 4. Thus, our stylized city has the following features:

1. Employment is at a single center, to which households commute along a direct line from their place of residence. Commuting costs k dollars annually per mile. The location of a household thus refers to its linear distance (d) from the employment center.
2. Households are identical, and the number of workers (commuters) per household is fixed. Household income (y) can be spent on commuting, all other goods (represented in dollars by x), and on housing.

3. Housing has fixed and uniform characteristics at all locations. Housing rent is an annual amount $R(d)$, which varies by location (commuting distance d).
4. Housing is provided by combining a fixed amount of land per unit of housing (acres, q) together with a fixed amount of housing capital (materials and labor) that costs c (no factor substitution). Residential density, therefore, is $1/q$.
5. Housing is occupied by households who offer the highest rent, and land is allocated to that use yielding the greatest rent.

The last assumption is crucial, for it implies that when the housing market is in equilibrium in this stylized city, decreased rents as one moves out from the city center must exactly offset the increasing commuting costs. Since the quality and density of housing is fixed across locations, the only variation possible in household or consumer welfare is the amount of income left for expenditure on other goods (x). If housing rents do not exactly offset commuting costs, then consumers who live at closer locations will have more income left to spend on other goods. In this case, consumers at farther locations would seek to move to closer ones and would offer greater housing rent than current occupants. Since housing is rented to the highest bidder, rents at the closer locations would rise, while those at farther sites would fall. When rents exactly offset commuting costs, households would no longer have an incentive to move, and the market is then said to be in equilibrium.[1] In effect, mobility due to consumer welfare differentials is not possible when a private market is at equilibrium. As long as all households are identical, household expenditure on other goods (x) must be constant across locations at some level x^0. Using the definition of consumer income and expenditure, housing rents follow in Equation (3.1) below.

$$R(d) = y - kd - x^0 \tag{3.1}$$

At the city center ($d = 0$), consumers will have no expenditure for commuting (at least in our stylized city), and so rent there, $R(0)$, will equal $y - x^0$. Moving outward, rent will decline dollar for dollar as commuting costs increase. At some distance, b, the city ends and housing rent will be at its cheapest. What determines this least expensive rent at the city's edge? The answer is the cost of constructing new units.

In many cities throughout the world, land beyond the edge of development is used for agriculture. In this use, it earns some rural rent *per acre* (labeled r^a).[2] In other situations, the land is simply held vacant with little or no meaningful rural rent. In the simple Ricardian model, assumption 5 implies that site or land owners seek the highest income from their land, just as housing is rented to those making the highest offer. Thus, as long as urban housing yields a rent for a site that exceeds that which the owner can receive from farming, land will be rented to urban households. We will see later that this simple criteria for the transition of land from rural to urban use still holds under a much more sophisticated analysis of the landowner's decision about when to develop land or convert it from farming to housing.

[1] This point is often referred to as a *spatial equilibrium,* since there is no incentive to change location.

[2] Throughout this book, capital letters will be used when referring to housing rent or price (R, P) while lower case letters will be used for land rent and price (r, p).

At the edge of the city (b), then, urban landlords can rent land for its agricultural or opportunity value of r^a per acre. With fixed density, a lot for each housing unit can be rented for $r^a q$. The rent for a housing unit at the edge of the city therefore has two components: the *land rent* $r^a q$ plus the *structure rent*, which is the *annualized* cost of constructing a unit (c). This structure rent could be measured by the annual mortgage payment necessary to cover the cost of constructing the unit. The sum of these two costs is the rent that is necessary to cover the creation of new housing at the city's edge. Combining this with Equation (3.1), we can determine that level of other expenditure x^0 that will prevail if a household that commutes from the city's edge must pay a rent for housing there equal to agricultural land rent plus the annual replacement cost of the housing structure:

$$x^0 = y - kb - (r^a q + c) \tag{3.2}$$

Moving in from the edge of the city, Equation (3.1) defines how rents must rise as commuting costs decrease in order for households to maintain the level of welfare (or expenditure) defined in Equation (3.2). Combining the two, housing rents at any location will equal replacement costs plus the difference between commuting costs at the urban edge and those at the location in question. The rent gradient for housing is:

$$R(d) = (r^a q + c) + k(b - d) \tag{3.3}$$

In effect, housing rent at any interior site absorbs the savings in commuting that result by moving in from the farthest location currently developed in the city. Only with these rents will (identical) households be willing to live at any location within the city.

Figure 3.1 traces the equilibrium rent gradient for housing (Equation (3.3)) as it varies along a radius (d) in our stylized circular city. There are three components to housing rent: (1) the rent necessary to convert a lot from farm land into urban land ($r^a q$), (2) the rent for the structure that sits on the lot, c, and (3) the location rent resulting from saved commuting costs, $k(b - d)$. It is important to note that both the agricultural rent and the structure rent are constant across locations. The slope of the housing rent gradient with respect to distance, $-k$, is due to the location rent; rents fall away from the city center (per mile) by exactly the amount of additional commuting incurred by each household.

Those components of housing rent that involve location and agricultural land are often combined into a hypothetical rent for just urban land $r(d)$. Urban land rent can be thought of as a residual: the land rent that is left after subtracting the rent for the housing structure from the total housing rent. From Equation (3.3), Equation (3.4) describes urban land rent. It is important to remember that housing rent is measured per unit, or per household, while land rent will be measured as rent per acre. Thus, to convert housing rent $R(d)$ into land rent, $r(d)$, we must first subtract the structure rent and then divide by land per unit (q). This is the same as multiplying housing-minus-structure rent per unit by residential density ($1/q$).

$$r(d) = r^a + \frac{k(b - d)}{q} \tag{3.4}$$

Urban land rent has two components. The first is the rent (per acre) for its alternative use (agricultural), while the second is the savings in commuting costs *per acre* that result when housing is placed on the land. At a density of $1/q$, there are that many households per acre, each of which is saving $k(b - d)$ in commuting. The gradient for land rent with respect to distance has a slope of $-k/q$: the rent per acre of land falls by the increased total commuting of all who live in the $1/q$ units built on the land.

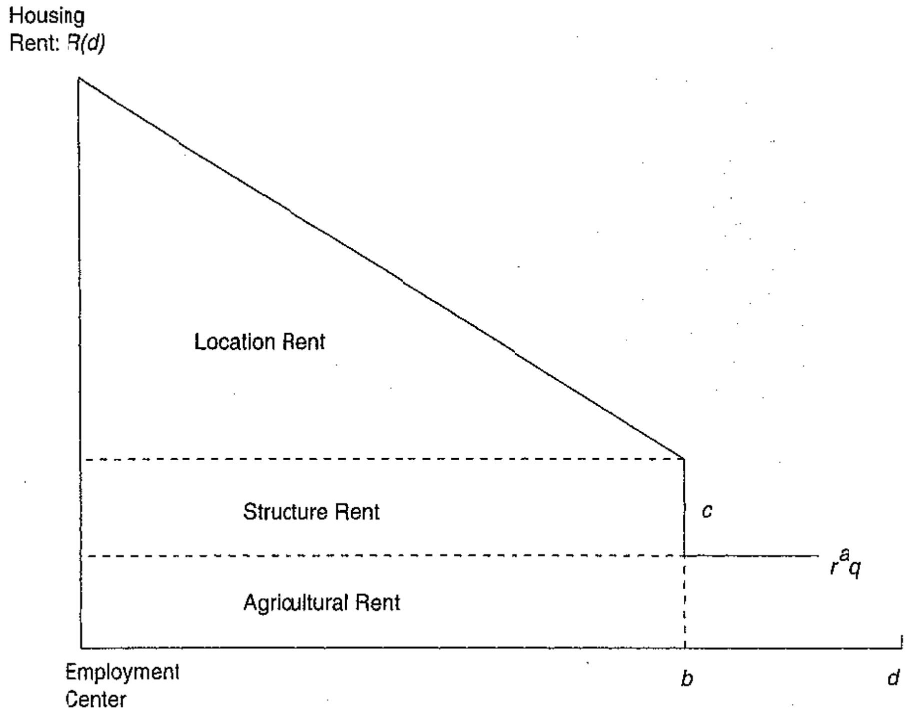

FIGURE 3.1 Components of housing rent.

Numerical Example

Let's consider a typical example of rents in a simple city in which urban density averages around 4 units (households) per acre, or 2,560 households per square mile (q = 0.0004 square miles). A structure costing $100,000 to build would rent for $7,000 per year using a 7 percent interest rate. Annual agricultural rental income from farming tends to fluctuate around $1,000 per acre, or $640,000 annually, per square mile. Given our assumption of 4 units per acre, the agricultural rent per year for a house lot would be $250. Finally, the cost of commuting for an average household with income of about $40,000 could be about $200 per year per mile. With these values, annual house rent at the urban border will be $7,250 ($7,000 in structure rent plus $250 in agricultural rent). If the city's border is 20 miles away from the center, housing rent would be $11,250 at the center ($7,000 in structure rent plus $250 in agricultural rent plus $4,000 in commuting cost savings). Land rent per acre will range from $1,000 at the urban fringe (4 units per acre times $250 per unit) to $17,000 at the center (4 units per acre times the sum of $250 per unit in agricultural rent plus $4,000 in commuting cost savings).

How do housing rent and urban land rent vary across cities or within one city over time? Equations (3.3) and (3.4) permit us to draw some fairly powerful comparative static conclusions:

1. When the edge of the city (b) is farther from the center and involves a greater commute, housing and land rent at all interior locations is higher since at these locations there is a greater savings in commuting cost.
2. When commuting is more costly (per mile), interior housing and land rents will be higher relative to edge rents because, again, the commuting cost savings at interior locations are higher.
3. When urban land has a more productive alternative use or higher agricultural rent (r^a), urban housing and land rents will also be greater, because this land rent component of housing rent is higher.
4. When the density of urban housing is greater, the gradient for urban *land* rents will become steeper, with higher rents at the city center relative to those near the edge. (Remember, the slope of the land rent gradient is $-k/q$. If the amount of land used per housing unit, q, decreases, then the land rent gradient becomes steeper.)

POPULATION, LAND SUPPLY, AND RICARDIAN RENTS

Since urban rents depend crucially on the distance to the city edge, the next logical step is to examine how far a given city's edge will be. This involves three considerations: the population of the city, the density of housing, and the role played by the area's topography in determining the supply of land. Again, we will take the density of housing development as fixed here; in Chapter 4, we will relax this assumption and explore the implications of varying density.

One of the original assumptions of our stylized city was that of linear, radial commuting. If the city is located on a featureless plain, then this implies that the city will be circular in shape, with land rents equal for all points at a given distance (d) from the center. On the other hand, consider a city whose employment center is located at the coastline of a lake or an ocean. Many cities developed historically to serve as ports, and therefore have a semicircular structure. Commuting and prices will still be equal for a given radius, but the city's circumference may extend only 270 degrees (Boston), 180 degrees (Chicago), or even only 30 degrees (the peninsula of Bombay). In an even more realistic world, mountains, lakes, and manmade obstacles also can reduce the supply of land at any given radius from the employment center.

Since we are dealing with a very stylized city, let us characterize land supply with a variable (v) that ranges from 0 to 1. When $v = 1$, the city is fully circular. At the other extreme, if $v = 0.1$, the city is constrained on a peninsula with a circumference of only

36 degrees. If $v = 0.5$, the city center is located on the straight edge of a body of water. Alternatively, a fully circular city could have $v < 1$ if it contains lakes and hills that limit the land available for development at various distances. We must also consider the number of households in the city (n), and from the previous section, the amount of land used per housing unit (q). Given these variables, the area of the city divided by land usage per housing unit must equal the number of households in the city. Recalling the simple formula for the area of a complete or partial circle, we can define the border, b, as:[3]

$$b = \left(\frac{nq}{\pi v}\right)^{\frac{1}{2}} \qquad (3.5)$$

Numerical Example

We can again use the density of four households per acre (or land consumption of 0.0004 square miles per household) to calculate the border for a large city with 2 million households (n). In most American cities, almost 20 percent of urban land area is used for streets, 10 percent is used for commercial uses, and 10 percent is used for open space. This suggests that the value of v in a circular city would actually be closer to 0.6. Following Equation (3.5), a fully circular city has an urban border of slightly more than 20 miles (20.6), while a semicircular city with these parameters would extend to 29.1 miles.

The implications of Equation (3.5) are quite clear. All else being equal, cities that have greater population (n), lower density (larger lots, q), and are less circular (smaller v) will be spatially larger with a border (b) at a farther distance from the center. From comparative static result 1 in the previous section, this implies that such cities will also have higher housing and land rents. In fact, we can summarize the implications of Equation (3.5) into a fifth comparative static conclusion:

5. A city with greater population, lower residential density, or that is less circular because of topography and land constraints will have a development edge that is at a greater distance from its center.

The combination of Equation (3.5) and Equation (3.3) or (3.4) provides a simple yet quite powerful model of urban land rents. Sometimes its implications are quite subtle. Consider two cities with equal populations and topography, but different residential densities. From Equation (3.5), the denser city will have a shorter distance to its edge, and from Equation (3.3) this will yield lower central housing rents. With respect to land, however, higher density increases the slope of the rent gradient at the same time it is shifted downward (from the nearer edge). Which of these two forces dominate? Will central land rents rise or fall with greater density? Combining Equations (3.4) and (3.5) with some

[3] The border, b, is the radius of a circular city. The land area of this circle is equal to πb^2. If the city is not fully circular, its area will equal $v\pi b^2$, where $0 < v < 1$. The land area of the city must be equal to the number of households, n, times the amount of land per household, q.

$$v\pi b^2 = nq \; ; \; b = \left(\frac{nq}{\pi v}\right)^{\frac{1}{2}}$$

mathematics, we are able to deduce that on net, central land rents should be higher in cities with greater density, even as house rents fall.[4]

It is important to remember that the conclusions of the model only compare equilibrium solutions, just as in Chapter 1 with our four-quadrant diagram. The model is static in that it ignores expected future growth and does not pretend to portray how cities adjust gradually over shorter periods of time. It indicates only what the city eventually should look like. To incorporate expected growth, our model must deal with prices as well as rents, which we will do later in this chapter. The model is also extremely simple in its assumptions of fixed density and identical individuals. In Chapter 4, we will relax the assumption of fixed density. Now let's turn to the question of different households or land uses.

COMPETITION AND SPATIAL SEPARATION

The second fundamental characteristic of urban land and housing markets is that they tend to naturally separate different households or land uses spatially. To illustrate this, we extend our stylized city model to consider the situation in which there are two categories of households. Initially, let's assume that these household groups differ only according to their costs of commuting (perhaps from their valuations of time). There are n_1 members of the first group (Group 1) who dislike commuting intensely, whereas the n_2 members of the second (Group 2) mind it much less. As a result, we subscript the cost of commuting and have $k_1 > k_2$. Consider two questions. First, will the two groups choose separate or intermixed locations, and which group will locate where? Second, what will housing rents look like in this two-household city? In all other respects, our city is the same as in the sections above. Thus, there are $n_1 + n_2$ houses that are identical, with fixed lot sizes. There also is radial commuting, and the cost of constructing new housing at the city's edge again involves a rent both for structure capital and for agricultural land. Most importantly, rational owners will still rent housing to that household offering the most for it.

When there are multiple groups of households, there is no requirement that the members of one group must enjoy the same welfare (or level of expenditure) as members of another group. Markets treat equals equally and unequals unequally. To see this, let's examine the housing rents that will leave the members of *each* group with identical expenditure levels. We note such expenditure on other goods as x_1^0 and x_2^0 and pose the

[4]From Equation (3.4), we differentiate central land rents $r(0)$ with respect to the direct change in lot size, and the indirect effect of the altered border:

$$\frac{\partial r(0)}{\partial q} = -k\frac{(b-d)}{q^2} + \frac{k}{q}\left(\frac{\partial b}{\partial q}\right); \quad \frac{\partial b}{\partial q} - \frac{1}{2q}\left(\frac{nq}{\pi v}\right)^{\frac{1}{2}} = \left(\frac{b}{2q}\right)$$

Combining the two expressions and evaluating the result at the city center (d = 0), we have:

$$\frac{\partial r(0)}{\partial q} = -\frac{kb}{q^2} + \frac{kb}{2q^2} < 0$$

question of what rent exactly compensates the members of each household group for commuting. These housing rents are determined in Equation (3.6):

$$\begin{aligned} R_1(d) &= y - k_1 d - x_1^0 \\ R_2(d) &= y - k_2 d - x_2^0 \end{aligned} \tag{3.6}$$

Suppose for the moment that $x_1^0 = x_2^0$; then at every location (d), the rent that Group 2 households would be willing to pay for housing would exceed that of Group 1 households. This follows simply from the fact that $k_1 > k_2$. In this situation, no landlord would rent housing to the members of Group 1. Clearly this is not an equilibrium. If, however, the incomes of Group 1 households were also greater than those of Group 2, it might by pure chance turn out that the two groups would have equal expenditure levels (x levels). This, however, is not a requirement of market equilibrium. The only equilibrium conditions are that rents must leave members within each group equally well off, and that all members of both groups must have housing. We can refer to the rent functions in Equation (3.6) as the *equilibrium rents* of each group.

The next question to examine is whether these two groups need not be spatially separated but could intermix over some range of locations. For this to be the case, it would have to be true that at such locations, the equilibrium rent levels for each group of consumers are the same so that landlords would be equally willing to rent to either group. Assume for the moment that one such location exists and call its distance from the center, m. To the right of m (at a farther distance), it must be true that the equilibrium rents of Group 2 will exceed those of Group 1. The slope of Group 2's rent gradient is less than that of Group 1, since Group 2 has milder distaste for commuting ($k_1 > k_2$). Moving in toward the city center, the equilibrium rents of Group 1 will exceed those of Group 2, because the members of Group 1 value more highly the commute savings from more central locations. Thus, in this model, there can be at most only one site where the equilibrium rents for the two groups intersect. Since landlords rent housing to the group with the maximum rent, there will be a spatially segregated occupancy pattern in either direction from such a common location, with Group 1 occupying houses closer to the center (Figure 3.2).

It is important to realize that one site of a common equilibrium rent must exist; for, without it, one group would have a higher equilibrium rent than the other group over *all* locations. Since there is enough housing for both groups, this would allow housing to be vacant while the members of the lower rent group go homeless. Because renting two houses brings no additional welfare, the higher-rent group would reduce its equilibrium rents, while the lower-rent group would raise its rents. Eventually an equilibrium necessitates an intersection point—the location (m). With this intersection point, spatial segregation necessarily results, with the more central houses being occupied by that group with the more steeply sloped housing rent gradient.

In Figure 3.2, houses from the center to a location (m) are occupied by Group 1. Houses further out are rented to Group 2. This makes economic sense, since Group 1 households find commuting more distasteful and, therefore, are willing to pay higher rents to be closer to the city center. As discussed above, any other pattern would violate one of the two central conditions of a market equilibrium: that housing must be rented for

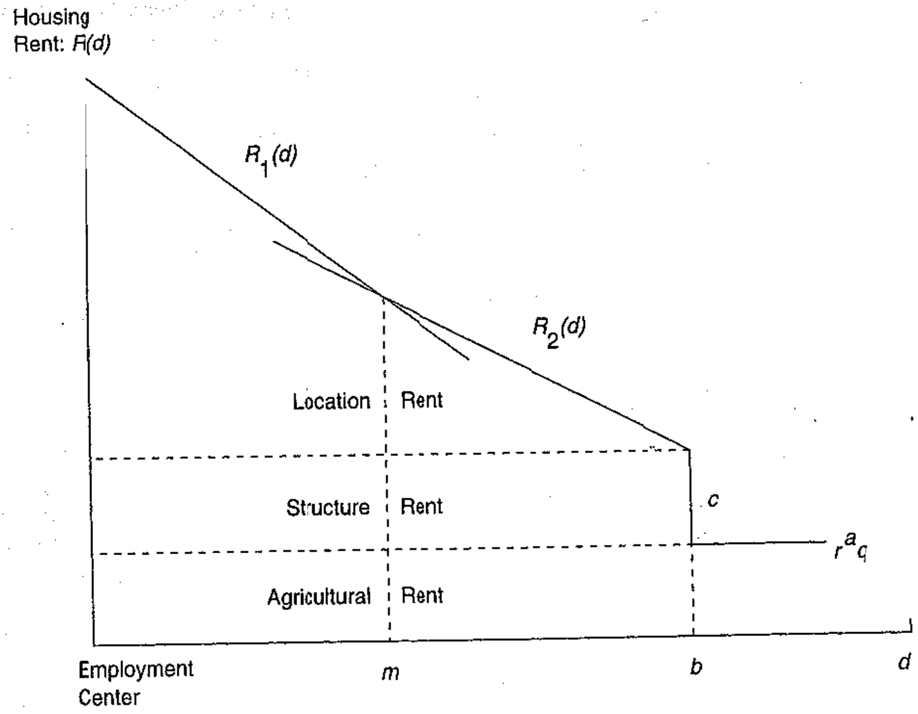

FIGURE 3.2 Housing rent gradient with two household types.

the maximum rent and that each household must occupy one house. To ensure the latter, the equilibrium rent at the edge of the city (the location b) must still equal the replacement cost of new units. As before, this covers a land rent per lot of $r^a q$ and a rent (c) for each housing unit's structure capital. This condition for the edge of the city, together with the definition of the intersection location (m), defines Equations (3.7) and (3.8):

$$R_1(m) = R_2(m), \quad \text{or,} \quad y - k_1 m - x_1^0 = y - k_2 m - x_2^0 \tag{3.7}$$

$$R_2(b) = y - k_2 b - x_2^0 = r^a q + c \tag{3.8}$$

The distance from the city center to the boundary between the two household groups (m) is determined from Equation (3.5) but is based on the number of Group 1 households (n_1). The urban border (b) is a distance such that $n_1 + n_2$ houses (with lot sizes q) can be built within that radius when the city circumference is v portion of a circle. These two conditions are used in Equations (3.9) and (3.10) to determine the intersection boundary and city edge distances (m,b):

$$m = \left(\frac{n_1 q}{\pi v}\right)^{\frac{1}{2}} \tag{3.9}$$

$$b = \left(\frac{(n_1 + n_2)}{\pi v} q\right)^{\frac{1}{2}} \tag{3.10}$$

With m and b determined by Equations (3.9) and (3.10), Equation (3.8) gives the level of consumption or welfare of the second group (x_2^0), whereas Equation (3.7) determines the same quantity for the first group (x_1^0). With each of these known, Equations (3.6) determine the pattern of rents as shown in Figure 3.2.

An important extension of our two-household model can be formulated to describe the longer-term development of land and the separation of different uses or households. Suppose, for the moment, that the housing density that each household group desires is fundamentally different. In particular, let Group 1 continue to have a higher cost of travel ($k_1 > k_2$), but now let that group also demand houses that have larger lot sizes ($q_1 > q_2$). This set of preferences could easily characterize households of different income levels. Households with higher income will certainly demand more land, since land is a normal good. Higher wages, however, also mean that the commuting time of Group 1 is more valuable relative to that of lower-income households. In most U.S. cities, higher-income households tend to live further from the city center, whereas the poor are concentrated in cities. It has been suggested that this pattern might simply represent a long-run equilibrium configuration in the U.S. land market.

In the long run (when existing housing deteriorates and is replaced), the pattern of development will be determined based on a competition between the two groups over land rather than housing. That is, land at each site in the city will be developed by that household group for which land rent is highest. We continue to assume that the housing structures of both groups are identical and can be built for a common annual cost of c. The land rents that emerge from development by each group are contained in Equations (3.11):

$$r_1(d) = \frac{y - k_1 d - x_1^0 - c}{q_1}$$
$$r_2(d) = \frac{y - k_2 d - x_2^0 - c}{q_2} \tag{3.11}$$

The pattern of land development that will emerge in this longer-run model is similar to that depicted in Figure 3.2. The crucial issue of which household group develops land at more central land sites now depends on the relative slopes of the land rent gradients in Equations (3.11). The slopes of these gradients represent the additional commuting costs incurred from an acre's worth of development: $-k_1/q_1$ and $-k_2/q_2$. Even if Group 1 has a greater cost of commuting than Group 2, its land rent gradient might not be steeper if it demands houses with much larger lots than those of Group 2. When the two groups differ because of income, the outcome depends completely on the income elasticities of land consumption as opposed to commuting costs. With income elastic land demand and inelastic commuting costs, higher-income households will have much larger lots and only slightly greater commuting costs. In this case, they will outbid other households at peripheral locations—a result that is consistent with observed location patterns in the U.S. With income inelastic land demand and elastic commuting costs, higher-income households have steeper bids and should locate centrally.[5] With these elasticities, we would have to incorporate other factors, such as concern over the quality of public services (e.g., schools), in order to explain the suburbanization of wealthy households in the U.S. We consider a variety of such other factors in Chapter 13.

In a model with multiple household groups, the comparative static conclusions 1 through 4 in the previous section continue to be true. Less land supply, larger populations (of either type), or greater lot sizes all generate a farther city edge and, possibly, intersection distance (m). This, in turn, leads to higher housing and land rents throughout the city. The main lesson from our two-household city has to do with the separation of occupancy patterns. Locations are always rented to that use (in this case household group) that is willing to offer the most. It is only by chance that two uses will offer identical or similar rents and that intermixing will occur. As a general principle, land-use segregation is a common and natural outcome in private housing or land markets, rather than a result of government regulations.

GROWTH AND RENTS

The jump from housing (or land) rents to housing (or land) prices is a complicated one. In Chapter 1, we discussed how rents that are determined in the property market get converted into asset values by the capital market. Four factors are central to determining the *rate* at which income is converted into value: (1) long-term interest rates, (2) the expected future growth of current rent, (3) the risk, or variance, associated with that rent, and (4) the federal tax treatment of real estate. Together these four factors determine the capitalization rate. Throughout this text, we will focus mainly on the role of interest rates, taxes, and expected rental growth, leaving detailed discussions of risk and its role to texts on real estate finance. In this chapter, we ignore the federal tax treatment of real estate but will explore its impact later in Chapter 8.[6]

The historic growth of urban housing (or land) rent can depend on a number of factors, but the Ricardian model suggests that the most important of these is the growth of a city's population. Cities grow and expand gradually as the population of a region or nation increases, and this growth is largely responsible for increases in locational or land rents. Capital markets look forward, however, and consider the likely future growth of rental income when determining capitalization rates. A dollar of rental income that is expected to grow is worth considerably more today than one which the market expects to remain constant. In Chapter 10, we will discuss a number of theories about how future expectations of rental growth are formed. At this point, we will simply assume that the

[5]Wheaton (1977) empirically estimated these elasticities for households in San Francisco and found an income elasticity of demand for land that was quite close to the income elasticity of commuting.

[6]The approach to prices and growth taken here builds on Capozza and Helsley (1989, 1990), who consider the impact of risk and uncertainty as well as urban growth on urban land prices.

market expects current or historic growth to continue into the future. Thus, to understand how housing prices are determined at a particular location, we must examine what is happening to rents at that location as a city grows over time. Just as with our discussion of rents, we will begin by analyzing house prices and then derive land prices as a residual.

Equation (3.3) makes it clear that the rent for housing at any location depends on the rental cost of structures, c, the agricultural rent foregone on the house lot, $r^a q$, and the cost of commuting from that location as opposed to commuting from the city's border at distance b. This border, in turn, is related to the population of the city through Equation (3.5). If we examine Equation (3.5) in detail, we can see that as the population of a city increases, the city's border grows by one-half of the growth rate of its population. If the population of a city is increasing at a rate of 4 percent annually, then the amount of developed land (area) should be expanding at the same rate, but circularity necessitates that the radius increase by only 2 percent annually.[7]

If the city population is growing smoothly at some constant rate and its edge is expanding at one-half that rate, then the border distance is growing exponentially over time. We will denote time with the variable or subscript (t), which runs from the current period ($t = 0$) out to infinity ($t = \infty$). We denote the current border or city edge as b_0 and use (g) as the constant rate at which this edge is expected to grow in the future. This is equivalent to assuming that population will increase at a rate that is twice (g). Thus, at time $t > 0$, the city edge b_t will be $b_0 e^{gt}$.[8] This expression for where the border will be at each future time period also allows us to determine the future time that any site (beyond the current border) can be expected to become developed. This is tantamount to asking at what time the border will reach a given distance. Inverting the border function, the time until development (T) for a parcel of land located a distance $d > b_0$ is $\log(d/b_0)/g$. Thus, in a city whose population is growing at a constant rate $2g$, Equations (3.12) give the border as a function of time b_t and the time at which the border will reach a particular distance $T(d)$.

$$b_t = b_0 e^{gt}, \qquad t > 0$$
$$T(d) = \frac{\log\left(\frac{d}{b_0}\right)}{g}, \quad d > b_0 \tag{3.12}$$

[7]If we take Equation (3.5) and differentiate it with respect to n, we get the following expression:

$$\frac{\partial b}{\partial n} = \frac{1}{2n}\left(\frac{nq}{\pi v}\right)^{\frac{1}{2}} = \frac{b}{2n}$$

Rearranging:

$$\frac{\partial b}{b} = \frac{1}{2}\frac{\partial n}{n}$$

[8]The Appendix at the end of this chapter provides a mathematical discussion of continuous compound growth and discounting.

Numerical Example

Returning to the example used throughout this chapter, consider a city with a current population of 2 million and a 20-mile border (b_0). If the population is increasing at 4 percent and the border is expanding by 2 percent each year, then Equation (3.12) indicates that after ten years, the border will stand at about 24.4 miles ($b_{10} = 24.4 = 20e^{0.2}$). At a 1 percent annual expansion rate, Equation (3.12) indicates that it will take almost twenty years for the border to increase from 20 to 24.4 miles: $T(24.4) = 19.9 = \log(24.4/20)/.01$.

As the city population grows at a constant rate, the increase in housing rents at any location will not be at a constant rate over time. Furthermore, the increase over any one time period will not be the same across different locations within the city. In fact, the largest percentage increases in housing rents will occur at the urban edge. To see this, we can examine Figure 3.3, where the housing rent gradient of the city is shown at two points in time, with the two city edges: b_0 and b_t. The percentage increase in rent at b_0, as the edge expands to b_t, is clearly greater than the percentage change near the center. Mathematically, we can take Equation (3.3) and examine how housing rents at any location (d) change in percentage terms as the city's edge expands at some constant percentage rate.[9]

Thus, in our highly stylized city, if the edge of development is expanding at some constant rate, the rate of increase of housing rents should be greatest at the city edge, and smallest at the city center. Furthermore, the percentage growth in rents at any given location declines over time as the border expands and that location becomes a more mature (interior) site. These features of rental growth only hold in a city where density is uniform and fixed.

Once the future border is known (or estimated) through Equation (3.12), we need only substitute this expression into Equation (3.3) to obtain the rent that should prevail at distance d at any future time period t: $R_t(d)$. Of course, housing will actually earn this rent only if it actually has been built; that is, if d is less than the (then) current border, b_t. For any site beyond the border, housing rent does not begin until construction takes place at the development date $T(d)$ specified in Equation (3.12). Thus, as the city grows over time and its boundary expands at the rate g, Equation (3.13) describes the rent for housing at any site that has already been developed, and at any time (t), when the urban border is b_t.

$$R_t(d) = r^a q + c + k(b_t - d), \qquad d \le b_t, \text{ for all } t \qquad (3.13)$$

Numerical Example

As we calculated earlier, housing rents today (which we always denote as time $t = 0$) ranged from \$7,250 (at the current edge of 20 miles) to \$11,250 at the center. Using

[9]If we differentiate Equation (3.3) with respect to the urban boundary we get: $\partial R(d)/\partial b = k$. Dividing both sides by $R(d)$, and then multiplying both sides by ∂b, yields the following:

$$\frac{\partial R(d)}{R(d)} = \frac{k\partial b}{R(d)} = \frac{kb}{R(d)}\frac{\partial b}{b}$$

For a given percentage increase in b on the right-hand side above, the percentage change in rents on the left-hand side will be largest at that location d where $R(d)$ is least: the border.

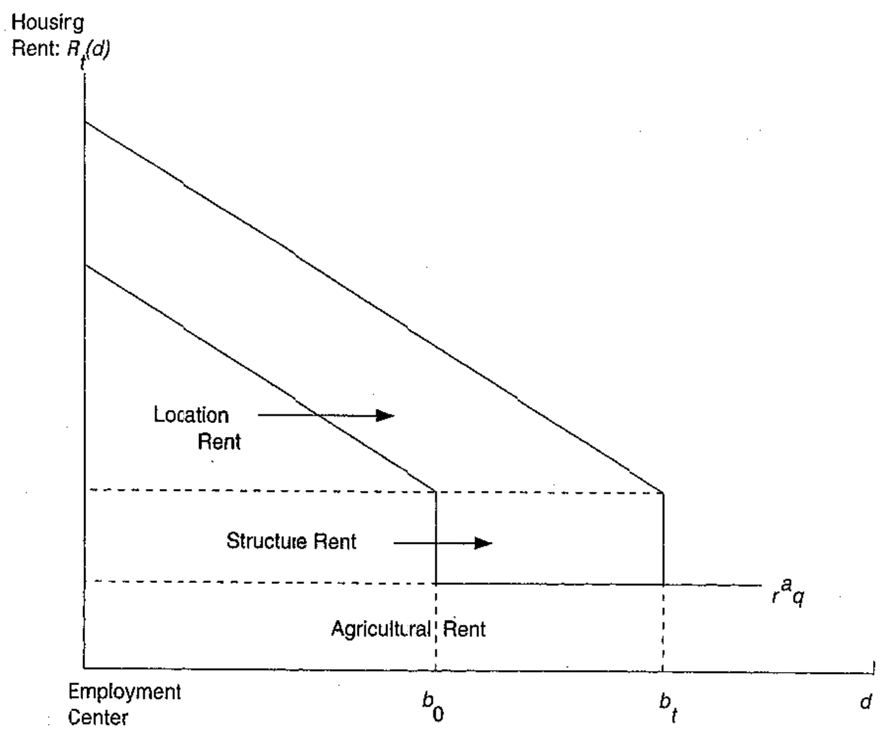

FIGURE 3.3 Housing rent gradient in a growing city.

Equation (3.13), we can calculate what the rent will be at these two locations in 10 years as growth expands the border outward. With the border expanding at 2 percent annually, Equation (3.12) says that in 10 years it will reach 24.4 miles ($24.4 = b_{10} = b_0 e^{0.02 \times 10}$). The rents at the old border will be \$8,130 (\$250 in agricultural rent plus \$7,000 in structure rent plus \$880 in location rent as a result of expansion of the border (880 = 200(24.4 – 20)). At the same time, rents in the center will be \$12,130 (\$4,000 in initial locational rent plus \$7,000 in structure rent plus the \$880 in location rent as a result of expansion of the border). As the border expands, the location rent of interior sites rises due to the increase in commuting savings from a more distant border. Note that the percentage change in rents is higher at the old border, where rents increased 12.1 percent, than at the center, where rents increased by 7.8 percent.

GROWTH AND PRICES

At any time, and in any location that is already developed, the price of housing will simply equal the present discounted value (PDV) of the rental income stream that is defined in Equation (3.13). Using an interest or discount rate of i, and applying the continuous discounting mathematics in the appendix, we can obtain Equation (3.14) for the price of existing housing

$$P_t(d) = \underset{t \to \infty}{PDV}[R_t(d)]$$

$$= \underbrace{\frac{r^a q}{i}}_{\text{Agric. value}} + \underbrace{\frac{c}{i}}_{\text{Structure value}} + \underbrace{\frac{k(b_t - d)}{i}}_{\text{Current location value}} + \underbrace{\frac{kb_t g}{i(i-g)}}_{\text{Future growth in location value}} \quad d < b_t,\ i > g \tag{3.14}$$

The four terms in Equation (3.14) are readily interpreted. The second term is the present discounted value of structure rent, or simply the cost of constructing the house. The first and last two terms together constitute the value of the house's lot. The first term is the present discounted value of the agricultural rent foregone on the lot when the house is developed. We call this the *agricultural value of land.* The third term is the present discounted value of the current location rent at the site (the commuting costs saved at the site (d) relative to the current urban border at time t (b_t). We call this the *current location value*. The final term is the discounted value of the increased commuting costs saved at the location in the future as the border grows (at the rate g).[10] This is the value of the expected increases in location rent, or the future growth in location value. The sum of these last two terms can also be thought of as the present discounted value of the location rent that will exist at the site from now on. We have separated the two so as to illustrate the following points. First, for a house located right at the border, the third term drops out. There is no current location rent at the border; the value of a house is composed only of structural rent, agricultural rent, and the expected future increases in location rent. Second, if the city is not expected to grow $(g = 0)$, then the final term vanishes, and location value is composed only of today's location rent discounted forever.

It is useful to compare the house price in Equation (3.14) with the current rental income from housing as defined in Equation (3.13). As discussed in Chapter 1, the capitalization rate is defined as the ratio of rent to price or the rate at which rents are translated into asset value. Assets with growing income streams always have lower capitalization rates than assets with fixed or declining income streams. In the Ricardian model, the fact that house rents rise fastest nearer the border and slowest at the center suggests that for developed land, the capitalization rate varies across locations. If we compare the ratio of current $(t = 0)$ rent-to-house price across locations, we get Equation (3.15).

$$\frac{R_0(d)}{P_0(d)} = \frac{i(i-g)R_0(d)}{(i-g)R_0(d) + kb_0 g},\ d \le b_0 \tag{3.15}$$

The expression for the capitalization rate in Equation (3.15) is complicated and somewhat difficult to interpret. To facilitate its interpretation, we can consider the inverse of a capitalization rate—a site's price-rent ratio, as expressed in Equation (3.16).

$$\frac{P_0(d)}{R_0(d)} = \frac{1}{i} + \frac{kb_0 g}{i(i-g)R_0(d)},\ d \le b_0 \tag{3.16}$$

The price-rent ratio in Equation (3.16) can be thought of as comparable to a price-earnings ratio for equity shares or common stocks (P/E). Since the P/E is the inverse of the capitalization rate, assets with growing income streams always have higher P/Es than assets with fixed or declining income streams.

How does the ratio of price to rent in Equation (3.16) change as the city grows? If the city is not growing $(g = 0)$, then the entire last expression vanishes and the price-rent ratio is simply the inverse of the interest or discount rate, or $1/i$. Hence, with no growth, the capitalization rate for housing is the interest or discount rate and is the same throughout the city. When the city is expected to grow, the price-rent ratio exceeds $1/i$ (the capitalization rate is less than i). At sites closer to the urban border, $R_0(d)$ is lower, and, hence, the second term is larger. Within an urban area, the price-rent ratio will be somewhat greater (capitalization rate lower) at the edge of the city, where, in this model, rents increase the fastest. Moving toward the center, the expected increase in rents is slowest, and, hence, the price-rent ratio falls (capitalization rate increases).

Numerical Example

Let's continue by illustrating how house prices vary across locations in our stylized city. Assume that the long-term interest rate is 7 percent and the border is expanding at 2 percent per year. Remember that housing rents at today's fringe of 20 miles are $7,250. Plugging these figures into Equation (3.16) yields a price-rent ratio at the border of 17.4, or a capitalization rate of 5.7 percent. The price of housing at the border is simply the current rent at the border divided by this capitalization rate, or $127,193 ($7,250/.057). At the city center, current housing rent is $11,250. Following Equation (3.16), the price-to-rent ratio at the center will be 16.3, or a capitalization rate of 6.1 percent. Dividing current rents at the center by this rate yields a price for central housing of $184,426.

There are several important lessons from this simple model about house rents, prices, and capitalization rates.

1. When there is no expected growth in the city as a whole, the capitalization rate is simply the interest rate and is constant across all locations within the city.

[10]In Equation (3.14), the last term can be derived mathematically as:

$$\frac{kb_t g}{i(i-g)} = \frac{kb_t}{(i-g)} - \frac{kb_t}{i}$$

The first term on the right-hand side above is the present discounted value of commuting costs to future borders—starting from the border at time t, b_t. We discount with the combined rate $(i - g)$ as described in the Appendix, because the border is growing at the rate g, assuming that $(i > g)$. The second term is the present discounted value of commuting costs to the (fixed) border that exists at time t. The difference is the present discounted value of the increase in commuting costs as the border grows (out from time t). These increased commuting costs will generate increases in locational rent.

2. As the city grows spatially (for example, from population increases), the capitalization rate becomes less than the interest rate and no longer is constant across locations. With faster rental growth at the urban edge, the capitalization rate is lower there than at the urban center. Over time, as the edge expands, the capitalization rate at interior sites rises.
3. House prices tend to grow over time at the same percentage rate as housing rents: faster at the urban fringe than at the urban center, and more slowly at interior locations.

Just as land rent was derived as a residual from housing rent, land prices are derived as a residual from housing prices. What this means is that land will absorb all of the anticipated increase in location rent as the city grows. If the structure cost is deducted from the price of housing, what is left is the value of the lot. Dividing by lot size (or multiplying by density) gives land price (per acre). This yields Equation (3.17) for the residual value of land that is already developed at location d and at time t (when the border is at b_t). The residual price of land at a developed site is composed of discounted agricultural rent (agricultural value), the discounted value of current location rent (current location value), and the discounted value of expected increases in location rent (future growth in location value).

$$p_t(d) = \left[P_t(d) - \frac{c}{i}\right]\frac{1}{q}$$

$$= \frac{r^a}{i} + \frac{k(b_t - d)}{qi} + \frac{kb_t g}{qi(i-g)}, \qquad d \le b_t \qquad (3.17)$$

Agric. value — Current location value — Future growth in location value

We now come to the question of how to value land that is today beyond the urban border ($d \ge b_0$) and, hence, vacant. At some future date T, this land will be developed. At that time, the land will collect its residual value according to Equation (3.17): $p_T(d)$. Until that time, vacant land will receive only agricultural rent. Thus, the price of vacant land today has two components. The first is the discounted value of agricultural rent collected between now and the time of development. The second is the residual value of the land at the time of development. Since the receipt of this residual value occurs in the future, it must be discounted back to today from the date of development. Combining these two components, Equation (3.18) gives the price of land today at a location that is currently vacant ($d \ge b_0$), and expected to be developed at the future date T. The value of vacant land is composed of its agricultural value and the expected increases in location value to be received in the future after development. For land that is today vacant, there is no current locational value.[11]

[11] The future value of land at the time it is developed $p_T(d)$ is simply Equation (3.17) evaluated at $t = T$, and at a distance equal to the future border: $d = b_t$.

$$p_t(d) = \underset{0\to T}{PDV}[r^a] + e^{-iT}p_T(d)$$

$$= \frac{r^a}{i} + e^{-iT}\frac{kb_T g}{qi(i-g)}, \qquad d > b_0 \qquad (3.18)$$

Agric. value — Future growth in locational value, discounted

Throughout this chapter, we assumed that the city develops compactly from the center outward. At any point in time, development occurs out to a border that is defined as that location where urban and agricultural rents intersect. This is based on simple rent-maximizing behavior by landowners. Now, however, we are dealing with prices and expectations about future growth. Thus, we must at least raise the possibility of some noncompact form of development. Would a landowner decide to speculate and keep a site vacant that was closer than the current border even though the rent on the site would be greater from development? Would an owner prematurely develop a site that is farther than the current border even if its urban residual rent was less than agricultural rent? Rational owners of vacant land will select that development time in Equation (3.18) that maximizes the current value of that land. If we maximize Equation (3.18) with respect to the development date T, it can be shown that this date is such that $b_t = d$; or, alternatively, $T = T(d)$, as defined in Equation (3.12). The current value of a vacant site is in fact maximized when the site is developed precisely at that time when the border has moved out to the location of the vacant site. A strategy of either speculating (waiting for the border to pass by the site), or, alternatively, developing the site before the border reaches the site only reduces the value of vacant land.[12]

We can combine Equation (3.17) for the residual value of developed land with Equation (3.18) for vacant land to obtain a complete picture of land pricing throughout the city. The various components of land prices are shown in Figure 3.4.

In Figure 3.4, the first component of urban land price (agricultural value) holds at any location and is the present discounted value of perpetual agricultural rent. Farm land actually receives this income, whereas for urban land, the loss of this rent is an opportunity cost. The triangular component represents the site's current location value, or the current commuting costs saved, again discounted forever. The rectangular component

[12] To demonstrate this conclusion, we must remember that Equation (3.18) was derived from Equation (3.17), with the assumption that at the time of development, the border has just reached the location in question, $d = b_T = b_0e^{gT}$. This allows the second term in Equation (3.17) to drop out. Without this assumption, the expression for the value of vacant land is:

$$p_0(d) = \frac{r^a}{i} + \frac{k}{qi}(b_0e^{gT} - d)e^{-iT} + \frac{kb_0e^{-(i-g)T}g}{qi(i-g)}$$

Maximizing this expression with respect to the development date, the derivative of p with respect to T must equal zero. This gives the following condition:

$$\frac{\partial p}{\partial T} = \frac{k}{qi}\{-i(b_0e^{gT} - d)e^{-iT} + gb_0e^{-(i-g)T} - gb_0e^{-(i-g)T}\} = 0$$

The last two terms above cancel, and in order for the remaining expression to equal zero, it must be true that: $[b_0e^{gT} - d] = 0$.

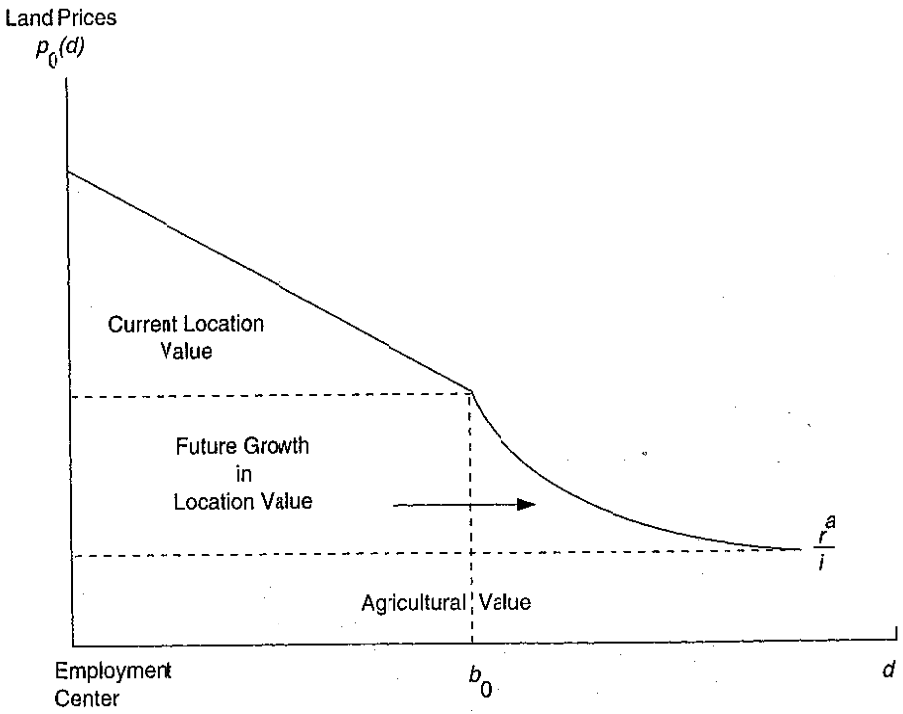

FIGURE 3.4 Components of land prices.

(future growth in location value) is the present discounted value of expected future increases in location rent at sites already developed. Beyond the current border, this value declines exponentially, since such increases in location rent will not begin until the date of future development, $T(d)$. As one considers ever more distant locations, the prospective date of development is so far into the future that, when discounted, these anticipated rental increases are worth close to zero. At this point, land has only its pure agricultural value.

Numerical Example

We can calculate the residual value of an acre of developed land today at the urban center. Remember that we assumed that there were four housing units per acre, which means that q is equal to 0.25. From Equation (3.17), we can calculate land prices by adding up the three terms. The first is the value of an acre of agricultural land: $14,286 ($1,000/.07). The current location value due to commuting cost savings is $228,571 (at $t = d = 0$; $(200 \times 20)/(0.25 \times 0.07)$). The future growth in location value as a result of anticipated increases in commuting cost savings is $91,429 $(200 \times 20 \times 0.02/(0.25 \times 0.07 \times 0.05))$. Summing the three terms yields a per-acre price for land at the center of $334,286. In

calculating the price of an acre of land at the current border, the middle term of Equation (3.17) drops out, so we get $105,715: the value of an acre of agricultural land ($14,286) plus the value of future commuting cost savings ($91,429). The price of land declines linearly from the city center to the current urban border (20 miles). From there out, vacant land prices decline exponentially, eventually reaching the value of pure agricultural land, $14,286. While these prices seem reasonable at the urban fringe, one might expect higher central prices, particularly for a large city with 2 million households that is growing at 4 percent annually. The explanation lies in our assumption throughout this chapter that density is uniform throughout the city at a level that is typical of U.S. suburban communities. We will see in the next chapter that densities tend to rise toward urban centers, which significantly increases central land prices.

It is important to point out how the various components in Figure 3.4 all change with market conditions. If the city is not growing, and is expected to continue not growing in the future, then $g = 0$ and both the rectangular area representing future value and its exponentially declining value (beyond b_0) collapse to zero. With no growth, there are no increases in location rent, and land beyond the current border will never be developed. In this case, land prices reflect only the discounted value of current location and agricultural rents. As the expected future growth rate of the market increases, the rectangular area rises and vacant land acquires considerable value because of its development prospects. This higher price for vacant land beyond the urban boundary gets reflected in higher housing prices and residual land values for sites already developed throughout the city.

This discussion of prices yields a final conclusion about the structure of land markets and how they may vary either between cities or within one city over time. Consider for the moment two cities whose current population, income, density, and transportation costs are identical. One city, however, has been at its current size for many years and is expected to remain so in the future. The other city has been growing rapidly and is expected to continue this growth in the future. The development edges of these two cities along with land and housing rents will be equal. The prices of land and housing, however, will be very different. The strong growth rate of the second city could yield prices that are significantly higher than those of its stagnant counterpart.

Within a city over time, economic or population growth often fluctuates with that of the surrounding region or the nation as a whole. As the current rate of growth fluctuates, the market may adjust its expectations about future growth. As these expectations adjust, there can be sharp changes in the value of urban land, even though current rents may not change at all. Returning to our simple numerical example, the value of an acre of land at the center and the border is considerably lower if our city of 2 million residents is not expected to grow at all (as opposed to 4 percent annually). In Equation (3.17), the third term drops out with no growth and central land is worth $242,857 rather than $334,286. Land at the border would plunge in value to $14,286 from $105,715. At the border, both the second and third terms of Equation (3.17) are equal to zero.

As we will see in Chapter 7, it not uncommon for urban growth rates to fluctuate in this range. Much of the variation of real estate prices over time is due to changes in the expected future growth of rental income, rather than to changes in the actual level of current rents. Faster expected growth of a city, all else equal, will yield lower capitalization rates. Thus, faster growing cities that are otherwise identical to slower growing cities (in population, income, and density) should have similar land and housing rents, but higher land and housing prices.

HOUSING PRICES: SOME EMPIRICAL EVIDENCE

How different are house prices across cities? Do the differences that exist correspond to those suggested by the simple Ricardian theories developed in this chapter? This is a question that a number of researchers have asked over the years but one that has proved elusive for two reasons: the lack of data and the confounding influence of other factors. In the stylized city developed in this chapter, housing structures were assumed identical and prices differed only because of the factors that affected land value: city size, geography, and residential density. In the real world, housing structures differ significantly because of history, climate, and varying regional preferences. In comparison to Minneapolis, many Los Angeles homes have swimming pools and air conditioning. Homes in the northeast have full basements—a feature rarely found in Texas. Thus, data on average sales prices by metropolitan area may not necessarily provide a good indicator of differences in house prices between markets when the typical unit can vary so much. In Chapters 4 and 8, we address the issue of quality-controlled house prices.

Keeping in mind these measurement problems, let's look at some data. Using 1990 Census data, Table 3.1 reports the median value of owner-occupied, single-family homes in each of 20 metropolitan areas.[13] We also include the three other variables that were used throughout this chapter to determine the price of urban land: the 1990 household population, the change in the number of households between 1980 and 1990, and a cost index for the construction of new homes.

Glancing at Table 3.1, it does seem the case that larger cities generally have higher house prices than smaller metropolitan areas (Los Angeles, San Francisco, Boston as opposed to New Orleans, Portland, Rochester, San Antonio). Within size categories, the relationship between values and growth is more difficult to discern. There also appears to be a relationship between house prices and the cost of residential construction. A more scientific approach to examining Table 3.1 involves a multivariate regression analysis. With the variables in Table 3.1, Equation (3.19) is estimated, with median house price (*PRICE*) as the dependent variable and metropolitan size (*HH*), growth (*HHGRO*), and construction cost (*COST*) as the independent variables.

$$\begin{aligned} PRICE = &-298{,}138 + 0.019\,HH + 152{,}156\,HHGRO + 1{,}622\,COST \\ &(10.0) \qquad (2.4) \qquad\quad (2.3) \qquad\qquad (4.2) \qquad R^2 = .76 \end{aligned} \tag{3.19}$$

[13]In Table 3.1, we use Consolidated Metropolitan Statistical Areas (CMSAs) for a number of cities. These are larger areas that often include more than one city. For example, the Los Angeles CMSA includes Anaheim, Santa Ana, Long Beach, Oxnard, Ventura, San Bernardino, and Riverside. The idea behind a CMSA is that these areas in many respects reflect a single urban area rather than a collection of smaller distinct areas. For our purposes, CMSAs better reflect the size and growth of the area as a whole.

TABLE 3.1 Median House Values and Construction Costs

	1990 Value	1990 Construction Cost Index	1980 HHs*	1990 HHs	% Difference
Boston CMSA	$176,400	248.8	1,219,603	1,547,004	26.8
Cincinnati CMSA	71,400	203.9	586,818	652,920	11.3
Dallas/Ft. Worth CMSA	78,700	187.9	1,076,297	1,449,872	34.7
Denver CMSA	89,300	198.4	609,360	737,806	21.1
Detroit CMSA	69,400	227.4	1,601,967	1,723,478	7.5
Houston CMSA	63,800	192.8	1,096,353	1,331,845	21.5
Kansas City MSA	66,500	209.7	493,485	602,347	22.1
Los Angeles CMSA	211,700	239.8	4,141,097	4,900,720	18.3
Miami CMSA	88,700	191.1	1,027,347	1,220,797	18.8
Minneapolis MSA	88,700	213.7	762,376	935,516	22.7
New Orleans MSA	70,000	188.2	418,406	455,178	8.8
Philadephia CMSA	102,300	230.5	1,925,787	2,154,104	11.9
Phoenix MSA	85,300	195.4	544,759	807,560	48.2
Pittsburgh CMSA	55,200	213.9	828,504	891,923	7.7
Portland CMSA	72,600	216.3	477,513	575,531	20.5
Rochester NY MSA	86,600	218.4	342,195	374,475	9.4
San Antonio MSA	57,300	182.6	349,330	451,021	29.1
San Francisco CMSA	257,700	267.3	1,970,549	2,329,808	18.2
Tampa MSA	71,300	191.3	638,816	869,481	36.1
Washington DC MSA	166,100	205.6	1,112,770	1,459,358	31.1

*HH, household.
CMSA, Consolidated Metropolitan Statistical Area.
MSA, Metropolitan Statistical Area.
Source: 1990 Census of Population and Housing, *Summary Population and Housing Characteristics*, CPH-1-1, Table 5; 1980 Census of Housing, *Detailed Housing Characteristics, U.S. State Summary*, HC80-1-B1; *Historical Cost Indices from Means Square Foot Costs 1994.*

The t-statistics are in parentheses below the coefficients. These results are consistent with the theory developed in this chapter. The size of the metropolitan area has a positive and statistically significant impact on house prices as suggested by Ricardian theory. Our theory also suggests that the expected rate of metropolitan growth positively affects house prices, which is confirmed as well. Finally, Equation (3.14) indicates that the cost of constructing structures should influence house prices, and, clearly, this is true in our statistical equation. With only three variables, our model is able to explain 76 percent of the variation in house prices across our sample of metropolitan areas, suggesting that sometimes even a very simple or abstract theory can provide powerful explanations of real-world phenomena.

SUMMARY

In this chapter we have shown that even simple models can be quite powerful in explaining how urban land and housing markets operate. The Ricardian model of location based on a monocentric city with fixed residential density suggests the following:

- Rents for housing, and the derived residual rents for urban land, vary by location within cities so as to exactly offset the value that households place on the advantages of those locations.
- Households and other land users such as firms compete with each other for locations within a land market. Locations are occupied by that use that derives the greatest benefit from the site's locational characteristics, and, therefore, offers the highest rent.
- Cities with greater population tend to expand horizontally, to farther or less desirable locations. This makes existing developed sites more valuable, and increases housing and land rents at all locations.
- The rate of population growth of a city is a prime determinant of housing or land prices. With faster growth comes increasing land and housing rents. The capitalized value of these increases form the basis for prices. The rent-to-price ratio or, capitalization rate for housing or land will vary both across locations at one point in time, as well as over time at one location.

APPENDIX: CONTINUOUS TIME DISCOUNTING[14]

A variable (such as rent r) that begins at r_0 and grows annually at a compound rate g, will after t years have the value: $r_0(1 + g)^t$. Now suppose that we compound the growth n periods annually, but use a growth rate of g/n. We get the expression below for the value after t years (tn periods), which in the limit (as $n = \infty$) yields the definition of a natural exponent:

$$\text{limit } (n \rightarrow \infty): \qquad r_0\left(1 + \frac{g}{n}\right)^{tn} = r_0 e^{gt}$$

Following the same procedure, a variable (such as rent r) received t years in the future can be discounted at a compound annual rate i. Today, that rent is worth: r/e^{it}, or re^{-it}.

The value (p) of a fixed annual income r, received each year from time 0 until time T, is the sum of the discounted values of this income over the received years: its present discounted value. When time is continuous, and interest or compounding figured likewise, the sum of discounted values is replaced by the integral of the discounted value function. Integrating re^{-it} from 0 to T yields:

$$p = \text{limit}(n \rightarrow \infty)\sum_{t=0}^{Tn} \frac{r}{\left(1 + \frac{i}{n}\right)^t} = \int_0^T re^{-it}dt = (1 - e^{-iT})\frac{r}{i}$$

If the income is received forever ($T = \infty$), then $p = r/i$.

[14]See Chiang (1984) for a complete discussion of discrete and continuous time discounting.

If an income stream starts out at an initial value (r_0) and then increases (from $t = 0$) at a continuous compound rate g, the discounted value of the income received at time t is:

$$r_0 e^{gt} e^{-it} = r_0 e^{-(i-g)t}$$

Following the same procedure as above, the current present value of this growing income stream (p) will be the integral of the discounted values:

$$p = \int_0^T r_0 e^{-(i-g)t}dt = (1 - e^{-(i-g)T})\frac{r_0}{(i-g)}$$

If the income is received forever ($T = \infty$), then $p = r_0/(i - g)$.

The formulas for the present value of a growing income stream are identical to those for a fixed income stream if the nominal discount rate (i) is replaced by a real discount rate ($i - g$). If the income stream is truly expected to grow at a rate that exceeds the discount rate ($i - g < 0$), then one has found the proverbial "free lunch," whose discounted value will be infinite.

The real discount rate ($i - g$) is the capitalization rate to be applied to a current dollar of riskless income expected to grow forever at the constant rate g.

REFERENCES AND ADDITIONAL READINGS

ALONSO, WILLIAM, *Location and Land Use.* Cambridge, Mass.: Harvard University Press, 1964.

CAPOZZA, DENNIS R., AND ROBERT W. HELSLEY, "The Fundamentals of Land Prices and Urban Growth." *Journal of Urban Economics* 26 (November 1989): 295–306.

CAPOZZA, DENNIS R. AND ROBERT W. HELSLEY, "The Stochastic City," *Journal of Urban Economics,* 28 (September 1990), 295–306.

CHIANG, ALPHA C., *Fundamental Methods of Mathematical Economics,* 3rd ed. New York: McGraw-Hill Book Company, 1984.

MILLS, EDWIN S., *Studies in the Structure of the Urban Economy.* Baltimore: The Johns Hopkins University Press, 1972.

MUTH, RICHARD F., *Cities and Housing.* Chicago: University of Chicago Press, 1969.

RICARDO, DAVID, *The Principles of Political Economy and Taxation* (1817; reprinted). London: J. M. Dent and Son, 1965.

STRASZHEIM, MAHLON, "The Theory of Urban Residential Location," in *Handbook of Regional and Urban Economics* (vol. 2), ed. Edwin S. Mills. Amsterdam: Elsevier Science Publishers B.V., 1987.

WHEATON, WILLIAM C., "Income and Urban Residence: An Analysis of Consumer Demand for Location," *American Economic Review,* 67 (September 1977), 620–631.

CHAPTER 4

The Urban Housing Market: Structural Attributes and Density

In Chapter 3, we began our examination of the urban housing market with a stylized and simple model of a city. In this city, all housing units were structurally the same and were built on identical plots of land. In reality, housing is a very heterogenous commodity. Lot sizes vary widely from single-family homes on several acres of land to apartment buildings where the land per unit is a small fraction of an acre. In most American cities, housing is more densely developed closer to the city center, where land values are high. Land parcels also vary widely in the amenities they provide. Lots may be level or sloped, wooded or cleared, with spectacular views, close to amenities such as a park or disamenities such as a polluting factory. The structural characteristics of housing also vary widely, from basic features such as the number of bedrooms and bathrooms, or the quality of kitchen facilities, to the more subtle features of architectural design and construction quality.

At any point in time, most houses available on the market are existing units, not newly constructed units. As a result, households rarely are able to assemble separately the individual attributes of a house into a custom package. Instead, they select that house that most closely meets their preferences and budget from a range of houses that offer different predetermined bundles of characteristics. While households must choose a complete unit, each individual characteristic of the house—from structural features to lot and location—is separately valued by the household. Hence, much can be learned about consumer preferences for individual housing characteristics by studying the patterns of choice and prices that emerge in a market of heterogeneous preexisting housing units.

The durability of housing units has a profound influence on the pattern of residential development in an urban area. In most U.S. urban housing markets, newly constructed housing units have tended to be larger over time, both in terms of interior space and lot sizes. This lower-density development pattern occurs at the fringe of the urban area, where land is more plentiful and less expensive, and stands in contrast with the older, higher-density development occupying central locations. As cities have spread out in this manner, average residential density has declined. Yet, at the same time, redevelopment has occurred at more central locations, frequently increasing residential density.

In this chapter, we build on the analysis presented in Chapter 3 by relaxing the assumptions of identical housing structures and lots to consider the heterogeneity of housing and how particular patterns of housing structures and density emerge over time. We set the stage for the chapter with a discussion of urban population density, focusing on how population density has changed over the last two decades within a particular metropolitan area. We then develop a way of statistically determining the value that households place on specific housing attributes. Finally, we present a simple model of how the density of development or redevelopment is determined over time and show how this model explains the patterns of density observed in most metropolitan areas.

URBAN DENSITY

Unlike the city in our simple model presented in Chapter 3, the density at which residential real estate is developed frequently varies widely within a metropolitan area. The underlying reason for variation in density is simple. As location rent increases the price of land, a substitution occurs between land and structure. Economists call this process *factor substitution.* At locations at which land is more valuable, development tends to use less land (the more expensive factor) per housing unit and relatively more structural capital (the less expensive factor). In other words, our models predict denser residential development at more central locations, where land is most expensive. As a result, we expect to find more multifamily buildings and fewer single-family homes as we move closer to the central city. As we will see, the relationship between density and land value moves in both directions: higher value not only encourages greater density, but greater density increases the value of land.

Casual observation in most cities suggests that residential density does indeed decline with distance from the central city. Empirically, this relationship is often measured by examining how gross population density (population divided by land area) varies with increasing distance from the center of the metropolitan area.[1] In Figure 4.1, we present the population density for 146 cities and towns in the Boston area based on the 1990 Census.[2] Together these 146 towns had a total population in 1990 of 3.86 million. The

[1] Alternatively, we could examine residential structure density—housing units per square mile. In these data, the pattern of residential structure density is very similar to that for population density.

[2] The 146 towns do not correspond to any Census definitions of the Boston metropolitan area. We defined the area to include towns within the region's outermost circumferential highway.

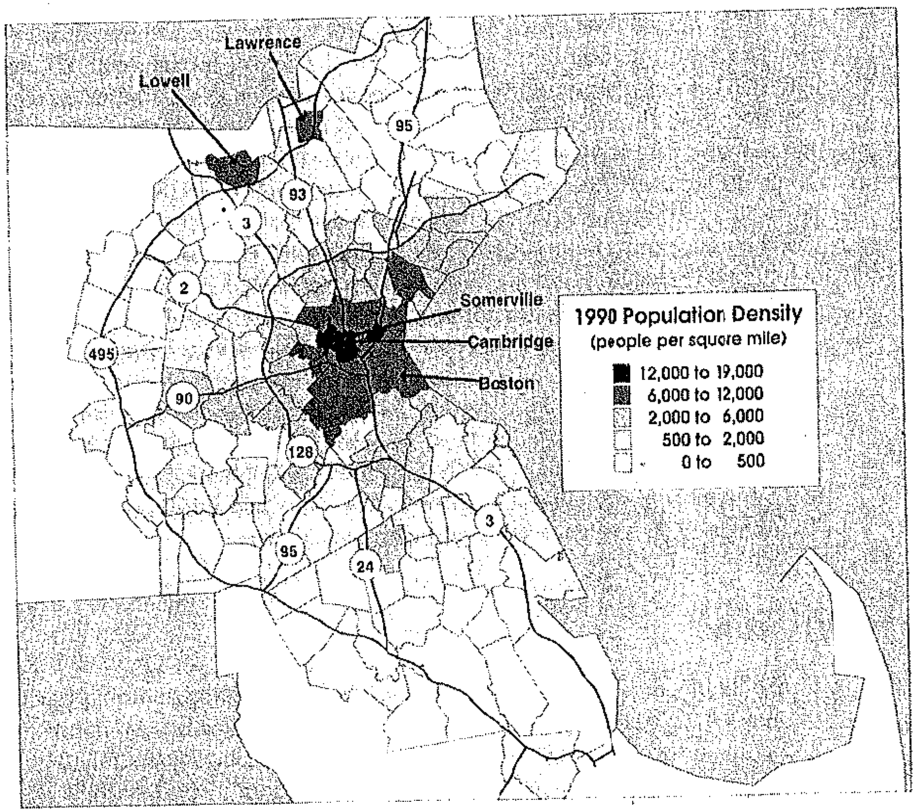

FIGURE 4.1 Population density for Boston-area cities and towns, 1990.

Source: 1990 Census.

average population density across these towns is 1,589 people per square mile, but density varies widely from 18,543 people per square mile in Somerville, which is just outside the city of Boston, to 157 in Bolton, to the far west of Boston. This wide variation and the general decline in population density with distance from the center are clear from the map.

There are some anomalies in our density data. For example, the city of Boston has a lower population density than several of its neighboring towns, such as Cambridge and Somerville. This is due to the fact that our measure of population density is based on the total land area of a city or town, not just land used for residential purposes. In the city of Boston, more land is used for commercial and institutional purposes (e.g., offices, hospitals, government, universities) than in some of its neighboring jurisdictions, a fact that drives down the population density as we measure it. Note also the two high-density towns to the northwest of Boston. These are Lowell and Lawrence, old manufacturing centers that developed historically as employment centers with dense residential development of their own.

The variation in population density with distance from the central city is often summarized through the estimation of a population density gradient. A *density gradient* is estimated by a simple two-variable regression model with population density as the dependent variable and distance as the independent variable. The standard specification is the negative exponential, which takes the form:

$$D(d) = D_0 e^{-\alpha d} \tag{4.1}$$

Here, the population density at distance d miles from the center of the city, $D(d)$, has two components. D_0 is the model's estimate of the level of density at the center and α is the estimate of the coefficient on distance, which represents the percentage reduction in density with each unit increase in distance (i.e., each mile) from the center. In order to statistically estimate Equation (4.1), it must be transformed into a linear expression by taking the natural logs of both sides of the equation, which yields:

$$\log(D(d)) = \log(D_0) - \alpha d \tag{4.2}$$

Equation (4.2) is estimated by an ordinary least squares regression, across observations consisting of towns, in which the dependent variable is the log of town density and the independent variable is the town's distance, d. In Table 4.1, we present estimates of Equation (4.2) for the Boston metropolitan area. For 1990, the coefficient on distance (α) is –0.09 and is statistically significant. This provides an estimate of the slope of the area's density gradient and means that with each mile increase in distance from the center, population density decreases by 9 percent. The constant in this model is the estimate of $\log(D_0)$, or the average density level at the center. In this sample, it is estimated at 6,634 people per square mile ($\log(D_0) = 8.8$, or $D_0 = e^{8.8} = 6{,}634$). The density 20 miles from the center is predicted to be 1,097 people per square mile. This simple model explains 53 percent of the variation in density in 1990.

In Figure 4.2, we present a scatter plot of actual town densities by distance from the center of the city of Boston with the line through the points representing our estimate of Equation (4.2). While the model fits the data well from about eight miles outward, it does considerably worse in fitting the data near the center. In fact, the actual population

TABLE 4.1 Population Density Gradients for Boston Metropolitan Area

			1990			1970		
	1990	1970	North	West	South	North	West	South
Constant	8.80	9.03	8.98	9.00	8.49	9.12	9.20	8.80
	(11.72)	(11.06)	(10.76)	(12.08)	(15.41)	(10.16)	(11.41)	(14.30)
Coefficient on Distance	–0.09	–0.11	–0.09	–0.11	–0.08	–0.10	–0.13	–0.10
	(–12.68)	(–14.14)	(–7.36)	(–8.42)	(–8.18)	(–7.80)	(–9.00)	(–9.79)
R^2	0.53	0.58	0.52	0.58	0.63	0.54	0.61	0.71
Observations*	146	146	53	54	41	53	54	41

Dependent variable: gross population density.
T-statistics are in parentheses.
*The city of Boston was used in each of the corridor density gradients.
Source: Authors' calculations, using the 1970 and 1990 Census, *Detailed Housing Characteristics*.

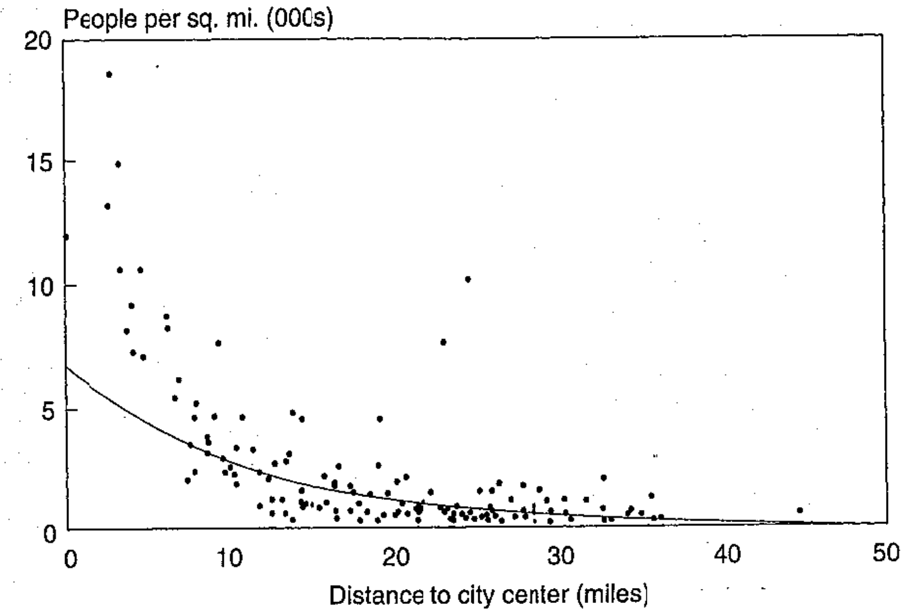

FIGURE 4.2 Boston-area population density, 1990.

Source: Authors' calculations, using the 1990 Census, *Detailed Housing Characteristics.*

density of the city of Boston in 1990 was 11,860, 79 percent higher than that predicted by the model. An inspection of the data suggests that the combination of the large number of cities and towns with relatively low density and the two major outliers between 20 and 30 miles from the city may be responsible for the flattening of the gradient at more central locations.

While it is clear that residential density tends on average to decline with distance, it is not necessarily true that density is the same at all points equidistant from the center. In Table 4.1, we also present density gradients estimated for three corridors in Boston, using only those towns to the north, west, and south. The coefficients on distance in 1990 vary from -0.11 in the west to –0.08 to the south. The flatter gradients to the north and south are largely due to older outlying employment centers (Lawrence and Lowell to the north; Brockton and Quincy to the south). The evidence presented in Table 4.1 suggests that densities vary considerably within a metropolitan area.

How does the pattern of residential density change over time? In Figure 4.3, we provide a graph comparing 1970 and 1990 estimated density gradients. The 1970 gradient is clearly steeper for the metro area. As shown in Table 4.1, the estimated gradients for the north, west, and south corridors for 1970 and 1990 all show a flattening effect over time. This illustrates a pattern of increased suburbanization of Boston's population over the past two decades.

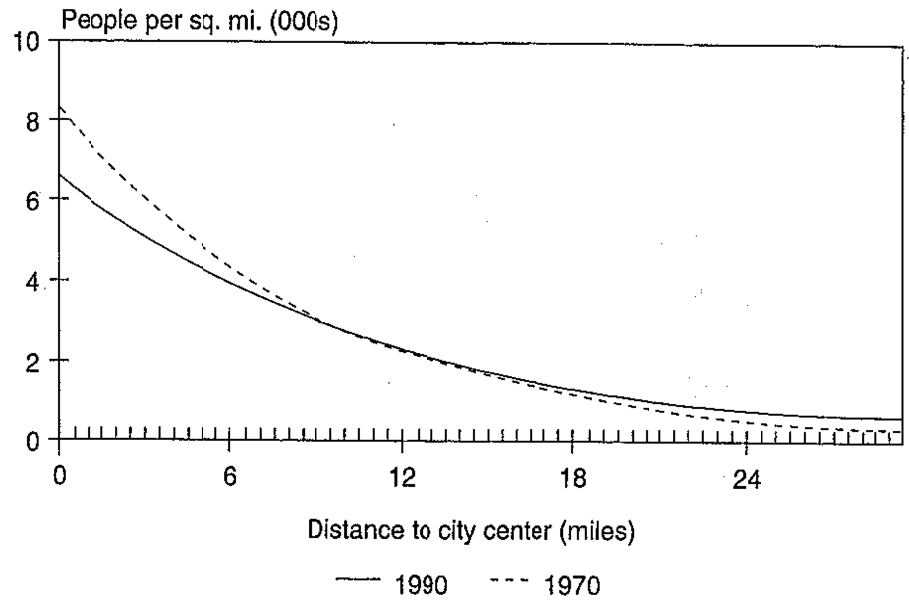

FIGURE 4.3 Boston-area population density gradients, 1970 and 1990.

Source: Authors' calculations, using the 1970 and 1990 Census, *Detailed Housing Characteristics.*

Between 1970 and 1990, the total population in these 146 cities and towns rose only 2.1 percent, or from 3.78 million to 3.86 million. However, Figure 4.4 illustrates that there have been dramatic shifts in the relative distribution of the population. The more central cities and towns lost population over the two decades, while towns at the fringe grew significantly.

While the density gradients of various cities can be quite different, in most cities researchers have observed the same trends over time that are occurring in the Boston area. During the last several decades, the population of most areas has become more decentralized in suburban communities, with a resulting flattening of the residential density gradient (Mills 1972).

HOUSING ATTRIBUTES AND HOUSEHOLD PREFERENCES

Consider two houses that are both purchased for $100,000. This purchase price reflects housing expenditures but does not imply that the quantity or quality of housing purchased is identical. In Houston, for example, $100,000 of expenditure might buy a three-bedroom ranch-style home with a large back yard in a middle-income suburban community. In New York City, the same $100,000 expenditure might purchase only a studio or one-bedroom condominium in a marginal neighborhood. Even within the same market,

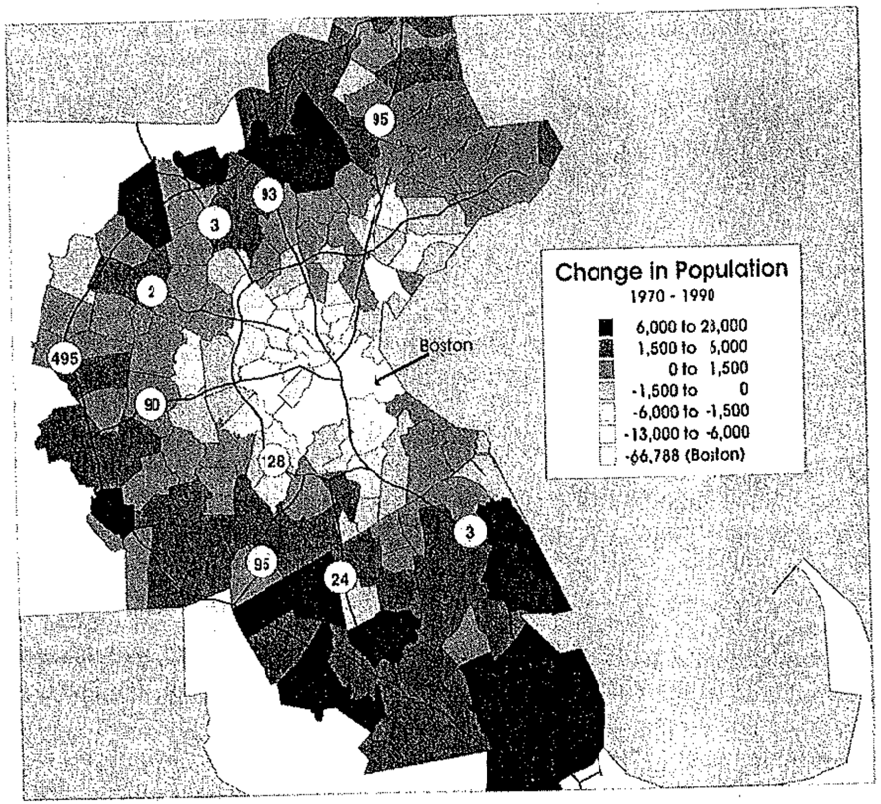

FIGURE 4.4 Change in population for Boston-area cities and towns, 1970–1990.

Source: 1970, 1990 Census.

$100,000 of expenditure might purchase a two-bedroom home with a large lot and swimming pool, or a four-bedroom house with zero lot line and no pool.

The point here is that housing is a heterogeneous commodity—houses differ in structure size and characteristics, as well as in the location and type of lot on which they sit. Because housing units are fixed in space, a household implicitly chooses many different goods and services when it selects a house, including neighborhood and school district, as well as the components of the structure itself. Identical houses on identical lots will often sell for very different prices depending on the quality of community schools or other public services. We will discuss the impact of local public services on property values in detail in Chapter 13.

In studying housing markets, it is important to distinguish between the expenditures that households make and a true measure of price. A true market price is defined for a fixed quantity of a good (e.g., price per pound of oranges or price per gallon of gasoline); expenditures are this unit price times the quantity purchased. In the housing market, we generally observe expenditures, not price per standard quantity (or quality) of housing. In

this respect, housing is quite different from other markets that economists study where standardized unit prices are directly observed.

In Chapter 3, we introduced the notion that housing rents or prices compensate consumers for locational advantage. In a competitive market equilibrium in which all households are identical, households should be as equally well off paying less at less advantageous sites as they are paying more to occupy more desirable locations. This principle of compensation across locations continues to hold as we consider housing as a heterogeneous good with many different attributes. Households examine each house in the market and choose that unit which, considering price, makes them best off. Assuming (as we did in Chapter 3) that households have similar tastes and incomes, the price of each house will have to compensate exactly for its varied attributes. In Chapter 3, housing rents were hypothesized to exactly compensate for the commuting costs associated with different locations. In this chapter, housing rents or prices will have to compensate for all of the desirable or undesirable features of each unit, such as density, size, number of bathrooms, or construction quality, as well as the locational advantages associated with the site such as commuting. What amount of money would make a household indifferent between a three- and a four-bedroom house? What is an extra bathroom worth to a potential homebuyer?

When households evaluate a housing unit they apply a valuation process that is based on the unit's various individual attributes. It is important for both sellers of existing units and builders of new units to understand this implicit valuation process of buyers, because explicit prices for individual attributes are never directly observed in the housing market. As with any economic commodity, we expect that the implicit valuation of individual attributes, like bedrooms and bathrooms, will follow the law of diminishing marginal utility: the added value of additional consumption of a commodity drops as more is consumed. To illustrate this economic principle, we turn to Figure 4.5, which shows the amount a household is willing to pay for units of different size (measured in square feet of floor area, *SIZE*). The solid line depicts how the household's total valuation for a house varies with its floor area, while the dashed line depicts the implicit valuation of each additional square foot. Both demonstrate that a household is willing to pay less per square foot as more floor area is acquired.[3]

What are the specific slopes and shapes of the curves in Figure 4.5? How much does additional square footage actually add to the price of a house? How can we quantitatively estimate the implicit price of individual housing attributes, given that these prices are not directly observed? We can measure such consumer valuations using multiple regression analysis to estimate what is called an *hedonic price equation.*[4] An hedonic price equation considers the market price paid for a house, P, to be a function of the levels of all observable characteristics of that house, X_i, $i = 1, n$. The dependent variable (housing price or rent) can be developed by tracking actual sale or lease transactions or

[3]The dashed line (or marginal value of a square foot) is the derivative of the solid line (or total valuation of the unit).

[4]According to Berndt (1991, p. 111), Court (1939) coined the term *hedonic pricing* from hedonistic philosophies found in utilitarianism. Court estimated the enjoyment that consumers receive from the individual attributes of an automobile such as speed, internal comfort, safety, and so on.

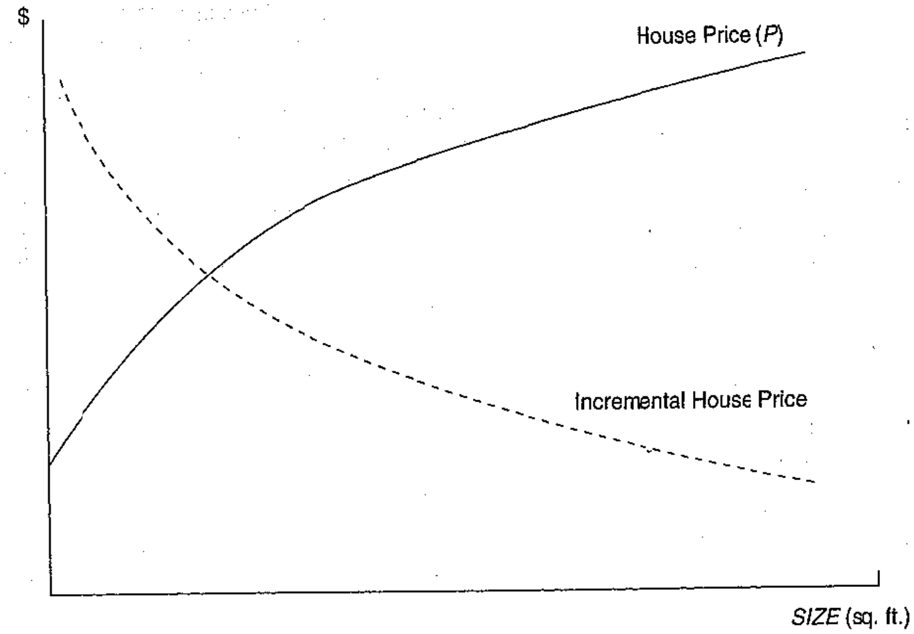

FIGURE 4.5 House prices and unit size.

by surveying current unit occupants and obtaining estimates of market price or rent. The characteristics used as independent variables include continuous variables such as square feet, integer variables such as number of baths, as well as discrete variables such as identifying whether the unit has a garage or a swimming pool. Often, qualitative judgments made by surveyors are included concerning the maintenance of the housing unit or the general quality of the neighborhood. Estimating such hedonic equations requires housing unit data that combines information on housing price or rent with a reasonably complete set of measures for the characteristics of the house and neighborhood. In its most simple form, linear hedonic equations look like the one shown in Equation (4.3).

$$P = \alpha + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_n X_n \quad (4.3)$$

In Equation (4.3), the estimated coefficients on the housing characteristics, β_i, may be interpreted as estimates of an implicit price that households are willing to pay for more of each attribute. A linear hedonic equation assumes that this price is constant and does not depend on how much of each attribute the unit has. In other words, it assumes that all of a unit's square feet or space add the same value and that there is no diminishing marginal utility with additional space.

To illustrate the linear hedonic price technique, we have taken information from the U.S. Census Bureau's 1989 American Housing Survey (AHS) for the Boston metropolitan area. The survey provides a detailed set of unit and neighborhood attributes for a sample of

TABLE 4.2 Average Attributes of 1989 Boston-Area Houses

Attribute	Mean	Standard Deviation
House Value	199,720	75,445
Number of Bedrooms	3.201	0.890
Number of Bathrooms	1.624	0.680
Age of Structure	27.281	9.354
Single-Family Attached*	0.014	0.117
Garage*	0.632	0.482
Poor-Quality Unit*	0.024	0.154
Fair or Poor Neighborhood*	0.069	0.253
Central City*	0.036	0.187

* indicates a dummy variable: 0 for no, 1 for yes.
Source: U.S. Census Bureau, *American Housing Survey*, 1989.

1,648 owner-occupied, single-family housing units.[5] In Table 4.2, the characteristics of the houses used in this analysis are listed along with the sample means and standard deviations for each characteristic (the X_i above). In this sample, the average house has a price of almost $200,000, has 3.2 bedrooms and 1.6 bathrooms, and is an average of 27 years old.[6] One percent of the units are single-family attached, and 63 percent of the units have a garage.

The AHS survey also provides various measures of the quality of the structure. The U.S. Department of Housing and Urban Development aggregates these measures to determine whether the unit meets government adequacy standards. In this sample, 2 percent of the units are structurally inadequate by these government standards. The survey includes detailed characteristics of the neighborhood as well, covering the presence of abandoned buildings in the area and the occupant's view as to the quality of the neighborhood. These variables are combined to create a neighborhood quality variable that takes on the value 1 if the neighborhood is of fair or poor quality and 0, otherwise. In the Boston area, 6.9 percent of the units are defined to be in fair- or poor-quality neighborhoods. A final variable considers whether or not the unit is located in the central city; 3.6 percent of the units are in the central city. Using these AHS data, we estimate the linear hedonic regression equation described in Equation (4.4):

$$\begin{aligned} P = {} & 61508 + 13935\ BEDRMS + 50678\ BATHRMS + 21681\ GARAGE \\ & (161.29)\ (146.40) \qquad (378.73) \qquad (134.75) \\ & - 60\ AGE - 3880\ SFA - 3425\ POOR\text{-}QUAL\ UNIT \\ & (-7.45) \quad (-6.15) \quad (-7.18) \\ & - 6175\ BAD\ AREA - 4997\ CENTRAL\ CITY \\ & (-21.18) \qquad (-12.62) \\ & N = 1168 \qquad R^2 = 0.38 \end{aligned} \quad (4.4)$$

[5] In this case, single-family units include both detached and attached units. An attached unit is a unit attached to another unit by a common wall (e.g., townhouses or row houses).

[6] The AHS is a survey of the housing stock, not a sample of housing transactions. As a result, the prices reported here are not transaction prices but rather the occupant's estimate of the value of the house. This raises questions about the accuracy of owners' estimates of value. In addition, the Census only reports value by ranges. Like other researchers, we assign each unit the midpoint value of that range. (See DiPasquale and Somerville 1995).

In Equation (4.4), t-statistics are in parentheses beneath the coefficients. More bedrooms, bathrooms, and the presence of a garage all increase the price of the home. Price declines if the unit is an attached rather than a detached unit (the variable *SFA*), or if the unit is of poor quality or located in a bad neighborhood. Price also declines if the unit is located in the central city. This result is the opposite of what is expected in the monocentric model presented in Chapter 3 in which the reduction in commuting costs closer to the CBD increases housing rents and prices. The negative impact of a central city location in Equation (4.4) may reflect other characteristics of the central city ignored in the monocentric model, such as high crime or poor-quality schools. The expected sign for the coefficient on the age of structure is unclear. Older structures could be expected to provide lower-quality services, but it is also true that older units often have unique features, such as charm or style, that home purchasers value. In this hedonic, age has a negative influence on house price. All of the coefficients are statistically significant.

Using Equation (4.4), we can explore the issue of how households value individual attributes of a housing unit, as well as how they value an entire house with a specific set of attributes. If, for example, we apply the estimated coefficients to the average characteristics provided in Table 4.2 and sum the terms, we can estimate the value of a house with those average attributes. That calculation produces an estimated average house price of $199,738, which is almost exactly the average value of houses in this sample.[7]

The coefficients in Equation (4.4) are estimates of the incremental value to be gained from more of each attribute. This hedonic equation indicates that a bathroom is the most valuable attribute, worth $50,678. A garage adds $21,681 to the value of a home. Holding all other characteristics constant, a single-family attached unit is worth $3,880 less than if the unit were free-standing.

While linear hedonic equations are frequently used in property valuations, they do have the unrealistic feature of assuming that each additional room or bathroom has the same value. As we discussed earlier in this chapter, it seems reasonable to expect that the law of diminishing marginal utility applies and that the value of additional bedrooms or bathrooms declines as more are added to a unit. By altering the specification of the hedonic model, we can permit the curvature between price and attributes implied by the law of diminishing marginal utility. A common model specification designed to address this issue takes the form:

$$P = \alpha X_1^{\beta 1} X_2^{\beta 2} \ldots X_n^{\beta n} \tag{4.5}$$

To statistically estimate the parameters of Equation (4.5), we transform it into a linear equation by taking the natural logs of both sides. This yields:

$$\log P = \log \alpha + \beta_1 \log X_1 + \beta_2 \log X_2 + \ldots + \beta_n \log X_n \tag{4.6}$$

[7]A fundamental principle in regression analysis is that when an estimated regression equation is evaluated at the mean values of the independent variables (the estimated coefficients are multiplied by the mean values of the independent variables), the model predicts the mean value of the dependent variable (see Pindyck and Rubinfeld 1991). The small discrepancy between the calculated house value and the actual mean house value is caused by rounding.

The coefficients in this model are obtained by estimating a linear regression equation in which the dependent variable is the natural log of price, and the independent variables are the natural log of the original attribute measures. Rather than determining the (constant) value of an additional unit of each attribute, X_i, the coefficients of Equation (4.6) represent the elasticity of price with respect to increases in the attribute: the percentage change in the dependent variable that results from a percentage change in the independent variable.

Using the same AHS data, we can estimate the following hedonic equation using the log-log specification.[8]

$$\begin{aligned} \text{LOG}\,P = {} & \underset{(3563.6)}{11.71} + \underset{(97.54)}{0.165}\ \text{LOG}\ \mathit{BEDRMS} + \underset{(326.09)}{0.473}\ \text{LOG}\ \mathit{BATHRMS} \\ & + \underset{(103.76)}{0.145}\ \text{LOG}\ \mathit{GARAGE} - \underset{(-5.06)}{0.004}\ \text{LOG}\ \mathit{AGE} \\ & - \underset{(-0.27)}{0.001}\ \text{LOG}\ \mathit{SFA} - \underset{(-29.70)}{0.122}\ \text{LOG}\ \mathit{POOR\text{-}QUAL\ UNIT} \\ & - \underset{(-40.77)}{0.103}\ \text{LOG}\ \mathit{BAD\ AREA} - \underset{(-4.34)}{0.015}\ \text{LOG}\ \mathit{CENTRAL\ CITY} \\ N = {} & 1168 \qquad R^2 = 0.286 \end{aligned} \tag{4.7}$$

How do we interpret the coefficients in Equation (4.7)? These coefficients are estimates of the exponents in Equation (4.5). For a discrete variable, such as the presence of a garage, we do not think about percentage changes; rather, we think only about a house either having or not having a garage. If a house has a garage, the variable has been coded with the value 2; if it does not, the value is 1.[9] Using Equation (4.5), we find that the presence of a garage adds 10.6 percent to the value of the house ($2^{0.145} = 1.106$). If the house is located in an undesirable neighborhood (another discrete variable), its value is reduced by 6.9 percent ($2^{-0.103} = 0.931$).

We predict a house's price by simply inserting values for each attribute in Equation (4.5), using the coefficients estimated in Equation (4.7) as the exponents in Equation (4.5), and then multiplying all of the terms. The price for a single-family detached house that is 27 years old; has three bedrooms, two bathrooms, and a garage; is in good shape; and is located in a good neighborhood in the suburbs would be:

$$\begin{aligned} P &= e^{11.71} 3^{0.165} 2^{0.473} 2^{0.145} 27^{-0.004} 1^{-0.001} 1^{-0.122} 1^{-0.103} 1^{-0.015} \\ &= 221{,}118 \end{aligned} \tag{4.8}$$

Using our house with the attributes defined above but varying the number of bathrooms, we can illustrate the law of diminishing marginal utility. A house with those average characteristics that has one bathroom is valued at $159,308. With two bathrooms, an

[8]The R^2 value of this equation is not strictly comparable to that of the linear Equation (4.4) and so cannot be used to judge which model fits the data best statistically. There are more advanced statistical procedures for making this comparison (see Pindyck and Rubinfeld 1991).

[9]Since we cannot take the log of 0, those variables taking on the values of either 0 or 1 are entered as 1 or 2 in estimating Equation (4.6).

otherwise identical house is worth $221,118, an increase of $61,810, or 38.8 percent. With three bathrooms, its value is $267,865, an increase of $46,747 or 21.1 percent. These calculations clearly illustrate the law of diminishing marginal utility, and may be contrasted with the linear equation (4.4) where each bath added a constant $50,678 to house value.

HOUSING ATTRIBUTES AND NEW CONSTRUCTION

Understanding the value that consumers place on specific housing attributes can provide the key to successfully developing residential real estate. It also helps to explain the evolution of the housing stock in any particular city. In the long run, it is primarily the preferences of consumers and their willingness to pay for those preferences that dictate the type and configuration of housing that gets built. As with any economic good, however, consumer prices must be judged against the cost of providing that good. Consider as an example the question of how large to make a single-family home. We might assume as a starting point that the construction cost per square foot of space is roughly constant. If this square foot cost is (C), then to construct a house of *SIZE* square feet costs $C \times SIZE$. As we have seen, however, the value of a house with *SIZE* square feet should follow the law of diminishing marginal utility. Using the hedonic model of the previous section, the value of a home with *SIZE* square feet, *P*(*SIZE*), holding all other attributes fixed, will look something like the price function pictured in Figure 4.6. The cost of constructing housing will be represented by the linear ray out of the origin. At the size indicated as $SIZE^*$, the difference between house price or value and construction cost is maximized. At that level, the incremental value of an additional square foot will exactly balance the cost of construction.[10]

If over time, more modern homes tend to be larger, for example, this could be explained by one or both of two market changes. First, the cost of construction (per square foot) could have come down through technological improvements. This would flatten the cost ray in Figure 4.6 and lead to a higher value of $SIZE^*$, the most profitable home size. Second, consumers could have changed their preferences and become willing to pay more for additional space. This would lead to a willingness-to-pay function (*P* in Figure 4.6) that is less curved. This also generates a greater $SIZE^*$ solution.

Most attributes of a house can be evaluated against their cost of construction or installation in a similar manner. Thus, as market suppliers learn about their clientele, the homes that are constructed generally are those with the most profitable configuration of attributes. In this way, the units added to the stock each year tend to reflect both the market preferences and the construction technology prevailing at that time. Thus, in any given year, the overall stock of housing is an aggregation of units constructed at different

[10]At the profit-maximizing size ($SIZE^*$), the derivative of price (or willingness to pay) with respect to square feet, or the slope of the price function ($\partial P/\partial SIZE$), will equal the marginal cost of construction (C) or the slope of the cost function.

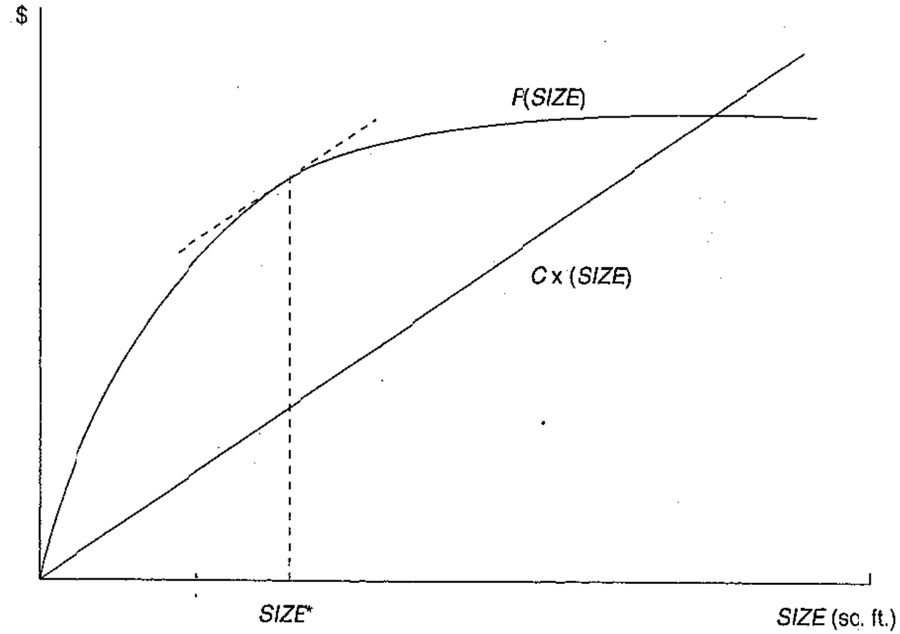

FIGURE 4.6 Willingness to pay and unit size.

periods with attributes based on the demand and costs in effect at the time each unit was built (Harrison and Kain, 1974). The one exception to the profit-maximizing principle just discussed involves unit density, or lot size. Determining the most profitable density for development is more complicated than the procedure for other attributes. Since density is the housing attribute that determines a city's overall pattern of residential development, we need to consider it in some detail.

RESIDENTIAL DENSITY, LAND VALUE, AND HIGHEST USE

When land is developed, its most central characteristic is the density of the development (i.e., the number of housing units per acre). As with any attribute, density is determined in a manner that maximizes the profit from development. As has been discussed both in this chapter and Chapter 3, a site's locational features and the structural characteristics of potential housing determine the amount of location rent that will accrue to any particular housing *unit*. In this respect, density is no different from any other attribute. As the first part of this chapter made clear, greater density should tend to reduce a unit's value through the loss of open space, green area, and privacy. However, density also determines the number of units to be developed per acre. In this respect, it is an important determinant of the location rent *per acre* that can be obtained from the site.

A landowner-developer contemplating the development of a site wants to maximize the residual profits to be obtained from the land, after construction costs. Thus, the density of development should be that which maximizes the potential residual value of the land. Let's assume here for simplicity that the only type of development is for residential use. In evaluating different densities of residential development, the developer must consider how the consumers' willingness to pay for units will vary with density, as well as how density increases the number of units to be placed on the site. Since consumers in general are willing to pay less per housing unit as density increases, there exists a tradeoff. Greater density reduces the value and, hence, profit from each unit, but increases the number of units that can be placed on the land. The former reduces site profits or value, while the latter increases it. A developer must balance these two forces in seeking the highest return to a site, rather than simply assuming that greater density yields higher profits and residual value. Let's work through this tradeoff more carefully.

Density can be measured in two ways: as the ratio of housing *units* to total land area of the property, or as the ratio of total housing *floor area* to total land area, commonly referred to as the FAR (*floor area ratio*). Throughout this chapter we will use the FAR measure, although our discussion can easily be recast using the units measure. We will refer to the FAR ratio with the variable F. As already discussed, we expect that, holding all other attributes of the unit and location constant, consumers will pay less for a housing unit that is in a taller building or a more dense development. In an hedonic model like those presented in the previous section, we would expect to get a negative coefficient on density. Using those hedonic models in the previous section, we can define an hedonic equation for the price per square foot of floor area (P) in a housing unit: $P = \alpha - \beta F$. The coefficient α represents the collective value of all other locational and housing attributes that can affect the price (per square foot of floor area) of a house, while β represents the marginal reduction in value that occurs as the house lot is reduced and its density or FAR ratio increased.[11]

We also expect that the cost (per square foot of floor area) of constructing housing units (C) will vary with the FAR of the residential development. In fact, the cost per square foot of construction tends to rise with greater FAR due to increased foundation work, greater structural support, or the necessity of elevators. For simplicity, we abbreviate the cost of construction as: $C = \mu + \tau F$. Here, μ represents a basic cost of construction (per square foot) and τ the incremental additional cost (assumed linear) as density is increased.

In the top panel of Figure 4.7, we present the price schedule and construction cost schedule per square foot of floor area as functions of FAR. The profit (per square foot of floor area) to be made from constructing a housing unit is simply $P - C$, or the vertical distance between the price and cost schedules. Clearly, profits per square foot of constructed floor area are reduced as the site's FAR of development is increased. Profits are

[11] As in Chapter 3, we will adopt the convention of using upper-case letters to refer to housing variables, lower-case letters for land variables, and Greek symbols for parameters or coefficients.

The assumption that the FAR reduction is linear is made only for ease of mathematical exposition. In fact, given that P is expressed per square foot, the negative relationship between P and FAR is actually somewhat curved.

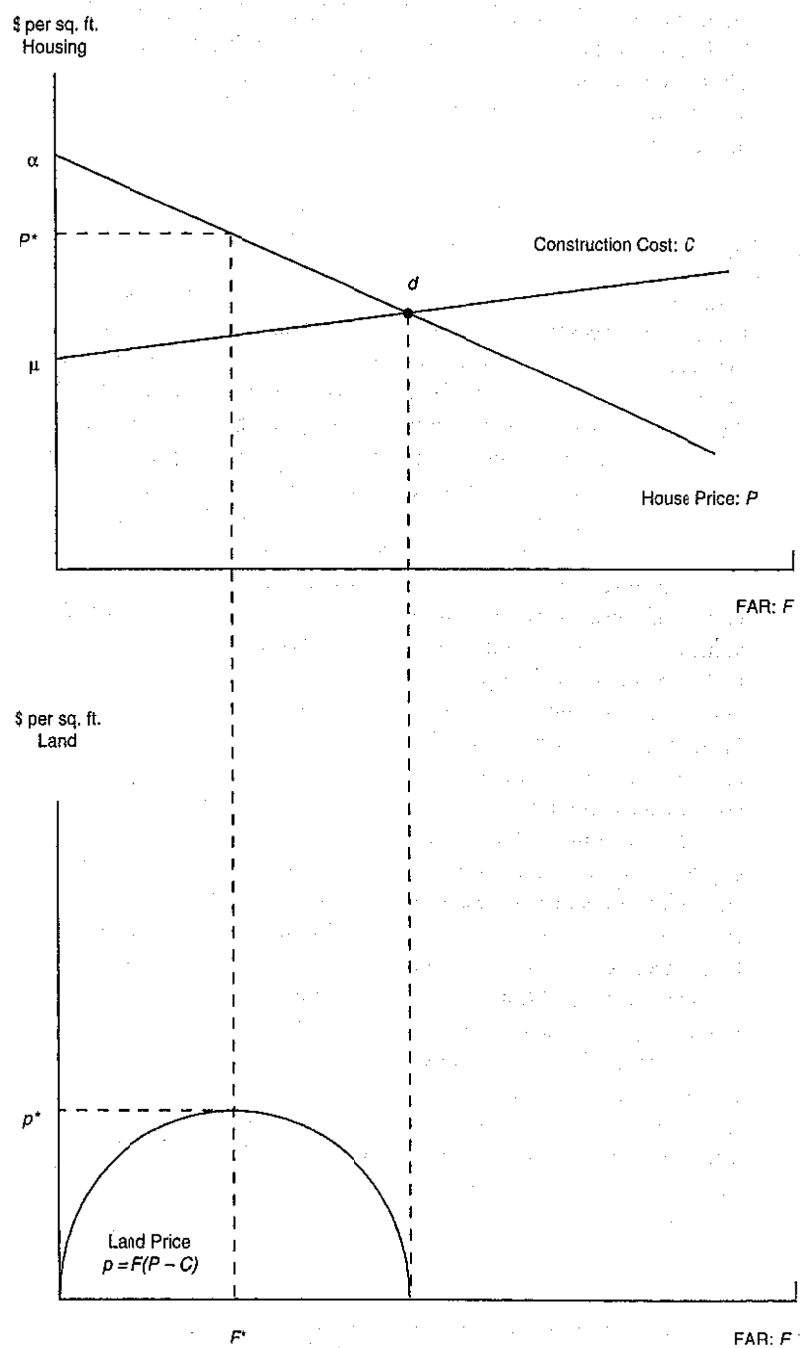

FIGURE 4.7 Optimal FAR.

zero at the point (d) where the two schedules intersect, positive to the left of (d) where price exceeds costs, and negative to the right where costs exceed price.

What is the profit that accrues to land? While both P and C depend on F, they each are measured in terms of square feet of *floor area*. The residual value of land will be measured as a profit per square foot of *land* area. To get the latter from the former is somewhat complicated, but is the key to determining a site's maximum profitability.

We begin by taking the housing profit (per square foot of floor area) illustrated in the top panel of Figure 4.7 as the difference between the price and construction costs schedules. We then multiply this profit per square foot of floor area by the value of F on the FAR axis (the ratio of floor area to land area). This yields a dollar value per square foot of land area—the residual profit $p = F[P - C]$. This residual profit per square foot of land is shown in the bottom panel of Figure 4.7. At the origin, the value of p must equal zero, since $F = 0$. At the point d where $P = C$ in the top panel of Figure 4.7, p must also equal zero, since with no profit per floor area, there can be no residual value to the land. In between, the value of p rises, reaches its maximum value at F^*, and then falls. Moving up from F^* and over to the vertical axis, p^* is the value (per square foot) of land evaluated at the FAR level F^*. Moving up from F^* to the house price function (P schedule) in the top panel of Figure 4.7, we can go over to the vertical axis and find the value of floor area at the optimal FAR, P^*. The mathematical expressions for F^* and p^* are:[12]

$$F^* = \frac{[\alpha - \mu]}{2[\beta + \tau]}$$

$$p^* = \frac{[\alpha - \mu]}{2} F = \frac{[\alpha - \mu]^2}{4[\beta + \tau]} \tag{4.9}$$

The development's maximum land profit per square foot is, in principle, an equilibrium value for what land is worth. Regardless of what was paid for the site, the density of development should follow Equation (4.9) even to minimize the losses from an overly high original purchase price. Of course, in the longer run, the maximum residual value from residential development at any site must exceed both the rural opportunity cost of land and the value that other uses might yield in order for residential development to make sense. Or, put another way, for land to be developed residentially in equilibrium, the FAR chosen for each site must not only be that which maximizes residual land profit, but this land profit in turn must exceed that from other uses.

To illustrate the important role that density plays in determining a development's profitability, we can consider the market for condominiums in a fashionable and historic area of central Boston. To do this, we use data on 578 condominium sales transactions that occurred in the Beacon Hill and the Back Bay sections of downtown Boston in 1984.[13] The housing structures in this area are predominantly brownstones or small brick walkup apartment buildings, but occasionally there is a larger elevator-serviced building. Our price data for these condominiums are all expressed as price per square foot. As can be seen in Table 4.3, the average price per square foot in 1984 was \$150, with prices ranging from a high of \$190 per square foot on prestigious Beacon Hill to \$144 on Commonwealth Avenue. The average number of both bedrooms and bathrooms in the sample is 1.4 per unit. An important consideration in deciding to locate in a downtown area is the availability of parking; 25 percent of the units in this database had on-site parking. While we do not have detailed data on locational amenities, we defined the center of the downtown area as the Boston Common and calculated the distance of each unit to the Common. On average, these units are 5,167 feet from the Common, or just under 1 mile. Finally, the average number of stories in each of the buildings is 7.6.

Using these data, we can estimate an hedonic equation for sales price per square foot (rather than total sales price). Again, to keep the mathematics simple, we will estimate a linear equation, even though we argued earlier that a nonlinear function is more consistent with economic theory. Using 578 sales observations in 1984, the following equation was estimated:[14]

$$\begin{aligned} PSQFT = {} & 181.0 + 6.2BED + 2.7BATH + 23.1PARK - 38.6BDUM \\ & (14.8)\quad (2.6)\quad (0.8)\quad (5.9)\quad (-3.2) \\ & - 40.5MDUM - 49.1CDUM + 0.0007CDIST - 1.48STORY \\ & (-3.3)\quad (-4.0)\quad (1.2)\quad (-3.8) \\ & N = 578 \qquad R^2 = .16 \end{aligned} \tag{4.10}$$

Beneath each coefficient, the t-statistic is presented in parentheses. Given the linear specification, the coefficients in Equation (4.10) should be interpreted as dollars per square foot. If the unit comes with parking, it adds \$23 per square foot to the value of the unit. As expected, bedrooms, bathrooms, and parking all have a positive impact on the price per

[12]The residual profit per square foot of land development (p) equals the floor area profit multiplied by the development FAR: $F[P - C]$. To find the value of F that maximizes land profit, we begin by substituting in the expressions for floor area price and construction costs:

$$P = \alpha - \beta F \quad , \quad C = \mu + \tau F$$

$$p = F[P - C] \; = \; F[\alpha - \mu] - F^2[\beta + \tau]$$

Setting the derivative $\partial p/\partial F$ equal to 0, we solve for the F value (F^*) where this condition holds:

$$\frac{\partial p}{\partial F} = [\alpha - \mu] - 2F[\beta + \tau] = 0$$

$$F^* = \frac{[\alpha - \mu]}{2[\beta + \tau]}$$

We can now obtain p^* by substituting our expression for F^* into that for p:

$$p^* = \frac{[\alpha - \mu]^2}{4[\beta + \tau]}$$

[13]The data for the Back Bay neighborhood encompassed only the three main streets in this area—Marlborough, Beacon, and Commonwealth.

[14]The variables in this equation are:

PSQFT: Sales price per square foot	*BATH*: Number baths
BED: Number bedrooms	*BDUM*: Unit on Beacon St.
PARK: Unit has parking	*CDUM*: Unit on Commonwealth Ave.
MDUM: Unit on Marlborough St.	*STORY*: Number stories in building
CDIST: Distance from Common	

Units not located on Beacon, Marlborough, or Commonwealth streets are on Beacon Hill.

TABLE 4.3 Characteristics of Boston Back Bay Condominiums, 1984

	Mean	Standard Deviation
Price per Square Foot	149.84	39.27
Beacon Hill	190.08	32.23
Marlborough St.	156.50	40.28
Beacon St.	150.90	41.22
Commonwealth Ave.	143.55	34.68
Parking	0.25	0.43
No. of Bedrooms	1.44	0.82
No. of Bathrooms	1.43	0.59
Distance to Boston Common (ft.)	5,167.41	3,112.47
No. of Stories	7.62	4.39

The following location distribution applied to the units used in this study: Beacon Hill, 1.73%; Marlborough St., 15.57%; Beacon St, 47.23%; and Commonwealth Ave., 35.47%.
Source: Greater Boston Board of Realtors.

square foot of the unit, although the coefficient on bathrooms is statistically insignificant. Locations on the three streets in the Back Bay has a negative impact on value relative to a location on Beacon Hill, which was the default location left out of the equation. Surprisingly, the coefficient on distance from the Boston Common is positive, but statistically insignificant. Finally, the taller the building that houses the unit, the lower the value of the unit.

The data in our sample suggest that the Boston Back Bay area has an average FAR of about 7.5, but this does not necessarily reflect the market demand for housing in this neighborhood. The older historic townhouses tend to be 4 to 5 stories tall, while many buildings constructed between 1920 and 1960 vary between 8 and 12 stories. Since 1960, the neighborhood has been an historic district in which a public board limits the FAR of new development to be compatible with the older townhouses. To determine the most profitable FAR (in 1984), we can take a prototypical two-bedroom, two-bath unit with parking and determine its base value for all terms in the hedonic equation above, except for number of stories. This becomes the value of the coefficient α in Equation (4.9). For Beacon Hill, the most prestigious location in the sample, these terms add up to a value of \$222.04 per square foot $(181.0 + (6.2 \times 2) + (2.7 \times 2) + 23.1 + (0.0007 \times 200) = 222.04)$. The contribution of FAR to the price equation (the coefficient β in Equation (4.9)) is \$–1.48 (for each additional story). Using these estimates, the price per square foot for the unit would be $\$222.04 - 1.48F$. If the unit was located on Commonwealth Avenue rather than Beacon Hill, the equation would be $\$176.89 - 1.48F$.[15]

For a construction cost equation, consultation with local architects suggests that in 1984, square-foot construction costs could be closely approximated with the function $C = 100 + 2F$. Following the steps in Equation (4.9), we find that the optimal FAR for new construction on Beacon Hill would be 17.5—more than twice the average height in our sample of buildings in this area $(F^* = (222.04 - 100)/2(1.43 + 2) = 17.5)$. Using Equation

[15]In this example, the Beacon Hill value is calculated with the three locational dummy variables set to zero and distance to Boston Common set at 200 feet. The calculation for the Commonwealth Avenue location is based on $CDUM = 1$ and the average distance for units on that street to Boston Common of 5,855 feet.

(4.9), we can determine that at the suggested FAR of 17.5, land is worth \$1,068 per square foot $(p^* = (222.04 - 100) \times 17.5/2)$. With 43,560 square feet in an acre, the value per acre of land (at an FAR of 17.5) is \$46.5 million. At a typical townhouse FAR level (four stories), land for new development would be worth only \$10.6 million. In historic districts, or any other area in which height restrictions are in place, the reduction in land value because of lower FAR levels (relative to the optimal FAR) might be viewed as the cost of imposing such land-use regulations. In some situations, however, this cost may be justified by a broader increase in values, a topic that we will discuss in detail in Chapters 13 and 14.

LOCATION AND RESIDENTIAL DENSITY

At the beginning of this chapter, we argued that density increases at sites with greater location rent because as land rents rise, developers substitute away from land and use more structure capital. More valuable land is used more intensely. As we consider residential density at different sites within a city, density normally tends to be higher at more valuable locations where location rent is greater. This is a pattern seen in cities throughout the world, at least in those in which private land markets determine development. Along oceans, lakes or rivers, there are often walls of high-rise apartments. Real estate is developed at similarly high densities in downtown areas. Even in the suburbs, clusters of apartments and townhouses tend to be developed near shopping areas or business districts, or at town centers. Let's explore in more detail the argument that it is more profitable to develop more desirable sites at higher density.

Referring back to the expressions in Equation (4.9), as well as to the information in Figure 4.7, we pose the question of how the relationship between FAR and land price would change as the analysis is done at different locations. As we move to a more desirable location, the intercept of the house price function (α) increases. In the diagram, this shifts the price line upward. Of course, this will change the solution value for the optimal FAR (F^*), and the resulting residual value for land p^*. From the first part of Equation (4.9), it should be clear that a greater α will increase the solution value for F^*. In the second part of Equation (4.9), this in turn increases the potential value of land p^*.

To see this result graphically, we can imagine Figure 4.7 with a higher housing price (P) schedule—presented as the dashed line in the top panel of Figure 4.8. Let's assume that construction costs have not changed. With this higher price schedule, the maximum FAR that can still cover construction costs will shift outward. Thus, at any FAR, the profit per unit will be greater and the triangular area will expand. Increased location rent always generates greater potential housing profits (per square foot of floor area). When this greater profit per floor area is multiplied by FAR, the result is that the semicircular land profit value in the bottom panel of Figure 4.8 both rises and shifts to the right. This rise and shift means that the maximum value of the land profit function will occur at a higher value of F^*, which, in turn, generates a greater residual land value p^*.[16]

[16]Returning to the example of Boston's historic district, we can calculate the optimal FAR at the less prestigious location on Commonwealth Avenue, where the price equation is $176.89 - 1.48F$. The result is an optimal FAR of 11.0 (rather than 17.5) and a land value of \$425 per square foot (rather than \$1,068).

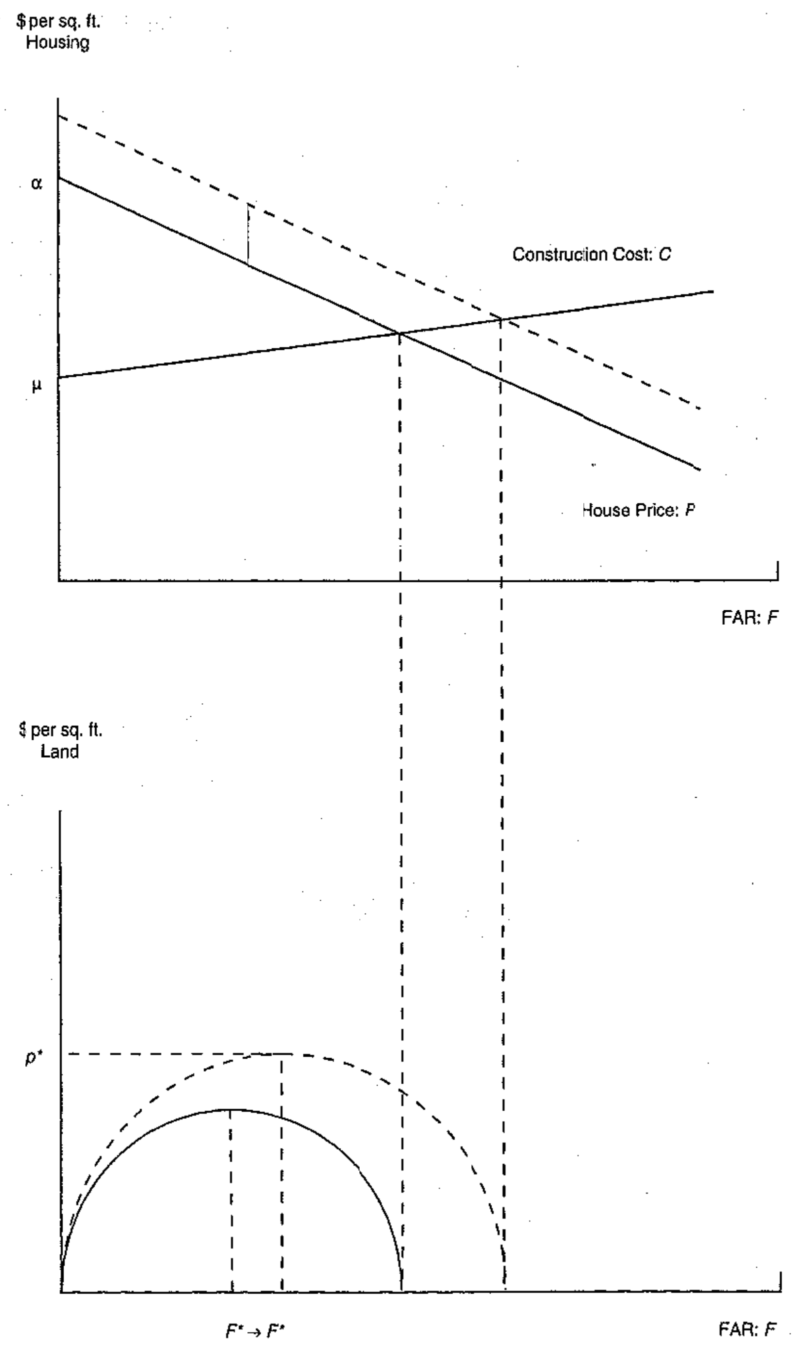

FIGURE 4.8 Optimal FAR: shift in housing price.

If this type of analysis were conducted at each location within the city, the result would be an ideal or potential market density gradient of residential development. Downtown sites would be more densely developed, as would sites at suburban transportation intersections. Sites near desirable natural or manmade amenities would also be developed at greater density, as well as would those locations with exceptional views or vistas. Wherever, the rule would apply: increased location rent encourages a substitution away from land usage, and this yields higher development density. Of course, in Chapter 3, we saw that greater density also increases the locational rent at desirable locations. When we compound these two effects, the result can be the extremely high densities and land prices often found in downtowns or along waterfronts.[17]

The question to ask is whether actual cities have patterns of residential density that look like this. In the beginning of this chapter, we presented some empirical density gradients for cities and towns that exhibited many of these traits (Figures 4.2, 4.3). In most cities, residential density is greatest both at the center and at strategic suburban nodes. It would be wrong to infer from this similarity, however, that cities generally have optimal or market development densities at most locations. Would the small single-family houses surrounding downtown Los Angeles be developed today if land were vacant, or does this pattern of housing belong to a period of 60 to 90 years ago when the structures in this area were built? Would Boston have the same extensive neighborhoods of row houses if that land were rebuilt today? In examining the determinants of actual residential density patterns, modem development at market density is always competing with the value of existing structures built years earlier. When little vacant land is available for development, the density observed at desirable sites may be more the product of market forces decades, or even centuries, ago. This is certainly the case in the Back Bay area of Boston that we examined in the previous section. Existing housing provides a large opportunity cost against which new development often cannot compete.

PATTERNS OF URBAN DEVELOPMENT AND REDEVELOPMENT

As cities grow over time, the location of new development tends to occur in distinct patterns. As a rule, new development mainly occurs on vacant land at the existing edge of the city. In this sense, the tendency is for population growth to increase urban development horizontally, much as in the simple model of Chapter 3. The density of development that occurs at the urban edge will be based on two considerations: the shape of the housing demand and construction cost schedules (in Figures 4.7 and 4.8) and the value of vacant fringe land (based on agricultural rent and the value of the development option).

[17]As an example of this compounding substitution effect, we can refer back to definition of p^* in Equation (4.9). As house prices increase at a better location (an upward shift in α), the residual land value rises not linearly but exponentially (to the power of 2) as α is increased.

In fact, the residual value per acre from new residential development at the fringe should, in equilibrium, exactly equal the price of vacant (agricultural) land.

For any city, the equality at the urban fringe between the residual value of developed land and vacant land prices serves to determine both the price of new housing and the density of new development. This condition for market equilibrium effectively links the results of Chapter 3 with those here. The price of new housing at the urban fringe will be such that when land is developed at its optimal FAR, its residual value [p^* in Equation (4.9)] will exactly equal the price for vacant agricultural land ($p_0(b)$) as defined by Equation (3.18).[18] This yields the important conclusion that the density of development at the edge of a city is *not* based on transportation costs, the size of the city, or other determinants of interior land prices. At the urban fringe, the price of land is fixed at its agricultural value plus the value of the development option; this dictates the density of fringe development.

It is important to reiterate that while the population of a city and its transportation system will surely determine how far the urban fringe extends, they do not have much impact on fringe land prices or density. Thus, the density of new fringe development in Los Angeles will be quite similar to that occurring in the tiny metropolitan area of Bakersfield, California, even though rents at interior locations in these two cities will be dramatically different. This important conclusion explains why there often exists low-density single-family housing at very central sites in many newer American cities. These interior sites often were the urban fringe just 50 years ago, and market conditions may not have changed over the intervening years. Thus, the housing built 50 years ago may use a similar density to that built at today's fringe, even though population growth has vastly expanded the city and pushed the fringe far past its original borders.

As cities grow horizontally, the price of interior housing, built years earlier, may rise or fall. If the housing is maintained, its price will normally increase because the growing city is creating more and more location rent for interior sites. If the housing is not maintained, its price could decline if the loss in value from the deteriorating structure exceeds the increase in locational site value. In some cases, older housing may completely deteriorate or be destroyed, creating pockets of vacant land. Such vacant land creates the opportunity for new development to occur within the city rather than just at the urban fringe.

When vacant land becomes available within the interior of a city, it is developed at a current market density, sometimes in stark contrast to the historic density of surrounding structures. This is easy to understand. The older structures were built at a time when the area was at the fringe of a much smaller city. Being at the (then) fringe, land had little location rent or value. Now in a much larger city, if the area has acquired location rent, the optimal FAR for modern development may have increased dramatically. These differences raise the question of whether the market must wait for housing to deteriorate or be

[18] In the housing demand equation, the price of housing is represented by $\alpha - \beta F$. In market equilibrium, and at the optimal density, this price must yield a residual land value ($p^* = F^*[\alpha - \beta F^* - \mu - \tau F^*]$), which, in turn, equals the price of vacant land. To meet this condition, α (the hedonic value of all housing attributes besides FAR) becomes endogenously determined, as was true of the models in Chapter 3.

destroyed before vacant land in the interior becomes available. Is it ever profitable to purchase and demolish perfectly sound housing simply because the value of a site as vacant land exceeds its value with the existing housing already on it?

When existing structures are purchased and demolished to create vacant land for new uses, we say that the site is being *redeveloped*. Historically, widespread redevelopment of urban sites has been an important force in shaping the land use patterns of many cities. Consider New York or Boston, for example. Around 1800, both of these cities had predominantly detached wood-frame housing. Over the next 50 years, most of this housing was replaced with three- to five-story row houses of masonry construction, at much higher FAR levels. Today, virtually none of the first-generation housing remains. While fires contributed some to this transition, historical documents clearly show that many owners simply demolished and rebuilt their sites as these cities grew very rapidly in the first half of the nineteenth century. After 100 or more years, these same cities experi-

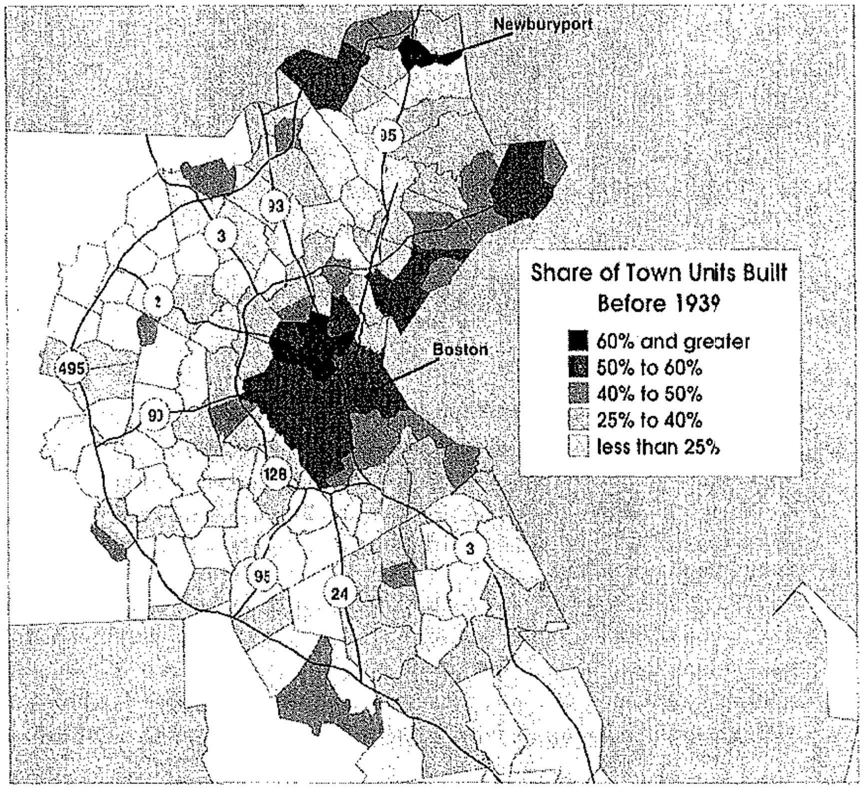

FIGURE 4.9 Town units built before 1939 in Boston-area cities and towns, 1990.

Source: 1990 Census.

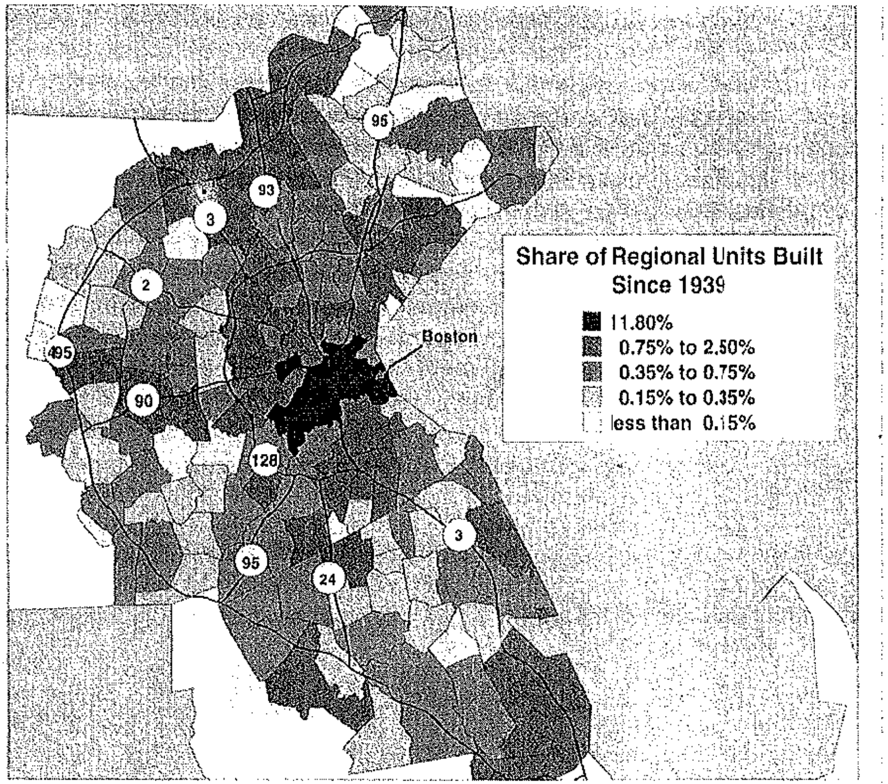

FIGURE 4.10 Regional units built since 1939 in Boston-area cities and towns, 1990.
Source: 1990 Census.

enced another wave of redevelopment to a third generation of housing. Particularly at very valuable sites such as those near parks and transportation terminals, row houses were replaced with elevator-serviced apartment buildings, involving a further increase in the site's FAR. Thus, in older cities, the higher-density structures that we often observe today are third-generation housing that resulted during more recent periods of vertical redevelopment.

To illustrate how such patterns of redevelopment can occur, we have displayed the fraction of housing units that were built prior to 1939 for each town in the Boston Metropolitan area (Figure 4.9). Clearly, there is a strong tendency for the average age of housing to decline as one moves from the center city outward.[19] To show the extent of redevelopment in the region, Figure 4.10 examines the locational distribution of housing

[19]Anomalies occur in some of the older outlying towns, which developed during the early colonial period.

built within the region since 1939. Almost 12 percent of all housing built during this period occurred on redeveloped sites within the fully developed city of Boston. Older center-city sites within the Boston metropolitan area have been the target of extensive redevelopment during recent decades.

Redevelopment represents an adjustment process by which housing capital is gradually replaced. In the process, historic density is upgraded to more modern market density. The speed and ease with which this adjustment occurs is very important. With rapid adjustment, cities will look quite modern, with vertical redevelopment occurring at the same time as horizontal development happens at the urban fringe. The result is that density throughout the city will be closer to those market levels dictated by today's economic conditions. With slow adjustment, the housing stock of cities will have an evolutionary character to it, with growth occurring primarily through horizontal expansion at the urban fringe. This suggests that we should closely examine the economic conditions that will generate redevelopment.

For redevelopment of a site to occur, the net residual value to land if developed optimally must exceed the gross value of *land and capital* that currently exists on the site plus the cost of demolishing the old capital. To illustrate the conditions under which this might happen, we can expand our previous discussion using the simple equations for housing and residual land prices. Let's suppose that the existing historic housing on a site has the value (per square foot of floor area): $P^0 = \alpha^0 - \beta F^0$. Here, the term α^0 represents the hedonic value of the existing housing capital on the site in contrast to the hedonic value of modern housing capital, the term α in Equation (4.9). Similarly, F^0 is the site's preexisting historic FAR in contrast to the optimal FAR, F^* in Equation (4.9). The total value (per square foot of land) of preexisting historic land use is: $p^0 = F^0\,[\alpha^0 - \beta F^0]$. Redevelopment can occur only if the net land return (per square foot) from new development, p^* in Equation (4.9), exceeds p^0 by more than the cost of demolishing the existing structure. We can approximate the demolition costs associated with clearing a square foot of land through the term δF^0, where δ is the demolition cost per square foot of floor area. Substituting in, the condition that p^* exceed p^0 by these demolition costs reduces to:[20]

$$p^* - p^0 > \delta F^0 \text{ implies: } \frac{F^*}{F^0} > \frac{\alpha^0 - \beta F^0 + \delta}{(\alpha - \mu) - (\beta + \tau)F^*} \qquad (4.11)$$

Examining Equation (4.11) more closely, it is possible to draw several conclusions. Suppose for the moment that the existing housing capital on the site has been completely maintained so that its value equals that of current new capital ($\alpha^0 = \alpha$). The right-hand side of Equation (4.11) then becomes the ratio of existing house value plus demolition

[20]In Equation (4.11) the right-hand side of the inequality is defined only for values of F^* where a positive profit from development exists; that is, $(\alpha - \mu) - (\beta - \tau)F^* > 0$. When density is so great that construction costs exceed floor prices, the return from development is negative and redevelopment obviously is unprofitable.

costs (per floor area) to newly constructed house profit (per floor area). The reader should be able to verify that for the inequality to hold when $\alpha^0 = \alpha$, F^* must be greater than F^0.[21] Thus, one way redevelopment can occur is when market conditions dictate an optimal FAR that is higher than that which already exists on the site.

How much greater does market FAR have to be for redevelopment to occur? Let's return to our example from Boston's Back Bay area. There, the average FAR among existing buildings was 7.6. In fact, the district is largely composed of brownstones and townhouse structures that are generally four stories tall and newer apartment buildings that are 8 to 12 stories tall. Consider two sites in the district: one with a four-story brownstone and the other with a 10-story apartment building. Using our previous estimates for α, β, μ, and τ and assuming that demolition costs (δ) are $20 per square foot (of floor area) redevelopment becomes profitable for the site with a current FAR of 4 when the new proposed FAR for new development exceeds 12.[22] Since the optimal FAR from our earlier calculations (F^*), is 17.5, the redevelopment of existing townhouses would clearly be profitable. For a site with an existing 10-story apartment building, redevelopment is not profitable. The opportunity cost of existing structures at this higher density cannot be overcome with a current optimal FAR of 17.5. These results suggest that without the restrictions that are in place to maintain the low-rise historic buildings in the Back Bay, the district would seem to be ready for a major land-use transition, eliminating many of the older brownstones and replacing them with higher-rise apartment buildings.

Redevelopment may also be profitable at relatively low FARs if the housing capital that exists on a site has become seriously deteriorated or outmoded. Using Equation (4.11), we can see that for redevelopment to occur at a similar density to that which already exists ($F^* = F^0$), the criteria requires that α must exceed α^0. Again, using the Boston example, we determine that redevelopment at the current FAR of 4 would technically be profitable if the existing structure capital (per square foot) were to deteriorate to about 42 percent of what new structure capital is worth.

These examples illustrate that vertical redevelopment can be economically feasible under a range of market changes that might occur over relatively short intervals of time (decades). This is particularly likely in faster-growing economies in which location rents quickly build up and housing capital easily becomes outmoded. The fact that redevelopment is not always that common might be explained by two widespread institutional constraints. First, in order to acquire a substantial site, land often must be assembled from varied owners, a time-consuming and difficult process. Second, in the twentieth century, zoning and other land-use regulations have become widespread, and these often prohibit or slow down the process of land use change.

[21] If $F^* = F^0$, then the right-hand side of Equation (4.11) is greater than one, while the left-hand side equals one. Why would anyone tear down a house only to rebuild the identical structure?

[22] With $\alpha = 222.02$, $\beta = 1.48$, $\mu = 100$, $\tau = 2$, and $\delta = 20$, a new FAR of 12 just meets the criteria displayed in Equation (4.11).

REDEVELOPMENT, OCCUPANCY, AND LAND-USE SUCCESSION

As interior land is redeveloped to new density levels, the occupants of that land will often change. Such combined changes in occupancy and density of use frequently are referred to as *land-use succession.* These changes sometimes can involve the large-scale relocation of different groups of households. Understanding such patterns can be important in determining when and where redevelopment will occur and for whom the new uses should be built. To illustrate how the location of different households can change as land is redeveloped, we can return to a city with two types of households such as we used in Chapter 3.

In a city with two groups of households (labeled 1 and 2), the willingness-to-pay for housing (per square foot of floor area) is described in Equations (4.12). We integrate Chapters 3 and 4 now by making the demand for housing depend on both distance from the city center (d), as well as FAR (F). As in Chapter 3, the parameter k represents the marginal reduction in house prices (per square foot of floor area) with greater commuting, and we continue to use the parameter β to represent the marginal reduction in willingness-to-pay as a result of greater density.

$$\begin{aligned} P_1 &= \alpha - k_1 d - \beta_1 F, \\ P_2 &= \alpha - k_2 d - \beta_2 F, \\ k_1 &> k_2,\ \beta_1 > \beta_2 \end{aligned} \tag{4.12}$$

In Equations (4.12), we make the assumption that members of household Group 1 find commuting more burdensome than households of Group 2, and that they also value open space more, as evidenced by their greater distaste for higher FAR levels. This pattern of preferences is often thought to hold across households of different income levels: those earning more have both a higher value for time (hence, distaste for commuting) as well as a stronger demand for lower-density single-family housing (hence, distaste for FAR).

As the city develops initially in a horizontal manner, we could assume that the historic housing stock has evolved relatively homogeneously, with a relatively flat FAR gradient. As our earlier discussion suggested, this pattern is common in newer, faster-growing cities. With this historic stock of housing (and its constant FAR), the question of which household type locates where (closer or farther from the city center) will depend exclusively on the different preferences for commuting, as it did in Chapter 3. With the assumption in Equations (4.12) that Group 1 households have greater distaste for commuting, they will outbid Group 2 households for central housing, and vice versa for peripheral units. The pattern of house prices and occupancy will resemble the P_1, P_2 schedules shown in Figure 4.11.

Over time, selected units within the existing stock may deteriorate and vacant land will become available. The question is not just at what FAR will this land be developed, but for whom. Will the highest use for vacant land continue to involve development for households of Group 1 at central sites and households of Group 2 at farther sites? The answer will depend on which group is willing to offer the most per acre for land at central as opposed to peripheral sites.

To determine the willingness-to-pay for land by each type of household, we must determine which FAR represents the most profitable development at a given location. Using the construction cost function from the previous section, and following the procedures

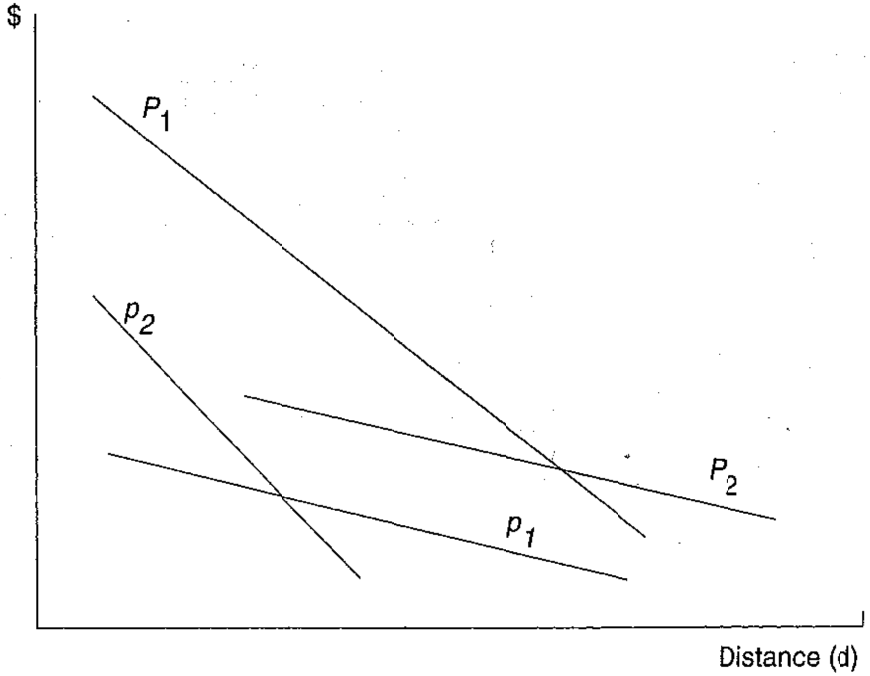

FIGURE 4.11 House and land price bids for two household types.

described with Equations (4.9), we solve for the optimal FAR levels for each household as a function of location (distance *d*). This, in turn, allows us to determine what is the maximum residual value for land, if developed for each type of household, and how this varies with distance from the urban center.

$$F_1^* = \frac{\alpha - k_1 d - \mu}{2[\beta_1 + \tau]}, \quad F_2^* = \frac{\alpha - k_2 d - \mu}{2[\beta_2 + \tau]}$$
$$p_1^* = \frac{[\alpha - \mu - k_1 d]F_1^*}{2}, \quad p_2^* = \frac{[\alpha - \mu - k_2 d]F_2^*}{2} \qquad (4.13)$$

At a common site (or location *d*), the optimal FAR of development for Group 1 households will always be less than that for Group 2 households, since $k_1 > k_2$ and $\beta_1 > \beta_2$. As discussed in Chapter 3, the issue of which group will outbid the other for land at more central sites hinges on the relative *slopes* of the residual land price functions. In Equations (4.13), the slope of p_1^* with respect to distance (d) is $-k_1F_1^*/2$, whereas that of p_2^* is $-k_2F_2^*/2$. Given our assumptions about household preferences, it is easily possible for Group 2 households to have a more steeply sloped residual land price gradient than Group 1 households. This will be the case if the FAR of development for Group 2 households is significantly greater than for Group 1 households, whereas the distaste for commuting by Group 2 is only slightly less. This case is shown in Figure 4.11, with the p_1,p_2 schedules.

With the pattern of preferences shown in Figure 4.11, when vacant land becomes available at more central sites within the city, it not only will be developed at a higher FAR than surrounding historic housing, but it will be developed for and occupied by a different type of household than those occupying existing housing in the area currently. In principle, after a long enough period of time, the pattern of land use might eventually change from one in which Group 1 households occupy central sites to one in which Group 2 households live centrally. This is the process of land use succession—one use gives way to another, with a combined alteration of density or FAR as well as occupancy.

This example illustrates the complexity of determining the true highest and best use for urban land. In principle, the maximum value for any given site can be determined only by considering the optimal FAR that should be developed for each potential group of occupants or type of land use and then comparing these potential residual values across types or uses. If the land market is functioning well, it should be the case that at current housing density levels, the present occupants of an area do indeed represent the highest-paying users. At different potential densities, however, it may well be the case that other groups would generate greater residual value to the land.

SUMMARY

In this chapter we examined how housing development, particularly density, varies across locations within cities, and evolves over time as cities experience population growth.

- The price of a housing unit can be decomposed into implicit prices for each of the attributes that make up the unit, such as the presence of a garage, the amount of square feet, and the total density. There are distinct patterns to residential density in cities throughout the world.
- At higher density, otherwise identical housing units tend to have lower prices. The cost of constructing units, however, tends to increase with the density of development. Thus, the profit per housing unit tends to be lower at greater density, although more units can be placed on a given parcel of land. Determining the most profitable density of development requires trading off these two opposing considerations.
- At more desirable locations, this tradeoff leads to a highest and best use that involves greater density. In effect, higher prices for housing units lead to greater land values, and thus land is used more sparingly per housing unit. Most cities have density gradients, with denser residential uses located near transportation centers, highways, parks, waterfronts, or other amenities.
- As cities grow in population, new development initially tends to occur horizontally at the expanding border of the city. Thus, at any time, existing units reflect the history of development, with the oldest units at the center and the newest at the fringe.
- At critical times in the development of a city, it can become profitable to demolish existing housing units, built years earlier, particularly at more desirable locations. At such sites, the net return to just the land from new development can

exceed the total price for existing housing. Normally, this occurs only when new development calls for much greater density than exists with the current housing.

REFERENCES AND ADDITIONAL READINGS

ANAS, ALEX, "The Dynamics of Residential Growth," *Journal of Urban Economics*, 5 (1978).

BERNDT, ERNST R., *The Practice of Econometrics: Classic and Contemporary.* Reading, Mass.: Addison-Wesley Publishing Company, 1991.

CAPOZZA, DENNIS, AND GORDON SICK, "Valuing Long Term Leases and the Option to Redevelop," *Journal of Real Estate Finance and Economics*, 4 (1991).

COURT, ANDREW T., "Hedonic Price Indexes with Automotive Examples," in *The Dynamics of Automobile Demand.* New York: The General Motors Corporation, 1939.

DIPASQUALE, DENISE, AND C. TSURIEL SOMERVILLE, "Do House Price Indexes Based on Transacting Units Represent the Entire Stock? Evidence from the American Housing Survey," *Journal of Housing Economics* 4, 3 (1995).

HALVORSEN, R., AND HENRY POLLAKOWSKI, "Choice of Functional Form for Hedonic Price Equations," *Journal of Urban Economics*, 10 (1981).

HARRISON, DAVID, AND JOHN KAIN, "Cumulative Urban Growth and Urban Density Functions," *Journal of Urban Economics*, 1 (1974).

MILLS, EDWIN S., *Studies in the Structure of the Urban Economy.* Baltimore, Md.: The Johns Hopkins University Press, 1972.

PINDYCK, ROBERT S., AND DANIEL L. RUBINFELD, *Econometric Models and Economic Forecast*, 3rd ed. New York: McGraw-Hill Book Company, 199[illegible].

ROSEN, SHERWIN M., "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition," *Journal of Political Economy*, 82, 1 (1974).

WHEATON, WILLIAM, "Urban Spatial Development with Durable but Replaceable Capital," *Journal of Urban Economics*, 3 (1978).

CHAPTER 5

FIRM SITE SELECTION, EMPLOYMENT DECENTRALIZATION, AND MULTICENTERED CITIES

In Chapters 3 and 4, we examined the location decisions of households and the operation of the residential land market within a metropolitan area. We now turn our attention to nonresidential land uses. How do firms choose their locations within a metropolitan area? Why do we tend to see spatial concentrations of commercial, industrial, or retail firms? How do we explain the decentralization of employment within metropolitan areas during the last half of the twentieth century? Our focus in this chapter is the operation of a land market in which space is used by commercial and industrial firms as well as by households.

Chapters 3 and 4 explained how residential land markets operate around a single employment center when workers must commute to that center. In Chapter 5, we extend that model to show why, historically, firms tended to concentrate in a single, central location. The traditional explanations for this phenomenon focus on the need of firms to transport both raw materials to their plants and goods to a central market through a transportation terminal such as a port. While this simple model does a fairly good job of describing firm location in nineteenth-century cities, employment today uses different production as well as transportation technologies.

Why has employment decentralized? Retail trade and personal service firms have simply followed their household clients as they have suburbanized over the last century. These patterns will be discussed in Chapter 6. Here, we focus on firms that produce and sell products on a broader national scale, and, therefore, are not tied to local residential clients as a market for their goods or services. In brief, production technologies changed in such a way that today manufacturing and wholesale firms must use more land per unit of output than they did in the previous era of employment centralization. This trend, in turn, has pushed these firms to fringe locations where land is cheaper. In addition, the rise in importance of truck transportation in getting goods to and from markets has made access to the interstate highway system important. For office-using commercial firms, the recent changes in telecommunication technology (e.g., e-mail and fax machines) have reduced the importance of physical proximity to both clients and other firms. Finally, for all types of firms, perhaps the most critical change has been the suburbanization of the major factor of production: labor.

While employment has been decentralizing, the spatial distribution of jobs in most cities is anything but uniform. Rather, metropolitan areas have become multicentered. Some of the forces that led to the original development of central cities are still at work and have led suburban firms to concentrate or cluster at peripheral locations. Such concentration results from the advantages of having similar and complementary firms in close proximity as well as from the importance of access to the highway system for getting products to market and employees to work. Thus, the modern multicentered city is the outcome of competing forces—the benefits of decentralization versus the benefits from concentration.

We begin this chapter with an examination of the spatial distribution of employment and how it has changed over time. Next, we discuss the traditional theories used to explain the existence of the central business district (CBD). Following that, we examine more modern explanations for why the forces that originally led to centralized employment have dissolved and have been replaced by forces encouraging employment dispersal. With this more current economic analysis, we find some close links between the urban land market and the labor market.

THE SPATIAL DISTRIBUTION OF EMPLOYMENT

As was the case with population density in Chapter 4, casual observation in most cities suggests that *employment density*—defined as employment divided by land area or jobs per square mile—declines with distance from the central city. Figure 5.1 displays employment density for 146 cities and towns in the Boston area in 1990. Together, these cities and towns had a total employment in 1990 of 2.07 million.[1] The average employment density across these towns is 1,025 jobs per square mile but varies widely from 16,062 in Cambridge and 11,104 in Boston to 29 jobs per square mile in Boxford at the far northern part of the region. This wide variation, as well as the general decline in employment density with distance, can be clearly seen in the map.

[1] In each state, a department is responsible for surveying businesses and tabulating data on jobs and wages as part of the U.S. unemployment insurance system. In Massachusetts, this information is collected by the state's Department of Employment and Training. The data are collected and organized by type of industry, as defined by the Standard Industrial Classification (SIC) system. For further explanation of SICs see footnote 6 in Chapter 7.

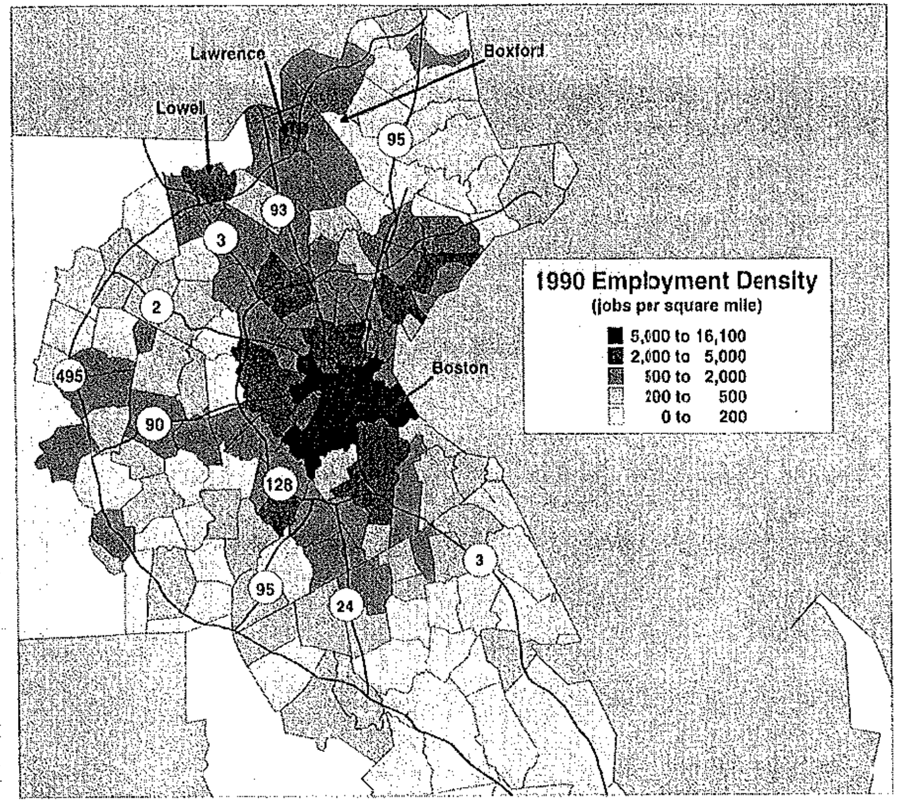

FIGURE 5.1 Employment density for Boston-area cities and towns, 1990.

Source: Massachusetts Department of Employment and Training, *Employment and Wages in Massachusetts Cities and Towns* (selected years), Boston, MA.

As expected, the anomalies in the employment densities closely match the anomalies in the population densities that were discussed in Chapter 4. The city of Cambridge has a higher employment density than Boston. Lowell and Lawrence to the far north also have relatively higher employment densities (2,911 and 3,360, respectively), given their more distant locations.

Empirically, we can measure the relationship between employment density and distance the same way we measured the relationship between population density and distance in Chapter 4. Using the negative exponential specification defined in Equations (4.1) and (4.2), we can statistically estimate a density gradient using the log of employment density as the dependent variable and distance as the independent variable:

$$\log [D(d)] = \underset{(8.1)}{8.05} - \underset{(-11.15)}{0.10}\, d \qquad R^2 = 0.46 \quad N = 146 \tag{5.1}$$

In Equation (5.1), $D(d)$ is the density of employment at distance d. Our estimate of the coefficient on distance of –0.10 means that with each mile increase in distance from the center, employment density decreases 10 percent.[2] The constant in Equation (5.1) is the model's estimate of the log of the density at the center of the metropolitan area defined here as the city of Boston; the model predicts that Boston's density should be 3,134 jobs per square mile ($e^{8.05} = 3,134$). This estimate is far less than the city's actual density of 11,104. Figure 5.2 uses a scatter plot to compare actual town employment densities; the line represents Equation (5.1). Clearly, the model does far better in predicting employment density past mile 5 than within 5 miles of Boston. Even with these problems, this very simple model still explains 46 percent of the variation in employment densities.

For most of the towns in our sample, commercial or industrial uses take up only a small fraction of total land area. Thus, our measure of employment density may be

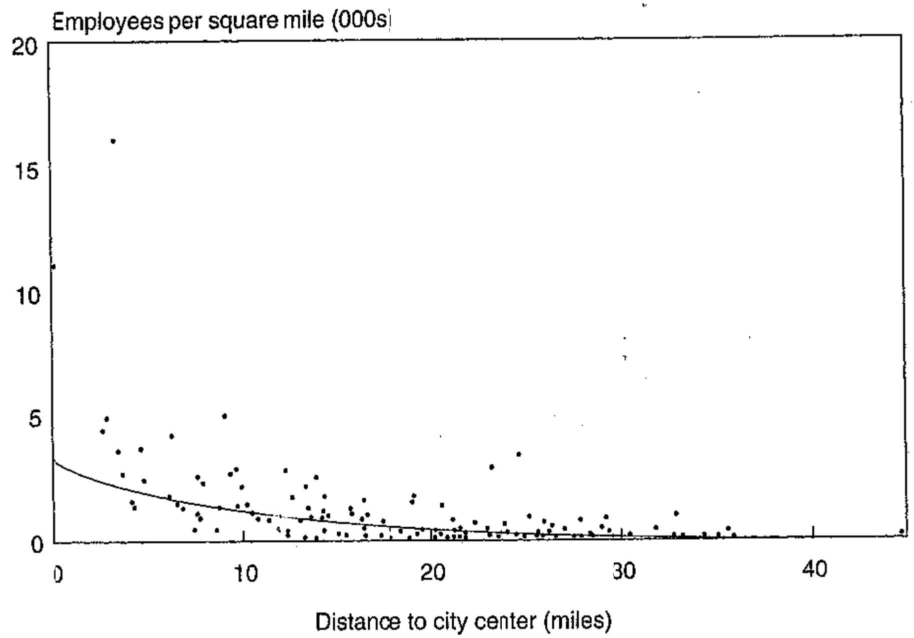

FIGURE 5.2 Boston employment density, 1990.

Source: Massachusetts Department of Employment and Training, *Employment and Wages in Massachusetts Cities and Towns* (selected years), Boston, MA.

[2]The employment density gradient in Boston has flattened over time. Estimating Equation (5.1) using 1967 employment densities yields the following equation, in which employment density decreases 12.6 percent for each additional mile from Boston:

$$\log [D(d)] = 7.75 - 0.126d \qquad R^2 = 0.47$$
$$(6.57) \quad (-11.35) \qquad N = 146$$

somewhat misleading. An alternative approach is to consider the town's share of total regional employment. For 1990, each town's share of the 2.07 million jobs in this region are displayed in Figure 5.3. While the general patterns are the same in Figures 5.1 and 5.3, there are some important differences. The city of Boston has far and away the largest share of the area's jobs, representing 25.9 percent of the total. Cambridge ranks second, with just 5.0 percent of total jobs, followed by Waltham and Newton, which have 3.0 percent and 2.2 percent, respectively. Lowell and Lawrence to the north have 1.9 percent and 1.1 percent of the area's jobs, respectively.

Over the past decades, employment has suburbanized in Boston. Figure 5.4 illustrates this by displaying the change in each town's share of the region's total employment between 1967 and 1990. The Boston share of total employment declined from 34 percent in 1967 to 25.9 percent in 1990, a decline of 8 percentage points. Declines in employment share were also realized by many of the cities and towns closest to the city of Boston as

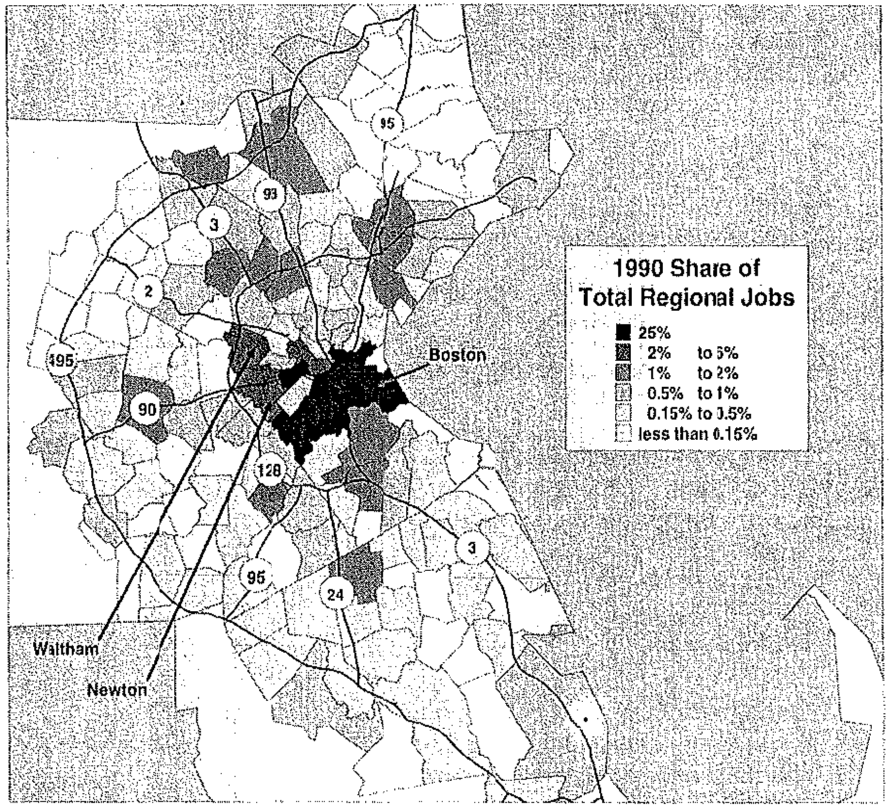

FIGURE 5.3 Share of regional jobs for Boston-area cities and towns, 1990.

Source: Massachusetts Department of Employment and Training, *Employment and Wages in Massachusetts Cities and Towns* (selected years), Boston, MA.

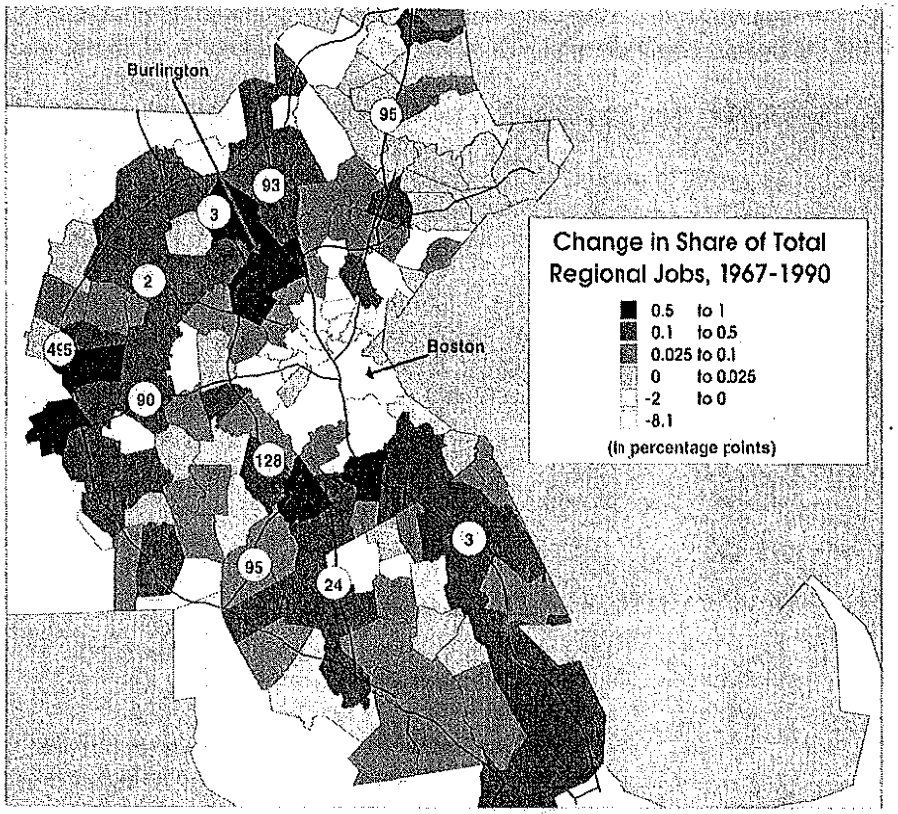

FIGURE 5.4 Change in share of total regional jobs for Boston-area cities and towns, 1967–1990.

Source: Massachusetts Department of Employment and Training, *Employment and Wages in Massachusetts Cities and Towns* (selected years), Boston, MA.

well as in older employment centers such as Lawrence, where the share fell from 2.1 percent to 1.1 percent. Growth in employment share over the period occurred mainly in suburban communities such as Burlington, a community in which the number of jobs more than tripled since 1967; its share of the area's employment went from 0.7 percent in 1967 to 1.6 percent in 1990. It should be noted that many of the towns that experienced growth in employment share are located at the intersections of major highways.

The patterns shown for the Boston region are quite typical of those in most older, mature, metropolitan areas. In newer metropolitan areas (e.g., Dallas, Los Angeles, Atlanta), the size and role of the core central city is much less than in Boston, and the extent of employment decentralization is far greater. We begin the task of explaining job decentralization by first considering the forces that originally created urban centers, focusing on the location decisions of firms that primarily export their products to clients or buyers outside of the region.

LAND MARKETS WITH A CENTRAL BUSINESS DISTRICT

In Chapter 3, we considered the location decisions of households in a city in which all workers commuted to a single employment zone located at the center of the city. The discussion in Chapter 3 suggests that for such a center to exist, firms must value central land more than households do. In a competitive, unregulated land market, a central commercial or manufacturing district can exist only if the land rent from these nonresidential land uses exceeds the land rent derived from housing for the surrounding workers. Let's begin our analysis of firm location by considering how the transportation system in eighteenth- and nineteenth-century cities and the shipping needs of firms might have been responsible for the creation of central business districts (CBDs). Our stylized city and the firms located within that city have the following characteristics:

1. The city has a single port or transportation terminal to which firms must bring their goods for export shipment and through which they receive their raw material or inputs from other cities. Similarly, imported consumer goods arrive through this facility. Within the city, the transport of goods (from the terminal) costs s dollars per unit per mile. Distance from this transportation center is denoted as d.
2. Firms produce an identical product using the same production process. The units of output produced per firm is fixed at Q.
3. There is no factor substitution. Both the lot size f and structure capital used by each firm are fixed. The rent for the structure used by firms is C, whereas the firm's residual land rent per acre varies with location, $r_c(d)$. With fixed land and structure, output per acre is fixed.
4. Output and input markets are perfectly competitive, with free entry into the industry. This means that each firm takes prices as given and economic profit is zero.
5. Land is allocated or rented to that use and to those plants or offices that yield the greatest rent.

The assumptions about our stylized city are, in many respects, analogous to the assumptions that we made at the beginning of Chapter 3. Using these assumptions, consider how firm profits (revenues minus all costs) would vary with location. If each firm sells Q units of its good and the unit price is ρ, then firm total revenues will be ρQ. Variable costs include wages and material production costs per unit, A, and transport or shipping costs to market per unit, sd. Firm fixed costs include the rent for the building, C, and land rent per acre, $r_c(d)$, times the number of acres used by the firm, f. Hence, profits, π, are:

$$\pi = Q(\rho - A - sd) - C - r_c(d)f \tag{5.2}$$

With competition between firms ensuring zero profits, land rent per acre, $r_c(d)$, may be determined as a residual:

$$r_c(d) = \frac{Q(\rho - A - sd) - C}{f} \tag{5.3}$$

In Equation (5.3), land rent is defined as that rent that gives firms the same zero level of profit regardless of location. Under the assumption that ρ, Q, A, and C, do not vary with location within a metropolitan area, land rents will exactly compensate firms for the increased transport costs that are associated with a farther distance to the transportation terminal. In this respect, the model is similar to that of Chapter 3, in which residential land rents exactly compensate households for the commuting costs associated with more distant locations. In this model, the transport costs at the city center ($d = 0$) are 0, and $r_c(0)$ equals $[Q(\rho - A) - C]/f$. Moving out from the center, land rents per acre decline by exactly the increase in transport costs per acre to firms: $-sQ/f$. If rents decline by more (less) than this, then firms could realize excess profits by moving out from (in towards) the urban center. In equilibrium, incentives to move must not exist; this equilibrium is analogous to the notion of *spatial equilibrium* that we defined for residential use in Chapter 3.

Returning to the main question of this section, we can see how a central business district comes into existence: the slope of the firm rent gradient must be steeper than the rent gradient of households. Recalling the results of Chapter 3, this means that as we move out from the transportation terminal, the shipping costs of firms (per acre of commercial use) must increase by more than the worker's commuting costs to the firm (per acre of residential use). This is likely to be the case if goods or materials are more expensive to move around than people (people frequently walked in nineteenth-century cities) and if commercial uses are more dense than residential. Under these conditions, firms constitute the dominant land use from the center to some intermediate boundary, m, and from m to the urban border, b, housing makes up the dominant use. At the intermediate boundary, the land rent from firms will equal residential land rent (labeled $r(m)$, as in Chapter 3). This pattern of rents is shown in Figure 5.5.

At the urban border, residential land rent must equal that from agriculture [$r(b) = r^a$]. Moving inward, land rent rises as commuting costs (per acre of residential development) fall until reaching $r(m)$ at the boundary between residential and business use (m). With the condition of equal rent at the boundary between business and residential land use, business land rent has two components: the opportunity rent from residential uses (who would occupy the land if firms did not) and the location rent or savings in transport costs associated with moving in from this boundary, $sQ(m - d)/f$. As a result, the land rent gradient for firms is:

$$r_c(d) = r(m) + \frac{sQ(m-d)}{f} \tag{5.4}$$

As with the discussion of residential land use in Chapter 3, we can relax the assumption that all firms have identical production processes and use identical amounts of land for their facilities. With different production processes and land requirements, there will be a systematic location pattern by type of firm. Suppose we identify types of industrial or commercial uses with the subscript i, and allow all firms within a type to have a specific output level Q_i, land usage f_i, and shipping or transport cost per mile to the terminal s_i. Firms that produce a great deal of output per acre—output which is also difficult or expensive to move around—will have a high ratio of s_iQ_i/f_i. Firms whose goods are easy to ship and who produce little output per acre will have a low s_iQ_i/f_i ratio. Since the slope of a firm's zero-profit rent gradient is directly equal to this ratio, in a locational equilibrium within the central business district, firms will naturally tend to separate themselves spatially according to their ratio value.

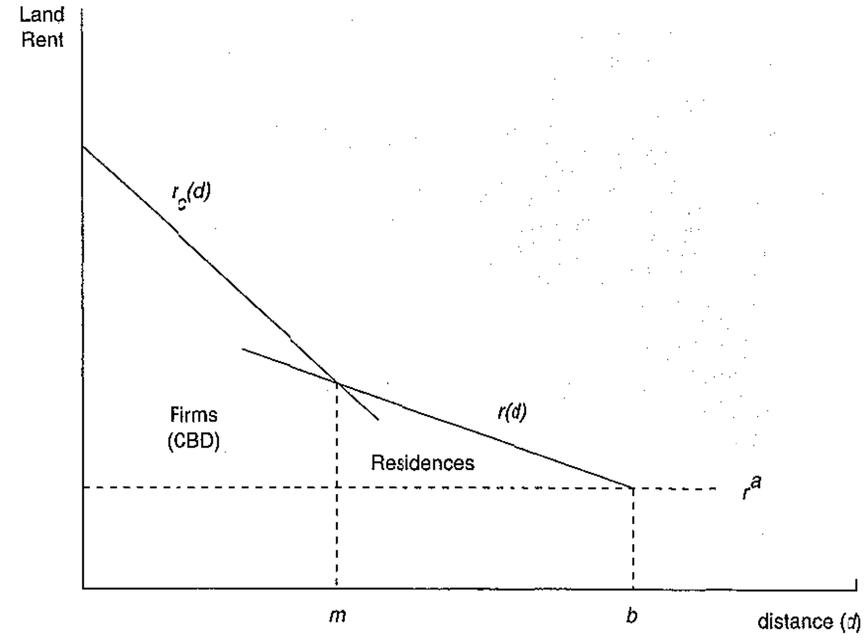

FIGURE 5.5 Spatial separation and land rents: firms and residences.

Within the CBD, those types of firms that produce high output per acre and whose products are expensive to ship will have a steeper rent gradient. In a competitive land market, they will offer the highest rent for the most central sites. Firms with lower output per acre and/or firms whose goods are easier to ship will have a rent gradient that is somewhat flatter. They will tend to occupy sites farther from the transportation terminal and nearer to residential uses. The land market will allocate different commercial or industrial uses to sites within the CBD just as it allocated different types of households in Chapter 3. Land is always allocated to that use offering the highest rent and, in this case, central rents will be highest for those uses with the most output per acre.

The separation of different industrial or commercial uses within the CBD also leads to the emergence of a business density gradient. Since the level of output or production per acre is closely (and inversely) related to firm land consumption, f, those firms locating most centrally will tend to be those whose facilities are most dense. This pattern holds today in many cities in which the tallest office buildings tend to be developed on the choicest sites, adjacent to subways or other transportation terminals. At the edges of the business district, density is much lower.

TECHNOLOGY AND THE DECENTRALIZATION OF MANUFACTURERS

In the nineteenth-century city, offices, warehouses, manufacturing facilities and stores were all located in the central business district. Over time, the first firms to decentralize were industrial firms—those engaged in manufacturing or warehousing. The traditional explanation for this movement is that industrial technology changed to make the zero profit rent gradient of industrial uses much flatter across space. Before discussing this argument in detail, let's examine the extent of industrial decentralization within the Boston metropolitan region.

In the 146 cities and towns in our Boston-area sample, firms in the manufacturing and wholesale SIC categories provide 466,017, or 22.5 percent, of the area's jobs. While

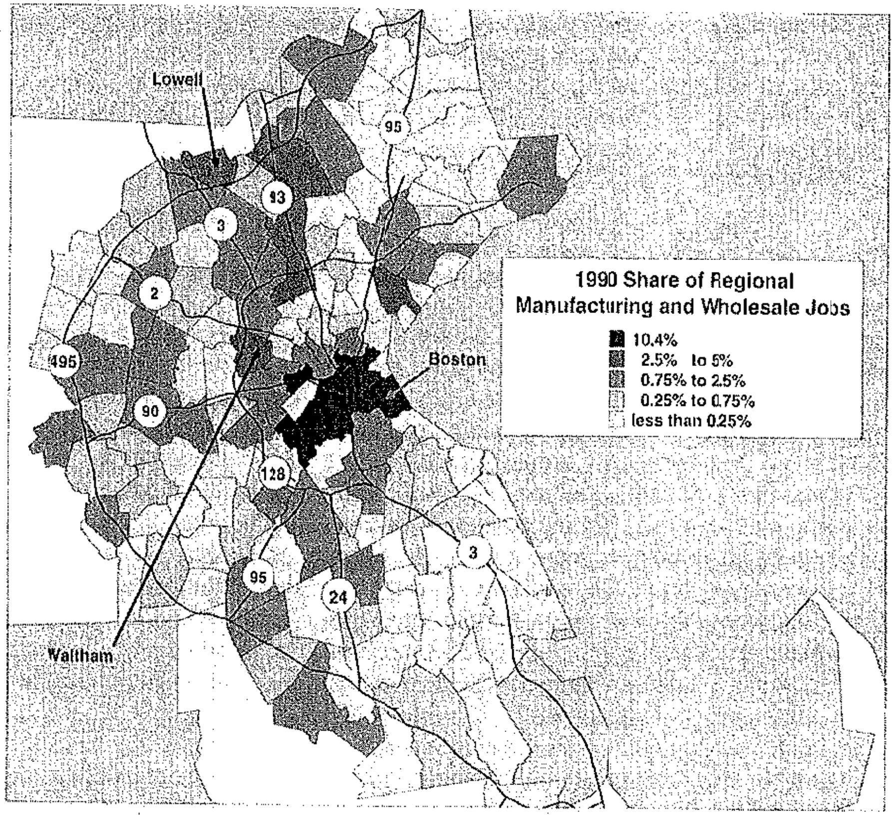

FIGURE 5.6 Share of regional manufacturing and wholesale jobs for Boston-area cities and towns, 1990.

Source: Massachusetts Department of Employment and Training, *Employment and Wages in Massachusetts Cities and Towns* (selected years), Boston, MA.

the city of Boston has the largest share of manufacturing and wholesale jobs (10.5 percent), these jobs are widely dispersed, with other cities or towns also having high shares (e.g., Waltham and Lowell have 4.7 percent and 3.5 percent, respectively). Figure 5.6 displays the share of metropolitan manufacturing and wholesale jobs by city and town. Outside of the city of Boston, those towns with a larger share of industrial jobs tend to be located near major highways.

The number of manufacturing and wholesale jobs in the Boston area has declined 8 percent since 1967. The spatial distribution of those jobs, however, changed more dramatically between 1967 and 1990. In Figure 5.7, we display the change in the share of regional manufacturing and wholesale jobs by city and town. The city of Boston's share dropped 12.3 percentage points over this period, from 22.8 percent in 1967 to 10.5 percent in 1990. It is clear from the map that Boston and nearby towns lost in their share of jobs, while more distant towns, particularly near major highways, gained considerably in their shares.

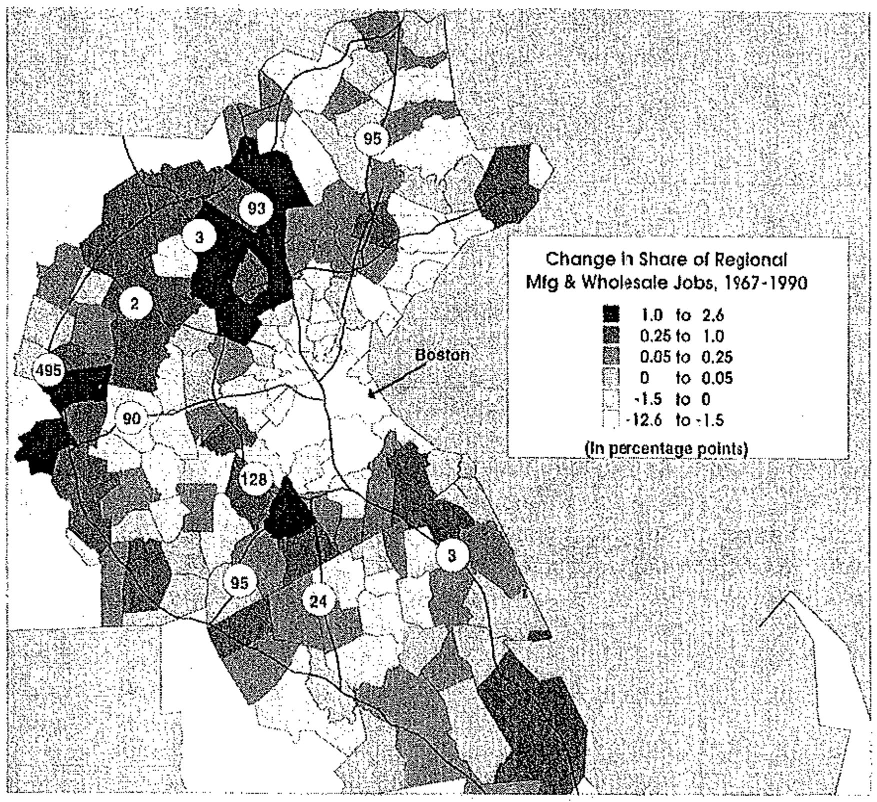

FIGURE 5.7 Change in share of regional manufacturing and wholesale jobs for Boston-area cities and towns, 1967–1990.

Source: Massachusetts Department of Employment and Training, *Employment and Wages in Massachusetts Cities and Towns* (selected years), Boston, MA.

Explanations for why manufacturing and wholesale jobs decentralized focus on two technological developments of the twentieth century. The first is the evolution of the transportation system. With the advent of railway lines, the central transportation terminal typical of the nineteenth century (e.g., a port) gradually gave way to a much more dispersed pattern of freight terminals. This pattern, in turn, evolved into an even more dispersed pattern of highways as manufacturers relied increasingly on truck transportation. The widespread adoption of rail and truck transportation meant that firms no longer needed to move their products to or receive their materials through the center of the city. The very rationale for a centrally oriented rent gradient by manufacturers virtually disappeared during the twentieth century.

In addition to changes in transportation technology, the methods of industrial production and storage technology were also evolving during this same period. Changes in both production and storage methods greatly increased the amount of land used per unit of output by industrial firms. During the later stages of the industrial revolution, manufacturers increasingly adopted production processes that were based on integrated horizontal assembly lines. It is argued that horizontal assembly increased the amount of land needed per unit of output. Modern inventory technology also has a high land requirement, because it requires manufacturers to store goods in large, single story horizontal structures.

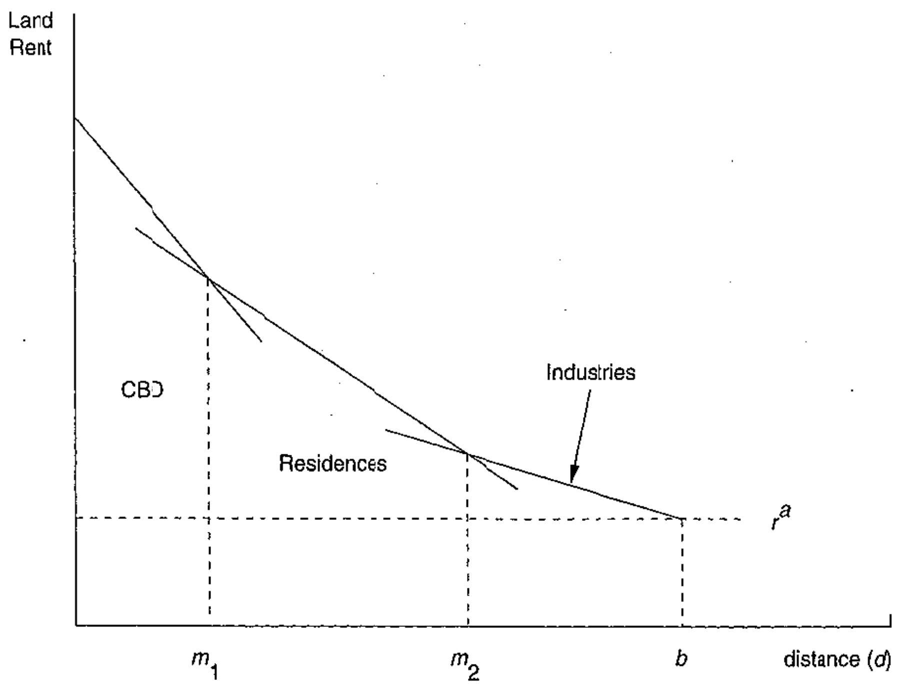

FIGURE 5.8 Spatial separation: land rents for commercial, residential, and industrial uses.

As a result of these changes, the zero-profit rent gradient of industrial firms became quite flat with respect to distance from the traditional urban center. Industrial firms thus were willing to pay less per unit of land for central locations than were most other uses. In the competitive land market, it became most profitable for them to choose more distant locations. Today, new industrial structures are almost always constructed at the very edge of cities in which land competition from other uses is minimal. This pattern of location and the land rents that result are shown in Figure 5.8. Commercial and office uses occupy a CBD (up to a distance m_1), followed by residences from distance m_1 to m_2. Beyond residences, land is used by low-density industries and warehouses.

The extensive land required for industrial uses has led some to argue that industrial land rent gradients are, in fact, virtually flat. As a result, industrial location may be determined more by where other land uses decide not to locate rather than by the explicit attractiveness of certain locations for industrial firms. For example, sites located right on major highways or near airports may be undesirable for many other uses because of noise or other aesthetic considerations. With little competition for these sites, industrial firms can find large tracts of cheap land.

WAGES, THE LABOR MARKET, AND OFFICE DECENTRALIZATION

While industrial land uses in the U.S. have been decentralizing from center cities for many decades, financial and service firms that use office space have remained much more centralized, at least until recently. Clearly, many American cities still contain large central business districts with numerous high-rise office buildings. However, since 1980, there has been an explosion of office construction outside of CBDs. Within the Boston metropolitan area, we can see this change in Figures 5.9 and 5.10.

In Figure 5.9, the 1990 share of regional jobs in the FIRE (financial, insurance, and real estate) and service SIC categories is still quite concentrated within the center city of Boston. In fact, Boston contains 35 percent of these jobs while having only 18 percent of the region's population. However, the map also shows that by 1990, several suburban communities adjacent to the inner circumferential highway (Route 128) had high concentrations of office employment (Newton 4 percent, Waltham 5 percent, Quincy 4 percent). If we look at Figure 5.10, in which the changes in regional service jobs shares between 1967 and 1990 are displayed, the decentralization of services becomes clearer. Boston's share dropped 12 percentage points (from 47 percent to 35 percent), while shares of numerous suburban communities rose. There was still absolute growth in office jobs within the CBD over this period, but it was far less than the growth of such jobs in the suburbs.

Within the framework that we have used to study the decentralization of manufacturing jobs, the movement to the suburbs of office jobs may seem somewhat puzzling. Don't offices tend to have high FAR (floor/area ratio) levels, which means that they should outbid other uses for desirable locations? While this is true, the key to understanding the decentralization of offices lies in defining what types of locations such firms find desirable. In fact, since office firms do not ship products or receive inputs, there was little

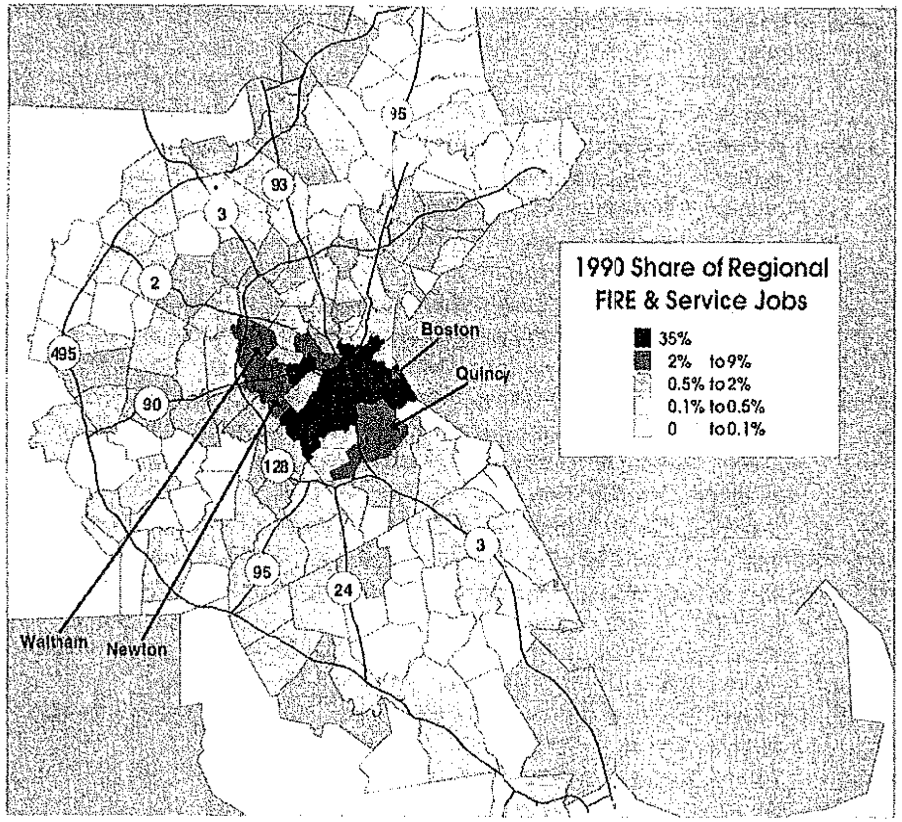

FIGURE 5.9 Share of regional FIRE and service jobs for Boston-area cities and towns, 1990.

Source: Massachusetts Department of Employment and Training, *Employment and Wages in Massachusetts Cities and Towns* (selected years), Boston, MA.

reason, historically, for them to use the central locations surrounding regional port or transportation facilities. Rather, for office firms, labor is by far the dominant and almost exclusive factor used in production. The locational incentives for office firms revolve around the costs associated with assembling their workforce.

When the location of firms is fixed at one urban center, we have seen how workers are compensated for commuting to that center through variation in land rents (or prices). But what happens if a firm currently located at that single center contemplates moving to some alternative location nearer to the residences of their employees? This question was first posed by L. N. Moses (1962), who argued that such a firm could hire those workers that lived next to it for a lower wage rate. Carrying the argument further, Moses hypothesized that there should be a location-specific wage rate at each distance from the center which a decentralizing firm could pay local workers. Since the only alternative

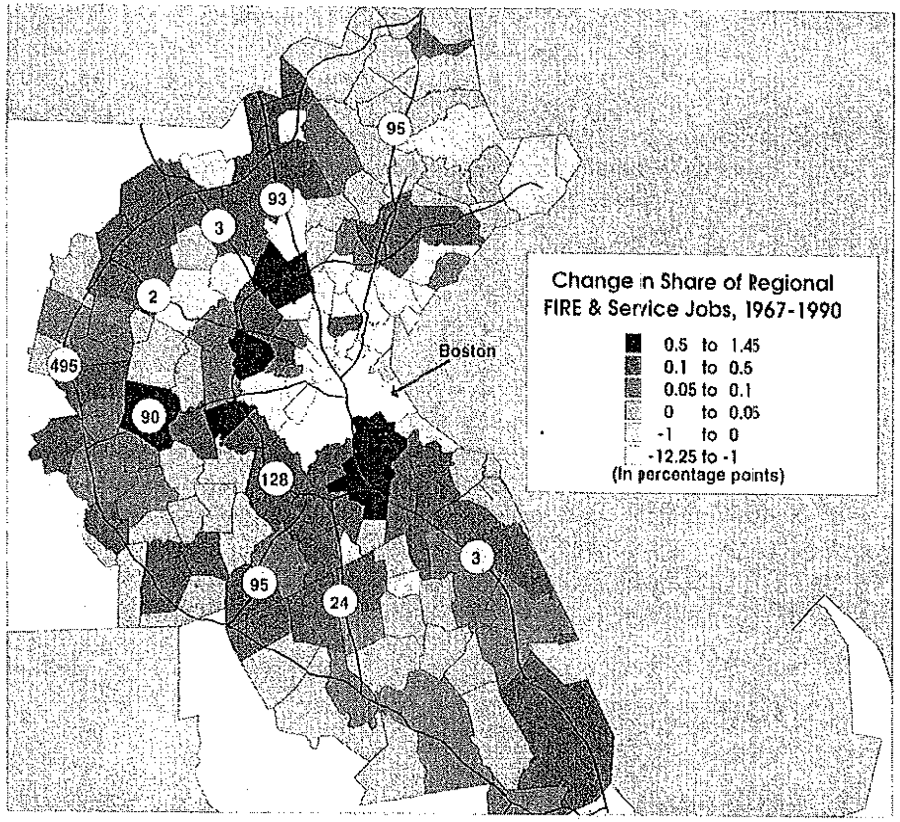

FIGURE 5.10 Change in share of regional FIRE and service jobs for Boston-area cities and towns, 1967–1990.

Source: Massachusetts Department of Employment and Training, *Employment and Wages in Massachusetts Cities and Towns* (selected years), Boston, MA.

source of employment is at the center, this decentralized firm could pay a wage that is lower than the wage paid at the center, reflecting the decrease in commuting costs to workers. In effect, there should be an urban wage gradient in addition to a land rent gradient. As a result, the urban labor market is closely intertwined with the market for land. Let's look at this argument in more detail by examining Figure 5.11.

In contrast to Figure 5.5, Figure 5.11 shows the CBD with a horizontal land rent from firm or business uses, r_c. We show this under the assumption that within the CBD, commuting costs are zero, so that workers must only pay to commute to the edges of the CBD (at distances d_1 or d_6). Within this city, a firm could decentralize its facility or office to some location d_i, using the land from d_2 to d_3. Remembering that in this model, all firms and workers are homogeneous, the decentralizing firm will be able to employ those

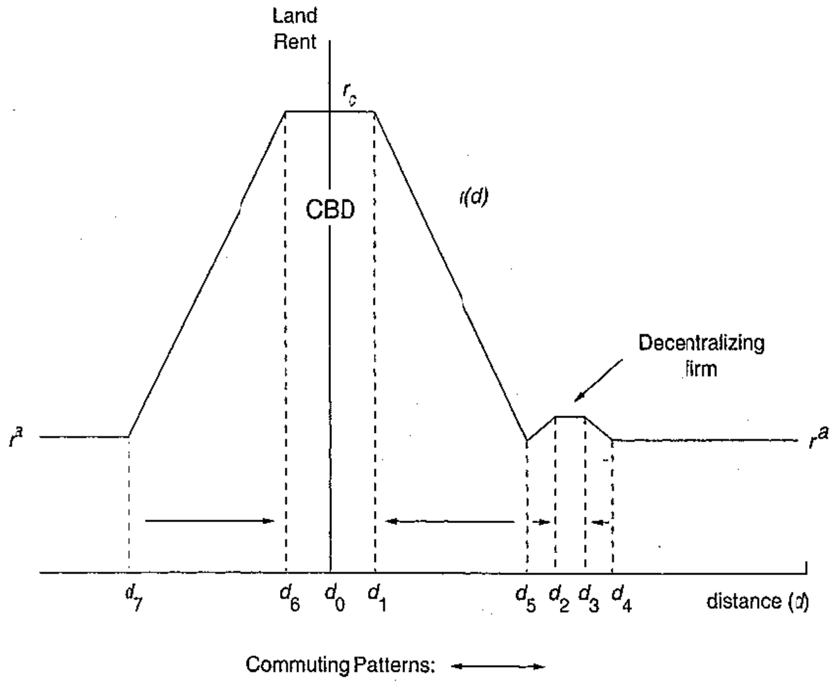

FIGURE 5.11 Decentralizing firms.

workers who live nearest its new location (to the right of d_3 and to the left of d_2). Since these workers are willing to pay more in rent as their commuting costs decrease (towards d_3 or d_2), workers employed at the decentralizing firm will bid up nearby residential land rents. Hence, we get the bump in the rent gradient between d_5 and d_4. At these locations, workers at the decentralized firm will outbid CBD workers for land. In order to actually occupy the land between d_2 and d_3, the land rent paid by the decentralizing firm must match the residential land rents paid by its workers. Let's first focus on the impact of firm decentralization on wages and commuting costs; we will examine the impact on firm land rents later in the chapter.

If workers at the decentralized firm are paid the same wages as CBD workers, then clearly they will have higher net wages since, on average, they have considerably shorter commutes. For example, at the distance d_5, workers commuting leftward to the CBD or rightward to the decentralized firm will both pay the same for land $[r(d_5)]$. Those working at the decentralized firm would have the shorter commute $(d_2 - d_5)$, versus the longer CBD commute $(d_5 - d_1)$. Since workers are identical and free to change jobs, a decentralizing firm will be able to pay less in wages. How much less? Just enough to leave workers indifferent between commuting to the decentralized firm and to the CBD. As in Chapter 3, we can assume that workers value their net income (gross income minus commuting expenses and rent for a standard house). Thus, the suburbanizing firm need pay

only a wage w_2 that yields the same net income to its workers as the wage w_1 at the CBD. If the cost (per mile) of commuting is k, this requires that for workers living at d_5:

$$w_2 - r(d_5) - k(d_2 - d_5) = w_1 - r(d_5) - k(d_5 - d_1); \text{ or, } w_2 = w_1 - k(d_5 - d_1) + k(d_2 - d_5) \qquad (5.5)$$

Equation (5.5) is the spatial equilibrium condition for workers which permits us to identify some important implications. First, if the firm chooses to locate farther from the CBD, d_2 and d_5 increase, while the distance $d_2 - d_5$ remains fixed. This enables the firm to pay increasingly lower wages. Thus, there should be a wage gradient as one moves from interior to peripheral sites within a metropolitan area. Second, a firm contemplating decentralization with no costs other than wages should seek a location near the current edge of the city, where half of its workers can be accommodated on land between d_3 and the city's border, d_4. This is where its wages will be lowest.

While Equation (5.5) defines a spatial equilibrium with respect to workers, it does not address the question of whether firms are in a spatial equilibrium. Given our assumptions so far, the suburban firm with its lower wages has a clear cost advantage over firms in the CBD. Why won't all firms benefit from decentralizing and, in turn, move? Wages will be lowest when the commuting distances of workers are shortest and land rents are lowest. This argument leads to an ideal picture of the overall dispersal of firms, in which each firm is surrounded by its own workers. The metropolis becomes completely decentralized, with a sprawl of individual firm-worker subcenters, such as depicted in Figure 5.12.

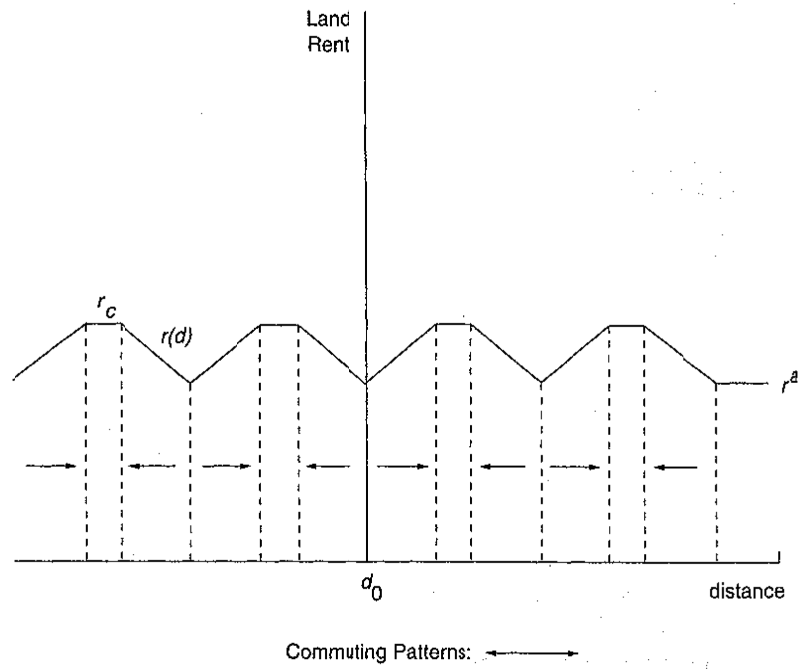

FIGURE 5.12 Subcenter sprawl.

As we have described it, the fully decentralized city in Figure 5.12 would have a number of advantages over the traditional monocentric city that we developed in Chapter 3 and that characterized nineteenth-century cities. To begin with, the average commuting time of residents would be very low, which, in turn, would generate lower average land (and housing) rents. Furthermore, as the population of the city grew over time, the commuting expenses of its workers (and, hence, land rents) would not necessarily increase, as was always the case with the monocentric city. In fact, larger metropolitan areas would simply have more subcenters rather than a larger single center. Do cities with more decentralized employment patterns have lower average commuting times relative to more centralized cities? (See (White 1988)). Gordon, Kumar, and Richardson (1989), for example, find strong statistical evidence that they do.

While this theory of employment dispersal is logically consistent and economically rational, in the real world, decentralization runs into several impediments that can severely limit its occurrence. First, our theory assumes that commuting in any direction is equally easy. Clearly, this need not be the case. Extensive transit infrastructure in many older cities has created a transportation system with far greater capacity to move people to the CBD than out or around to suburban subcenters. Even within the suburbs, the recent addition of circumferential highways concentrates highway capacity at a limited number of intersection points. Thus, it might be argued that older cities will never completely decentralize, whereas employment dispersal in newer cities will be determined by the extent and character of each region's transportation network.

Second, our theory was formulated with a homogeneous labor market in which all households/workers were identical. Clearly, this is not the case, and firms must employ workers of diverse skills at different wage rates. Such workers will also tend to live in different types of housing units at various density levels. Suppose for the moment that a firm contemplating decentralizing has a different mix of workers from other firms located at the CBD. Will the workers who currently live in the suburbs and work at the CBD be appropriate for the new firm? If not, the firm may have to hire workers from a much farther distance, which will limit its ability to pay a lower wage. Alternatively, could the firm's present workers, who currently live at other locations, move and buy the land or housing that is adjacent to the firm's new suburban site? Only if the housing on that land has the attributes desired by those workers. In short, assembling a workforce around a suburbanizing firm in theory is a simple task, but in reality can be quite difficult. Suburban workers will accept lower wages only if they can successfully switch jobs or housing in order to reduce their commute significantly.

During the last decade, researchers have begun to assemble enough empirical evidence to document that within many metropolitan areas, different wages are paid by location and that wage gradients do exist. Eberts (1981) found that comparable municipal workers are paid less in more distant suburbs. Ihlanfeldt (1992) found that suburban-based firms pay lower wages for a wide range of occupations in a sample of four metropolitan areas. Finally, Madden (1985) showed that when a worker changes jobs, a longer (shorter) commute is statistically associated with receiving an increase (decrease) in wages. All of this research seems to suggest that many workers are able to cluster close enough to their firms so that commute reductions are in fact being realized.

SUBCENTERS AND URBAN AGGLOMERATION

In the discussion so far, the only limiting factors that prevent firms from completely decentralizing are the impediments of the housing stock, local government institutions, and irregularities in the provision of infrastructure. There seem to be no economic limits to decentralization; or, alternatively, no economic forces that operate to encourage firm clustering in more centralized location patterns. But few, if any, metropolitan areas approach this pattern of complete decentralization. Instead, in most metropolitan areas, decentralizing firms tend to be clustered in any number of subcenters. To help explain this phenomenon, over the years, economists have developed several arguments suggesting that firms clustering in larger subcenters might receive some form of *agglomeration* benefit that could improve productivity, increase innovation, or reduce production costs. Let's examine these arguments in more detail.

One argument is based on the notion that, at least historically, there may have been productivity gains to an individual firm which are realized when all of its divisions or functions are located at the same point. Thus, if production, marketing, research, and administration are all carried out adjacent to each other, the ability to communicate face-to-face helps the firm operate more efficiently. This would clearly lead to the common clustering of facilities for one firm. In recent years, however, this argument has become less convincing as many firms seem to locate their various divisions in widely different locations. With the advent of vastly improved telecommunications, each division of a firm can be in constant contact with other units, operating with common information databases. Today, it may be quite economical for a firm to have its headquarters in downtown Manhattan, its production facilities in a New Jersey suburb, its clerical support in a Connecticut suburb, and its telemarketing unit in North Dakota!

A second explanation, also emphasizing the ability to easily communicate face-to-face, argues that *different* firms might desire to cluster together. Historically, urbanists have advanced this as a primary reason for the existence of a downtown CBD. It is unclear in the argument, however, whether the communication being undertaken is between similar firms (e.g., information sharing), between firms and their suppliers (e.g., informal vertical integration), or between firms and their clients (e.g., marketing). The distinction is important, for it suggests which kinds of firms may gain advantages from co-locating in the same cluster: firms in the same industry as opposed to firms in different industries vertically linked in business relationships.

These contact theories of urban agglomeration have been subject to little empirical research over the years. Surveys done in the 1960s and early 1970s did suggest that direct business contacts were then particularly important in the more skilled managerial occupations (Goddard 1973; Thorngren 1970). More recent interviews (Clapp 1980; Archer 1981) reveal that office firms located in the CBD explain their locational choice in terms of communicating with clients or other businesses. Those firms selecting suburban sites voice a greater concern for labor availability and accessibility. While these studies are suggestive, clearly more evidence is needed that *direct* business contacts play a significant role in determining firm locations, especially as new telecommunications media are more widely adopted.

If we assume for the moment that some form of agglomeration economy does exists within subcenters, we can develop a simple explanation of how the number of subcenters and their size is determined. In Figure 5.13, we illustrate how the cost of business for firms might vary with the size of the subcenter. The increasing schedule shows the rise in wages that must be paid at larger subcenters to compensate workers for increased commuting costs. This schedule must intersect the vertical axis, since a positive wage is required even at a single firm subcenter. Its rise should be linear if the density of residential development is fixed. The declining schedule shows a hypothetical agglomeration cost of production based on the informational or other benefits discussed above. It, too, must intersect the vertical axis (at A_0), but, from there, must eventually run parallel to the horizontal axis. This gives it a distinct nonlinear shape. The sum of the two schedules is the total cost of production. It will initially fall, reach a minimum, and then rise as the beneficial effects of better information wear off while increased commuting continues to drive up wages.

In Figure 5.13, it is interesting to estimate the impact of improvements in telecommunications on the shape of the schedules. The agglomeration cost schedule should not only be lower (shifting the intercept from A_0 to A_0'), but also flatter, as direct-contact costs are replaced by a medium that is almost insensitive to physical proximity. The end result is that with improved telecommunications, total production costs become lower, less U-shaped, and with a minimum point that is closer to the origin. The dashed schedules in Figure 5.13 show the impact of this technological change.

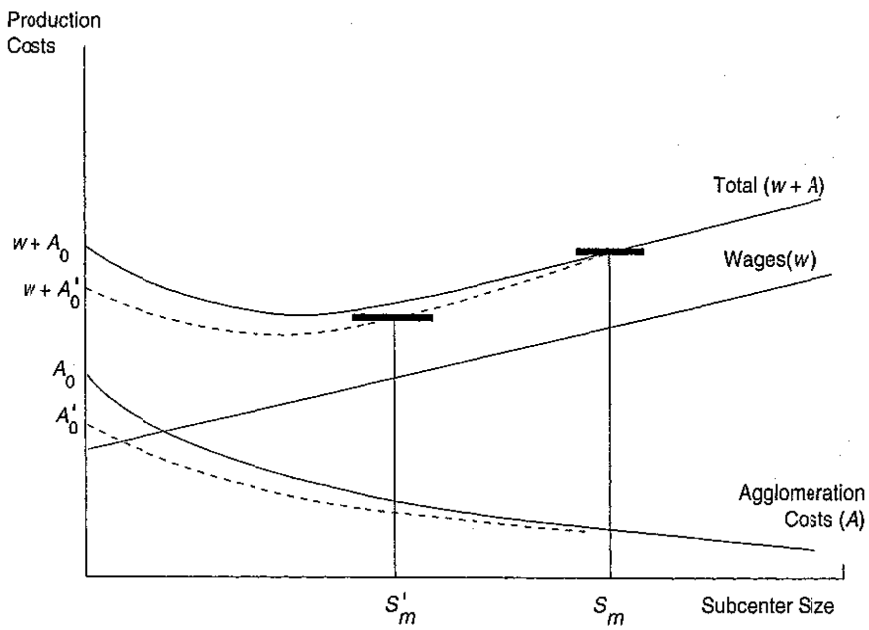

FIGURE 5.13 Subcenter size and firm costs.

Given firm production costs such as those in Figure 5.13, Helsley and Sullivan (1991) developed a simple model of how and why subcenters are formed. It begins with the assumption that as a region grows, new firms enter the region, select a location with the cheapest cost of production, and invest in a plant or facility. The plant or facility becomes a sunk cost which is assumed to make it prohibitive for the firm to move, even if its production costs (wages and agglomeration) would be lower at another site. Initially, there is one small center with low wages but high agglomeration costs (A_0). As the region grows, this center expands and total production costs move down at first. After the center reaches a certain size, total costs begin to rise, but a new center cannot yet form since total costs there ($w + A_0$) would still be higher. Eventually, as the center reaches some maximum level (S_m), its total costs of production rise to the same level as an isolated firm would experience ($w + A_0$). As the region grows, a second center is formed, and new firms are attracted to this center, where production costs are now falling, as opposed to the original center, where they remain at the higher level dictated by the size S_m.[3] As this second subcenter grows and reaches the size S_m, a third new subcenter forms, and so on. The end result is that after many years, there are a number of older centers, each of size S_m, together with one new center of some size less than S_m.

Within this framework, a technological improvement that reduces information costs (e.g., telecommunications) does have a decentralizing impact. The dashed schedules lead to lower agglomeration costs for the isolated firm (A_0'), but little cost change for firms at larger centers. For a city with the same number of total firms, the changing technology eventually dictates a larger number of smaller-sized subcenters (with size S_m'). As cities grow over time, a shift to the dashed schedule will also lead to the earlier formation of a second or third subcenter. As agglomeration or information costs become less important, cost-minimizing firms are better able to take advantage of the potential wage savings that come with more decentralized locations.

This discussion of urban agglomeration has omitted an important issue which will be raised in Chapter 14. If the location of a firm in a cluster contributes some distinct information benefit to the other firms located there, is the private market sending a sufficient signal to firms (through lower production costs) about the full benefit of clustering? A firm will certainly capitalize on the benefit of clustering, but will it consider the impact that it has on others? How can we take account of the benefits to other firms in a cluster as an additional firm enters? We discuss these external benefits, or externalities, in detail in Chapter 14.

SUBCENTERS, WAGES, AND THE URBAN LAND MARKET

Our theory of employment decentralization by firms has been depicted as a tradeoff between lower wages (caused by reduced worker commuting) and higher agglomeration or information costs (caused by the absence of many nearby firms). So far, we have not

[3]Firms at the older center do not move to the lower-cost, newer center because that would involve investing in a new plant.

fully integrated land rents into our discussion of firm decentralization. Residential land rents will vary with location as will wages, and to occupy land in a decentralized city, firms will have to match the land rents of residential users. In Figure 5.11, a decentralizing firm would not only pay lower wages, but would pay lower land rents as well. To understand more fully how businesses and households reach a spatial equilibrium when rents as well as wages vary across locations, we examine a city with two centers, a CBD and a suburban subcenter, and study the operation of the land market. We can generalize to many subcenters after completing our two-center example.

Our model has a CBD and a suburban subcenter located in a linear city, as illustrated in Figure 5.14. We choose a simpler, linear city rather than a circular one to keep our mathematics to a minimum.[4] The overall population of the city is fixed and is measured as the number of one-worker households N. N_1 of these workers are employed by firms in the CBD, while $N_2 = N - N_1$ work at firms in the second center. As in our previous discussion, we will assume for simplicity that all workers and all firms are identical. In this city, firms in the CBD occupy land between locations d_6 and d_1. To the left of d_6, land is occupied by residents who work in the CBD, out to the edge of the city at d_7. In the subcenter, firms use land between d_2 and d_3. To the right of the subcenter, land is used by subcenter workers, extending to the city's border farthest to the right, d_4. In between the CBD and the subcenter, there is residential land occupied by workers from both centers, and a boundary location d_5, to the left of which CBD workers live, and to the right of which subcenter workers reside.

Households-workers in our city have identical houses, as in Chapter 3, and a fixed amount of land per house: q. The firms that employ these workers at subcenters also build facilities (plants or offices) that occupy land at a fixed FAR. We combine the fixed firm FAR with a given amount of floor space per worker into a fixed overall usage of land per worker by firms: f. Thus, the density of workers at their place of work is $1/f$ while density at their place of residence is $1/q$. Workers commute to their firm along the line at an annual cost per mile of k dollars. Agricultural land to the right of d_4 and the left of d_7 is valued at r^a. Residential land rents will exactly compensate workers for the cost of commuting per acre.[5]

While land rents make the households who work at each subcenter equally well off at different residential locations, it is wages that will make mobile households equally well off *between* centers. Consider for the moment the workers who occupy land at the urban borders, d_4 and d_7. Both of these workers have land rent expenses equal to $r^a q$. They may, however, have quite different commuting expenses, depending on the commuting distances to the two centers. Denoting the wages paid at the CBD and subcenter

[4]In such a city, we must assume that there is actually a strip of land with some width, and that commuting occurs only linearly along this strip. To keep the mathematics simple, we assume that the width is 1 mile.

[5]Following the methodology presented in Chapter 3, residential rent, $r(d)$, can be defined as:

$$r(d) = r^a + k(d - d_7)/q, \quad d_6 > d > d_7$$
$$r(d) = r^a + k(d_4 - d)/q, \quad d_4 > d > d_3$$
$$r(d) = r(d_1) - k(d - d_1)/q, \quad d_5 > d > d_1$$
$$r(d) = r(d_2) - k(d_2 - d)/q, \quad d_2 > d > d_5$$

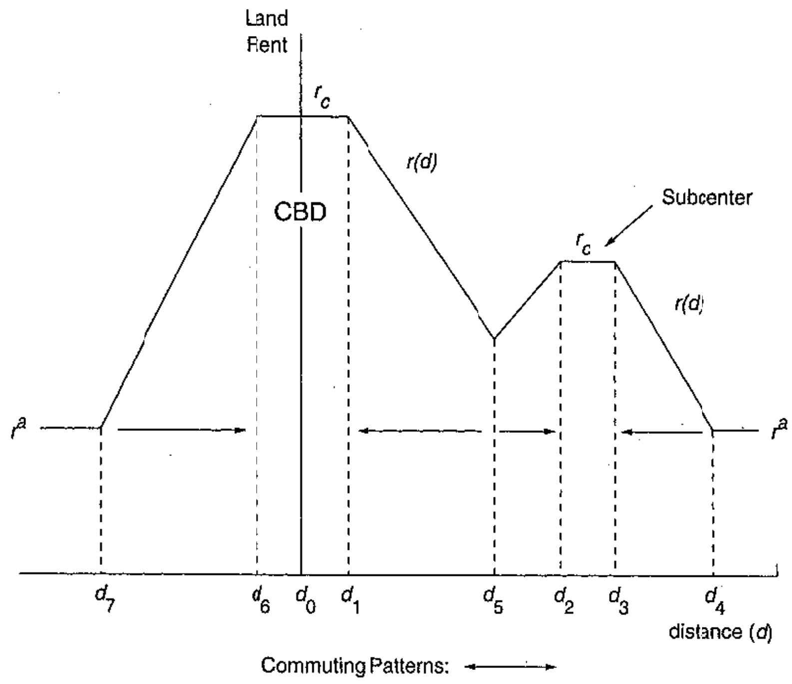

FIGURE 5.14 CBD and subcenter land markets.

as w_1 and w_2, the net income (after rent and commuting) of these two groups of workers must be the same:

$$w_1 - r^a q - k(d_6 - d_7) = w_2 - r^a q - k(d_4 - d_3), \text{ or,}$$
$$w_1 - w_2 = k[(d_6 - d_7) - (d_4 - d_3)] \tag{5.6}$$

Thus, as long as wages compensate workers for the commuting expenses made by the farthest worker at each subcenter, workers throughout the city will all be indifferent to location. Why? Because rents around each subcenter make workers at closer locations as well off as their colleague who commutes the farthest. Equation (5.6) is quite important, for it says that a system of different-sized subcenters can easily coexist within an urban land and labor market if higher wages at larger subcenters compensate workers employed there for the longer commute and higher land rents that inevitably result from greater size.

The population or employment of each subcenter is sufficient to determine all of the boundaries (d_1 through d_7), given the fixed location and distance between the center inner edges. Given the N_1 workers at the CBD (subcenter 1), its left boundary must extend far enough to accommodate that center's firms at the fixed worker density: $(d_1 - d_6)/f = N_1$. With a given overall workforce, the boundary to the right of the second center is determined from a similar condition: $(d_3 - d_2)/f = N - N_1$. The number of workers in each center also determines the various residential boundaries, given the fixed residential density. Remember that each household has only one worker in this model. In the interior between the centers, the boundary d_5 occurs where residential rents from each center are equal. The left extreme boundary, d_7, is determined such that enough land is developed for center 1 (CBD) workers: $(d_6 - d_7 + d_5 - d_1)/q = N_1$. Similarly, the right extreme boundary, d_4, is determined by $(d_4 - d_3 + d_2 - d_5)/q = N - N_1$. Thus, a center that has greater employment (i.e., the CBD) will have a larger land area to accommodate both its firms and households. These larger areas give rise to greater commuting distances and higher overall land rents, which necessitates paying higher wages.

Firms, like workers, are assumed to be all of the same type, exporting their products to world markets. Thus, as discussed earlier in this chapter, firms are price takers, and their profits depend only on the costs of producing their products. Some firm production costs will vary with location: wages, land rents, and any costs associated with center size or agglomeration. Other costs of business are exogenous, and do not vary appreciably across locations within the metropolis; for example, intermediate inputs and the capital costs of plant and equipment. We also assume that the cost of transporting products to world markets is exogenous, unlike in the monocentric model presented earlier in this chapter. In choosing a location, firms seek to minimize those costs that vary with location. Thus, any exogenous costs can be ignored in our model.

Let's now turn our attention to those costs that influence firm location. We define production costs due to agglomeration as $A(N)$ and we assume that these decline with center size, as was shown in Figure 5.13. Thus, firms operating at the CBD (center 1) have wage and agglomeration costs equal to $w_1 + A(N_1)$ dollars per worker, while those at the subcenter (center 2) have costs of $w_2 + A(N_2)$ dollars per worker.

With a competitive land market, commercial land within each center will have a rent equal to residential land rent at the edge of the subcenter. If we assume that there is effectively no cost of commuting within each subcenter, business land rents will be the same within the subcenter, and firms will have a uniform rental cost per worker equal to $1/f$ times this residential (and commercial) land rent.[6]

Just like households, firms in equilibrium must be equally well off, and this requires that the total costs of production be the same across centers. Aggregating the three components of business costs per worker (rents, wages, and agglomeration), we have the firm equilibrium condition:

$$A(N_1) + w_1 + r_1 f = A(N_2) + w_2 + r_2 f, \text{ or,}$$
$$A(N_1) - A(N_2) = w_2 - w_1 + f(r_2 - r_1) \tag{5.7}$$

Equation (5.7) can be used to identify the range of feasible sizes for each subcenter. We know that wage differences will simply equal the difference in commuting expenses to each subcenter's edge. We also know that firm rent differences will equal residential

[6]Business land rents at each center are r and r_2. Under the assumption of a competitive land market, we have:

$$r_1 = r(d_1) = r(d_6) = r^a + k(d_6 - d_7)/q$$
$$r_2 = r(d_2) = r(d_3) = r^a + k(d_4 - d_3)/q$$

rent differences, which, in turn, also equal this same commuting expense difference to the two subcenter edges. As a result, when firms are in equilibrium, the difference in production costs from agglomeration between the two centers must exactly balance with the differences in firm rents and wages:

$$A(N_1) - A(N_2) = \underbrace{k[(d_4 - d_3) - (d_6 - d_7)]}_{\substack{\text{wage difference} \\ \text{center 2} - \text{center 1}}} + \underbrace{(f/q)k[(d_4 - d_3) - (d_6 - d_7)]}_{\substack{\text{firm rent difference} \\ \text{center 2} - \text{center 1}}} \tag{5.8}$$

It should be clear from Equation (5.8) that long-run differences in center size (employment and households) can exist only if production costs vary because of agglomeration in a manner that exactly offsets the increases in wages and rents that result from differences in center size. This is equivalent to saying that in Figure 5.13, the sum of agglomeration and wage-rent costs is relatively flat over some range of subcenter sizes. Let's consider more carefully how this balancing can occur.

If in Equation (5.8) we assume for the moment that there are no agglomeration impacts on production costs $[A(N_1) = A(N_2) = 0]$, then equilibrium requires that the centers must have equal commuting distances, land areas, and, hence, workers (households): $d_4 - d_3 = d_6 - d_7$. On the other hand, suppose that agglomeration does effect production costs and $N_1 > N_2$, so that $A(N_1) < A(N_2)$. In this case, center 1's land area and commuting distance $(d_6 - d_7)$ must exceed that of center 2 $(d_4 - d_3)$.[7] Center 1's larger size will generate lower agglomeration production costs, but the corresponding greater commuting costs will lead to higher land rents and wages than at center 2. For the larger center to be in long-run equilibrium with the smaller, the production advantage from agglomeration to firms must be exactly offset by the combination of higher land rents and worker wages. We can summarize this result:

1. Larger subcenters can coexist with smaller subcenters within a metropolitan area if they offer firms some production advantage through agglomeration. In equilibrium, the agglomeration advantage of greater size must be balanced by the larger subcenter's farther commuting distances, which increase land rents and necessitate paying higher wages to keep workers.

Numerical Example

Let's illustrate this conclusion with a specific example. A reasonable worker density for office firms would be $1/f = 500$ workers per acre (320,000 workers per square mile).[8] Our metropolitan area will begin with *two even-sized subcenters* of $N_1 = N_2 = 1$ million workers each and we assume that all development takes place on a strip of land that is

[7]Given the conditions in footnote 5 and Equation (5.6), it should be clear that the wage differences and rent differences between centers must be of the same sign to ensure that workers are paid. Thus, if the left-hand side of (5.8) is negative, both expressions on the right-hand side must be also.

[8]This worker density would exist with FAR levels of between 2 and 3, together with roughly 250 square feet of floor area per worker. These are very typical figures for suburban office space.

10 miles wide throughout.[9] Given our worker density, the total land area required for each center is 3.125 square miles. With a land strip 10 miles wide, each center must be 0.3125 miles long ($d_1 - d_6 = d_3 - d_2 = 0.3125$) to accommodate the workers at their firms. If the center inner edges are 20 miles apart ($20 = d_2 - d_1$), and the worker density of housing is $1/q = 4$ per acre (2,560 per square mile), then the 200 square miles of land area between the centers can accommodate a total of 512,000 workers (or 256,000 workers for each center). On the outer edges, each center will have to extend about 29 miles ($d_6 - d_7 = d_4 - d_3 = 29.06$) for there to be enough land to house each center's remaining 744,000 workers ($29.06 \times 10 \times 2560 = 744{,}000$).

Land rents in this city can be calculated using the same numerical examples as in Chapter 3. If agricultural land rents, r^a, are \$1,000 per acre at the edges of the subcenters (d_4 and d_7), and the cost of commuting, k, at each center is \$200 per mile annually, then at the core of each subcenter land will rent for \$24,248 per acre to both firms and households (\$1,000 in agricultural rent plus 4 households per acre times \$5,812 in commute savings to each household).

Suppose that we allow some form of agglomeration to affect the production costs of firms. For illustration, assume that a center that is twice the size of another has a 5 percent increase in labor productivity over the other city's, or a \$2,500 cost advantage per worker. The question is, how will this level of agglomeration alter the center sizes, land rents, and transportation expenses of our example? Will a \$2,500 cost advantage in a center that is twice the size of another be eroded by the additional commuting expenses and land rents that such a size increase would entail? The right-hand side of Equation (5.8) says that to balance a \$2,500 agglomeration cost advantage, the difference in edge commuting distances between the two centers must reach 12.4 miles [$2{,}500/(k + kf/q) = 2{,}500/201.6 = 12.4$]. Thus, if center 1 is the larger, it must expand so that its outer edge, d_7, is 35.26 miles from the center [$29.06 + 1/2(12.4)$]; while center 2's border, d_4, contracts to an edge of 22.86 miles from its center [$29.06 - 1/2(12.4)$]. Since commuting costs \$200 per mile annually, the travel expenses of the farthest worker at center 1 will increase by \$1,240 ($1{,}240 = 200 \times 6.2$), while maximum commuting expenses at center 2 decrease by this same amount. For workers to be indifferent between employment at each center, Equation (5.6) says that wages at center 1 will have to rise by \$2,480 relative to center 2.

With the new outer edges of residential development established, we can recalculate central land rents at each subcenter. With center 1's edge at 35.26 miles (d_7), central rents will be \$29,208 per acre, while center 2's 22.86-mile edge at d_4 creates central rents of \$19,288 per acre. In a competitive land market, as the central rents rise in center 1, it will expand to take more of the land between the two centers, and center 2 will shrink, taking less of this land. The boundary between the centers (d_5 in Figure 5.14) moves from halfway between the two when they were the same size to 16.2 miles from center 1 and 3.8 miles from center 2. Center 1 now has 6.2 miles of additional interior land and another 6.2 miles from border expansion, while center 2 has lost this much through contracting its edge and share of interior land. The net result is that center 1 now has 93 percent more workers (households) than center 2. Center 1's total land area [now $(16.2 + 35.26) \times 10 = 514.6$ square miles] multiplied by 2,560 households per square mile yields 1,317,376 workers at that center, with a residual 682,624 at center 2. This is almost exactly equal to the difference in center sizes that was assumed to generate the agglomeration advantage. In other words, with the *assumed* level of agglomeration, a center that is twice as large as another will require additional commuting expenses and rent that roughly match the agglomeration advantage.

In this example, households at the edge of the now-expanded center 1 will incur additional commuting expenses of \$1,240 annually, while those living next to that center's employment zone will find that their annual rent for a lot has increased by \$1,240. At center 2, both the commute from the edge and central rents fall by these amounts. Firms at center 1 must pay their workers \$2,480 in additional wages relative to center 2 in order to successfully compete in the labor market. But what of the additional \$20 that remains between the original \$2,500 in agglomeration benefits and the \$2,480 in higher wages? This is exactly made up (with rounding error) by the additional land rent that firms must pay (per worker) at center 1 versus center 2. With a worker density of 500 per acre, the \$9,920 rent differential between center 1 and center 2 translates into a difference of \$19.84 per worker ($9{,}920/500 = 19.84$).

Given Equation (5.8), it is not surprising that almost all of center 1's \$2,500 agglomeration production advantage winds up getting absorbed by the difference between center wages. To make households equally well off, rents per acre at the core of center 1 rise from \$24,248 to \$29,208, while those at center 2's core are reduced to \$19,288. The difference between core land rents per acre is \$9,920, or \$2,480 for each household's quarter-acre lot. This difference in lot rents is what must be compensated for in wages. The difference in core land rents also impacts firms, but, in this model, it does so only trivially. The \$9,920 rent difference per acre at the core increases the firm's annual cost of land per worker by only \$20, because firms use such little land per worker. A \$9,920 difference in land rents per acre between centers translates into a difference of only \$0.23 per square foot. If firm facilities have a FAR of 2.0, the floor space rent for those facilities should differ by less than \$0.12 per square foot. This example illustrates an important conclusion that holds generally when there is a competitive land market in which firms and residents need only match each other's rents to occupy land.

2. At larger subcenters, with agglomeration advantages, workers will have to commute farther, increasing residential and commercial land rents per acre. Because households use much more land for their residence than workers use in their workplace, the impact of greater land rents is primarily to generate wage differences rather than differences in the rent for facility space to firms.

As we discussed earlier in this chapter, there is considerable empirical research which documents the existence of urban wage gradients; wage rates are higher at large urban CBDs and lower at suburban subcenters or more peripheral sites. These wage

[9] In footnote 4, we reminded the reader that our linear city must have a width. To simplify the mathematics we assumed that width was 1 mile. Here, to make our example more realistic, we assume a width of 10 miles.

differences presumably exist because of differences in the commuting expenses associated with working downtown as opposed to working at various suburban locations. This is quite consistent with our model above, as long as agglomeration productivity factors compensate firms for the higher wages at CBDs. The model's conclusion that differences in the rent for facility space should be minimal across subcenters is more problematic. Let's examine some data on urban subcenters to see how they vary in size and rental rates.

SUBCENTERS AND RENTS IN GREATER BOSTON

At the beginning of this chapter, we focused on the changing employment patterns of the Boston metropolitan area, illustrating how manufacturing jobs and office and service jobs have been rapidly decentralizing. We now consider information from commercial real estate brokerage companies about the location of industrial or office buildings and the rents that they currently command in the marketplace. These kinds of data are widely available within many U.S. cities. In Table 5.1, we list those towns or subcenters within the Boston area that, in 1993, contained contiguous office buildings with a total of at least 1.5 million square feet (1.5 percent of the region's total office space). There are three such clusters within the city of Boston and 11 subcenters in the suburban towns that ring the center city. Included in the table are average market rent levels per square foot for that office space within each center that is in buildings over 50,000 square feet in size that were built during the last 15 years.

TABLE 5.1 Office Area, Buildings, and Asking Rents, Boston-Area Towns, 1993

Town (Cluster)	Square Feet (000s)	No. of Buildings	Rent
Boston (Back Bay)	10,675	66	25.19
(Financial District)	26,754	141	26.73
(South Station)	3,053	21	23.50
Andover	1,438	10	16.25
Burlington	3,498	43	18.90
Cambridge	11,103	116	18.64
Framingham	3,196	39	14.06
Lexington	2,320	38	19.41
Natick	1,518	19	15.50
Newton	1,973	38	18.32
Quincy	4,797	44	15.90
Waltham	5,843	60	19.60
Wellesley	1,774	36	19.45
Westborough	1,664	15	12.50
Residual	26,793	548	15.21
MSA	106,399	1,234	20.74

Source: Special tabulation from The Property Information Management System, CB Commercial, Inc., August 1993.

In the Boston office market, the three center-city clusters account for 35 percent of the region's office space, while the 11 suburban subcenters combined house another 40 percent of the market. The remaining 136 cities and towns contain only 25 percent of the office space, which clearly illustrates the tendency for office buildings to cluster. Across the 14 towns and clusters, average rental rates vary by almost 114 percent, from $26.73 in Boston's financial district to $12.50 per square foot at a distant subcenter. While the larger central-city areas clearly have higher rents than the suburban subcenters, within the suburbs, there remain large differences in rents that seem unrelated to the size of the center. For example, Wellesley has higher rent than Cambridge, despite the fact that Cambridge has over six times more space; there is a $3.70 rent difference between Waltham and Quincy, which have similar stocks of space.

In Table 5.2, we examine similar information for clusters or subcenters of industrial space within the Boston metropolitan area, using the same minimum-size criteria of 1.5 percent of the market. (For industrial space, this minimum would be 4.2 million square feet. The average rent levels are for warehouse space within each cluster that is less than 15 years old.) Comparing the data provided in Tables 5.1 and 5.2 yields two

TABLE 5.2 Industrial Area, Buildings, and Asking Rents, Boston-Area Towns, 1993

Town	Square Feet (000s)	No. of Buildings	Rent
Boston	8,969	116	NA
South Boston	5,170	92	5.00
Andover	5,192	49	5.56
Bedford	4,126	61	6.65
Billerica	8,430	161	4.26
Braintree	4,510	79	4.50
Burlington	4,353	106	7.12
Cambridge	7,351	128	NA
Canton	5,026	91	3.83
Chelmsford	4,565	58	1.87
Framingham	4,520	72	4.64
Lawrence	6,548	39	6.96
Lowell	11,350	111	2.54
Lynn	6,795	54	NA
Mansfield	5,432	53	5.53
Marlborough	5,567	86	5.73
Peabody	4,074	48	6.91
Quincy	4,197	18	NA
Waltham	5,877	149	5.50
Westborough	5,677	75	4.04
Wilmington	8,013	128	4.83
Woburn	12,233	249	5.51
Residual	139,658	2,490	3.94
MSA	277,633	4,513	4.40

NA, No available rental quotations.
Source: Special tabulation from The Property Information Management System, CB Commercial, Inc., August 1993.

important differences between office and industrial space in Boston. First, there is far less clustering of industrial space. The two industrial clusters within the city of Boston contain only 5 percent of the region's industrial space, and two suburban submarkets (Lowell and Woburn) have somewhat more than Boston's share of the market. Using the 1.5-percent-of-market minimum, there are more clusters (22, versus 14 for the office market), and, combined, they account for just under one-half of the region's industrial space (versus 75 percent for office). Thus, industrial space is far more dispersed and less likely to cluster than office space. Second, industrial rent for comparable space exhibits even more variation than was true for office buildings. Industrial rents range from $1.87 up to $7.12 per square foot, a 280 percent difference. Like the office market, there seems to be no association between subcenter size and rent level.

These Boston metropolitan area data suggest two important conclusions. First, the formation of subcenters is quite different across the two uses: office buildings tend to form fewer larger centers, while industrial buildings are more widely dispersed in a larger number of smaller clusters. If we rely on the arguments and models of the previous section, this systematic difference between the two uses suggests that whatever agglomeration economies exist must clearly be stronger for office firms than for manufacturers. Second, the tables also reveal significant differences for both types of property in rental rates across subcenters. This result seems to contradict a conclusion of our simple model, in which wages rather than rents varied most with different-sized subcenters. However, the variation in rents within Tables 5.1 and 5.2 does seem to be largely unrelated to the size of the subcenter. With the exception of Boston's higher office rents, larger subcenters do not systematically have higher rents than smaller ones. Perhaps wages, rather than rents, really do vary most across subcenters of different sizes, but other factors largely unrelated with subcenter size exert an influence over commercial and industrial rents.

SUBCENTERS, LOCAL GOVERNMENTS, AND LAND-USE RESTRICTIONS

Throughout this chapter, we have operated under the assumption of a competitive land market in which sites are eventually occupied by the use offering the highest return or rent. In reality, many cities and towns limit the area that certain uses may occupy. Thus, the size, growth, and existence of subcenters can be keenly influenced by the willingness of selected local governments to allow industrial or commercial uses within their boundaries. In Chapters 13 and 14, we will discuss the incentives and rationale for local governments to limit, regulate, or encourage the location of particular uses. At this point, we want to indicate only that the pattern of clustering within Boston or other metropolitan areas may be determined as much by land use regulation set by the local political process, as by underlying economic motives.

If local governments interfere with the land market through local zoning or other development regulations, then the rents for office or industrial space might vary far more across subcenters than is indicated by our simple model. If one use is not allowed to encroach or expand into the territory occupied by another, then there is no economic necessity for the land rents of these two uses to compete with one another. In the models above, we assumed that the rent for commercial land within each subcenter would equal or just exceed the rent for surrounding residential land. If the boundary between residential and commercial or industrial use is fixed institutionally rather than by market competition, this assumption no longer holds. Commercial rents within each subcenter can exceed or be less than residential land just across a border that is defined by regulation. Cities or towns that severely limit commercial or industrial development might have higher commercial land rents even if these developments are of smaller size. With land-use regulations, larger centers and suburban subcenters may coexist even if there are no agglomeration benefits.

Our simple model was also based on the notion that the cost of residential land varied only with subcenter size because of worker commuting. But suppose that residential density is not fixed, and, as in Chapter 4, larger subcenters tend to develop at higher residential density. With variable density, larger subcenters could have much higher residential land rents, and this, in turn, would generate a more pronounced variation of office or industrial rent with subcenter size. This might well explain the noticeably higher office rents within the city of Boston. The high historic urban densities in older cities can generate a powerful opportunity cost that commercial development must match if it is to acquire land.

There has been some limited empirical research on the role that local land-use regulations play in the commercial or industrial land market. A study within the Minneapolis metropolitan area demonstrated that the probability of industrial development occurring within a town was strongly influenced by the town's land-use policies (Wasylenko 1980). For land-use restrictions to have this effect, of course, they must not only be present but be binding.

More recent research has begun to examine directly the determinants of rent for comparable office space in the greater Los Angeles area. Sivitanidou (forthcoming) finds that office rents within the Los Angeles area vary by even more than the Boston data of Table 5.1. In order to explain statistical differences in subcenter office rental rates, she grouped explanatory variables into three categories: locational factors that firms would find desirable (e.g., proximity to the airport), factors that would influence the opportunity rent for residential land (e.g., school quality), and land-use regulations. Land-use regulations included measures of the FAR limits on commercial development, the scarcity of commercially zoned land, and the existence of various development moratoria. All of these factors had strong impacts on the rent for commercial space.

Throughout this chapter, we have seen that the location of industrial and commercial development has been influenced by long-run economic and technological change. Elevators for office buildings and horizontal assembly lines in factories have redefined the FAR levels of these two uses and given them distinct location patterns. The declining cost of transportation relative to labor has refocused business location decisions around their most important resource—labor. In this process of change, the public sector plays an increasingly important role. Would industries have decentralized in the early twentieth century without the initial construction of roads? Would office development have remained more centralized recently if urban centers had better transit systems rather than circumferential highway networks? Will local governments cooperate with business to provide the diverse housing needed to accommodate the business workforce locally? To what extent have some industries sought to flee the tax burdens of inner cities? In Chapters 13 and 14, we will further examine the role played by the public sector in development and track the evolution of public sector development policies.

SUMMARY

In this chapter, we have examined the location decision of firms within metropolitan areas, in particular those firms that export their products or services. These jobs have undergone extensive decentralization in most metropolitan areas since the early 1900s. As a consequence, the spatial structure of metropolitan areas is changing from one of monocentricity (single employment center) to polycentricity (many employment centers).

- Industrial firms have largely moved to suburban location because modern production and storage technologies make them extensive users of land. In part due to the spatially diffuse character of truck and rail transportation, industries decentralize because they are less willing to compete with denser (higher rent paying) land uses.
- More recently, many office-using firms are decentralizing to be closer to their workforce and, consequently, can pay lower wages. While land rent compensates workers for the commuting costs to a common employment center, variation in wages must exist to compensate workers for any differences in the average commuting costs that exist when there are multiple employment centers.
- Several factors can limit this trend toward decentralization. Local governments can hinder the ability of firms to match locations with their workforce by regulating the type of housing that is built. Also, firms may derive economic advantages from locating in larger, more central clusters, which facilitate communication and information sharing. With new telecommunications technology, this agglomeration effect may be diminishing, helping to encourage decentralization.

REFERENCES AND ADDITIONAL READINGS

Archer, Wayne "Determinants of Location for General Purpose Office Firms Within Medium Size Cities." *AREUEA Journal* 9 (1981): 283–297.

Brennan, T. P., R. E. Cannaday, and P. F. Colwell. "Office Rent in the Chicago CBD." *AREUEA Journal* 12 (1984): 243–260.

Clapp, John. "The Intra-Metropolitan Location of Office Activities." *Journal of Regional Science* 20 (1980): 387–399.

Eberts, R. W. "An Empirical Investigation of Intra-Urban Wage Gradients." *Journal of Urban Economics* 10 (1981): 50–60.

Goddard, John B. *Office Location in Urban and Regional Development.* Oxford: Oxford University Press, 1973.

Gordon, P., A. Kumar, and H. Richardson. "The Influence of Metropolitan Spacial Structure on Commuting Time." *Journal of Urban Economics* 26 (1989):138.

Hartnett, H. "A Location Analysis of Manufacturing in Chicago." *AREUEA Journal* (June 1973).

Helsley, R., and A. Sullivan. "Urban Subcenter Formation." *Regional Science and Urban Economics* 21, 2 (July 1991).

Ihlanfeldt, K. R. "Intra Urban Wage Gradients: Evidence by Race, Gender, Occupation and Class." *Journal of Urban Economics* 32 (1992): 70–91.

Madden, J. F. "Urban Wage Gradients: Empirical Evidence." *Journal of Urban Economics* 18 (1985): 291–301.

McMillen, D., and L. Singell. "Work Location, Residence Location and the Intraurban Wage Gradient." *Journal of Urban Economics* 32 (1992): 195–213.

Moses, L. N. "Towards a Theory of Intra-Urban Wage Differentials and Their Influence on Travel Patterns." *Papers of the Regional Science Association* 9 (1962): 53–63.

Schmenner, R. "The Rent Gradient for Manufacturing." *Journal of Urban Economics* 9 (1981).

Sivitanidou, Rena. "Urban Spatial Variations in Office Commercial Rents: the Role of Commercial Zoning." Forthcoming in the *Journal of Urban Economics.*

Sivitanidou, Rena, and William C. Wheaton. "Wage and Rent Capitalization in the Commercial Real Estate Market." *Journal of Urban Economics* 31, 2 (1992), 206–229.

Struyk, R., and F. James. *Intra-Metropolitan Industrial Location.* Lexington, MA: Lexington Books, 1975.

Thorngren, B. "How Do Contact Systems Affect Regional Development?" *Environment and Planning* 2 (1970): 409–427

Wasylenko, M. "Evidence of Fiscal Differentials and Intra-Metropolitan Firm Relocation." *Land Economics* (August 1980).

White Michele. "Firm Suburbanization and Urban Subcenters." *Journal of Urban Economics* 3 (1976): 323–343.

White, Michele. "Location Choice and Commuting Behavior in Cities with Decentralized Employment." *Journal of Urban Economics* 24 (1988): 129–152.

KING, JOHN L., AND THOMAS E. MCCUE. "Office Building Investment and the Macroeconomy: Empirical Evidence 1973–85." *AREUEA Journal* 15, 3 (1987): 234–255.

ROSEN, KENNETH T. "Toward a Model of the Office Building Sector." *AREUEA Journal* 12, 3 (1984): 261–269.

SHILLING, JAMES D., C. F. SIRMANS, AND JOHN B. CORGEL. "Price Adjustment Process for Rental Office Space." *Journal of Urban Economics* 22, 1 (1987): 90–100.

VOITH, RICHARD, AND THEODORE CRONE. "National Vacancy Rates and the Persistence of Shocks in U.S. Office Markets." *AREUEA Journal* 16, 4 (1988): 437–458.

WHEATON, WILLIAM C. "The Cyclic Behavior of the National Office Market." *AREUEA Journal* 15, 4 (1987): 281–299.

WHEATON, WILLIAM C., AND RAYMOND G. TORTO. "Office Rent Indices and Their Behavior Over Time." *Journal of Urban Economics* 35 (1994): 112–139.

The Market for Industrial Space in Philadelphia

ABEL, ANDREW B., AND OLIVIER J. BLANCHARD. "The Present Value of Profits and Cyclic Movements in Investment." *Econometrica*, 54, 2 (1986): 249–273.

FELDSTEIN, MARTIN, ed. *The Effects of Taxation on Capital Accumulation*. Chicago: University of Chicago Press, 1987.

KING, JOHN L., AND THOMAS MCCUE. "Stylized Facts about Industrial Property Construction." *Journal of Real Estate Research* 6, 3 (1991): 293–304.

WHEATON, WILLIAM C., AND RAYMOND G. TORTO. "An Investment Model of the Demand and Supply for Industrial Real Estate." *AREUEA Journal* 18, 4 (1990): 530–547.

CHAPTER 13

LOCAL GOVERNMENTS, PROPERTY TAXES, AND REAL ESTATE MARKETS

In most countries, local governments are the primary authorities that directly regulate how land is developed. Local governments control the pattern of development through type-of-use zoning, the character of buildings through subdivision regulations, and the location of development through the provision of infrastructure, roads, and utilities. Local governments are also the major providers of public services to residents and local government taxation of real estate assets is the dominant source of revenue for funding these services. In this chapter, we examine the provision of and payment for local public goods and services. In Chapter 14, we analyze the impact of local land use regulations on the pattern of development and the market for land.

In the United States in particular, a wide range of local jurisdictions holds the legal power to tax (mostly real estate) and the responsibility to provide certain public services. Municipalities or towns, counties, and school and special districts all operate with property taxes as their major source of tax revenue. In turn, these governments provide a range of public services that includes education, police and fire protection, water, sewage, and other infrastructure. In many instances, federal and state governments will financially assist local governments with funding for these services. While many government services, such as highway construction and welfare, are administered by state and local governments, often the federal government transfers money for these services and sometimes regulates their administration. In the U.S., federal and state governments are the direct providers of only a small number of services, most notably defense, national infrastructure, and social insurance programs such as social security.

The relationship between local governments and real estate markets is one of mutual interdependence. Since each local government may provide a very different package of services and taxes, household and firm location decisions are strongly influenced by local government behavior and the spatial distribution of local jurisdictions within a metropolitan area. Households, for example, will often commute extraordinary distances to live in a community with an acclaimed school system. Likewise, real estate values can be profoundly affected by a neighborhood's level of crime or its tax rate. In short, the system of local governments that exists within a metropolitan area exerts a major impact on the location of firms and households as well as the resulting pattern of prices for real estate assets.

The impact of local governments on property rents or real estate prices can be so strong that it dominates many other factors. Consider, for example, the simple land-price gradient in Chapter 3 that emerged as a result of commuting. This is shown as the dashed line in Figure 13.1. In many metropolitan areas, however, the actual pattern of land prices may resemble that of the solid line. As a result of deteriorating services, crime, and high taxes, many inner cities have depressed land prices. Frequently, there are established affluent communities adjacent to central cities, incorporated many years ago to provide their own services and escape the financial burdens of the city. Within literally a few blocks, land prices can change dramatically. Moving further out, there are numerous other suburbs for which land prices also vary in response to zoning regulations, taxes, and the quality of services. The net result is that better services are frequently found in communities located at greater distances from the city center, and this can override the effect

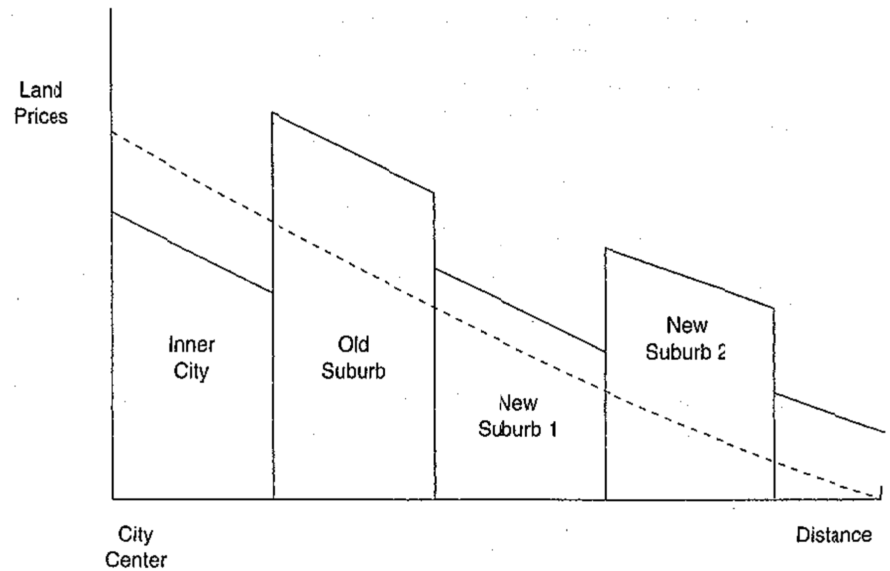

FIGURE 13.1 Land-price gradient with local jurisdictions.

of commuting, leading to a price or rent gradient that resembles the solid line rather than the dashed line in Figure 13.1.

While the services that local governments provide influence the real estate market, the opposite is also true: the real estate market can determine the character of communities and their financial ability to provide services. The link operates both ways. The composition of a town can change over time, altering the real estate tax base that is needed to raise the town's revenue. Often, different towns can wind up competing with each other over land use developments that provide (or are perceived to provide) more tax revenue than they use in services. Through zoning or other mechanisms, towns can try to block real estate developments that will use more services than they yield in new tax receipts.

In this chapter, we explore the different links that exist between the real estate market and the system of local governments that provides many of our services. We begin by documenting how local governments operate in the U.S. and describe the patterns of service responsibility and taxation between the federal, state, and local level. This is followed by a careful examination of different town budgets within a metropolitan area to explore the role that real estate markets play as a town's resource base. In the next section, we look at how property taxation creates a number of financial incentives for towns to regulate development and to exclude certain uses. The final sections of the chapter tackle the question of how firms and households select a town in which to reside or set up business. This leads to a theory of *fiscal capitalization* in which the value of services received and taxes paid across different local governments may be reflected exactly in the real estate values of those communities.

GOVERNMENTS AND PUBLIC SERVICES

In the United States, three levels of government provide public services: federal, state, and local. The Statistical Abstract of the United States recorded the following numbers for the five types of legal local jurisdictions in 1992: 3,043 counties, 19,296 municipalities, 16,666 townships or towns, 14,556 school districts, and 33,131 special districts. County governments cover almost all the land in the U.S., while municipalities, townships, and school and special districts cover only a small fraction of the nation's land area. Furthermore, counties, towns, and districts are rarely mutually exclusive. Thus, in the urban portions of California, for example, it is quite common for a household to pay a share of property taxes to a county (for roads and social services), to a town (for police, fire, and sanitation services), to a special district (for water and power), and to a school district. In Massachusetts, on the other hand, districts rarely exist and counties have virtually no responsibilities. Almost all local public services are provided by a system of municipalities or towns that covers the state completely. The most common arrangement in other states is for rural residents to be served by county governments only, whereas urban residents are most commonly served by some mix of towns and districts.

In Table 13.1, we group the expenditures made by all local governments and districts into a single category and compare those figures to expenditures made by the federal and state governments for FY 1990–1991. The first column, "Federal, Direct,"

TABLE 13.1 Government Expenditures,[*] 1990–1991

Expenditure	Federal		State		Local		All Governments	
	Direct	Transfer[†]	Direct	Transfer	Direct	Transfer	Expenditure[‡]	Transfer
Defense and international relations	366,112	—	—	—	—	—	366,112	—
Health, welfare and social insurance[§]	523,071	101,472	207,986	32,781	94,301	3,111	825,358	137,364
Education	20,192	24,537	80,468	116,180	228,834	429	329,494	141,146
Infrastructure and natural resources[**]	54,801	18,382	52,808	11,985	51,107	779	158,716	31,146
Law enforcement and fire protection[††]	8,111	736	22,592	2,154	51,332	103	82,035	2,993
Sewage, solid waste management and utilities	0	0	8,050	761	99,802	96	107,852	857
Other[‡‡]	347,077	15,018	70,391	22,679	92,133	883	509,601	38,580
Total	1,319,364	160,145	442,295	186,540	517,509	5401	2,379,168	352,086

[*]In millions of 1991 dollars.
[†]Transfer columns represent intergovernmental transfers to all other levels of government. Total expenditure per category for each level of government is the sum of Direct and Transfer columns.
[‡]Excludes duplicative intergovernmental transactions.
[§]Includes social services and income maintenance, insurance trust expenditure, housing, and community development.
[**]Includes natural resources, parks and recreation, highways, air transportation/airports, and other transportation.
[††]Includes police protection, fire protection, and correction.
[‡‡]Includes other general expenditures such as space research and technology, postal service, and libraries; government administration; and interest on debts.
Source: Government Finances: *1990–91, U.S. Census of Governments,* U.S. Bureau of the Census, Washington, D.C., vol. 4, no. 5.

gives the direct expenditure outlay for each budgetary category made by the federal government. The next column to the right, "Federal, Transfer," gives the federal expenditure that was disbursed in the form of grants to lower levels of government. The sum of the two is the total federal outlay for each Expenditure category. For example, the federal government takes full responsibility for defense and international relations and spent $366 billion in this category in 1990–1991. Under infrastructure and natural resources, the federal government spent $73 billion, of which $18 billion was in the form of grants to state and local governments. The remaining $55 billion was in the form of direct federal purchases for the interstate highway system, national parks, and so on. State governments spent nearly $65 billion within this category, roughly $53 billion directly, and $12 billion in the form of intergovernmental transfers. Local governments, in turn, spent $52 billion. Total spending for all governments on infrastructure and natural resources was $159 billion, excluding the $31 billion of duplicative intergovernmental transfers.

Table 13.1 has some limitations. Of the $52 billion that local governments spent on infrastructure and natural resources, for example, we don't know how much of that derived from their own funds and how much derived from grants from states and federal government. The Census of Governments data does not specify the different recipients of federal aid expenditures—all intergovernmental transactions are lumped together.

Therefore, we can't be certain what portion of the total $18 billion in federal aid for infrastructure eventually went to local governments as opposed to states. What we can be certain of, however, is the portion of total expenditures each level of government directly controlled.

The categories of expenditure that are used in Table 13.1 are chosen to illustrate the different responsibilities of each level of government. The largest category of federal expenditure (accounting for 42 percent of the federal budget) covers all transfer payments: health, welfare, and social insurance. The major items in this category are social security payments and health insurance. Other items include direct welfare payments, unemployment insurance, veterans benefits, and housing assistance. Total state expenditure (direct and transfer) in this category is much smaller than that made by the federal government, but it accounts for a similar share of total state expenditure (about 38 percent). State expenditure on transfer programs reflects mainly the state share of federally mandated programs (Aid to Families with Dependent Children, Unemployment Insurance, Medicare); local governments spend only 16 percent of their budget on this category. Some local governments operate their own hospitals and health clinics. Overall, the federal government is responsible for more than three-fourths of spending on health, welfare, and social insurance.

The category of infrastructure and natural resources is more evenly split between the three levels of government. The federal government gives only a fraction (25 percent) of what it spends to lower levels of government; the states transfer 18 percent of their expenditures in this category to local governments. As a share of each level's budget, infrastructure represents between 5 percent and 10 percent of total expenditure.

The "Education" category covers higher education as well as primary and secondary schools. Federal government expenditure is mainly for scholarship support with a small amount for primary education programs such as school lunches. State governments spend directly on their systems of public colleges and universities, but the largest item of state expenditure is financial aid to school districts for primary and secondary schools. In the U.S., local governments are responsible for primary and secondary education, spending almost 70 percent of the education budget for all governments. However, nearly two-thirds of local government expenditures on education may come from higher level government funds.

The final two categories of services are all primarily local functions: law enforcement, fire protection, and sanitation and utilities (water, solid waste management). There is virtually no federal role in any of these services.

The differences between the three levels of governments in the U.S. are reflected in revenue sources as well. Table 13.2 divides up the revenue of each level of government by several common categories that distinguish between types of taxes, user fees, and grants.[1] At one extreme, the federal government collects 79 percent of its revenue from personal and corporate income taxes and only 14 percent from direct charges or user fees. At the other extreme, local governments receive 20 percent of their revenue from

[1]Excise taxes cover both general sales taxes as well as specific sales taxes such as that on cigarettes, lodging, and gasoline. Fees include any charge for a service and cover such items as postal service revenue, water and sewer fees, publicly operated utility charges, and license or registration fees.

TABLE 13.2 Government Receipts,* 1990–1991

Revenue Source	Federal	State	Local	All Governments
Personal income and wages†	856,170	201,031	26,229	1,083,430
Corporate income	98,086	20,357	1,886	120,329
Excise‡	58,495	160,009	86,299	304,803
Property	0	6,228	161,772	168,000
Fees§	167,123	97,627	125,126	389,876
Other taxes	17,574	31,163	9,039	57,776
Receipts from intergovernmental transactions	3,234	143,534	201,833	348,601
Total revenue	1,200,682	559,949	612,184	2,124,214**
Employment (000s)	3,091††	4,115	10,076	17,281

*In millions of 1991 dollars.
†Includes individual income, and insurance trust revenue.
‡Includes sales, gross receipts and customs, and utility and liquor store revenue.
§Includes charges and miscellaneous general revenue.
**Excludes receipts from intergovernmental transactions.
††Civilian. Includes employees outside the United States.
Source: Revenues, *1990–1991 U.S. Census of Governments,* U.S. Bureau of the Census, Washington, D.C., vol. 4, no. 5; employment *Statistical Abstract of the United States,* Washington, D.C., 1991.

user fees and another third from federal and state grants. Only 47 percent of local government income is accounted for by all forms of taxation. This breaks down to 5 percent from personal or corporate income taxes, 14 percent from excise taxes, and 26 percent from taxes on property (the remainder comes from miscellaneous other taxes). State governments lie between these two extremes, with the most diversified revenue sources: 34 percent from income taxes, 24 percent from excise taxes, 15 percent from fees, and 22 percent from intergovernmental transfers.

Several conclusions emerge from Tables 13.1 and 13.2 about the structure of government in the U.S. Local governments collect the least revenue from their constituents but use that money to provide services directly. Thus, local governments employ more than 10 million workers, far more than either state or federal governments. The main civilian role of the federal and state governments is to provide a system of health, welfare, and social insurance with transfer payments to individuals and to contract with the private sector for infrastructure. Grants from state and federal governments provide one-third of the funding for direct services provided by local governments. The largest local revenue source is a tax on property or real estate assets.

Economists have developed a number of theories about the advantages and consequences of assigning public services to different levels of government. In this literature, for example, it is argued that income redistribution and social transfer payments belong at the federal level. If local governments were responsible for welfare programs, not only would there be uneven benefits across local jurisdictions, but governments that provided higher-than-average benefits would attract eligible recipients. The possibility of attracting program recipients from other jurisdictions could lead local governments to provide less of such a public service than the country as a whole might find desirable. In addition, if local governments were responsible for funding social transfer programs, wealthy citizens might move from jurisdictions with poor citizens to avoid paying the higher local taxes needed to provide such programs. Federal funding of transfer programs makes it difficult for wealthy citizens to avoid paying for these services (Ladd and Doolittle 1982).

In the case of services such as air quality and interstate highways, there is a different argument for federal provision that involves the benefits that the national population receives from locally produced services. Without assistance, local governments might ignore these external benefits, and again would underprovide services relative to what the nation as a whole would deem appropriate.

The primary economic argument favoring local service provision is that it introduces variety and competition into government, and this allows the quality and breadth of services to be custom-tailored to the demands of local constituents. Matching the provision of services to differences in the preferences of local residents can enhance efficiency. For example, national systems of primary and secondary schooling (such as exist in Japan or France) typically have a more uniform educational program throughout the country. By contrast, it is argued that a system of locally run schools can better tailor curricula to the needs and demands of local residents. This efficiency argument assumes that there are widespread differences within a nation in the kinds and degree of services that local residents need or desire. Since the level of education, law enforcement, fire protection, and sanitation services often vary widely across localities, perhaps there are substantial differences in the service demands of local residents.

LOCAL GOVERNMENTS AND PROPERTY TAXES IN METROPOLITAN AREAS

Within metropolitan areas, significant differences in the patterns of both expenditure and revenue exist between central cities and the numerous smaller suburban communities that normally surround them. These differences are particularly important because they may influence the location of resources across towns within a metropolitan area. That is, the movement of firms and households between communities can put strong competitive pressures on local governments. To illustrate these differences, we examine local expenditure in 1990 for a sample of towns in the Boston metropolitan area, where municipalities are virtually the sole form of local government. With a single system of local governments we avoid the problem of trying to add up budgets over different layers of local governments whose boundaries often overlap.[2] In Table 13.3, the city of Boston is compared with an affluent suburb (Concord), two middle-class growing suburbs (Burlington and Needham), and an older working-class suburb (Quincy). These communities are reasonably representative of the different types of towns found within the broader metropolitan area.

[2]While Boston's single system of local governments makes it ideal to examine local variation of public services, the state government has imposed a uniform maximum limit on the property tax rate of local jurisdictions. Although not always binding, this provision can limit some of the variation that might occur across towns in revenue collection and spending.

TABLE 13.3 Profiles of Selected Massachusetts Cities, 1990*

Item	Boston	Burlington	Concord	Needham	Quincy
1989 Median HH income	30,757	58,975	73,695	63,618	37,795
Households	250,683	8,054	4,764	10,405	37,732
Population	574,283	23,302	17,076	27,557	84,985
Unemployment rate	5.5%	5.0%	2.7%	3.2%	5.8%
Expenditures†					
Education/pupil	6,679	5,501	7,179	6,053	5,836
Education/HH	1,438	2,340	3,156	1,876	992
General government/HH	280	244	339	214	136
Police & fire/HH	836	773	641	544	590
Other public safety/HH	224	36	41	52	39
Public works/HH	284	577	300	416	313
Health & welfare/HH	704	76	64	56	19
Culture and recreation/HH	138	165	234	101	67
Debt service/HH	328	233	256	370	205
Other expenditures/HH‡	955	816	553	855	902
Total expenditures/HH	5,184	5,260	5,584	4,495	3,263
Revenues					
State aid/HH	1,846	707	554	360	942
Local receipts/HH§	1,425	993	434	954	461
Total property tax levy/HH	2,071	3,768	4,535	3,167	1,749
Other revenue/HH**	45	362	349	314	415
Total revenue/HH	5,389	5,830	5,872	4,795	3,567
Residential tax rate	0.85%	0.88%	0.97%	1.00%	1.02%
Percent of total levy	30.1%	36.0%	81.7%	73.0%	60.0%
Commercial & industrial tax rate	2.39%	1.73%	1.08%	1.22%	2.29%
Percent of total levy	64.3%	61.9%	16.5%	25.3%	37.4%
Assessed residential value ($ billion)	20.6	1.2	1.8	2.4	3.9
Total assessed value ($ billion)	35.8	2.4	2.2	3.1	5
Residential taxes/HH	623	1,358	3,706	2,313	1,050
Estimated total payments/HH††	1,052	1,716	4,061	3,010	1,327
Avg. single-family property tax bill‡‡	1,377	1,577	3,535	2,647	1,608

HH, Household.
*In 1990 dollars.
†Expenditures are from general fund only; special revenues, enterprise capital projects, and trust funds are not included.
‡Other Expenditures includes court judgments, municipal employees' health insurance and retirement plans, medicaid, liability insurance, payments to school, water, and fire districts; and expenditures not otherwise classified.
§Local Receipts includes excise taxes, interest on investments, fines, permits, rentals, forfeits, and water and sewer funds surpluses.
**Other Revenue includes capitalization and stabilization funds and municipal light surpluses.
††Total Payments/HH are estimated by multiplying the percent of tax levy from residential property by local receipts/HH and adding residential taxes/HH.
‡‡Average single-family property tax bill is computed by multiplying the single-family assessed value by the residential property tax rate and dividing by the number of single-family parcels.
Source: All data from Massachusetts Department of Revenue Data Bank/Online Service, except Median HH income; Median HH income, U.S. Bureau of the Census.

The first row in Table 13.3 shows the wide differences that exist in the average income of their residents. Median annual household income ranges from $31,000 in Boston to $74,000 in Concord, with Burlington and the other towns in between. The next few rows illustrate some of the demographic differences between the towns. In Boston, households have only an average of two persons, whereas in Concord and Burlington, the average is closer to four. Turning to "Expenditures," when measured on a per-pupil basis, educational expenditure is only weakly related to town income: Concord spends more than Needham, which spends more than Burlington or Quincy, but Boston (the community with the lowest median household income) has the second-highest per-pupil spending. When the comparison is made on a per-household basis, however, the Boston anomaly vanishes. Because Boston has fewer pupils per household than its surrounding bedroom suburbs, it is able to match their pupil expenditure with a lower cost per taxpayer or household. When measured per household, the 140 percent difference in income between Boston and Concord results in a 119 percent difference in educational expenditure. This suggests (as many studies have found) that spending on educational services is somewhat income inelastic.

The next several rows show that expenditures on services besides education also do not seem related to a town's income level. Police and fire outlays are again highest in Boston, whereas among the suburban communities, these expenditures show little relationship to town income. The wealthiest and poorest suburbs (Concord and Quincy) spend almost the same, while Burlington spends the most. A similar pattern among the suburbs emerges with public works expenditures, whereas the city of Boston spends noticeably less per household. In the health and welfare category, the social problems of inner cities have hit hard in Boston, resulting in a high level of expenditure there, whereas each of the suburbs spends virtually nothing.

The net result of these expenditure patterns is that the city of Boston spends nearly as much or more per household in the aggregate than all but the wealthiest suburb, despite having only half of the income per household of many of these towns. The 95 percent income difference between Concord and Quincy leads to only 71 percent greater per household expenditure.

Two explanations have been offered for why the simple link between town expenditure and average town income appears to be so weak. The first is that there are other important demand factors that generate public expenditure besides a town's resources. For example, when a town has a needy, low-income constituency or a land-use pattern comprised of dense, wood-frame houses, it may simply need to spend more on social services or fire protection. A second explanation emphasizes that a town's resources, at least as viewed by its decision makers, often consist of far more than its residents' income. Additional resources include the grants that are made available to local governments and the ability to tax business, tourists, and other uses or activities.

The row that depicts state aid suggests that at least some of the spending paradox may be explained by the grants that each town receives from the state government. Such grants make up more than 34 percent of Boston's per-household revenue but between only 7.5 and 12 percent for the wealthier towns of Concord, Needham, and Burlington. The poorer suburb of Quincy receives slightly more than 26 percent of its budget in state aid. Thus, public spending with only local resources is more strongly related to income than is total spending. Would the city of Boston spend as much as it currently does if those grants were withdrawn? In many situations, towns have the flexibility to use grants in one of two ways. They can either augment services (spending) and leave their own tax effort fixed, or they can reduce local taxes and leave service levels the same.[3]

The next row, labeled "Local Receipts," gives the amount of revenue per household that each town receives from fees (e.g., water and sewer), permits, excise taxes, and other sources. While some of this revenue comes from town households, firms and businesses make significant payments as well. This explains why this entry is so large for the city of Boston and also suggests that firms can represent an important potential source of revenue for a community.

If we take total expenditure per household and subtract state aid and other receipts, we are left with tax revenue. In Massachusetts, cities and towns can only tax property. Thus, total property tax revenue in Boston is $2,071 per household, and in Concord $4,535. In Boston's case, however, only 30 percent of this amount is raised from property taxes on households, whereas 64 percent ($1,325) is raised from property taxes on business. In Concord, property taxes on businesses are only 16.5 percent of the levy. To the extent that the owners of this business property do not live in the town in which their business is located, taxes on this property represent a potential source of revenue that does not come from town or city residents. In fact, property tax payments from residential property (houses and apartments) only account for $623 per household in Boston, a small share of local revenue. In Concord, residential taxes per household are $3,706.

If we assume that the same share of the category "Local Recipts/HH" comes from businesses as does property tax revenue, then in Boston, households pay an average of only $1,052 annually—for $5,184 worth of services! A full 80 percent of Boston's expenditures are financed from state aid and taxes or fees on businesses. In Concord, households pay $4,060 for $5,584 worth of services, so only 27 percent comes from sources other than town residents. In Quincy, 60 percent of expenditure is financed from nonresidents, while in Needham the share is 33 percent. Thus, there is a strong negative correlation between town income and the share of a town's budget that is financed from nonresidents. The exception is Burlington, a middle-income suburb that has undergone an explosive growth of suburban office and retail development over the last two decades. Here, two-thirds of total expenditures are from sources other than town residents.

If we take the average estimated total payments made by town households (next to last row) and divide it by the town's median household income (top row), we get some

[3]An economic argument can be made that the effect of grants on a town's budget depends very much on the terms of the grant. *Matching grants* reimburse a town for some fraction of its own expenditure (on certain budget items). Since the dollar value of the grant depends on how much the town spends, this form of aid tends to encourage higher spending, and the grant mostly augments services. *Block grants* are given in a fixed amount per capita or per pupil. Since block grants are not tied to local expenditures, they do not encourage increased local spending. Local governments may use block grants more as a substitute for local taxes.

Different state or federal programs use different grant formulas, but most grant programs are targeted toward cities or towns with fewer resources, and/or many needy constituents. Thus, the choice by Massachusetts to give more aid to Boston than to Concord is clearly one of social policy.

idea of the tax burden paid by residents. The residents of Boston wind up with taxes and fees that are about 3.4 percent of their income, even though total expenditure per household is close to 17 percent of income. In Concord and Needham, residents contribute 5.5 percent and 4.7 percent of their incomes, respectively, but without the extra nonresidential resources of Boston, this yields lower service percentages (7.6 and 7.1 percent of income, respectively). Thus, when measured as a percentage of town income, public expenditure is far less in wealthier towns. On the other hand, tax payments as a percentage of income rise slightly with town income. Of course, all of this would look quite different if the pattern of state aid or the distribution of nonresidential property were different across towns.

The last row in Table 13.3 gives the average tax payment made per single-family house, as opposed to a household. Comparing this number to average residential property taxes per household provides an indication of the pattern of housing in each town. In Boston, the average residential payment is almost half of the average single-family tax bill. This confirms that the city has many apartments and that their average unit value is far less than the city's typical single-family home. A similar story occurs in Quincy. However, in the middle- and upper-income suburbs, where single-family housing predominates, the two numbers are much closer. If the single-family tax bill is divided by the residential tax rate, the result is an average assessed value for the town's single-family housing.[4] Average assessed housing values—$162,000 in Boston, $265,000 in Needham, and $364,000 in Concord—tend to be roughly proportional to average town household income.

The data in Table 13.3, together with numerous studies, suggest a series of conclusions about the comparative fiscal behavior of cities and towns within metropolitan areas:

1. The demand for services is quite inelastic with respect to town income. In part, this results because education, police, and fire protection are viewed as necessities, rather than luxuries. It also results because expenditure on some services is generated by need (crime, poverty), rather than by demand.
2. Since the demand for housing has an income elasticity near to one, residential property value tends to be roughly proportional to town income. With no other resources, this would potentially lead to much lower property tax rates in higher income communities.
3. Nonresidential property provides an important tax resource for many communities, tending to be concentrated in inner cities or in selected moderate income suburban towns. This helps to offset the tax burden of providing services in these communities and makes property tax rates more equal across communities.

[4]The town's statutory residential tax rate, t_s, multiplied by the average assessed value of single-family property, AV, equals the average single-family tax bill, T, or $T = t_s AV$. The town *effective rate* t is the tax bill divided by the true "market" value of homes, P, or:

$$t = \frac{T}{P} = t_s \frac{AV}{P}$$

4. The progressive distribution of state and federal aid also is crucial in helping lower-income communities to provide comparable levels of public services without unusually high property tax rates.

The conclusions about how a town's residential property tax rate is determined and how this rate tends to vary across towns with different household income can be illustrated with a simplified town budget as shown in Equation (13.1).

$$t = \frac{(G - A - N)s}{P} \qquad (13.1)$$

Where: t = town residential effective tax rate
G = total town expenditure/household
A = state aid received/household
N = nonproperty tax revenue
P = average market value of houses in the town
s = share of property value that is residential

Using Equation (13.1) and the figures provided in Table 13.3, we can calculate the effective residential tax rate for Boston and Concord. For these calculations, we take total expenditures per household and subtract revenues per household from state aid and nonproperty tax revenue per household. We then take that sum times the portion of the tax base that is residential and divide by the average market value of housing units in the town. The average market value of housing units is determined by taking assessed residential value of houses and apartments in Table 13.3 and dividing by the total number of households.[5] Boston's average unit value, P, is \$82,175, whereas Concord's average unit value is \$377,834. For Boston, Equation (13.1) would yield an effective residential tax rate of:

$$t = \frac{(5184 - 1846 - 1470)0.301}{82175}$$
$$= 0.68 \text{ percent}$$

This calculated effective tax rate is reasonably close to the actual residential tax rate given in Table 13.3 of 0.85 percent. Using Equation (13.1), we calculate an effective residential tax rate for Concord of 0.92 percent. While Concord's household income is 140 percent that of Boston, Concord's per household expenditures are only 7.7 percent higher than Boston's. However, Concord's effective residential tax rate is 35 percent higher than Boston's.

How does Boston spend close to what Concord spends per household given its lower income and effective residential property tax rate? The answer lies in the large amount of state aid that Boston receives relative to Concord, Boston's large nonresidential property tax base, and the city's nonproperty tax sources of revenue. To illustrate the point, let's assume that the level of state aid, the nonresidential property tax base, and nonproperty tax revenues were equal across towns. For example, assume that Boston continued to spend \$5,184 per household but now had the same state aid, nonproperty tax revenues, and residential portion of its property tax base as Concord. Assuming that Boston's average house value would be the same, the city's effective residential property tax rate would skyrocket to 3.8 percent, or 4.1 times Concord's effective tax rate.

[5] For these calculations, we are assuming that average assessed value is equal to average market value.

The balance that seems to exist between cities and towns within a metropolitan area raises several questions about how resources are distributed across these communities. Why is there so much variation in town income? Why is it that nonresidential real estate tends to be located in communities that have lower incomes and thus are most pressed for tax resources? Why is it important for tax rates not to differ widely across communities? To answer these, we need first to consider the motives and incentives that households face when selecting a community to live in. It is equally important to examine the fiscal incentives that towns face when regulating their residential development. Finally, we look at the fiscal and environmental benefits to towns of nonresidential development.

HOUSEHOLD LOCATION AND COMMUNITY STRATIFICATION

The observation that communities vary widely in terms of average household income can be at least partially explained by extending the models developed in Chapter 3. In that chapter, we determined how, in equilibrium, the price of housing would vary across locations with commuting costs so as to leave households indifferent between living at each location. We also determined that if there were different types of households, those whose time was more valuable would be willing to pay relatively more for housing nearer to work. This is a location pattern that represents a competitive market equilibrium. As we discussed in Chapter 3, this approach assumes that there is perfect mobility, so that a form of price arbitrage exists across locations. In this chapter, we adapt this model to consider how households select a home from among a continuum of towns. In effect, "town" replaces a location's commuting cost as the determinant of price. We then ask what different households are willing to pay for a house (assumed uniform) in towns with different levels of expenditures and tax rates. Following Tiebout (1956), we seek to ascertain how, in equilibrium, a competitive housing market will allocate households to towns. To be complete, we also must incorporate the fact that the prices households are willing to pay for houses ultimately will determine town property values and hence tax rates.

In Figure 13.2, the horizontal axis depicts the level of town expenditure on public services, G. Towns to the left spend less, while towns to the right spend more. Consider for the moment three types of households, subscripted $j = L,M,H$, for low-, middle-, and high-income levels. Each line in Figure 13.2 depicts the price of housing, P_j, that would make that type of household equally well off across towns with different public expenditure levels. There are two considerations that go into determining these house price indifference lines. First, there is the willingness-to-pay for public services. This locational rent is the product of the annual dollar value that the household places on a unit of town public services (R_j) multiplied by the level of service (G). Second, there is the cost that a household faces for these public services, which is measured by the taxes that the household will pay if it lives in the town. If the town tax rate is t, then the household's annual tax payment will be tP_j. With a discount or interest rate of i, the house price that makes

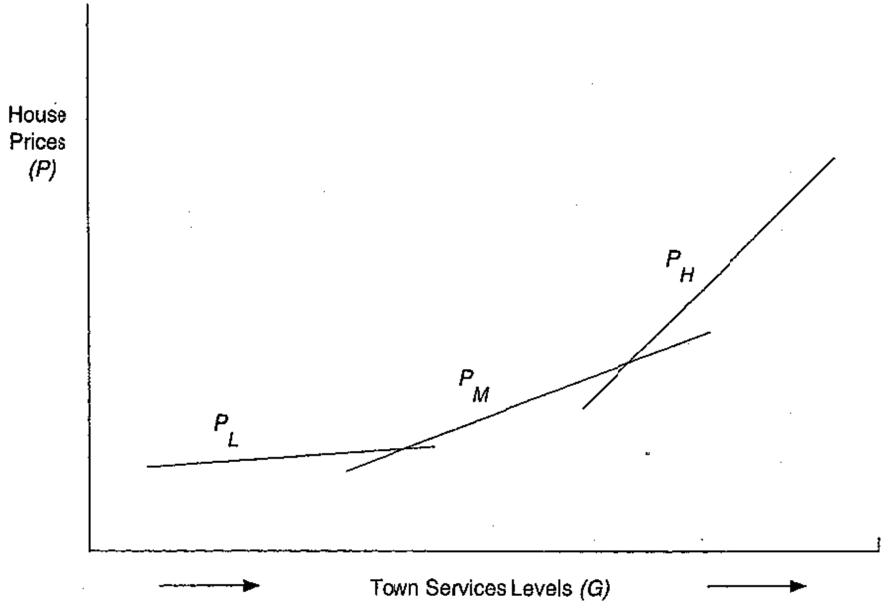

FIGURE 13.2 House prices and town service levels.

each household indifferent should be the present value of the household's valuation of services minus tax payments:

$$P_j = \frac{R_j G - tP_j}{i}, \; j = L, M, H, \text{ or, on solving for } P_j: \qquad (13.2)$$
$$= \frac{R_j G}{(i + t)}$$

As long as households with a greater income have a greater willingness-to-pay for public services (i.e., $R_H > R_M > R_L$), then the P_j lines will be as depicted in Figure 13.2.[6] At the right-hand extreme, high-income households will outbid those with moderate or lower incomes for the right to occupy housing in towns with high levels of public services. This conclusion holds even when such households consider that by bidding more, their taxes will also be (relatively) greater. Moving from right to left, middle-income households will likewise outbid lower-income households for units in middle-service towns, and so on through any number of household-income categories. If towns with higher levels of public services are those with greater incomes (which, all else being equal, they should be), then households will self-sort by income and maintain a pattern of income stratification. The importance of this simple argument cannot be overemphasized. In a competitive

[6]If the annual willingness-to-pay for services is ordered by income, then the derivative of P_j with respect to G, $R_j/(i + t)$, will also be ordered by income: P_H rises with greater services G, more than P_M does, and so on. This ensures that the P_j are as depicted in Figure 13.2 and that stratification by income prevails.

housing market, income stratification across towns by public service levels would seem to be a natural outcome of market locational forces.[7]

In the real world, the housing market that exists across towns is not uniform, as assumed in this simple model. Larger homes, for example, which are often preferred by high-income households, tend to be located in wealthier, high-service towns. This only increases the tendency for wealthier households to segregate into their own communities, outbidding lower-income households for units in these communities. If the housing in each town has been built primarily for the existing residents, then those residents are more likely be the highest bidders for such units. Self-sorting by income is only reinforced in this situation.

The picture changes quite dramatically when we consider the development of new land: how do the offered prices made for land by different income groups vary across towns with different service levels? This is crucial for determining how a town's housing pattern changes over time. A wealthy community will be able to maintain its character only if wealthy residents are willing to pay more (per acre) than other moderate- or lower-income households for the right to develop raw land in the town. We again return to Chapters 4 and 5 to examine how household bids for unimproved land are determined. We continue to assume that houses are identical except for lot size and denote C as the capital cost of constructing this identical unit. If q_j represents the size of lot demanded by households of each income level, then we can define an equation for that price of raw land, p_j, that makes each type of household indifferent across towns with different service levels. Land prices again will be a residual value per acre from house prices minus construction costs.

$$p_j = \frac{P_j - C}{q_j} = \frac{R_j G}{q_j(i + t)} - \frac{C}{q_j}, \; j = L, M, H \qquad (13.3)$$

The pattern of raw land prices across towns, p_j, need not resemble at all the pattern shown in Figure 13.2. It will certainly be true that the lot sizes demanded by households are ordered with income ($q_L < q_M < q_H$), as will be their values for government services. In this case, it is not clear that high-income households will be willing to offer more per acre for raw land in their town than will households with less income. If low-income households are willing to live on much smaller lots than wealthier residents, then they may be

[7]The reader may notice that in this model of income stratification, the tax rate in each town is assumed to be independent of the town's level of expenditure and its property values. It ignores how tax rates will eventually depend on expenditures and house prices. This simultaneous relationship is easily incorporated into the model, with no change in the model's conclusions. With no federal or state aid or nonresidential property, the town's budget relationship between services (measured in dollars), house prices, and the tax rate is simply:

$$t = \frac{G}{P}$$

Incorporating this into Equation (13.2), we get the revised house price equations:

$$P_j = \frac{G(R_j - 1)}{i}, \; j = L, M, H$$

As long as the annual willingness to pay for services, R_j, is ordered by income, then the house price lines above will still resemble those in Figure 13.2.

willing to pay more *per acre* than their wealthier competitors. The price *per acre* for raw land depends not only on the value of the community services obtained by living there, but also is based on how many households per acre are experiencing these advantages.[8]

If lower-income households are able to offer more for raw land in higher-income communities, then these communities will not remain high-income for very long. In effect, income stratification across towns might not be sustainable by a private market. Yet, there are strong incentives that encourage middle- and higher-income communities to try and maintain themselves, even when the market would not sustain them. Given these incentives, communities have developed a system of regulating land use through *zoning*. Zoning laws allow towns to set minimum lot sizes that are equal to or often greater than those that already exist in the community. This effectively prevents lower-income residents from consuming the smaller quantities of land which, in turn, would allow them to compete with higher-income residents for town land at offered prices. By requiring equal or minimum land consumption, towns are able to ensure that offered land prices resemble the house prices in Figure 13.2. Thus, income stratification across towns is maintained. But why are towns so interested in maintaining such stratification?

ZONING AND THE FISCAL IMPACTS OF RESIDENTIAL DEVELOPMENT

The determination of a town's property tax rate through Equation (13.1) provides some powerful incentives for communities to regulate the development of new land. Under current state enabling legislation, zoning laws permit towns to control several aspects of development, primarily density and land use. Since local governments are generally directly elected, the presumption is that such land-use regulations will be designed to achieve the maximum fiscal benefit or fiscal surplus for a town's current residents.

As a city or town grows, development inevitably requires the expansion of services, particularly in the long run. Studies suggest that police, fire, and school services often exhibit roughly constant returns to scale. Thus, as long as the character of a community does not change, long-term growth tends to require a proportional expansion of municipal budgets.[9] New development, however, can often affect the character of a community and always adds to the community's base of taxable property. In Equation (13.1), new development can alter a town's tax rate in three possible ways. It can change the level of public expenditure per household, G, the average value of houses, P, or the percentage of residential property, s. If the net effect of these changes leads to a lower tax rate, then existing residents will benefit from the development's fiscal surplus. A higher tax rate means that residents must pay more for the same services, since the new development creates a fiscal deficit.

The citizens of a community are the primary users of public services. The quality of services that residents receive depends on how much is spent per household (G) relative to how much "need" there is per household. "Need" is a tricky concept and is best illustrated with some examples. Two communities will have very different school systems if they spend the same per household but the average number of children per household (i.e., need) is quite different. Likewise, two communities that spend the same amount on police and fire services per household can have quite different risks depending on their density levels, building types, or underlying social problems (i.e., need). If a new residential development in a community adds an average number of new school children per housing unit and requires average per unit servicing by the police and fire departments, then in Equation (13.1), G does not have to change to keep service quality at existing levels. On the other hand, if community growth occurs with developments that are safer than average or that contain households with fewer children, average town expenditure on services can fall, while service quality remains intact.[10] Thus, the character of new residential development can alter the public expenditure that existing inhabitants must pay to maintain their current level of services.

On the revenue side, the character of new residential development also alters a town's tax rate by directly affecting the average value of houses (P). New development that adds houses which are above average in value causes the tax rate to fall, while the opposite will be true for development that is below average in value. The fiscal surplus of a development is its net effect on town finances considering both its expenditure and revenue impacts. A large mansion inhabited by a widow normally will generate a significant fiscal surplus, while a small apartment occupied by a family with four children should create a fiscal deficit.

As communities attempt to encourage residential development which generates a fiscal surplus, they have a limited number of policies at their disposal. The major one is the requirement that land be developed with a minimum lot size (MLS) or at a maximum density. Such zoning, if enforced, is almost always able to ensure that new development is of a minimum taxable value. While this may not directly prohibit someone from buying a 2-acre, $150,000 lot and then parking a $10,000 trailer on it, those with $160,000 to spend on housing generally prefer a better balance between land and capital. If binding, MLS zoning requires that a development use more land than it otherwise would and makes it more likely that the houses developed will be of higher value. This yields greater tax revenue (per new household).

[8] In order for income stratification to occur across communities, it must be the case that p_H increases more than p_M and p_L as one moves from left to right in Figure 13.2. This will occur only if the derivative of p_j with respect to town services, G, is ordered by income. This derivative is equal to $R_j/q_j(i+t)$, so stratification happens only if:

$$\frac{R_L}{q_L} < \frac{R_M}{q_M} < \frac{R_H}{q_H}$$

In effect, if high-income households value public services only slightly more than lower-income households, while the latter are willing to live on much smaller lots, then low-income households will outbid high-income residents for raw land in high-income (high-service) towns.

[9] Numerous statistical studies have examined the issue of returns to scale in municipal services. The most common approach analyzes how local public expenditure on a particular sector (e.g., schools) varies with the size of the population (e.g., children), holding other factors that might influence costs (e.g., salaries) or demand (e.g., town income) constant. There is some evidence of scale economies in fire protection, but little in education or law enforcement.

[10] Alternatively, the existing citizenry can keep expenditure the same and achieve a fiscal surplus through an improved level of service quality.

While MLS zoning encourages higher value houses, it can also reduce the value (per acre) of unimproved land. This effect can be seen by referring back to the model of development density used in Chapter 4. In the top frame of Figure 13.3, the downward-sloping willingness-to-pay schedule (P) gives the market price (per square foot of house floor area) as a function of that house's FAR (floor-to-land area ratio). It slopes downward because, all else equal, households prefer a larger yard area or lower residential density. The construction cost schedule for the house, C, may rise slightly, as more dense development tends to have higher costs (per square foot of floor area). The profit per square foot of house floor area is simply the difference between the P and C schedules. Beyond the FAR level labeled d, construction costs exceed house prices and development is no longer profitable. In the lower frame, the hump-shaped p schedule represents the profit per square foot of land area. Recalling Chapter 4, we derive this by multiplying the FAR level (F) by the profit per square foot of house floor area in the upper frame: $p = F[P - C]$. At the FAR level of F^*, development yields the maximum profit per square foot of land area, p^*. With no regulation, F^* would be the FAR or density level privately chosen by developers, and p^* would be the competitive market price for land.

With MLS zoning, the FAR allowed for development is constrained to be F^0. If the zoning limit is binding, then $F^0 < F^*$. At F^0, land profit (and, hence, land price) is only p^0, which must be less than p^* since p^* is the maximum land price achievable. Moving back to the upper frame, the greater required lot area yields a finished house price (per square foot) that is higher: $P^0 > P^*$.

What minimum lot size should a town choose in its zoning laws? If the community is interested only in the fiscal surplus of new development, the larger the required lot size, the greater will be the fiscal surplus (per developed housing unit). The problem with requiring FAR values that are near to the origin in Figure 13.3 is that they diminish the value of raw land. While the existing residents of the town obtain a fiscal surplus, the owners of unimproved land suffer losses in value. The legal tradition in the United States is that regulations in the public interest are allowed within reason until they deny property owners so much of their value as to be considered an effective seizing of property by the government. In practice, courts have upheld the right of towns to enact minimum lot sizes that are somewhat but not too much larger than the existing patterns of land use.[11]

The widespread use of minimum lot size zoning provides a compelling explanation for how income stratification across towns is maintained over time. Each town is effectively able to exclude residents whose income is significantly lower than the town's average from developing land. When forced to compete over a common-sized lot, lower-income households simply cannot match the offers of higher-income residents. When applied to towns at each level of income or public services, town stratification by income results.

[11] In some states like Massachusetts and New Jersey, state government has intervened to restrict the local application of MLS zoning. This is done to encourage the development of more affordable, higher-density housing.

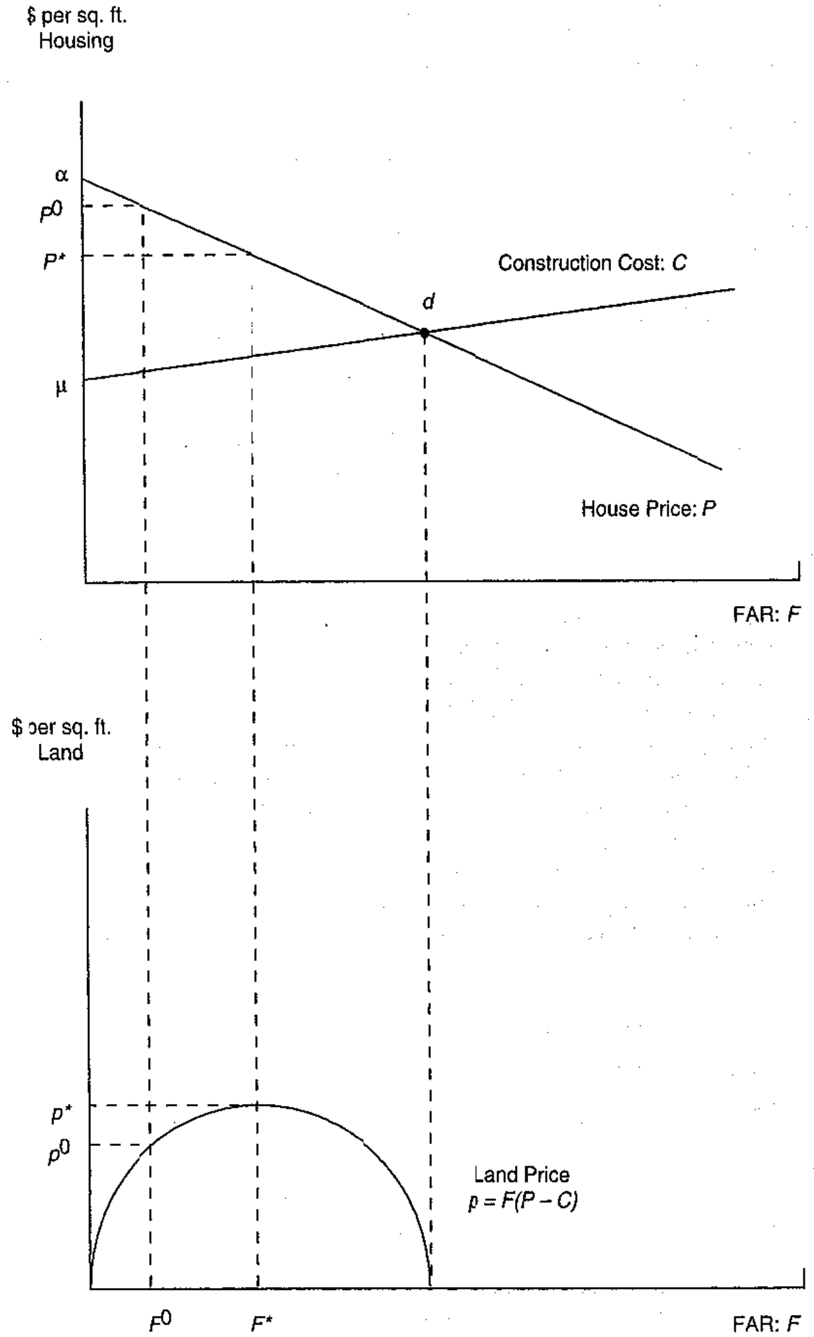

FIGURE 13.3 FAR model with zoning.

THE FISCAL COSTS AND BENEFITS OF NONRESIDENTIAL DEVELOPMENT

The zoning laws that enable communities to regulate the density of residential development also provide considerable power over the nonresidential use of land. Towns that wish to encourage commercial or industrial development designate large tracts of land for that use and then provide the access to infrastructure necessary to service such development. Communities that wish to remain residential need only refrain from these policies. There is little precedent in the law that prevents a town from excluding commercial development—even if such uses have existed historically in the community. As communities decide on their local zoning plans, a common issue of concern is the fiscal surplus or deficit that arises from nonresidential development.

The net fiscal impact of commercial or industrial uses involves three considerations. As with residential developments, a town must assess the likely impact of the commercial use on service demand as well as the likely tax payments to be made by the development. In the case of nonresidential uses, however, a third concern also arises frequently—the environment. The term *environment* here covers not only the obvious problems of noise and pollution, but also a range of more broadly defined issues such as congestion, changing town character, or the visual impact of the development. Let's examine these three fiscal impacts in turn.

It is now generally agreed that commercial and industrial uses generate very little demand on public services. The children of those workers employed in a town or of those shoppers using a store in a town do not attend the town's schools unless they are town residents. It is the children of *residents* who are educated in a town. In a similar manner, residential neighborhoods normally place greater demand on police and fire services than commercial and industrial areas. As for sanitation, most communities require that firms privately dispose of refuse—only residents are provided with free municipal collection. Businesses can use large quantities of water and produce commensurate waste, but in most cases they pay directly for these with a combination of hook-up and usage fees.

The primary public service demand that firms impose on towns is for infrastructure. The demand for transportation infrastructure by commercial or industrial development can create large service burdens that must be publicly financed. New industrial parks, shopping centers, and office buildings all act as traffic "attractors." When undertaken at large scale, such development can easily overwhelm existing roads or highways. The costs of adaptation can range from installing a small intersection to a major arterial or highway expansion. While commercial development can strain a town's transportation system, aggregate outlays on such infrastructure usually represents between only 5 and 10 percent of a community's budget. Furthermore, only a fraction of these outlays is generally attributable to commercial development. Thus, while a particular development may necessitate significant infrastructure upgrades, on average, the cost of infrastructure that is attributable to such development is only a small share of most municipal budgets.

On the revenue side, some forms of nonresidential development provide lucrative tax resources. A 3-acre site with a 40-story office building can easily be appraised at

$200 million and generate $4 million annually in tax revenue. The same size site with a warehouse in the suburbs, on the other hand, might yield only $50,000 in taxes. In most cases, the present discounted value of such revenue streams significantly exceeds the infrastructure or other tax-based service costs that such developments impose.[12]

The final consideration for towns in their zoning deliberation concerns the broadly defined environmental costs of nonresidential development. Does the development add to congestion, pollution, or noise? Does it create a visual blight on an otherwise peaceful country landscape? How important are such considerations to the town's residents? What dollar value do the residents of the community place on the loss of such amenities, if they are to be sacrificed? A town must compare the value of such losses against the net fiscal surplus of the development. This involves the following comparison of development impacts:

$$\text{Development Environmental Loss} \times \text{Community Value of Environment} + \text{Development Service Usage} > \text{or} < \text{Development Tax Revenue} \quad (13.4)$$

It is important to realize that the outcome of the comparison in Equation (13.4) will depend not just on the specific features of the development being proposed, but also on the community making the decision. It is the community's current tax rate that helps to determine the likely revenue from the development and the community's residents that value any environmental impacts. Two communities can evaluate the same development quite differently.

In Equation (13.4) it is clear that communities with low tax rates and where residents value the environment most highly will be those least likely to approve of nonresidential development. Conversely, those towns that place a lower dollar value on the environment and who have higher tax rates will often find the tradeoff worthwhile. The notion of valuing the environment is an economic concept exactly like that discussed in Chapter 4 for valuing a feature or an attribute of a house. It depends partly on an individual's own utility for or from the environment, but it also depends on an individual's income. All else equal, households with greater income are willing or able to sacrifice more of that income to purchase something.

The discussion in the previous section illustrated how wealthier towns tend to have at least modestly lower tax rates than poorer communities. When combined with the observation that wealthier communities may also be willing to sacrifice more to preserve the environment, the implication of Equation (13.4) is that such towns will opt to exclude commercial development. On the other hand, poorer communities with higher tax rates and lower environmental valuation will tend to embrace commercial development. Since central cities frequently have long histories of industrial development, it is perhaps not fair to compare their industrial land-use policies with those of newer suburbs. To avoid

[12]As with other examples in this book, a real discount rate, r, should be used in this present value calculation. Let i be the nominal interest rate, and g the likely rate of growth of the town's expenditure per household. The latter should at least equal the CPI inflation rate and reflects the speed with which property tax collections will rise. Adding in the rate of physical decay or obsolescence in the structure, δ, the appropriate discount rate would be $r = i - g + \delta$.

this bias, Table 13.4 examines some of those suburban towns within the Boston metropolitan area that adjoin Route 128, the region's major circumferential highway. Such communities have roughly comparable transportation access, which Chapters 5 and 6 showed was of importance for commercial and industrial firms. The table shows town income and the number of employees in the town relative to the town's resident population. Also depicted is the percent of the appraised property tax base that comes from nonresidential uses. There is at least a partial inverse relationship between town income and commercial-industrial activity. With the exception of Burlington (a more affluent community with a high ratio of jobs to population), the suburbs with higher incomes tend to have around 20 percent or less of their tax base in the nonresidential category (Concord, Lexington, Needham, Reading). These towns also have job-household ratios that are between 0.8 and 2.4. The three lower-income towns (Braintree, Peabody, Waltham) have nonresidential tax base shares that range from 27 to 39 percent, and the job-population ratios for Braintree and Waltham are between 2.5 and 3.0.

Table 13.4 shows that there can be systematic patterns in the way communities respond to the choice of whether to allow or encourage nonresidential development. It is interesting that there is also a relationship across industries or types of development between the magnitude of taxes paid and the general severity of environmental impacts created. With property as the tax base, the local taxes received from a development are determined largely by the value of the capital improvements placed on the land. A petrochemical complex, steel factory, or electric power plant can involve billions of dollars of capital investment. If appraised at such values, these facilities can generate tens of millions of dollars annually in tax revenue. This is equivalent to the entire budget of many smaller-sized suburbs. At the same time, such uses often expose a community to environmental risks and costs.

At the other extreme, suburban office parks and light industrial facilities normally pose little environmental risks and, if properly designed and located, create few environmental costs. The value of capital improvements from such development, however, is normally only a fraction of that from heavier industries. Thus, commercial uses tend to

TABLE 13.4 Distribution of Employment in Boston-Area Communities, 1989

	Braintree	Burlington	Concord	Lexington	Needham	Peabody	Reading	Waltham
Households	11,978	8,054	4,764	10,515	10,405	17,556	7,932	20,728
Employment	29,510	33,103	11,643	18,527	18,449	21,692	6,060	63,087
Jobs/HH	2.5	4.1	2.4	1.8	1.8	1.2	0.8	3.0
1989 Median HH income	47,151	58,975	73,695	71,030	63,618	41,950	55,635	40,595
Nonresidential tax base ($ billion)*	0.7	1.1	0.3	0.8	0.7	0.9	0.2	1.9
Nonresidential share of total tax base	31.70%	45.90%	15.10%	21.20%	21.80%	26.80%	10.80%	39.30%
Nonresidential tax rate	17%	1.57%	1.08%	1.59%	1.22%	1.38%	1.18%	1.46%

HH, Household.
*Nonresidential properties include assessed commercial and industrial values.
Source: Employment, Massachusetts Department of Employment & Training; Households, income, 1990 U.S. Bureau of the Census; Tax information, Massachusetts Department of Revenue.

provide tax revenues that are commensurate with their lower social costs. But what of a light industrial facility that processes hazardous waste and generates little tax revenue? Even more problematic, consider a public prison that yields no tax revenue at all. At the other end of the scale, a complicated telephone switching center may contain hundreds of millions of dollars worth of equipment, generate lucrative tax revenue, and be virtually invisible to the community. None of these uses is likely to generate tax revenue that in any way match the social costs of these uses.

Many economists have argued that the siting of major commercial or institutional uses should be determined according to a process of bargaining or bidding. Consider the hazardous waste facility mentioned above. The proposal is to require such a facility to pay communities whatever amount is necessary until some community accepts the compensation payment in exchange for allowing the facility to locate there. The community willing to accept the facility with the least compensation payment would presumably be that town where environmental issues were of less concern and where the facility would not create much inherent risk (because of topographical or locational factors). A facility that was viewed as an economic blessing might *receive* payments from communities and thereby wind up in the community that offered to *make* the largest payment. In either case, the siting decision would seem to occur at an efficient location where its harm is minimized or its advantages maximized.[13]

The argument to allow facilities and towns to bargain over the rights of firms to locate there has considerable merit. The payment of property taxes is based on the capital improvements that the facility requires, not the true economic costs or benefits that the development imposes on its host community. At the same time, it may be overly simple to assume that the impacts of a large facility are limited only to its host town. A hazardous waste plant may well impose risks on adjoining communities. These communities would not receive any tax revenue from the facility under current law, nor any compensation payments under the bargaining proposal above. The bargaining proposal works only if the boundaries of towns happen to reflect the boundaries of the impacts of a facility siting.

[13]An alternative to requiring facilities to pay towns for the right to locate would be to require towns to pay for the right to exclude facilities. The argument, according to Ronald Coase, is that the facility winds up at the same location with either scheme. To see this, consider the following table, in which a firm contemplates locating in either town A or town B. With no tax payment or compensation, the site minimizing production costs is in town B. Production plus environmental costs, however, are lowest in town A. If firms must pay for the right to locate, towns will require a compensation payment at least equal to the environmental costs imposed by the firm. The firm's total production plus compensation payment cost is minimized by locating in town A. Alternatively, towns could make a payment to the firm not to locate there, which would not exceed the environmental cost caused by the firm. If the firm locates in town A, it gets an *exclusionary* payment from town B of $20; if the firm locates in town B, it gets a $10 payment from town A. With exclusionary payments, locating in town A would also yield the lowest total production cost (net of exclusionary payment).

	Town A	Town B
Firm production costs:	$15	$10
Town environmental costs:	$10	$20
Prod. costs + "compensation":	$25	$30
Prod. costs – "exclusionary":	–$5	$0

CAPITALIZATION AND THE INCIDENCE OF LOCAL TAXES AND SERVICES

In the discussion about how households select communities, we adopted the approach of Chapter 3 in which we assume that house prices fully reflect both the value of services received and taxes paid at different locations. This approach assumes perfect capitalization, and it hinges on an assumed high degree of mobility by households. But what if households, particularly over a shorter time period, are not free to move for some reason? In Chapter 9, for example, we discussed how the high transaction costs of buying a home might limit mobility. Will house prices still reflect the value of services and taxes in this case?

In the sections above, the households or firms who reside in a town benefit from the community's public services. We have assumed that these town residents also pay for the services they enjoy through their property taxes. With rental real estate, the legal liability for property taxes rests with the owner and not its occupant. This is true whether the occupant is residential or nonresidential. For rental real estate, there arises the complicated question of whether the beneficiary of town services (the tenant) is actually paying the taxes for what is consumed. Imagine a town composed exclusively of tenants (residents) who live in buildings owned exclusively by absentee owners. Since only town residents have the power of voting, why wouldn't the tenants decide to greatly expand public services, at the expense of landlords? Under what conditions could (or would) the property owners be able to raise tenant rents by the amount of the taxes necessary to finance the expansion of services?

The economic issue of who pays for taxes on property when the owner and occupant are different is referred to as the question of *tax incidence*. Figure 13.4 explores tax incidence for the property market in a particular town. On the horizontal axis is the stock of space or housing in the community, and the level of rent is on the vertical axis. The downward-sloping demand curve, *D*, represents how many households (or firms) would choose to live in the community, depending on the town's average rent. The assumption earlier in this chapter, and in Chapters 3 and 4, was that mobility is sufficient so that this demand curve should be very elastic (nearly horizontal). With many communities to choose from, higher (lower) rents in one town should lead to sharp reductions (increases) in demanded space.

The supply schedule, *S*, represents how much space property owners would choose to develop at given rent levels. Alternatively, the schedule depicts the cost of supplying a given amount of space to the market. In Chapter 10, however, we demonstrated that space supply is inherently a dynamic concept. In the short run, the supply of space is largely fixed, as the number of units or buildings in any market declines only very slowly through depreciation or scrappage. Even the upward expansion of space takes time and is limited by the scarcity of land and its rising cost as the market grows. Thus, the notion in Figure 13.4 of a single supply schedule is perhaps overly simplistic. We might better think of a short-run supply schedule that is quite vertical (inelastic), as the stock is largely fixed and has trouble expanding quickly. In the longer run, the schedule is sometimes assumed to be more horizontal as the stock of space can expand or contract more elastically in response to changes in rent.

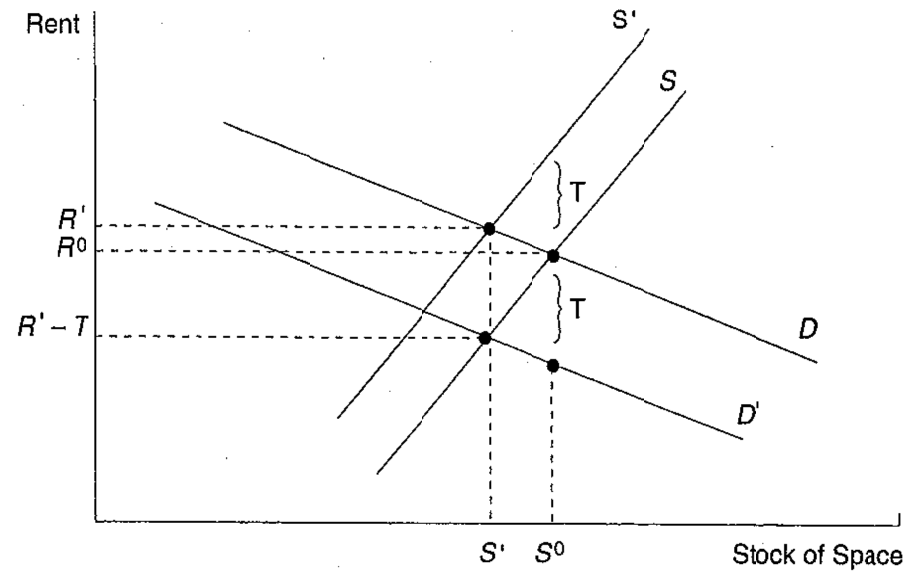

FIGURE 13.4 Incidence of real estate taxes on rental property.

Initially, the rental property market in this town can be assumed to be in long-run equilibrium at the point (S^0, R^0). This means that at the rent level R^0, and given the level of services in the town, S^0 amount of space is demanded. We now imagine that state aid suddenly contracts (for this town alone) by an amount equal to T per unit of space. The town can either raise property taxes by this amount (and leave services intact), or cut services by T and leave taxes unchanged.

If services remain fixed while taxes rise, then the cost of supplying space has risen by the amount of the tax increase, T. This shifts the space cost or supply schedule up to S', with a new equilibrium rent at R'. The rent actually received by the landlord moves from R^0 to $R' - T$, and this eventually causes a decline in the operational stock from S^0 to S'. The share of the tax increase that is accounted for by the upward movement of tenant rents, $(R' - R^0)/T$, is called the *incidence* of the tax that is paid for by tenants. Clearly this share will be small when demand is horizontal (elastic with respect to rents) while supply is inelastic (a vertical schedule). As a result, the tax is capitalized into the value of the property, which means that the owner pays the tax. This is the capitalization assumption that was made in earlier chapters. If supply is elastic, however, while demand is for some reason captive or inelastic, then capitalization will be incomplete and the owner can pass a significant portion of the tax to the tenant.

If taxes remain fixed while services fall (by an amount T), then the demand schedule must shift down by the tenant valuation of this loss in services. Assuming that services are valued at their cost, the demand schedule would shift down to D' by the amount T. In this case, however, the difference between gross and net rents has not been disturbed—both are lower by T. If supply is inelastic, however, and eventually contracts from S^0 to S', there is a partial recovery in rents. The ultimate equilibrium in either case is the same, as long as services are valued by tenants at their cost of provision.

The reader might reasonably question why in the very long run the supply schedule for space is not close to horizontal (perfectly elastic). Why would any landlord supply space at a lower rent than was obtainable elsewhere in other towns? Were this to be the case, the full long-run incidence of the tax would always be carried by tenants. The complete answer to the issue of long-run tax incidence involves a more complicated model that separately considers the capital and land inputs that are necessary to provide housing or commercial space. In the long run, the supply of capital is highly mobile (between towns), whereas the supply of land is almost, by definition, fixed. Thus, if higher taxes cause landlords to abandon (and not replace) their structures, the demand for land shifts and its price falls. With lower land prices, development can then profitably resume.

With this theory, the long-run incidence of the tax increase always lies with landowners, not tenants. This is consistent with the view in Chapter 3 that eventually the land market fully capitalizes locational advantage. With a highly elastic long-run demand curve, the rent for space in one community will always reflect only the valuation (by tenants) of the services they receive. It is the price of land that will eventually absorb the cost of providing those services. If the cost of provision and the value of the benefits received move with each other dollar for dollar, then tenants will pay for the services they receive. It is when cost and value move differently that landlords may receive a windfall gain or loss. If two towns have identical public services, but one receives significantly more state aid or has a sizeable base of taxable commercial property, then its lower taxes will be reflected in higher land prices. Rents, however, will be the same in the two communities. If the "fortunate" town opts to spend more on services, leaving taxes the same, then rents will be higher, and this (rather than lower taxes) will generate greater land prices.

SUBURBAN COMPETITION, TAXES, AND THE PLIGHT OF CENTRAL CITIES

The information in Table 13.3 illustrates a common problem that has plagued all major metropolitan areas of the U.S. during the last half-century. Sparked by a decaying housing stock and zoning constraints in adjoining suburban communities, lower-income households have been able to afford housing only in inner cities. Once concentrated there, the "need" for public services becomes especially acute. Greater crime, the need for special education programs, and social services all place great strains on central-city budgets. City governments find themselves "between a rock and a hard place." To meet these services in a manner that is even partially comparable to that of the suburbs will require greater expense, and, hence, higher taxes. These taxes, in turn, will lower property values. The alternative is to leave taxes at competitive levels, but allow services to decay. This will also lead to lower property values and also possibly exacerbate the city's difficult social problems.

The dilemma creates a vicious cycle. A declining real estate market only serves to attract or maintain the city's lower-income population and its resulting demands on public services. There is no way to strengthen the real estate market and attract a greater mix of households into an inner city without simultaneously lowering taxes, improving services, and solving city social problems. Two sources of assistance to cities have been discussed in this chapter: grants from state or federal governments and the tax subsidies that come from the city's commercial land uses.

Chapters 5 and 6 suggest that central cities may be in danger of losing the commercial land uses that for so long have helped finance their necessary services. Many manufacturers have already left for cheaper suburban land, while retailers are moving out to be nearer the greater purchasing power of suburban households. In the last decade, there has been a tendency for service firms, once centralized in inner city CBDs, also to cluster in newer suburban subcenters. Unless this loss of employment and the resulting erosion of inner-city tax bases is reversed, America's central cities increasingly will be forced to depend on higher-level governments for their financial viability.

SUMMARY

In this chapter, we examined several important connections between the real estate market of a metropolitan area and local governments. Not only do the taxes and services of local governments influence the locational choice of firms and households, but the impact this has on property prices greatly affects the financial resources of local governments.

- In the U.S., local governments provide most of our direct public services. In addition to assistance from states and the federal government, the property tax is the major source of local revenue to provide these services.
- Within metropolitan areas, suburbs, with their wealthier residents, tend to have similar total expenditure budgets as inner cities with poorer households. Greater assistance from state and federal governments, together with an extensive nonresidential property tax base, are additional resources that enable inner cities to provide similar service expenditures with property tax rates comparable to the suburbs. The greater density and social problems of inner cities dictate a different mix of services and impact the quality of services obtained from comparable spending.
- Local governments face strong financial incentives to regulate new development so that the combination of a project's environmental impact and its use of public services is more than offset by its tax payments. Zoning laws are used by local governments to achieve this kind of fiscal surplus.
- In a market equilibrium, the housing market tends to sort households by income level. Thus, the existence of communities with different income levels is a natural outcome of a competitive housing market. In the long run, however, this pattern often would not be sustainable without imposing zoning restrictions. Minimum lot size zoning, for examine, can block the development of higher density housing that would enable lower income households to live in more affluent towns.

REFERENCES AND ADDITIONAL READINGS

General

MUSGRAVE, PEGGY B., AND RICHARD A. MUSGRAVE. *Public Finance in Theory and Practice.* New York: McGraw-Hill, 1989.

Governments and Public Services

LADD, HELEN, AND FRED DOOLITTLE. "Who Should Take Care of Poor People?" *National Tax Journal* 35, 3 (September 1982): 323–336.

OATES, WALLACE E. *Fiscal Federalism.* New York: Harcourt, Brace, Jovanovich, 1972.

U.S. DEPARTMENT OF COMMERCE. *Census of Governments.* Washington, DC: U.S. Department of Commerce, 1982, 1987, 1992.

Local Governments and Property Taxes in Metropolitan Areas

FELDSTEIN, MARTIN S. "Wealth Neutrality and Local Choice in Public Education." *American Economic Review* 65, 1 (March, 1975): 75–89.

GRAMLICH, EDWARD M. "Intergovernmental Grants: A Review of the Empirical Evidence." pp. 219–239, in *Political Economy of Fiscal Federalism,* ed. Wallace E. Oates. Lexington: 1977.

GRAMLICH, EDWARD M., AND DANIEL L. RUBINFELD. "Micro Estimates of Public Spending Demand Functions and Tests of the Tiebout and Median-Voter Hypotheses." *Journal of Political Economy* 90, 3 (1982): 536–560.

LADD, HELEN F. "Local Education Expenditures, Fiscal Capacity, and the Composition of the Property Tax Base." *National Tax Journal* 28, 2 (June 1975): 145–158.

MUELLER, DENNIS C. "Public Choice: A Survey." *Journal of Economic Literature* 14, 2 (June 1976): 395–433.

Household Location and Community Stratification

TIEBOUT, C. "A Pure Theory of Local Public Expenditure." *Journal of Political Economy* 64 (October 1956): 416–424.

WHEATON, WILLIAM C. "Land Capitalization, Tiebout Mobility and the Role of Zoning Regulations." *Journal of Urban Economics,* 34 2 (1993).

YINGER, JOHN. "Capitalization and the Theory of Local Public Finance." *Journal of Political Economy* 90, 5 (1982): 917–943.

Zoning and the Fiscal Impacts of Residential Development

HANUSHEK, ERIC A. "The Economics of Schooling: Production and Efficiency in the Public Schools." *Journal of Economic Literature* 24, 3 (1986): 1141–1177.

KEMPER, P., AND J. QUIGLEY. *The Economics of Refuse Collection.* Cambridge, MA: Ballinger Press, 1976.

SCHWAB, ROBERT M., AND ERNEST M. ZAMPELLI. "Disentangling the Demand Function from the Production Function for Local Public Services: The Case of Public Safety." *Journal of Public Economics* 33, 2 (1987): 245-260.

WHITE, M. "Fiscal Zoning in Fragmented Metropolitan Areas," in *Fiscal Zoning and Land Use Controls,* ed. Mills and Oates. Lexington, MA: D. C. Heath, 1975.

The Fiscal Costs and Benefits of Nonresidential Development

COASE, RONALD H. "The Problem of Social Cost." *The Journal of Law and Economics* (October 1960).

FISCHEL, W. "Fiscal and Environmental Considerations in the Location of Firms," in *Fiscal Zoning and Land Use Controls,* ed. Mills and Oates. Lexington, MA: D. C. Heath, 1975.

Capitalization and the Incidence of Local Taxes and Services

DUSANSKY, RICHARD, MELVIN INGBER, AND NICHOLAS KARATJAS. "The Impact of Property Taxation on Housing Values and Rents." *Journal of Urban Economics* 10, 2 (1981): 240-255.

MIESZKOWSKI, PETER. "The Property Tax: An Excise Tax or a Profits Tax?" *Journal of Public Economics* 1 (1972): 73–96.

OATES, WALLACE E. "The Effects of Property Taxes and Local Public Spending on Property Values: An Empirical Study of Tax Capitalization and the Tiebout Hypothesis." *Journal of Political Economy* 77, 6 (1969): 957–971.

WHEATON, WILLIAM C. "The Incidence of Inter-Jurisdictional Differences in Commercial Property Taxes." *National Tax Journal* 37, 4 (1984): 515–527.

Suburban Competition, Taxes, and the Plight of Central Cities

BRADFORD, DAVID F., AND HARRY H. KELEJIAN. "An Econometric Model of the Flight to the Suburbs." *Journal of Political Economy* 81, 3 (1973): 566–589.

CHAPTER 14

Public Goods, Externalities, and Development Regulation

The quality of the urban environment often is more than simply the sum of the qualities of each individual parcel of land that makes up that environment. The quality of urban life is influenced by the overall aesthetics of the physical environment including the architectural compatibility of buildings in the community and the use and design of public spaces. Environmental conditions such as air and water quality and the treatment and disposal of waste are also important determinants of the livability of cities. Why do some communities provide a high quality of life while others do not?

In this chapter, we argue that quality of life depends on two important economic issues which must be acknowledged and appropriately managed. First, there exists a wide range of public goods which create collective enjoyment for all property owners. Being collectively enjoyed rather than individually consumed, these public goods are difficult to provide when individual property owners act independently. Thus, the individual decision making that characterizes private land markets does not work well for such goods. Institutional mechanisms are often necessary for owners to act cooperatively in determining what public goods ought to be provided and how they should be financed.

A second distinctive feature of real estate is that the independent actions of nearby property owners can have a dramatic impact on the value of a specific property. In other words, the real estate fortunes of one property owner are critically linked to the actions of others. This is a classic example of what economists call externalities—the action of one individual results in external costs or benefits to another individual. This interdependence between real estate sites can lead to fundamental coordination problems among owners:

all owners might be better off by acting in a certain way, but none has an individual incentive to do so. In an attempt to remedy this coordination problem, local governments have developed mechanisms for interfering with private property decisions. The most common of these is zoning laws. The land market, however, has also evolved at least some partial remedies for externalities. Privately written and enforced restrictive contracts or covenants often can coordinate the actions of individual property owners. Furthermore, when land is owned and developed simply at a larger scale, certain problems of coordination become better internalized or controlled.

In this chapter, we present a number of simple economic models that illustrate this class of problems and suggest a number of alternative solutions. Many of these remedies can involve governmental interference in the land market, suggesting that public planning has an important role to play in securing private interests. To illustrate the influence of public goods and external effects on the land market, we begin by illustrating some examples of how important these can be in determining property values.

PROPERTY VALUES AND PUBLIC AMENITIES

In Chapters 3 to 5, we demonstrated that market equilibrium and the indifference principle can help to uncover the value that consumers place on housing attributes, location, and commuting. In Chapter 4, we presented hedonic pricing models that provided estimates of the value placed on specific attributes of a house, including location. This same approach has been used often in recent years to obtain estimates of the value that consumers place on public goods or amenities. The fundamental principle is the same: if markets are in equilibrium, the value of otherwise identical housing or land in a location that offers more amenities should be priced just high enough to make buyers indifferent; lower prices in areas with poor amenities should just compensate buyers. Let's look at some specific examples.

Certain uses of land such as airports and heavy industrial plants are known to produce noise and air pollution as part of their normal operation. Since the early 1970s, economists have studied the impact that such facilities have on nearby property values (e.g., Mieszkowski and Saper 1978, Nelson 1978, and Smith 1978). In general, this research suggests that poor air quality can reduce residential property values by roughly 15 percent. Similarly, moderate noise under flight paths has been shown to reduce housing values by 5 to 10 percent, whereas more severe noise nearer to airports has a 20 to 30 percent negative impact. Grether and Mieszkowski (1980) found that residential property values increase with greater distance from disruptive land uses.

The siting of waste facilities often presents a problem because residents fear health risks and a potential negative impact on their property values. Kohlhase (1991) examined the impact on property values when the Environmental Protection Agency (EPA) places a toxic waste site on the Superfund list, meaning that the site becomes a federal priority for clean up efforts. Using data from Houston, she estimated that property values within a 6.2 mile radius of the site were decreased by as much as $3,310 for each mile closer to the site following the EPA announcement.

Publicly provided land uses, such as highways and open space, can also create external effects on nearby sites. Highways, for example, impact the land market in two ways. As part of a transportation network, highways increase property values by providing accessibility. At the same time, sites in close proximity to highways are exposed to greater noise and air pollution, which, according to Waddell, Berry, and Hoch (1993), can depress adjacent property values by 5 to 10 percent. Households want access to highways, but they don't want to be so close that they notice the noise and poor air quality when they are at home. The public preservation of land as open space likewise has two impacts. By reducing the supply of developable land, open space can raise surrounding land prices, as discussed in Chapter 3. Proximity to open space, however, also is a distinct public benefit to nearby private landowners. Frech and Lafferty (1984) found both of these effects in analyzing the impact of the California Coastal Commission on property values near the coast. The California Coastal Commission could veto any development within 1,000 yards of the high tide line as part of a policy to keep the land open for public use. Frech and Lafferty estimated that the Commission's control over development increased property values on the coast by 4 percent as a result of the reduced land supply and an additional 4 to 9 percent from increasing amenities.

On a more local level, there is growing evidence that the design of buildings and overall manner in which land is developed can significantly impact property value. There is clear evidence that households are willing to pay price premiums for houses of different architectural styles (Asabere, Hachey, and Grubaugh 1989). Vandell and Lane (1989) indicate that firms also seem willing to pay higher rents to occupy buildings with good architectural design as judged by professional panels. The design of one's building, however, is not only a private good to the occupant, but also a public good to the immediate environment. Are consumers willing to pay more to have *other* houses designed or laid out in a particular manner? Here, the results are new, but quite suggestive. Asabere (1990) found that a winding street/cul-de-sac layout in subdivisions, which creates visual diversity and green space, increases property prices by 25 percent over a simple grid layout. Several studies of house prices in historic districts where facades are strictly regulated also indicate that consumers value architectural compatibility and are willing to pay a price premium for it in their neighborhoods (Ford 1989).

Thus, there is substantial empirical evidence that the value of each particular site depends not just on its own intrinsic characteristics, but is also strongly influenced by the uses that occur on other nearby sites, the overall design of the neighborhood, and by the way streets, infrastructure, and open space are provided throughout the community.

PUBLIC AMENITIES AND FREE RIDERS

The problems associated with providing public amenities can be seen by considering the following simple example. Suppose a neighborhood has n individual property owners with identical lots that are already developed and one centrally located lot that is undeveloped. Each owner contemplates the development of the one remaining lot and decides that leaving the lot undeveloped as open space would increase the value of their property by \$$MV$ (the marginal value of having the open lot). In effect, MV represents the increase in house value to each of the n property owners that comes from having adjacent open space. The market value of the lot if it were sold for private development is \$$p$. To preserve the land for open space will require that it be purchased. We will assume that the price of the land exceeds its private value as open space to any one individual ($p > MV$), but that the aggregate or collective value of having the land remain open might justify its purchase ($p < nMV$). Two conclusions can be drawn from the facts in this example:

1. No individual owner will unilaterally purchase the lot for use as open space.
2. If the group were to agree to purchase the lot jointly, with individual owners contributing p/n, each owner would seek to abandon the group, allowing the others to purchase and split the cost (as long as $p > MV$).

Thus, there is a fundamental divergence between short-run individual interest and the longer-run collective good. This divergence results from three features of this simple example. First, we have assumed that the benefit of preserving the remaining lot is *nonexcludable*—none of the n property owners can be denied the advantage of the open space. Second, the benefit to each property owner of leaving the land open is *nonexhaustible*—it does not depend on how many owners share in the expense of acquisition. Finally, we have not introduced any contractual, legal, or institutional mechanism that *enforces participation* by all n property owners in the decision.

Consider the following illustration of this last point. Suppose all owners of the n lots were part of a majority-rule government that could propose only a "take it or leave it" vote with *all* owners required to pay p/n if the acquisition passed. Given the values of MV and p ($p > MV$) and the homogeneity of open space valuations, the vote would be unanimous in favor of acquisition. Of course, this would not work if owners could be individually excluded from the acquisition decision and cost. The nature of this public amenity problem is that each owner hopes to *free ride* on others. Without mandatory participation, it is difficult, if not impossible, to get the n members to participate in the common effort.

It is important to point out that this decision about open space would be made very differently if done prior to the sale of the n lots to individual owners. Consider the decision faced originally by a single owner-developer of $n + 1$ lots. If only n lots are developed and one left open, then the value of the n lots will increase by nMV, assuming the market would have valued open space the same then as the current owners do now; the proceeds lost from not developing the remaining lot are only p. Since $p < nMV$, the single developer-owner would clearly decide to keep the lot vacant. This example raises the broader question of exactly how much open space should be included in a development.

Consider a slightly more complicated example in which some agent has the power to require property owners to participate. We will again assume that there are n owners of existing developed 1-acre lots, who now face the possibility of acquiring any amount of adjacent land for open space. Let's further assume that the owners are divided into two groups with quite different views about open space. Suppose $n_1 < n$ owners value open space highly and feel that each acre of open space acquired (A) is worth $MV_1(A)$ to them.

With the law of diminishing marginal utility, we will expect that MV_1 declines with more A, which means that the value of additional open space is less as the amount of open space in the development increases. The remaining owners $n_2 = n - n_1$ have open space valuations of $MV_2(A)$, which are lower than those of the first group. Thus, $MV_2(A) < MV_1(A)$ for all ranges of A. Finally, the market value of a lot for private development, again, is p, which is the price that will have to be paid to acquire the adjacent land.

What is in the collective interest of all n owners? For each additional acre of open space purchased, the total benefit to all is simply the aggregate benefit to both types of owners: $n_1MV_1(A) + n_2MV_2(A)$. This total marginal benefit declines as more A is acquired. The collective interest of all is highest when the total marginal benefit of additional acreage just equals the marginal cost of acquisition (p).[1] Thus, the net value to all owners will be maximized with a purchase of open space acreage (A^*) such that:

$$n_1MV_1(A^*)+n_2MV_2(A^*)=p, \text{ or, } \left(\frac{n_1}{n}\right)MV_1(A^*)+\left(\frac{n_2}{n}\right)MV_2(A^*)=\frac{p}{n} \quad (14.1)$$

The second part of Equation (14.1) defines the optimal purchase of open space as occurring where the *weighted average* valuation of additional open space by the two types of households equals the per-household share of the cost of acquiring additional land.

While this ideal solution to the acquisition of open space is easy to define in principle, it is difficult to imagine a process by which it can be realized. To begin with, each individual property owner (of either type) would still like to free-ride or opt out of the group acquisition, thereby saving the cost (p/n) while still reaping the benefits of the nonexcludable, nonexhaustible open space. If we could enforce participation, we would have the additional problem of *social choice*—getting the two groups to agree on what action to take.

Suppose that all of the n property owners are required to participate in a collective decision and that all will be required to pay an even share of the amount of open space the group decides to acquire. In effect, we are requiring complete participation by all in a social choice process with the power to enforce the group's decision. Within this institutional mechanism, we ask what level of open space acquisition would each type of household like the group as a whole to undertake. With a unit acquisition cost to each household of (p/n), the level of open space that each household would prefer that the group select (A_1^*, A_2^*) will be such that its own marginal benefit equals its cost share:

$$MV_1(A_1^*) = \frac{p}{n}, \quad MV_2(A_2^*) = \frac{p}{n} \quad (14.2)$$

In Figure 14.1, we depict the three solutions (A^*, A_1^*, A_2^*). Each occurs at a level of open space (A) at which the marginal benefit schedule intersects the horizontal unit cost line. Of course, since the first household type values open space more than the second, its marginal benefit curve, $MV_1(A)$, is to the right of the second household type's

[1]Rational owners would seek to acquire that level of open space at which the total value to them net of acquisition cost is maximized. As discussed in Chapter 4, this occurs at that level of open space at which marginal benefits (the MV functions) equal marginal costs (p). At open space levels less than this, marginal benefits exceed costs, so further purchases are warranted. With more open space, the net gains are negative.

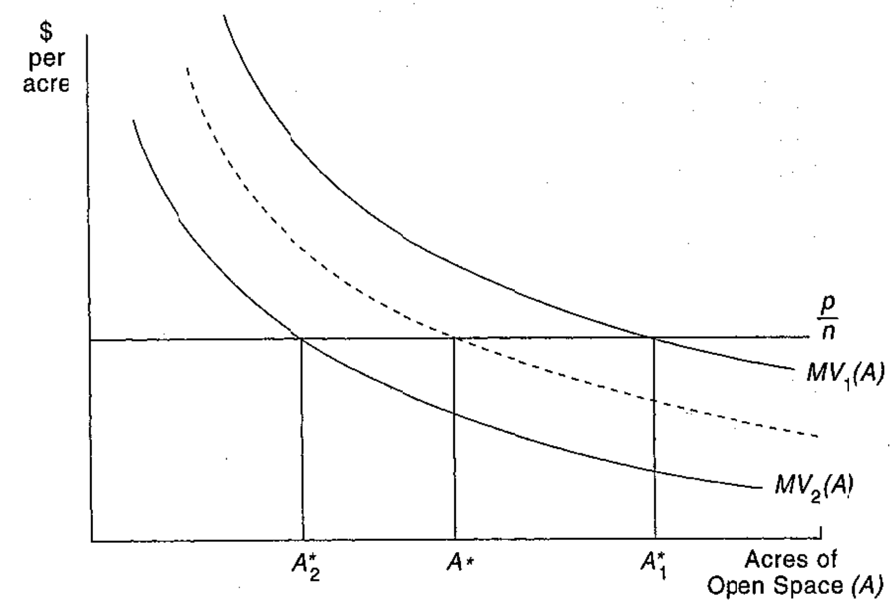

FIGURE 14.1 Social choice with two household types.
The dashed line represents the weighted average of open space valuations $MV_1(A)$ and $MV_2(A)$.

curve, $MV_2(A)$. Hence, the preferred outcome for group 1, A_1^*, is greater than A_2^*. We have defined the optimal amount of open space as A^*, a point at which the weighted average of the benefit curves for the two household types (the dashed line in Figure 14.1) intersects the cost line (p/n).

It is important to note that if the collectively optimal level of open space is somehow selected, each group of households would still prefer a different outcome: the 2s would want less and the 1s more. Such conflict is inherent in collective choice problems in which the "best" solution may not be preferred by anyone.

While having determined what each type of household would prefer to select, we still need to specify a process for arriving at a group decision. The most common mechanism used in such collective choice problems is voting. If a majority rule vote is adopted, then the group would select either A_1^* or A_2^*, depending on which group has the most members (whether n_1 is greater or less than n_2). Thus, direct voting normally will not yield the open space policy that is in the best collective interest of the group.[2]

[2]Other social choice processes have often been suggested by economists. For example, consider a binding mechanism that requires unanimous support but in which votes can be bought and sold. To see how this would work, imagine the group as a whole considering the point A^* as opposed to A_2^*. Only the 1s favor this change (since their preferred level is greater than A_2^*). The n_1 households of type 1 could all contribute funds equal to their incremental gain and then give $1/n_2$ of this amount to each of the type 2 households. Since A^* is defined as the point at which the weighted average gains are maximized, each group would be better off. In effect, the 1s have bought the support of the 2s.

It is important to point out that in most cases, any level of open space acquisition between A_1^* and A_2^* makes the property owners as a whole better off than having no open space at all. In this sense, even an imperfectly operating group decision process is likely to be better than simply leaving the n property owners alone to try and bargain in private. With the free-rider problem present, voluntary cooperation becomes more difficult, and private decision making will often fail to provide any level of public amenities, particularly when the number of property owners is large.

We should point out again that the problem of open space acquisition would, in principle, be far easier to handle by an original single developer-owner of the lots. In that case, we might assume that the overall market values adjacent open space at the average valuation of the two groups that will eventually purchase the lots, and that the owner knows this market value. As long as the aggregate value of the n lots owned by the developer can be increased by more than the cost of acquiring and preserving adjacent land, then the acquisition is in the developer's self-interest. This criteria is simply the same as Equation (14.1) for the optimal amount of open space (A^*).[3]

The kinds of public goods and amenities that we have illustrated in this section involve the creation of collective value on third-party land. Open space or parks, waterways, or sidewalks all occur on land outside of each owner's individual parcel. Often, individual actions on one's own property enhance or diminish the value of others' property.

EXTERNAL EFFECTS ACROSS PROPERTIES

Property owners are often interested in the actions of other property owners in their communities because those actions may directly impact their own properties or quality of life. For example, a homeowner may be justifiably concerned when an adjacent property is abandoned or permitted to deteriorate, because it is aesthetically unappealing and this may negatively impact the value of his property.

In Chapter 4, we argued that the density of development is an important attribute of a property to its occupant, and that, all else equal, households desire to live in structures of lower densities. With a demonstrated willingness-to-pay for lower-density development, profitable land development involves trading this off against the ability to put more houses on each parcel of land. Particularly in large urban areas, density takes on an additional dimension and becomes an important quality of life variable. In this setting, the density of a particular parcel not only affects its own value, but impacts the value of surrounding properties as well. When a neighborhood is dense, light is reduced, wind tunnels are created, and there is little public space. Older historic neighborhoods with four-story townhouses, such as the Back Bay area of Boston discussed in Chapter 4 and found in other cities as well, are often preferred to neighborhoods with rows of apartment

[3]If the buyers are representative of the market, the market average increase in lot value for an acre of adjacent open space would be $(n_1/n)MV_1 + (n_2/n)MV_2$. The owner would acquire open space until the aggregate value of his developed land net of acquisition cost for the adjacent land was maximized. This would occur when n times the market valuation equals the acquisition cost, as in Equation (14.1).

towers. This preference has as much to do with the atmosphere of the neighborhood (e.g., its openness and human scale) as it does with the private value of townhouses as opposed to apartment buildings. Current residents of these historic areas often band together and fight politically to oppose newer, denser development in order to preserve the character of the neighborhood. Surely this indicates that the neighborhood density of surrounding buildings has an "external" impact on each individual property. It also means that individual decisions about density can have broader ramifications that the developer may ignore. The conflict between the individual and the neighborhood can be resolved either cooperatively or individually, with quite different results.

To illustrate the distinction between individual and cooperative choices, let's return to the example of the Back Bay area of Boston presented in Chapter 4. Remember that this area of Boston is dominated by four-story townhouses with an occasional 8-to-12 story building. We defined the price per square foot of housing as $P = \alpha - \beta F$, where α is the collective value of all other locational and housing attributes that can affect the price per square foot of a house and β is the marginal reduction in value that occurs when the FAR of the building, F, increases. At what FAR will an individual owner develop his site? As we showed in Chapter 4, the developer will choose the FAR that maximizes the residual profit per square foot of land, defined as $p = F(P - C)$, where C is per-square-foot construction cost. Our innovation to the model here is to incorporate the loss in value to each property that results from denser development of the neighborhood. Let's define f as the FAR of the neighborhood (as distinct from individual property FAR, F) and γ as the incremental loss in value to each house with increased neighborhood density. The developer of any individual property, then, takes the neighborhood density as given and chooses that FAR, F, that maximizes p:[4]

$$p = (\alpha - \beta F - C - \gamma f)F$$

$$F = \frac{(\alpha - C - \gamma f)}{2\beta} \qquad (14.3)$$

Over time, as all developers act this way, the neighborhood FAR will begin to evolve towards that chosen by each developer. As each parcel is developed at F, neighborhood density, f, moves towards F. Thus, in the long run, $f = F$. Using $f = F$, we can define the long-run equilibrium level of density that will prevail when all property owners act individually (F^m):

$$F^m = \frac{\alpha - C}{2\beta + \gamma} \qquad (14.4)$$

[4]The residual profit per square foot of land development (p) equals the floor area profit multiplied by the development's FAR: $p = (\alpha - \beta F - C - \gamma f)F$. We find the F that maximizes p by setting the derivative $\partial p/\partial F$ equal to 0:

$$\partial p/\partial F = \alpha - 2\beta F - C - \gamma f = 0$$

Solving for F, we obtain:

$$F = \frac{\alpha - C - \gamma f}{2\beta}$$

At this level of density, we can also determine the value of land (p^m).[5]

$$p^m = \beta\left(\frac{\alpha - C}{2\beta + \gamma}\right)^2 \quad (14.5)$$

In Chapter 4, we estimated α at $222 per square foot and β at 1.48. To continue this numerical example, we assume for simplicity that construction costs are constant at $120 per square foot. If we assume a value of γ of $4 per square foot (the incremental loss in value for each house in the neighborhood as the neighborhood FAR increases), $F^m = 14.7$, considerably higher than the four-story brownstones in the area. When neighborhood concern over density is lower (i.e., lower γ), the neighborhood FAR and land prices will be higher. For example, in our Chapter 4 exercise, γ equalled 0, F^m was 17.5, and p^m was $1,068. With $\gamma = 4$, $F^m = 14.7$ and $p^m = \$318$ per square foot.

It should be clear in the example so far that while each property owner considers the external cost imposed on him by other property owners, each property owner is ignoring his external impact on others. What happens if, rather than having individual owners make density choices independently, these choices are made collectively? Suppose owners cooperate to choose a common FAR for their entire neighborhood, considering all of the costs incurred. In this case, they jointly select the density they want *both* for the neighborhood *and* for their individual properties. Assume that landowners are constructing a new neighborhood on vacant land and must decide on the density of all development. With cooperation, the neighborhood FAR gets defined in advance as equal to that of each property, $f = F$. As a result, the collectively optimal FAR, F^*, is that which maximizes:[6]

[5]We can determine land value, p^m, by using Equation (14.3):

$$p^m = (\alpha - \beta F - C - \gamma f)F$$

Substituting $f = F$ yields:

$$p^m = (\alpha - C)F - (\beta + \gamma)F^2$$

Substituting $F = F^m = \frac{\alpha - C}{2\beta + \gamma}$:

$$p^m = \frac{(2\beta + \gamma)(\alpha - C)^2 - (\beta + \gamma)(\alpha - C)^2}{(2\beta + \gamma)^2}$$

$$= \beta\left(\frac{\alpha - C}{2\beta + \gamma}\right)^2 = \beta(F^m)^2$$

[6]Here, cooperation means that we begin with $F = f$, so we maximize:

$$p = (\alpha - \beta F - C - \gamma F)F$$

Following footnote 4: $\partial p/\partial F = \alpha - 2\beta F - C - 2\gamma F = 0$

$$F^* = \frac{\alpha - C}{2(\beta + \gamma)}$$

$$p = (\alpha - \beta F - C - \gamma F)F$$

$$F^* = \frac{\alpha - C}{2(\beta + \gamma)} < F^m \quad (14.6)$$

Clearly, cooperation leads to a lower density of development. Using the parameter values outlined above, $F^* = 9.3$. What is the impact on the value of land? Our expression for land value, p^*, becomes:[7]

$$p^* = \frac{(\alpha - C)^2}{4(\beta + \gamma)} > p^m \quad (14.7)$$

Cooperation winds up yielding a higher value for land. Again using the parameter values outlined above, p^* equals $475 per square foot, considerably higher than the p^m value of $318 when there was no cooperation. In the long run, cooperative behavior will yield higher property values and is in the best interest of all parties.

As a practical matter, how are cooperative decisions made? There may be neighborhood or community associations that exert influence over certain aspects of development such as density. But more often, these issues are dealt with by local governments through land-use controls such as height restrictions, minimum setbacks, and open space requirements. Under the assumptions used in this model, there can be a lot gained from such regulations.

CONTRACTS, COOPERATION, AND GOVERNMENT REGULATION

How can we encourage individual lot owners to develop their parcels at densities and use architectural designs that are compatible with the community? How can we best prevent individual property owners from introducing an adverse use into an otherwise harmonious neighborhood? In seeking to remedy the externality problem discussed above, economists and lawyers have developed three general approaches. None is foolproof and the usefulness of each depends on the situation.

Private Bargaining with Contracts

Particularly when the number of participants is small, it has been argued that private agreements among property owners sometimes will work. While each owner may have

[7]Following footnote 5, we now determine land value, p^*, by substituting $F = F^* = \frac{\alpha - C}{2(\beta + \gamma)}$ into the equation for p in Equation (14.6):

$$p = (\alpha - C)\left(\frac{\alpha - C}{2(\beta + \gamma)}\right) - (\beta + \gamma)\left(\frac{\alpha - C}{2(\beta + \gamma)}\right)^2$$

$$= \frac{(\beta + \gamma)(\alpha - C)^2}{4(\beta + \gamma)^2}$$

$$= \frac{(\alpha - C)^2}{4(\beta + \gamma)}$$

an incentive to shirk, through education and enlightenment, owners may agree to act in the group's interest *if* the others do so as well. To ensure this, the group can attempt to agree on a contract or accept a common deed restriction on their properties. Either will provide both the initial commitment and the ongoing enforcement that is necessary to ensure the cooperative outcome. This voluntary approach to the externality problem is frequently attributed to the University of Chicago School of Economics.

Private contracts or deed restrictions are quite common in a few American cities, generally those with little or no government land-use regulations. Such deed restrictions frequently require minimal home maintenance, prevent obtrusive activities, and limit any change in land use. Any individual owner might find such restrictions privately undesirable but recognizes that all will be better off with them. For example, Speyer (1989) found that Houston subdivisions containing such restrictions have higher property values than those with no controls, providing evidence that cooperative solutions can indeed be superior.

The problem with private contracts is that when the number of participants is large, they rarely work. In Houston, for example, deed covenants most frequently are placed on properties at the time of development when there is only a single owner-developer. To continue to remain valid, courts often require that deed covenants be resigned by the current occupants after a number of years. Researchers have documented the difficulty of getting a large number of property owners to recommit to the original agreement that was imposed by a single-minded developer (see Siegan 1972).

Larger Scale Developments

The problem with public goods is caused by having multiple owners of a common resource. With externalities, the problem is getting multiple owners to acknowledge the economic links among their properties. Both of these problems cease to exist when a single owner is making decisions. This suggests that when land is developed at a larger scale, it is more likely that compatible design, landscaping, infrastructure, or other public goods will be provided, assuming that these features really do add to the net collective value of the property being developed.

There are two problems with relying on a larger scale of development. First, there are limits to this solution. It is sometimes difficult both to assemble large tracts of land and to acquire the capital resources necessary for such large-scale development. Second, this solution works only to internalize externalities and provide initial public goods at the time of development. What about subsequent changes in market demand? How is the ongoing management of these problems to be handled after the land is subdivided, developed, and sold to many individual owners?

Government Policy

Many economists and planners argue that public goods and the management of external effects among properties rightly belong in the domain of the public sector through local governments. Two types of public policies are often advocated: regulation and economic incentives.

One solution to the public good–externality problem would be for some public entity simply to require that individual property owners or developers behave in the collective interest. In principle, the cooperative solution in the examples discussed above could become dictated by law. Thus, subdivision regulations frequently require certain levels of infrastructure, landscaping, density, and house layout, and zoning laws limit the range of uses allowed initially on property as well as future alterations. Such laws can be viewed as trying to ensure that minimum public goods are provided and that negative use externalities among property owners do not occur. Of course, this view of regulation assumes that the regulating authority has the information necessary (e.g., about individual or market valuations) to make a reasonable estimate of the collective solution. It is not at all clear that such information is widely known, particularly about the value of many types of externalities and public goods.

A second form of government intervention is to have some public authority tax undesirable activity and financially assist or subsidize behavior that is in the collective interest. Thus, in the externality example above, the private individual developer might be taxed for a FAR level above the level chosen by the neighborhood. In principle, there is a tax rate which will be sufficient to induce individuals to build at the cooperative FAR level even if individuals act in their own self-interest.[8] In practice, during the 1980s some U.S. cities created linkage programs under which developers were permitted to build above the FAR limits set by the city if they paid a surcharge. Often, the revenues from the surcharge were earmarked for affordable housing or other community development projects. Alternatively, tax credits have been used widely to encourage development that creates a positive externality, such as restoring properties of historic importance. Of course, determining the proper tax or credit again requires that a public agency be able to determine how much individuals value such public amenities.

In the U.S., the regulatory approach is far more prevalent than the use of financial incentives. Even here, the government limits its regulations primarily to restricting use and requiring infrastructure provision. There is virtually no public design review power like that found in some European countries. It has been argued that in the U.S. there is little consensus about which public goods should be provided or which externalities should be addressed by intervention. As a result, it may be reasonable for a representative government to take a more cautious approach and intervene only with minimal (and hence more widely accepted) regulations.

[8] The tax on FAR would be equal to the social impact of the higher FAR ($-\gamma F$). Thus, individuals would not develop their parcels to the point where private benefits equal cost, but rather to where private benefits equal cost plus the tax.

HISTORY, EXTERNALITIES, AND LAND-USE PATTERNS

Sometimes external economic impacts occur not only across properties within neighborhoods, but between development at one location in a city and land throughout the remainder of the metropolitan area. When these broader impacts are not fully internalized by the private land market, then private development decisions may need to be planned, guided, or regulated in order to yield outcomes that are in the collective interest. The incompatibility between major land uses is a widespread issue that can affect not only local property values but may also lead a private land market into a regional development pattern that is distinctly suboptimal and determined more by history than by current economic forces.

To illustrate this problem, let's again return to a simple, single-centered city with two uses, residential and industrial, much like the model discussed at the beginning of Chapter 5. To keep matters simple, we assume that both uses occupy 1 acre of land per household or industrial firm. Further, both uses incur travel costs to the city center—residents for commuting and firms to acquire materials and ship products. Annual commuting costs reduce the residential land rent gradient by $-k_H$ per mile, whereas shipping costs reduce the industrial rent gradient by $-k_I$ per mile. The central land rents for each use are r_H, r_I.

Our innovation to this model is to assume that industrial noise and pollution create an annual cost for residents that decreases with greater distance between households and industrial firms. Thus, residential land rents depend not only negatively on distance to the urban center but also positively on the distance between a residence and the nearest industrial use. The annual value that households place on each mile of separation will be γ, and, for the moment, we may assume that the two uses occupy separate areas of the city with a boundary between them occurring at a distance of m from the city center. The two gradients for land rent, therefore, are:

$$r_I(d) = r_I - k_I d$$
$$r_H(d) = r_H - k_H d + |m - d|\gamma \qquad (14.8)$$

In Equation (14.8), it is important to note that residential land rents depend on the absolute value of the difference $m - d$. When residents live closer to the center than industries, their rents decrease with distance from the center due to two forces: commuting to the center and proximity to industries. The farther from the center a resident is located, the closer she comes to the nearest offending industry. In this case, the slope of residential land rents (with respect to d) is $-[k_H + \gamma]$. On the other hand, when residents live beyond industries, these two forces act in opposite directions; greater distance reduces land rents because of commuting and increases them as distance generates greater separation. In this case, the residential land rent gradient has a slope of $-[k_H - \gamma]$.

To make this model more realistic, we may assume that residential commuting costs are more important than firm shipping costs ($k_H > k_I$), as Chapter 5 demonstrated was likely to be the case in modern cities. We also assume that the incremental externality cost is large relative to commuting costs, so much so that $\gamma > [k_H - k_I]$. With these parameters, our city has *two* land market equilibria as shown in Figure 14.2.

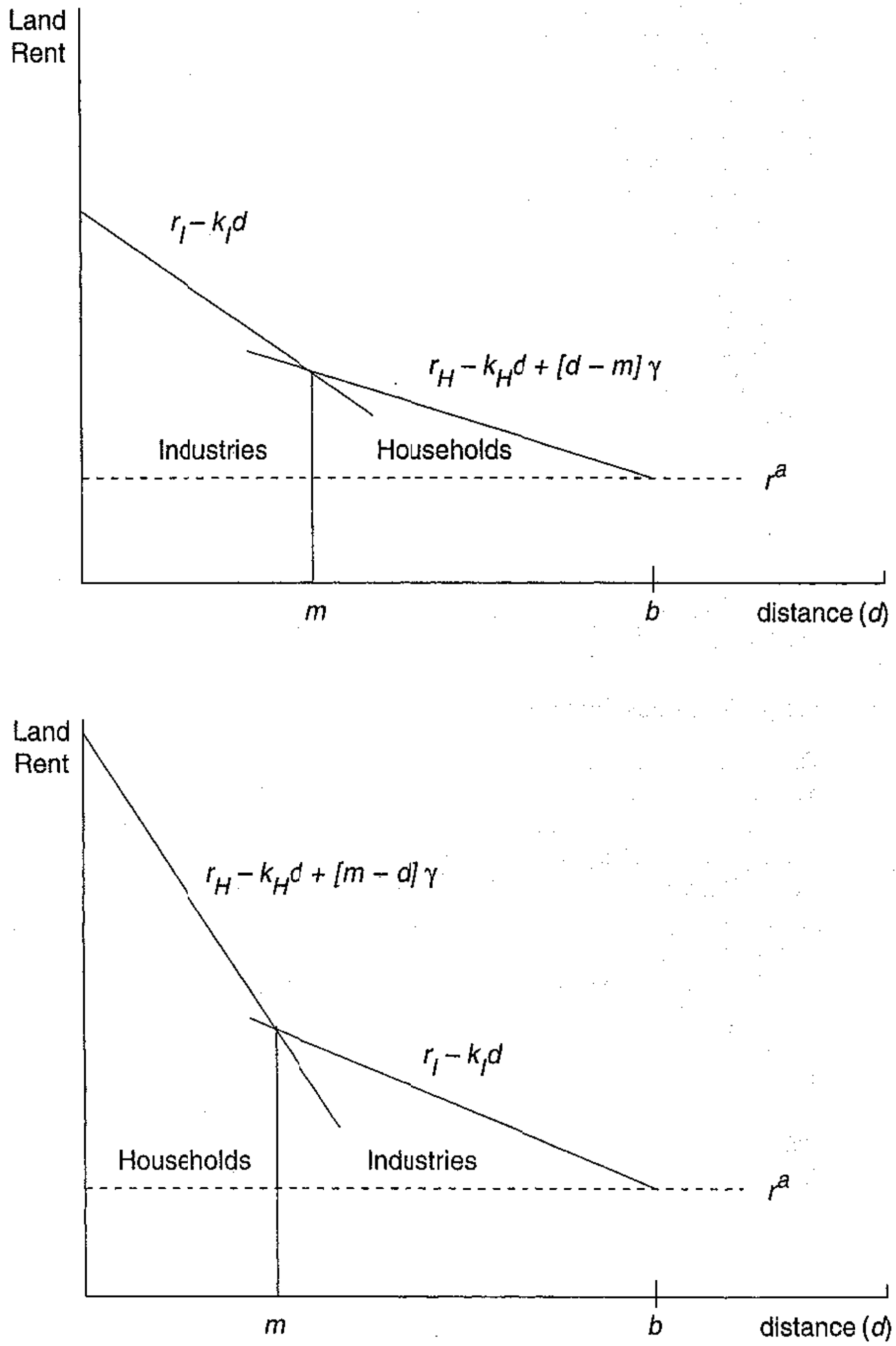

FIGURE 14.2 Alternative locational equilibria with residential and industrial uses.

Two market equilibria exist because being separate from each other is more important to the two uses than the absolute location pattern by which that separation occurs (center versus periphery locations). Even though residents prefer central locations more than industries do, their *net* preference for centrality *given that industries are already located there* is less. As a result, a location pattern in which industries occupy central sites becomes a stable, sustainable market equilibrium. Of course, if industries are already located in the more peripheral locations, then residents will clearly outbid them for central sites, since centrality now offers not only less commuting, but freedom from noise and pollution as well.

One of these market solutions is clearly preferred to the other. From the point of view of separating land uses due to the externality placed on residents by firms, either equilibrium is equally satisfactory. However, firms do not value centrality as much as residents do, which means that the second solution in Figure 14.2 is preferred. In fact, if we were to calculate the aggregate value of land rent in the two solutions, we would also find that it is larger in the second equilibrium.[9] This difference in the aggregate value of rent is illustrated in Figure 14.2, in which the total area under the rent gradients is greater in the second graph than in the first.

When markets have more than one equilibrium, we need to explain which one actually occurs and why. In many cases, history provides the answer. In Chapter 5, we argued that in the nineteenth century, industrial uses might well have valued access to some central transportation terminal more than residents, leading to the initial development of a central industrial district in many cities. In the twentieth century, however, industrial needs changed along with their rent gradients. In many cities, however, central waterfront locations are still occupied by industrial enterprises even though denser residential uses might offer higher rents than industries for such locations *if there were no other industries already there.* It is the externality that prevents residents *individually* from offering enough rent to begin the transition of central sites from industrial to residential use. In effect, the land market is, at least for a while, locked into the first solution in Figure 14.2 because of historical precedent. Economically, it would be more efficient and profitable to reverse the location pattern. However, rents are based on individuals and firms acting *unilaterally.* Thus, no market participant finds it individually advantageous to make the switch in location.

The existence of multiple market solutions also raises the broad issue of whether public sector intervention is needed. In the example above, market participants in the aggregate would be better off with the second solution even though no individual participant wishes to be the first to make the change. Several public policies could accomplish the transition from the first to the second equilibrium. A government or planning agency

[9] In the outer ring of development the slope of industrial rents in the second equilibrium ($-k_I$) is greater than the slope of residential rents in the first equilibrium solution ($-k_H + \gamma$) because of our assumption about the magnitude of the externality. Therefore, at the boundary m, rents are higher in the second solution, whereas they are equal at the urban border, b. From the boundary inward, the slope of residential rents in the second solution ($-k_H - \gamma$) is greater than industrial rents in the first ($-k_I$), since $k_H > k_I$. As a result, land rent also is higher over the inner ring in the second solution.

might initially subsidize those residents that agreed to make the first move into the industrial areas or those firms that agreed to move outward. Alternatively, industries and households that remained in the inefficient pattern could be taxed or fined. Rezoning each part of the city for the alternative use might also accomplish the change. Once existing industrial plants or houses deteriorated, owners could only rebuild by relocating. Given the long life of structures, this approach could take a long time to accomplish the goal.

It is important to remember that public intervention in this example has not been necessary to ensure the separation of the two uses. Given the model's parameters, this will occur automatically in the private market. The public role here is to bring about a more efficient location pattern. The possibility that some form of central planning may be essential to proper market functioning can be further illustrated with the issue of traffic congestion.

CONGESTION EXTERNALITIES, REGIONAL PLANNING, AND THE LAND MARKET

Throughout this book, we have argued that the accessibility of the site is an important determinant of the value of a development. But suppose as the development occurs, the traffic generated dramatically decreases that original accessibility? More importantly, what if the development decreases (through greater congestion) the accessibility of other land in the area that was already developed or adjacent land that has yet to be developed? Should the development still occur at this location? Since increased traffic congestion commonly accompanies larger scale development, let's examine this important problem in more detail.

We begin with the most simple of our single-centered cities developed in Chapters 3 through 5. Let's make the city linear with some width, meaning that the city is developed on a long, narrow strip of land, and note location along this strip as distance d. The city is composed of identical households who work in the center (distance d_0) and live at a constant residential density $1/q$, where q represents the acres of land consumed per household. Our innovation to this model will be to assume that residents of the east side of the city (at distances $d > d_0$) use cars and a road network for commuting to the center, whereas residents on the west side (at distances $d < d_0$) use a rapid transit subway system.

The subway system on the west side has considerable excess capacity, as is frequently the case with transit systems. As a result, any number of commuters can travel at a fixed cost of k_w per mile per year. For households to be in locational equilibrium, land on the west side of the city will have a rent gradient that declines over distance with a slope of $-k_w/q$. Development on the west extends to d_w miles. On the east side, drivers using the road system experience some degree of congestion. The level of congestion increases with the number of road users, which depends directly on how far the city extends to the east (distance d_e). Thus, for those on the east side, the cost per mile of travel will not be constant. Transportation costs per mile will depend on the amount of development there, which means that k_e is a function of d_e, that is, $k_e(d_e)$. For locational equilibrium in the east, land rents must decrease from the center with a slope of $-k_e(d_e)/q$.

Given the difference in transportation technologies on the east and west sides and the resulting differences in commuting costs per mile, we expect that the land rent gradient on the east will have a different slope than the land rent gradient to the west. Let's explore further the east and west land rent gradients.

Beyond the development borders d_e and d_w, land is devoted to agricultural use that yields a land rent, r^a. The land rent gradients to the east, r_e, and the west, r_w, are:

$$r_e(d) = r^a + k_e(d_e)\left[\frac{d_e - d}{q}\right]$$
$$r_w(d) = r^a + k_w\left[\frac{d - d_w}{q}\right] \tag{14.9}$$

From Equations (14.9), we know that land rents at the east and west borders of the city, $r_e(d_e)$ and $r_w(d_w)$, are equal. At the borders where $d = d_e$ or $d = d_w$, the second term of both equations in (14.9) vanish and rents at both borders equal the agricultural land rent, r^a.

What do we know about land rents at the center of the city? In order for the land market to be in spatial equilibrium, land rents at the center must be equal $[r_e(d_0) = r_w(d_0)]$. The equality of rents at the center is a key condition. If land rents at the center were not equal, then a household or firm could be better off by making a slight move to the east or west where rents are lower. In Equations (14.9), land rents are equal at the center (d_0) only if total commuting expenses on the east and the west sides are equal. If the per-mile commuting costs vary because of the differences in transportation systems on the two sides of the city, total transportation expenses can only be the same if east and west borders of the city are at different distances from the center. This distance difference must offset the per-mile commuting costs variations between the east and the west. Figure 14.3 illustrates this asymmetric land market in which rents are equal at the center, and the eastern border is a greater distance from the center than the western border, yielding a flatter gradient in the east.

How do the east and west land rent gradients change as more development occurs? If the city is small, and, hence, the distance to both borders (d_w, d_e) is short, the slope of the east-side rent gradient would be quite flat since use of the road system would not be sufficient to cause congestion. In this case, the city would extend to the east much farther than to the west. As the city expands from population growth, the west-side rent gradient shifts proportionately outward and d_w increases, just as with the cities in Chapter 3. In the east, however, more development increases road use, congestion sets in, travel speeds fall, and the slope of the land rent gradient begins to steepen as auto commuting costs rise. As a result, the expansion of the east side of the city begins to slow.[10]

As d_e expands and extra vehicles are added to the road system, the expense of commuting at any given closer location, $k_e(d_e)[d_e - d]$, increases. As one more person develops land to the east, the commuting cost per mile for all other east-side residents is made

[10]For simplicity, we assume that the subway system to the west and the road to the east extend beyond the borders.

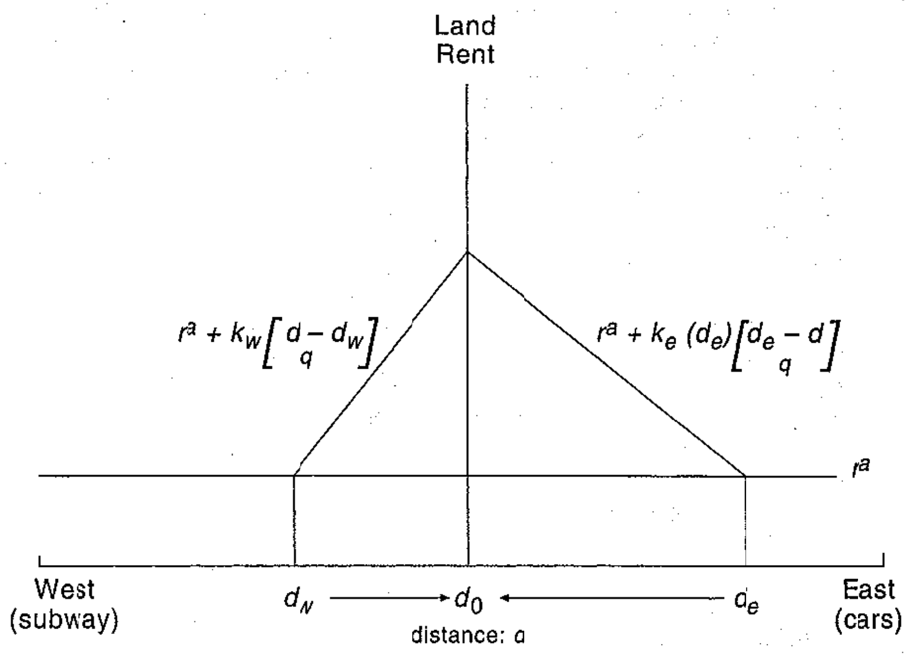

FIGURE 14.3 Land market equilibrium with two travel modes.

worse. On the west side, where k_w is fixed, this is not the case. Growth on the west side imposes no costs on existing west-side residents.

Consider the city to be in equilibrium as illustrated in Figure 14.3. Suppose that there is enough development on the east side so that there is some road congestion. The spatial equilibrium means that no individual can improve his situation by moving. However, suppose such a move would make others better off. In this model, the overall welfare of the city is determined by aggregate commuting expenditures. A city planner examining this situation could reduce aggregate commuting expenses and improve the city's welfare beyond that resulting from the equilibrium in Figure 14.3. The planner contemplates somehow moving a household from the east to the west, slightly expanding d_w and reducing d_e. The aggregate commuting expense incurred by the city's population will rise by $k_w d_w$ from the added trip on the west. In the east, one less trip will reduce aggregate commuting by $k_e(d_e)d_e$ *plus the savings to all the other east-side commuters who now have one less car to contend with on the roads.* The reduction in aggregate commuting expenses on the east is greater than the increase in the west, and thus this exchange would be in the interest of all. The household making the move is indifferent to living at either border as implied by the original spatial equilibrium. However, all the other residents of the east side will be better off.

Suppose the planner continues to make this exchange, moving residents from the east to the west, expanding d_w and contracting d_e. As these changes are made, the expense of commuting from the western edge grows, while that from the eastern fringe shrinks. Residents making the move are now personally worse off even though the remaining east-side residents are made better off. From the city's perspective, the moves should continue as long as the benefits to the remaining east-side residents exceed the costs to the individual making the move. When should the planner stop moving residents from the east to the west? Eventually, the west edge grows so much and the east contracts so much, that *the incremental resident making the move incurs losses that are greater than the gain for all the east-side residents that remain. At this point, the planner should stop making the switch since the city as a whole can be made no better off.*

In this example, the initial spatial equilibrium is reached with individual commuters considering only their own commuting costs, not the costs that they impose on other commuters by being on the road. What the city planner is doing by moving residents is considering the costs to all commuters. The problem with the city planner's solution is that it is not a market equilibrium. What we mean by this can be most easily seen by turning to Figure 14.4. If development of the west end is extended out beyond the market solution d_w to d_w^*, then for west-side households to be in locational equilibrium, land rents

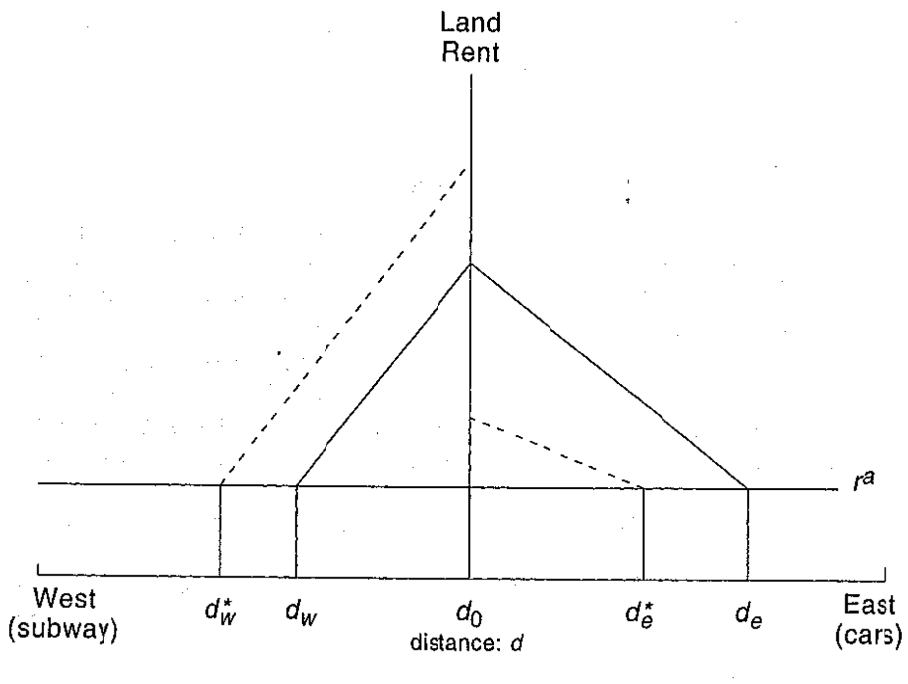

FIGURE 14.4 Market disequilibrium at planner's development solution.

will have to rise to the dashed line on the west side in the figure. On the other hand, with development on the east end curtailed, locational equilibrium among its residents will require both a lower and flatter land rent gradient. The gradient will be flatter because with less development on the east side, congestion will be less and the cost per mile of travel will be lower. With these land rents, eastern residents at the center pay far less than do western residents at the center. All residents on the east side therefore are better off than their west-side counterparts who therefore will want to move. Land rents which are set in the private market cannot maintain the city planner's more efficient development pattern.

In this example, the city planner correctly perceives that the overall welfare of the city is improved by minimizing transportation expenses. But how can this increase in efficiency be sustained in the private land market? Suppose that instead of the planner forcing households to move from east to west, commuters on the east side considered not only their commuting costs but the external costs that they impose on other drivers by being on the road. This could be done by having the city impose a congestion toll on each user of the road that exactly equals the costs that a commuter's use of the road imposes

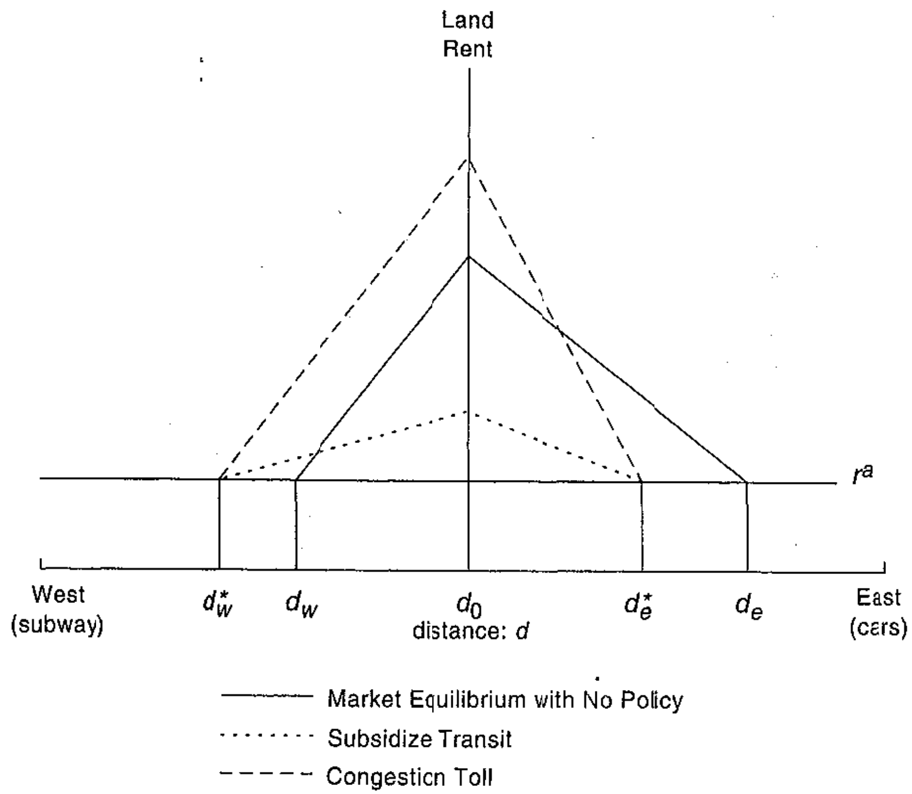

FIGURE 14.5 Policies to achieve market equilibrium at planner's development solution.

on other commuters. The toll revenues would be rebated to all residents. The congestion toll would force east-side commuters to pay the full costs of their travel by internalizing the externality that arises from their commuting.

What's the impact of this toll on the land rents that commuters are willing to pay? As shown in Figure 14.5, the increase in commuting costs as a result of the toll would increase the slope of the land rent gradient on the east side to the point at which east-side rents equaled west-side rents at the center. In the long run, there would be less growth on the east side with the toll in place than there would be without the toll, so the east boundary would be closer. The important point here is that with the toll, the land market would move to a new spatial equilibrium, with the same boundaries as the planner's solution.

There are alternative public interventions that also would lead to a land market equilibrium in which the borders match the planner's solution. For example, on the west side, the city could subsidize the public transit system, reducing transit costs for commuters. With this reduction in commuting costs, the land rent gradient on the west side will become flatter. With sufficient subsidy, it can be flattened to the point where west-side rents at the center are identical to those from the east. This transit subsidy solution is also illustrated in Figure 14.5. The reduction in transit costs could take the form of either a fare subsidy and/or an expansion of the system to provide uneconomically fast service. The gains to city residents from this policy of less expensive transit for west-side residents and less congestion for east-side residents could exceed the cost of the taxes needed to finance it.

Economists have long supported the implementation of congestion tolls over other approaches to solving congestion because this mechanism solves the problem by correctly pricing road use. Other policies can be less precise in correcting the problem. In our example, if the city instead subsidizes transit on the west side rather than charging the congestion toll on the east, then both the roads on the east and the subway on the west are incorrectly priced with the possible effect of encouraging more travel. Fuel taxes are often proposed as a method of dealing with congestion problems. But these taxes are blunt instruments since they tax fuel whether the travel occurs on a congested urban road or a rural uncongested road. The tax is also the same if the travel takes place during peak congestion times or at off-peak times. Congestion tolls that vary by time and location would force the private market to price correctly commuting by car.

The creation of congested transportation facilities is one of the most common problems in land development. The lesson of this simple example is that this congestion can require significant public intervention. The intervention not only alters the transportation system but can also redirect the entire pattern of development. Such government intervention, if appropriately done, can lead to significant improvements in the quality of urban life over that which results from an unregulated private market.

SUMMARY

In this chapter we discussed how certain economic goods and services are collectively enjoyed by a wider community of adjacent property owners. Public goods like open space

and environmental quality can be highly valued by property owners, but must be provided collectively. Individual owners also contribute to the community environment through their own design and development decisions, creating additional economic externalities.

- Individual decisions about the provision of public goods can lead to a misallocation of resources. Providing the "right" level of public services often requires contractual arrangements, larger scale developments, or the formation of government or quasi-governmental organizations. Unregulated private markets generally fail to adequately account for public goods.
- In a similar manner, decisions about individual property development, such as density, can often be detrimental or beneficial to adjacent properties. Coordinated decisions made by a number of property owners often can lead to a better development plan. Such coordination, however, again requires some form of organization or institutional assistance outside of private market behavior.
- At the metropolitan level, longer-term land-use patterns also can lead to externalities which the private market ignores. The noise and environmental impact of large industrial districts or the congestion that follows new development are just two examples. Public planning offers the potential of controlling these externalities through, for example, the use of metropolitan-wide land-use regulations or infrastructure fees. In principle, such policies could lead to land-use configurations that improve urban living.

REFERENCES AND ADDITIONAL READINGS

ARNOTT, RICHARD, A. DE PALMA, AND R. LINDSEY. "A Structural Model of Peak Period Congestion." *American Economic Review* 83, 1 (1993): 161–178.

ASABERE, PAUL K. "The Value of a Neighborhood Street with Reference to the Cul-de-Sac." *Journal of Real Estate Finance and Economics* 3 (1990): 185–193.

ASABERE, PAUL K., GEORGE HACHEY, AND STEVEN GRUBAUGH. "Architecture, Historic Zoning, and the Value of Homes." *Journal of Real Estate Finance and Economics* 2 (1989): 181–195.

DOWNS, ANTHONY. *An Economic Theory of Democracy*. New York: Harper & Row, 1957.

FORD, DEBORAH ANN. "The Effect of Historic District Designation on Single-Family Home Prices." *AREUEA Journal* 17 (1989): 353–362.

FRECH III, H. E., AND RONALD N. LAFFERTY. "The Effect of the California Coastal Commission on Housing Prices." *Journal of Urban Economics* 16 (1984): 105–123.

GRETHER, DAVID M., AND PETER MIESZKOWSKI. "The Effects of Nonresidential Land Uses on the Prices of Adjacent Housing: Some Estimates of Proximity Effects." *Journal of Urban Economics* 8 (1980): 1–15.

KOHLHASE, JANET E. "The Impact of Toxic Waste Sites on Housing Values." *Journal of Urban Economics* 30 (1991): 1–26.

MIESZKOWSKI, PETER, AND ARTHUR M. SAPER. "An Estimate of the Effects of Airport Noise on Property Values." *Journal of Urban Economics* 5 (1978): 425–440.

NELSON, JON P. "Residential Choice, Hedonic Prices, and the Demand for Urban Air Quality." *Journal of Urban Economics* 5 (1978): 357–369.

SIEGAN, BERNARD. *Land Use Without Zoning*. Lexington, MA: Lexington Books, 1972.

SMITH, BARTON A. "Measuring the Value of Urban Amenities." *Journal of Urban Economics* 5 (1978): 370–387.

SPEYER, JANET. "Effects of Land Use Restrictions on Single Family Housing in Houston." *Journal of Real Estate Finance and Economics* 2 (1989): 117–230.

VANDELL, KERRY D., AND JONATHAN S. LANE. "The Economics of Architecture and Urban Design: Some Preliminary Findings." *AREUEA Journal*. 17 (1989): 235–260.

WADDELL, PAUL, BRIAN J. L. BERRY, AND IRVING HOCH. "Residential Property Values in a Multinodal Urban Area: New Evidence on the Implicit Price of Location." *Journal of Real Estate Finance and Economics* 7 (1993): 117–141.

INDEX

A

B

C