ECON 325: URBAN ECONOMICS
THE MONOCENTRIC CITY MODEL

Last time, we discussed some of the economic forces that give rise to cities. Understanding these forces is one of the central goals of urban economics and the broader field of economic geography. Another key goal is understanding how economic forces influence the internal structure and organization of cities. We'll start with a discussion of one of the earliest formal attempts to model economic activity in space, the von Thünen model. Then, we'll discuss the basic version of the Alonso-Muth-Mills model, popularly known as the monocentric city model. This model formalizes the tradeoff between transportation costs and housing consumption to understand how consumers will choose where to reside relative to a central economic hub. Despite its simplicity, the model generates several clear predictions that can be tested empirically. We'll look at some evidence for the validity of the model, and highlight some shortcomings that will be addressed next class.

In 1826, Johann Heinrich von Thünen published *The Isolated State*, one of the earliest treatises on economic geography. His goal was to understand the spatial organization of agriculture around a city. He employed an abstract, stylized theoretical framework that would go on to inspire many other models of urban land use. We'll outline his model here and study its implications, as this will introduce some key concepts that we'll be using later.

He imagined a city located on a featureless plain, perfectly flat and with a uniform soil quality and climate in all directions. The city is isolated from others, requiring it to be economically self-sufficient (in a word, autarky). These abstractions, while unrealistic, allow us to focus on the underlying economic dynamics without distraction. He assumes the city is surrounded by farmland, with farmers growing crops for sale in the city. There are different types of crops available to grow, such as fruits, vegetables, and cereal grains. These differ in several key respects:

(1) The cost of transporting them to market,
(2) the price they command at market, and
(3) the amount of space required for cultivation.

So, where will different crops be cultivated around the city? Clearly, farmers of all types would prefer to locate closer to the city, as this will reduce the transportation costs they bear. It's equally clear that this is impossible - land used for one crop cannot be used for another. So, the spatial organization will end up depending on the tradeoff between the cost of transporting goods to market, and the rent farmers are willing to pay for land. Fruits are more valuable on a per-unit basis, can be cultivated more intensively, and potentially have higher transportation costs (due to the risk of damage/spoilage). Thus, fruit farmers will be willing and able to pay a premium to locate very close to town. The price fruit farmers are willing to pay for farmland at different distances from town can be represented using a *bid-rent curve*. The fruit farmers' bid-rent curve is depicted in red in the top panel of figure 1. Close to the city, these farmers' willingness to pay for land is very high, and it declines with distance, eventually reaching zero once the distance is so great that the fruit could not be transported to market economically.

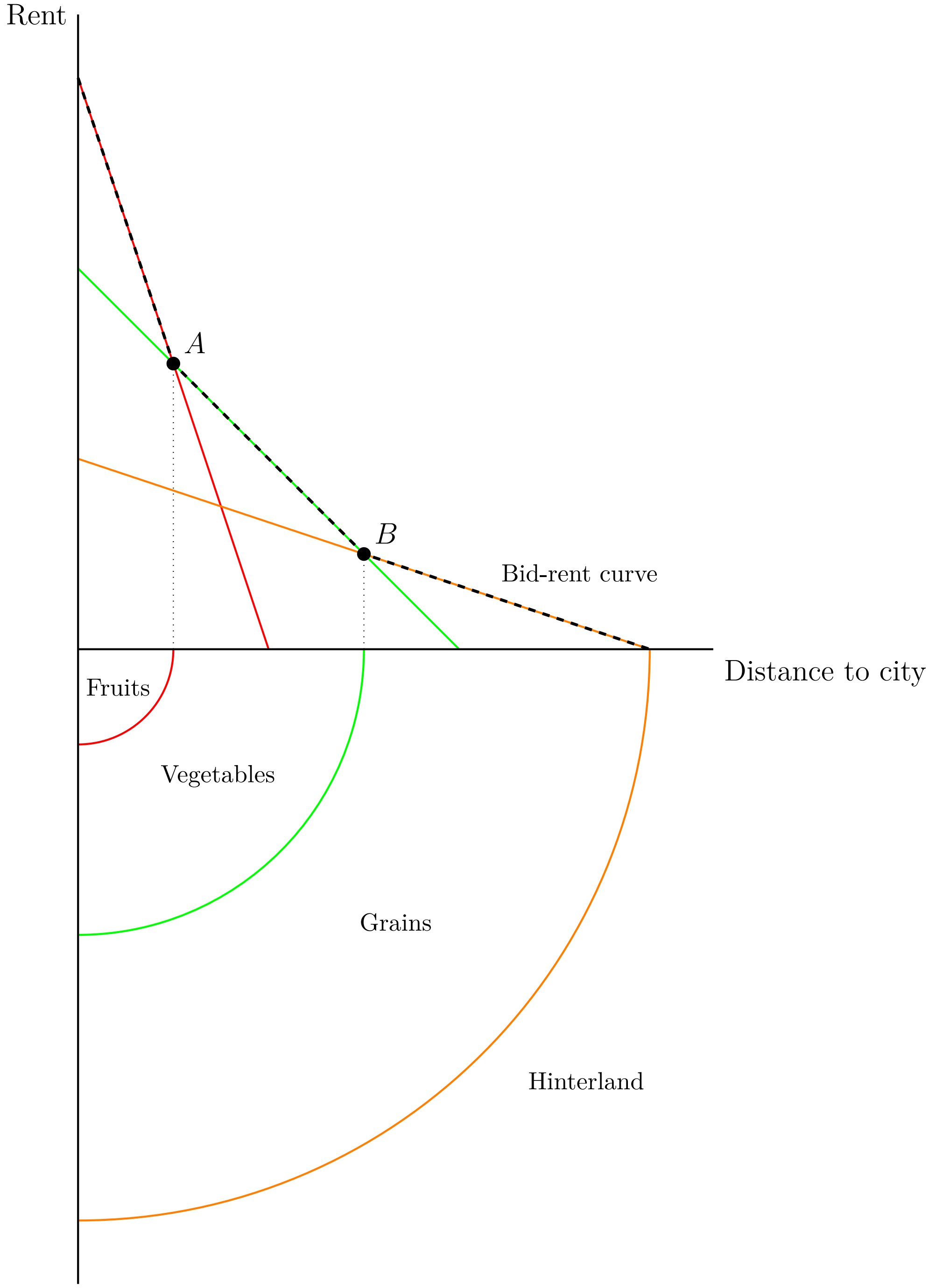
What about vegetables? These are less valuable on a per-unit basis, and require more land to cultivate, so the maximum willingness to pay for land by farmers is correspondingly lower. They're also hardier, and so easier to transport over distance. Thus, the vegetable farmers' bid-rent curve (depicted in green) starts at a lower point but is not as steep, reaching zero at a greater distance. Lastly, grains are much less valuable on a per-unit basis, and require substantially more land to cultivate. The associated bid-rent curve (depicted in orange) starts out very low, but doesn't reach zero until a great distance from the city, as this crop is the hardiest.

We can now see how agriculture will be organized around the city. From a distance of zero out to the distance marked by the point A, fruit will outbid other crops and occupy this valuable land, depicted in the bottom panel of figure 1. Beyond A and out to the point B, vegetables will outbid both fruit and grains - the former lose too much value in transit, and the latter are simply not valuable enough. Past B, grains become the dominant crop until their bid-rent curve reaches zero - at this point, it's no longer profitable to cultivate these and transport them to market. Beyond this point is uncultivated wilderness, too distant from the city to be economically useful. The dashed black line in figure 1, which represents the upper envelope of the overlapping bid-rent curves for each crop, reflects the actual land rent that will be extracted at each point around the city.

The pattern of agriculture determined by the model is the outcome of farmers competing for limited and valuable space. The concentric ring pattern that emerges could do so entirely through the forces of competition, without the direction of any central plan. So naturally the question emerges: Is this market-driven outcome *efficient*? Indeed, it is. Since the bid-rent curves reflect the preferences of the town's consumers balanced against the tradeoffs inherent in transportation costs, this is the economically efficient outcome, in the sense that it maximizes the total economic value of agriculture around the city.

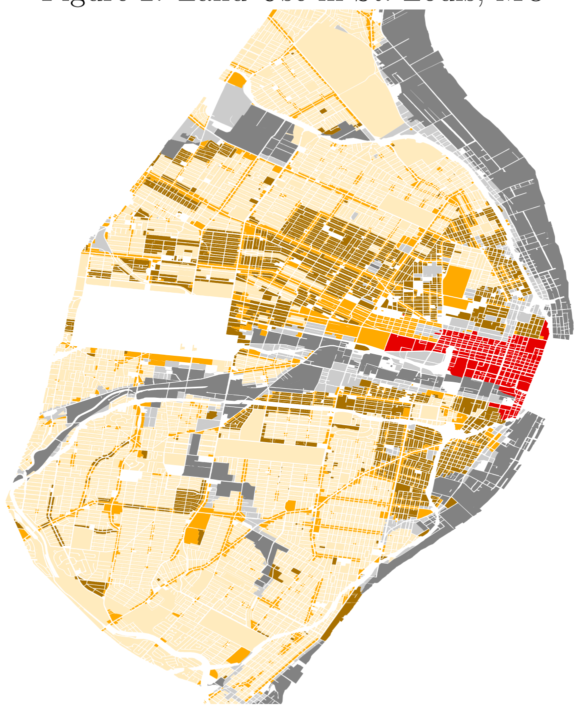
Another question may have occurred to you - why do we care about an abstract theory of nineteenth century agricultural organization? It turns out that the logic inherent in the model is quite general, and can apply to many different types of land use. Von Thünen's approach was built upon by William Alonso, who developed the initial form of the monocentric city model in his 1964 work *Location and Land Use*. To analyze modern land uses in a city,

Figure 1: The von Thünen Model



we can simply replace fruits, vegetables, and grains with commercial offices, manufacturing, and residential uses. In figure 2, we can see land uses in contemporary St. Louis, MO. Here, we see a downtown primarily focused on commercial office space (red), surrounded by manufacturing (gray), less intensive commercial uses (orange), and multifamily residential uses (brown). Farther out, the land is primarily devoted to less intensive single-family residential activity (beige).[1] The key idea here is that there is a tradeoff between proximity and rents, and that in a *spatial equilibrium* we will observe more intensive uses and higher rents the closer we are to the economic hub of the area. This is the motivating principle behind the monocentric city model.

Figure 2: Land Use in St. Louis, MO



Let's develop this idea formally. We'll focus first on a model of residential location choice. We'll assume that we have a circular city, with a dense radial road network that individuals use to commute to the *central business district* (CBD), where all jobs are located. (Hence the name "monocentric city.") The city is populated by residents, all of whom are identical in terms of income and preferences. Individuals can choose to locate anywhere in the city, and can relocate costlessly. We're making a lot of assumptions here, but these can all be relaxed at the cost of increased complexity. We'll look at some relaxations later.

Individuals consume two goods, housing q and non-housing goods c. You can think of q as square footage of housing. Non-housing goods c can be read as "money left over after you've paid rent." This good c is what we call a *numeraire good*, i.e., it has a price of 1 (basically, it's money). The price-per-square-foot of housing is p. A combination (or bundle) of housing and non-housing goods consumed is denoted (c, q).

Wherever individuals locate within the city, they will be some distance from the CBD. We'll denote this distance as x, with $x = 0$ being the location of the CBD. There is a

[1]Of course, we also see bands of gray extending far away from downtown - manufacturing following railroads and the Mississippi River. Cities are not, in general, situated on a featureless plain, and we'll have more to say about the importance of geography and transportation networks later.

commuting cost of t per round-trip mile, so the consumer will have to pay tx to locate x miles from the CBD. Housing costs p will be a function of x, as they may differ throughout the city. Otherwise, we assume that the city is homogenous, so that there are no non-pecuniary benefits to choosing different locations - it's another featureless plain.

The benefit an individual receives from consuming some amount of c and q is given by the individual's *utility function*, $u(c, q)$. This is essentially just a function that takes c and q as inputs and returns a number. We assume that larger numbers are better, so that if a consumer prefers the bundle (c, q) to a different bundle (c', q'), then $u(c, q) > u(c', q')$ (and vice versa). So, the utility function ranks combinations of housing/non-housing goods based on consumers' preferences.

How will the consumer choose which bundle to consume? We assume they'd always prefer both more housing and more non-housing goods, but would also like to strike a balance between the two. You don't want to consume just housing or just non-housing goods. In other words, $u(c, q)$ is increasing in both c and q, but at a decreasing rate. In addition, the consumer also faces a *budget constraint* that limits what they can afford, and thus which bundles are feasible. How do we define the budget constraint? We assume each individual has a fixed income y. From this, they must pay pq in housing costs (remember, p is the price-per-square-foot of housing, and q is the square footage). They will also spend c on non-housing goods, and they will pay the commuting cost tx. So, their total spending will be $pq + c + tx$, which must equal y.[2] Thus, their choice of q and c must satisfy the budget constraint

$$pq + c + tx = y.$$

The consumers' choice then boils down to solving the following optimization problem:

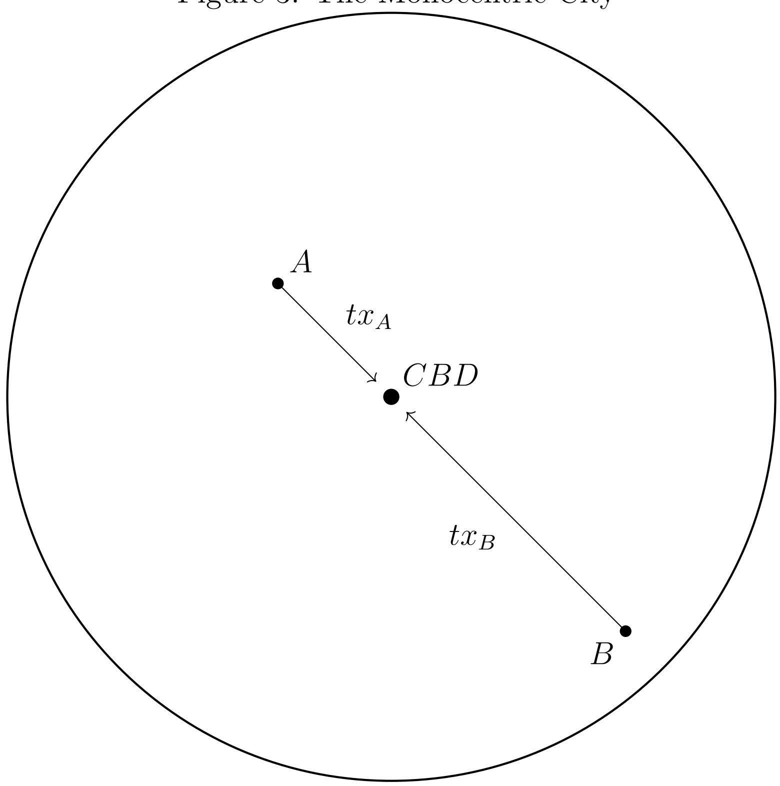
$$\begin{aligned} &\underset{c,q}{\text{maximize}}\ u(c, q) \\ &\text{subject to } c + pq + tx = y. \end{aligned}$$

In words, this says that they should choose a bundle (c, q) that maximizes their utility, subject to the condition that their total spending on housing, commuting, and non-housing goods does not exceed their income. What does a solution to this look like? Without a specific function for $u(c, q)$ (and some more sophisticated tools), we can't find an exact solution. However, it turns out that we don't need to - we can work out important implications of this model graphically using the *spatial equilibrium* condition. What does that mean in this context? Consider two city residents, A and B, living in two different locations, as depicted in figure 3. A lives close to the CBD, while B lives farther out in the suburbs; formally, $x_A < x_B$, where the subscripts are used to distinguish distance to the CBD for

[2]You might object on the grounds that they could save money instead of spending it, but we'll assume that "savings" are non-housing consumption, so they would be subsumed into c.

each individual. Because of this, A bears a lower commuting cost than B. Does that mean that A is better off, i.e., receives a higher utility than B? Suppose she does - what does that mean for B? It means that she'd prefer to relocate to A's position to receive a higher utility - remember, we said that relocation is costless. Thus, these location choices are not in equilibrium. In order for us to have a spatial equilibrium, it must be the case that no residents in the city wish to relocate. In other words, they must all be receiving the same level of utility.[3] If this were not the case, residents in low-utility locations would be incentivized to move to high-utility locations.

Figure 3: The Monocentric City



How is it possible for all city residents to have the same level of utility when commuting costs differ by location? Something must adjust to equalize utilities across the city. Since we've assumed that the price of non-housing goods is constant, and that individuals can choose whatever q they can afford at any given location, only one margin of adjustment remains - the price per square foot of housing, p. In order for a spatial equilibrium to hold, it must be the case that p is decreasing with x, i.e., the price of housing must fall as we move farther from the CBD. This is the first important conclusion we can draw from this model: Like von Thünen, it predicts a downward sloping bid-rent curve, also referred to as a *rent gradient*. Basically, higher housing prices close to the CBD act as a compensating

[3]Remember, everyone has the same income, so no one can out-spend anyone else.

differential offsetting the benefits of a shorter commute; likewise, low housing prices near the edge of the city compensate those residents for their more costly commute.

What else can we learn from this model? Let's think again about our two consumers A and B. Given their respective locations, we can rewrite the budget constraint for A as

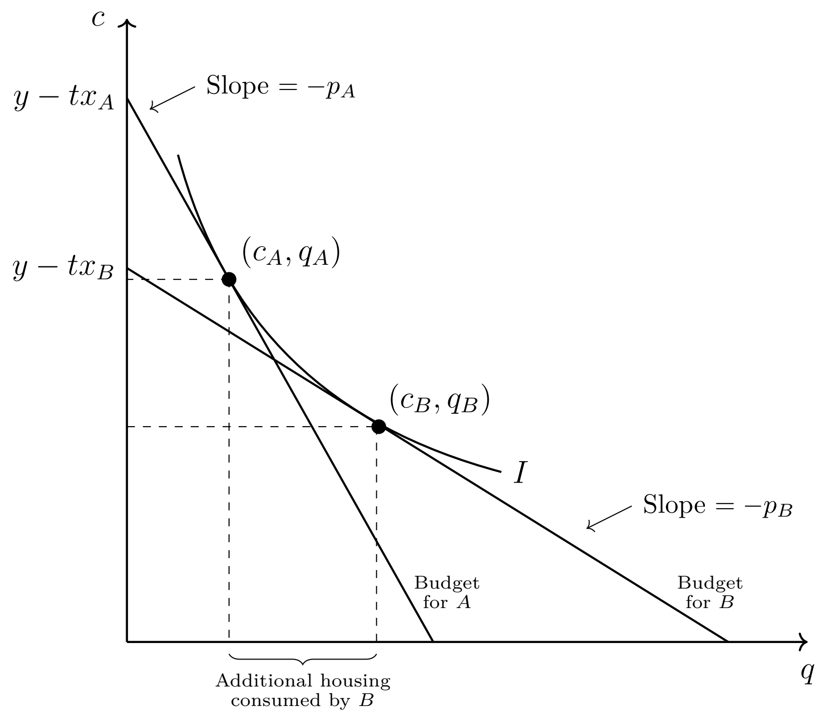
$$c + pq = y - tx_A,$$

where we think of $y - tx_A$ as A's disposable income available to consume c and q. We can do the same for B:

$$c + pq = y - tx_B.$$

We can graph both of these budget constraints as we've done in figure 4. Since A is closer to the CBD, they have a higher disposable income than B when $q = 0$, i.e., when they're spending all of their disposable income on c. However, since they face a higher p at location A, which we'll call p_A, their budget constraint declines more steeply than B's - each additional square foot of housing costs more. B's budget constraint starts out lower (higher fixed commuting cost reduces disposable income), but is less steep because they face the lower housing price p_B. Whatever bundles (c_A, q_A) and (c_B, q_B) end up being chosen, they will necessarily lie on the corresponding budget constraint line.

Figure 4: Tradeoff between housing and other goods under different budget constraints



Once we have the budget constraints graphed as in figure 4, we only need one more ingredient to see how A and B's consumption choices will differ. Recall that both consumers have the same preferences, incomes, and must achieve the same utility in a spatial equilibrium. This means that the bundle chosen by each consumer will lie on the same *indifference curve*, which maps out all of the bundles that achieve a given fixed utility level. (Hence the term "indifference" - the bundles may have different combinations of housing and non-housing goods, but those with more non-housing goods will have less housing and vice versa, in a manner that assures you will be indifferent between the different bundles.) If our spatial equilibrium utility level is $\bar{u}$, then the associated indifference curve will capture every possible bundle (c, q) such that $u(c, q) = \bar{u}$. If our consumers are optimizing, they will choose bundles that lie on the highest possible indifference curve that still intersects their budget constraints, i.e., they will pick the bundles that deliver the highest utility possible given their budgets. Because our consumers have different budget constraints, these intersection points will differ, as you can see in figure 4.

What do we learn from this exercise? We can see immediately that B consumes more housing space than A. Why? The price of non-housing goods is the same for both consumers, but housing is cheaper for B. So, in relative terms, it's cheaper for B to get more housing, and so they will. This is the second key prediction of the model: Individuals will consume more housing space as they move farther from the CBD. This replicates the well-known stereotype of tiny apartments in the city and large houses in the suburbs.

This second finding actually allows us to say more about the first finding (that prices decline with distance). It turns out that, even without a specific equation for the utility function, this model generates a mathematical expression for the relationship between housing prices and distance that is easily interpreted. Imagine a small increase in distance from the CBD, Δx. This will result in a change[4] in prices Δp. The change in p relative to x is just the ratio of these, $\frac{\Delta p}{\Delta x}$. It can be shown that this change in prices relative to distance is given by the following equation:

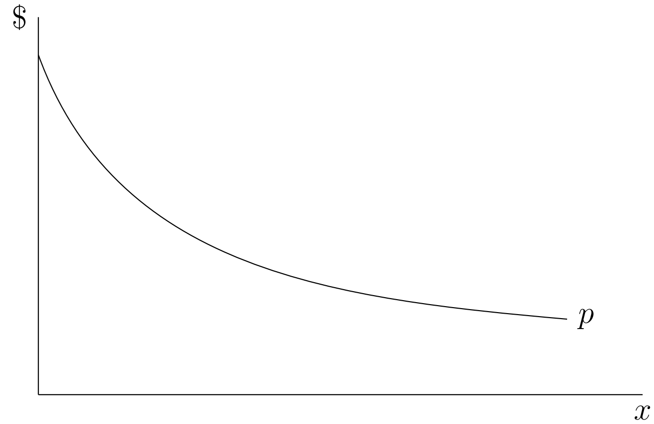
$$\frac{\Delta p}{\Delta x} = -\frac{t}{q}.$$

What does this mean? It communicates our first finding: Prices decline with distance, because both t and q are positive, so $-\frac{t}{q} < 0$. It also gives us an idea of what this relationship looks like qualitatively. Remember that we assumed a fixed per-roundtrip-mile cost t. And since our second finding was that q increases with distance, this implies that the impact of an increase in x on prices declines with distance (since q is in the denominator). So, prices fall faster with distance near the CBD, and slower once we get farther out, so the curve starts to flatten. Another way to say this is that the rent gradient is *convex*; figure 5 gives an example of just such a curve.

So, that's the basic version of the monocentric city model. Even this extremely simple model, featuring very few moving parts, generates several interesting predictions about urban

[4]We're employing the common notational convention of using a capital delta (Δ) to denote "change in."

Figure 5: Housing price gradient



location choice (and how they will translate into land use patterns). Essentially, all of these results are the simple product of transportation costs. Of course, many of our simplifying assumptions are unrealistic, and it will be worth checking to see how relaxations of these assumptions change our results. We'll see how to do this soon.

For now, we could ask the question, are these stylized predictions about cities generally true? Many studies have examined this and typically find a high level of support for the model's predictions, both across countries and across time. An example can be seen in figure 6, which reproduces a graph from Haughwout, Orr, & Bedoll (2008) depicting the log of land prices per square foot from many property transactions in New York City, as a function of distance from the Empire State Building (a proxy for the CBD). Here, we see that the rent gradient is both negative and convex, consistent with our model.[5] What about density? Our model predicts that people will choose smaller dwellings closer to downtown. This suggests that population density may be highest near downtown, and will decline as we move farther out. It turns out that this intuition is correct, both theoretically and empirically.

We now have a basic model of residential location choice with positive transportation costs, homogenous income, and costless relocation. The latter two assumptions implied that utility must be uniform across the city if we are in a spatial equilibrium, i.e., no one wants to move. Spatial equilibrium and positive transportation costs then implied that housing prices must fall with distance to the CBD. Because of substitution effects, this also implied that the quantity of housing consumed would increase with distance. Since housing is cheaper relative to other goods in the suburbs, those residing far from the CBD will consume more housing than those residing closer. This was basically a discussion of housing demand; we did not explicitly address the issue of housing supply, or what the buildings that consumers live in will look like. We turn to that now.

[5]These are log prices, but because of the properties of the log function, it would actually appear even more convex in terms of raw prices.

Figure 6: Land rent gradient in NYC

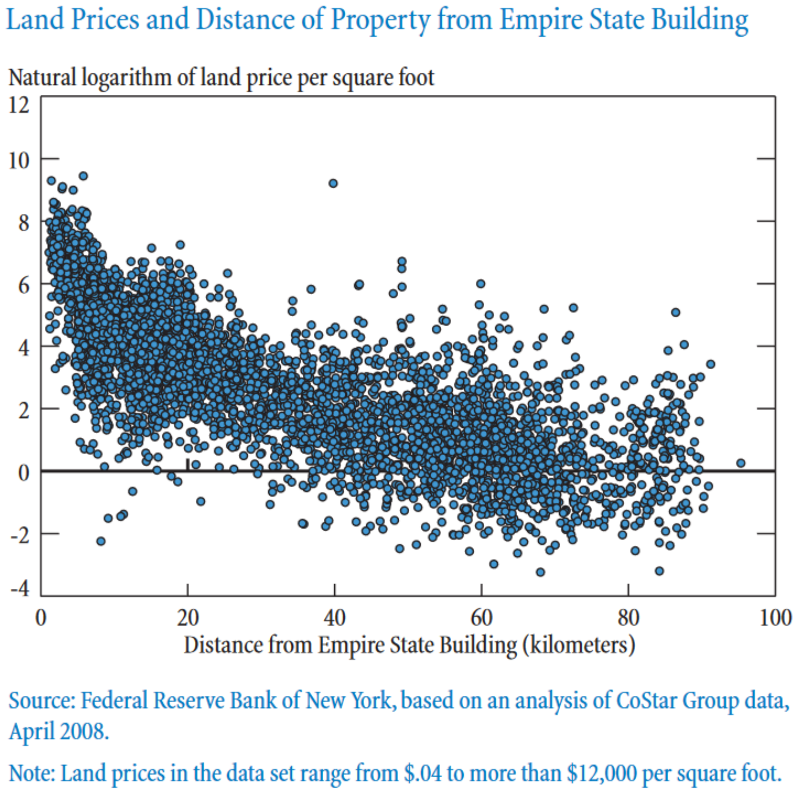
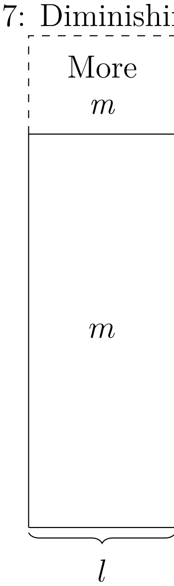


Figure 7: Diminishing returns



We'll assume that buildings are constructed by developers, who then rent the space to consumers.[6] Developers use two inputs to create housing, land l and building materials m. (We'll think of building materials in a broad sense, including all of the labor and capital required, in addition to raw materials like wood, glass, steel, etc.) Housing floor space q is produced according to a *production function* that translates the amount of land and building materials used into the residential floor space created. We'll denote the production function g, so that output is given by $q = g(l, m)$.

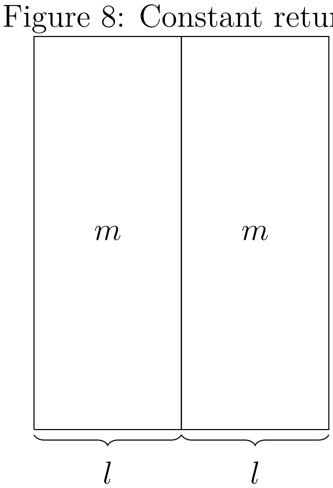
This production function has a few important properties. First, it is increasing in both inputs - more land and more building materials yield more floor space. Second, it is subject to diminishing returns in each individual input. This means that adding more m will yield progressively smaller increases in q when l remains

[6]In reality, of course, developers often sell property to landlords who act as intermediaries, but that distinction isn't important here. Also, we're assuming a population of renters and ignoring the ownership vs. renting decision, but we'll address that later.

fixed. Think of it this way: When land is fixed, the only way to get more floor space is to go up. When a building gets taller, more and more of the building materials must be used for construction aspects that do not increase floor space, such as elevators, deeper foundations, and mechanical floors. Figure 7 illustrates this idea of diminishing returns. This applies to land as well - for a fixed amount of building materials, increasing land area yields diminishing increases in floor space.

The third important property pertains to returns to scale. We discussed the idea of increasing returns to scale in industrial output in lecture I. We generally think of housing production as exhibiting constant returns to scale. This is illustrated in figure 8. If we double both land and building materials, we can imagine simply constructing an identical building next door, thus doubling floor space.

Figure 8: Constant returns



We assume that developers are profit-maximizers in the same way that consumers are utility-maximizers. They'd like to earn as much profit as possible off of their developments. For any given output q, a corresponding amount of l and m will be required. Their revenue from this production is then $pq = pg(l, m)$, where p is the market price of floor space. What about their costs? Suppose developers rent land from landowners at a rate r per square foot.[7] The per unit cost of building materials is c, and we'll assume that materials cost the same everywhere in the city. Then, a developer's total input cost will be $rl + cm$.

The question now becomes, given this setup, where will developers build, and how much will they build there? Developers have the largest incentive to build close to the CBD. Why? Because p is much higher there, meaning they can sell each square foot they produce for more money. So why would they ever choose to build in the suburbs? Absent variation in land rents r, they would have little incentive to do so. But naturally, landowners are profit-maximizers too, and will try to extract higher rents from the more valuable land near the CBD. Because developers are competing to build in the most profitable locations, they will drive up rents closer to downtown. Thus, we obtain a downward sloping land rent-distance gradient that is similar to the price-distance gradient we found earlier. In a spatial equilibrium, land rents will adjust to ensure that developers earn zero extra profits at any given location, so there is a uniform return to development throughout the city.

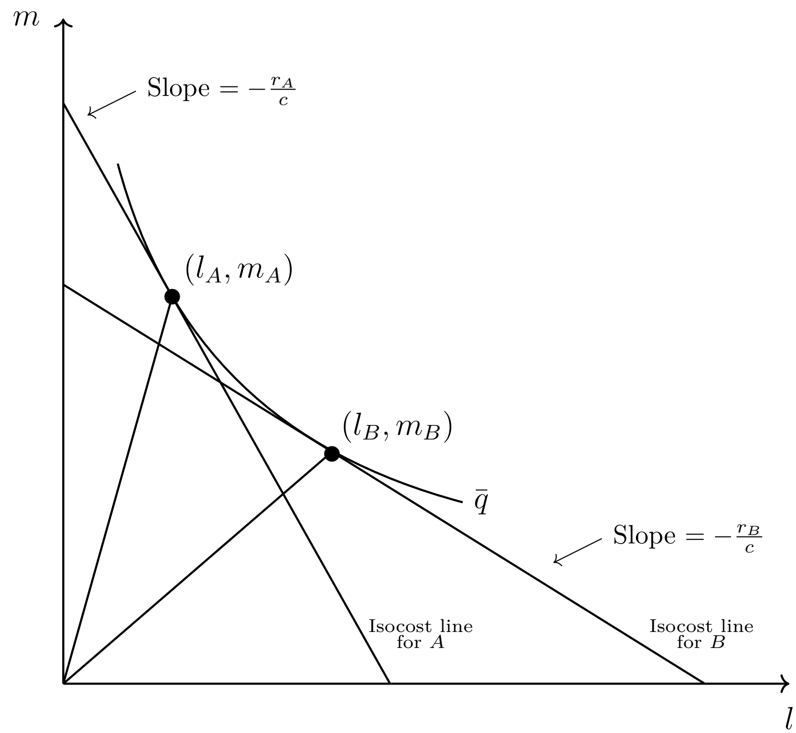
So, the first major prediction of this model is that land rents vary across the city, and decline with distance to downtown. How does that affect what *kind* of development occurs at

[7] They could of course just buy the land instead, but that distinction is not too important here.

different locations? Much like consumers facing a tradeoff between housing and non-housing consumption, developers face a tradeoff between using more land and more building materials to construct the same quantity of housing. Part of maximizing profits involves minimizing costs - the profit-maximizing quantity of housing can be produced using different mixtures of inputs, and developers should try to use the lowest-cost mix available. Let's work out the implications of that.

Suppose we have two developers, A and B, each constructing a fixed total quantity of housing $\bar{q}$ at two locations, x_A and x_B. We'll assume that x_A is close to the CBD, while x_B is farther out in the suburbs. This implies that the rental price of land faced by A, r_A, will be higher than that faced by B, i.e., $r_A > r_B$. We can graph the cost tradeoff between l and m using the notion of an *isocost line*. This is just a line that depicts all input bundles with the same total cost, analogous to the budget constraints discussed earlier. In figure 9, we can see both of these isocost lines. Since l is more expensive for A, their isocost line has a steeper slope; the slope for B is more gradual, because giving up more building materials will pay for a greater quantity of extra land.

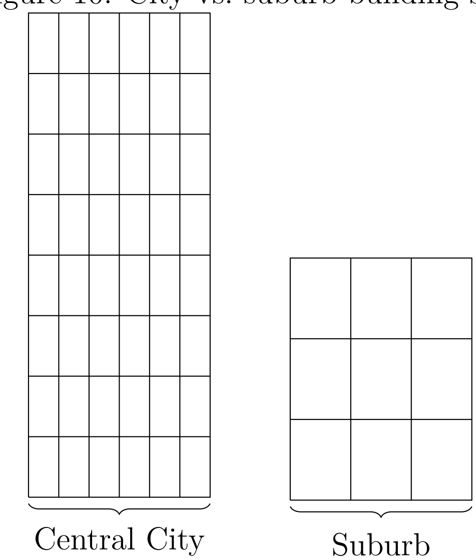
Figure 9: Tradeoff between land and material use in construction



To complete the analysis, we'll introduce an *isoquant*. This is a curve depicting different bundles of inputs that are sufficient to produce exactly the amount of housing we need, $\bar{q}$. To produce $\bar{q}$, the profit-maximizing developer will choose the bundle (l, m) that lies on lowest possible isocost line that still intersects with the $\bar{q}$ isoquant. Since we are in a spatial equilibrium, both developers will be incentivized to the build the same amount of housing at each location, so their choices will each fall on the same isoquant and each will have the same total cost (as is evidenced by the fact that the isocost lines intersect). Here, we see that the developers choose (l_A, m_A) and (l_B, m_B), respectively, with $l_A < l_B$ and $m_A > m_B$.

So, how do we interpret this result? Since land is relatively expensive near the CBD, the cost-minimizing developer wants to economize on that input, using as little as possible and making up for the difference with more building materials. With less land, this means a taller building is needed to get the same amount of floor space - this is depicted by the steeper ray emanating from the origin. In the more suburban location, land is relatively cheap, so more of it is used, meaning shorter buildings. This is the second major prediction of this model: Building heights should decline with distance to downtown.

Figure 10: City vs. suburb building style



What else can we learn from this basic model? There's a third major implication here, and that relates to population density. Density is typically defined in terms of persons per unit area of land (square mile, square kilometer, acre, etc.). How will density vary across our city? We already know that the central city will have taller buildings. And we already found that individuals in the central city will consume less housing square footage. That implies that these buildings will house more dwellings that are each individually smaller, as depicted in figure 10. It necessarily follows that population density must be declining with distance to the CBD, i.e., we have a negative density gradient. This prediction, along with the other two discussed earlier in this lecture, have been subjected to numerous tests using data spanning many countries and time periods. These studies overwhelmingly confirm the accuracy of these predictions, broadly speaking, to the point that we can call them empirical regularities (or, "stylized facts").

To summarize our findings thus far, the core assumptions of our model are positive transportation costs and the ability of consumers and developers to choose different locations. This latter assumption implies that both utility and profits should be uniform across space in equilibrium. Transportation costs and uniform utility together imply that housing prices must fall with distance to the CBD. This fact, along with spatially uniform profits, implies that the underlying cost of land itself must be falling with distance. Lower housing prices in the suburbs lead to a substitution effect that encourages consumers to buy more housing;

lower land rents in the suburbs leads developers to substitute land for building materials, leading to shorter buildings. Both of these facts then jointly imply that population density is falling with distance.