NBER WORKING PAPER SERIES

MARKUPS AND ENTRY IN A CIRCULAR HOTELLING MODEL

Robert J. Barro

Working Paper 32660
http://www.nber.org/papers/w32660

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2024

I have benefited from comments by Ed Glaeser and Elhanan Helpman. I have no sources of funding to disclose. The views expressed herein are those of the author and do not necessarily reflect the views of the National Bureau of Economic Research.

Markups and Entry in a Circular Hotelling Model
Robert J. Barro
NBER Working Paper No. 32660
July 2024
JEL No. L1, L12, L13

ABSTRACT

The Hotelling locational model and its adaptations to a circular city provide a core framework for research in industrial organization. The present paper expands the explanatory power of this model by incorporating a continuum of consumers with constant-elasticity demand functions along with stores that have constant marginal costs of production. The stores are evenly spaced in equilibrium. The model generates a simple formula in which the markup of price over marginal cost depends on the spacing between stores and a transportation-cost parameter but is independent of the elasticity of demand. This result reflects pricing decisions by stores that factor in the threat of losing business entirely at the borders with neighboring stores. The free-entry solutions for the number of stores and their spacing approximate socially optimal values but quantities of goods consumed are inefficiently low.

Robert J. Barro
Department of Economics
Littauer Center 218
Harvard University
Cambridge, MA 02138
and NBER
rbarro@harvard.edu

Hotelling (1929) constructed a model of monopolistic competition in which stores were located along a finite straight line and consumers purchased goods from their nearest store. Because of locational advantages, stores had some degree of monopoly power, and this power entered into their pricing decisions. The present analysis uses a circular version of the Hotelling model, usually labeled as the Salop model and attributed to Salop (1979).[1] A major technical advantage of the circle framework is that it allows for symmetric outcomes by avoiding issues related to the end points of a finite straight line.[2]

The finite straight line is an attractive framework in some contexts because the end points represent extremes of characteristics such as preferences of political parties about the size of government (Downs [1957]) or of religions about the degree of strictness (Barro and McCleary [2005]). However, for most applications in industrial organization, these end-point features do not apply, and the circle model provides a more useful setting.

I. Circular Hotelling Model

The economy, which can be viewed as a city, features consumers distributed uniformly around a circle with circumference H, as shown in Figure 1. The parameter H represents the size of the city. A number N of stores service the consumers. Each store j produces goods at the same constant marginal cost, c. (Hotelling [1929, p. 45] assumed that the marginal cost was

---

[1]Salop (1979) cites Schmalensee (1978) for a prior application of the circular setting. Moreover, the circular framework was used earlier by Vickrey (1964, Ch. 8), who referred to "… a simple loop (which may or may not be circular) …"

[2]Salop (1979, p. 142) says: "The product space of the industry is taken to be an infinite line or the unit-circumference of a circle. While neither assumption is realistic, both allow the 'corner' difficulties of the original Hotelling model to be ignored and an industry equilibrium with identical prices by equally-spaced firms to obtain." Hay (1976, p. 243) relied on a very long straight line: "For most of the analysis we assume that the market is very long in comparison with the market size of the firm. This enables us to avoid the analytic complications of market intervals at the end points of the market."

zero.) Store j prices at P_j. In addition to paying P_j for each unit of goods, a customer pays an amount per unit that increases linearly with the distance, z, from its nearest store, as in Hotelling (1929, p. 45). This transportation cost per unit is represented by tz, where $t>0$ is a transport-cost parameter that depends on transportation technology and the value of customer time. As stressed by Hotelling (1929, p. 45), "transportation cost" should be viewed as a metaphor for many differentiating characteristics of goods and stores that cause customers to prefer one seller over another for a given price charged.[3]

As in Salop (1979), the stores are assumed to be evenly spaced around the circle, as shown in Figure 1. Store 1 is adjacent to stores 2 and N, as shown. Going to the right from its location, store 1 services customers out a distance h, which is the mid-point of the spacing, $2h$, from store 2. Similarly, going to the left, store 1 services customers out a distance h, the mid-point of its spacing from store N. The even spacing of stores is an equilibrium outcome in the present model.

Customers located exactly at distance h to the right of store 1 will be indifferent between buying from store 1 or store 2 and similarly to the left of store 1 for store 1 versus store N. Behavior at these borders will be crucial for the analysis but is neglected initially for expository purposes. The market distance, h, will satisfy

$$h = \frac{1}{2}\frac{H}{N}. \quad (1)$$

For now, the analysis takes the number of stores, N, as given. However, in a full equilibrium, N and, hence, h are determined from a free-entry condition.

[3]Hotelling (1929, p. 45) says: "… there will be many causes leading particular classes of buyers to prefer one seller over another, but the ensemble of such consideration is here symbolized by transportation costs."

Consider the choice of price, P_j, for store *j*, given the prices of other sellers (as in a Bertrand analysis) and the positions of the other stores. In equilibrium, each store will charge the same price, $P_j=P$, and each customer will patronize its closest store. Without loss of generality, take $j=1$ and consider only the market going to the right where store 2 is the relevant alternative to store 1. (The results going to the left, toward store *N*, will be analogous.) The effective price per unit faced by customers at position *z*, where $0 \leq z \leq h$, is

$$P^*(z) = P + tz, \tag{2}$$

where the store subscript, 1 in this case, is omitted. Equation (2) indicates that a customer located at *z* pays the price *P* set at store 1's headquarters and also pays the transport cost, *tz*.[4]

Denote by *q(z)* the quantity demanded of store 1's goods by customers located at distance *z* from the store. This demand is assumed to take a constant-elasticity form:[5]

$$q(z) = A \cdot [P^*(z)]^{-\eta}, \tag{3}$$

where $\eta \geq 0$.[6] The standard analysis requires $\eta > 1$, but this restriction is ultimately unnecessary in the present model. The parameter $A>0$ represents the scale of consumer demand.

---

[4]An issue with literal transportation costs is that some of these costs, such as for travel time, would be independent of the amount purchased. These fixed costs for shopping are difficult to incorporate into the present model because the quantity bought, *q(z),* is modeled as an infinitesimal flow. In this respect, the model may be more appealing when "transportation costs" are interpreted broadly to encompass costs from consuming a good with characteristics that deviate from a customer's ideal type. As noted by Schmalensee (1978, p. 309), " … the formal correspondence between Lancastrian models with two characteristics and one-dimensional spatial models is almost exact." See Baumol (1967) and Lancaster (1975) in this context.

[5]Equation (3) is assumed to reflect substitution effects from changes in *P*(z)* and neglects income effects. In fact, customers located at higher *z* will effectively have lower income because they spend more on transportation costs for a given *q(z)*. However, this income difference would be exactly offset by lower residential rents if the only dimension of residential location that matters is its distance from the nearest store. This result accords with the standard insight in urban economics, summarized in Brueckner (1987), whereby differences in rents represent compensating differentials for differences in commuting costs.

[6]Hotelling (1929, pp. 45, 56) effectively assumed $\eta=0$. He says (p. 56): "The problem … might be varied by supposing that each consumer buys an amount of the commodity in question which depends on the delivered price. If one tries a particular demand function the mathematical complications will now be considerable … " Smithies (1941) and Hay (1976) extended the Hotelling analysis to allow for elastic consumer demand, though in the form of linear demand functions. These analyses are concerned mostly with locational decisions by firms.

A modeling device that turns out to be highly convenient treats *q(z)* as contributing an infinitesimal amount to the total quantity of goods demanded, *Q*, from store 1; that is,

$$\frac{dQ}{dz} = q(z).$$

The total quantity, *Q*, sold by store 1 to customers to the right of the store's location equals the integral of *q(z)* for *z* going from *0* to *h*:

$$Q = \int_0^h q(z)dz.$$

Using the demand curve from Eq. (3) and making a change of variable from *z* to *P*(z)*, the integral can be evaluated to get

(4) $$Q = \frac{A}{t(\eta-1)}[P^{1-\eta} - (P+th)^{1-\eta}],$$

where $\eta>1$ is assumed at this stage.

The marginal effect of *P* on *Q* can be calculated from Eq.(4) as

(5) $$\frac{\partial Q}{\partial P} = -\left(\frac{A}{t}\right)\cdot\left[P^{-\eta} - (1+t\frac{\partial h}{\partial P})(P+th)^{-\eta}\right].$$

The essence of the Hotelling analysis is the term $\frac{\partial h}{\partial P}$, which indicates how an increment to *P* affects the location of the border between stores. However, to see the relation to standard results on markup pricing, the analysis assumes to begin that $\frac{\partial h}{\partial P} = 0$.

The profit flow for store 1 (for sales to the right in Figure 1) is

(6) $$\pi = (P-c)Q - f,$$

where $f > 0$ is the fixed cost of operating a store. (The fixed cost could correspond to the rent paid on the store's premises. In the subsequent analysis, this fixed cost also covers sales to the left of store 1 in Figure 1.) The first-order condition for choosing P to maximize π follows from Eq. (6) as

(7) $$\frac{\partial\pi}{\partial P} = Q + (P-c)\frac{\partial Q}{\partial P} = 0.$$

Substituting into Eq. (7) for Q from Eq. (4) and for $\frac{\partial Q}{\partial P}$ from Eq, (5) (assuming $\frac{\partial h}{\partial P} = 0$) leads, if $\eta>1$, to the first-order maximization condition:

(8) $$\left(\frac{1}{\eta-1}\right)\cdot[P^{1-\eta} - (P+th)^{1-\eta}] = (P-c)[P^{-\eta} - (P+th)^{-\eta}].$$

The results simplify if the maximal operative transport cost, *th*, is small compared to the price, *P*; that is, if the main component of the effective price, *P*(z)* in Eq. (2), is always the price charged, *P*, rather than the transport cost, *tz*. Ultimately, the required condition will be that the marginal production cost dominates over transport cost, so that $th \ll c$. This condition allows for simplifications of the terms involving *P+th* in Eq. (8) by means of first-order Taylor approximations. These approximations deliver the standard result, described originally in Lerner (1934), whereby *P* is a constant markup on marginal cost, *c*:[7]

(9) $$P \approx c\cdot\left(\frac{\eta}{\eta-1}\right),$$

so that the markup ratio is

(10) $$\frac{P-c}{c} \approx \frac{1}{\eta-1}.$$

Therefore, when the effect of *P* on the location of the border is neglected, the markup ratio is determined by the elasticity of demand, η, with a higher elasticity implying a lower markup ratio. Finite price, *P*, and profit, π, require $\eta>1$.

---

[7]The second-order condition for profit maximization is satisfied here if $\eta>1$.

The price, P, shown in Eq. (9) is the one chosen by store 1. However, the same price applies for each store around the circle in Figure 1 as long as the marginal cost, c, and the elasticity of demand, η, are the same in each territory. The analysis now turns to the crucial influence on pricing behavior from the interactions at the borders between stores.

II. Hotelling Effect at the Borders

At the border between stores 1 and 2, where $z=h$ in Figure 1, customers are indifferent between buying goods from store 1 or store 2 when the prices are the same. Starting from a position of equal prices and holding fixed store 2's price, a positive increment in store 1's price, $dP>0$, would motivate customers located at the border to make all of their purchases from store 2.[8] (Similarly, a negative change, $dP<0$, would induce the opposite shift.) Since $q(h)$ is the quantity bought, the amount $q(h)\cdot dP$ is the extra cost imposed on buyers located at the border. In response, the border has to shift leftward toward store 1; that is $dh<0$ applies. This change corresponds to a reduction by $tq(h)\cdot dh$ in the amount paid on transport costs by customers at the border. In order for persons located at the new border to be indifferent between buying from store 1 or store 2, the magnitude of $tq(h)\cdot dh$ must equal that of $q(h)\cdot dP$. That is, $tq(h)\cdot dh + q(h)\cdot dP=0$ must hold. Therefore, the key condition that connects the position of the border to the increment in price is:

(11) $$\frac{\partial h}{\partial P} = -\frac{1}{t}.$$

[8]Since the quantity $q(h)$ is modeled as a flow of infinitesimal size, it might seem that this effect is negligible. In fact, an infinitesimal increase in price results in the loss of the full infinitesimal quantity of sales at the border, and this overall effect is not negligible.

Substituting the result from Eq. (11) into the expression for $\frac{\partial Q}{\partial P}$ in Eq. (5) results in the cancellation of the term that involves $1 + t\frac{\partial h}{\partial P}$ on the right side to get:

(12) $$\frac{\partial Q}{\partial P} = -(\frac{A}{t}) \cdot P^{-\eta}.$$

Substitution of this result into the expression for $\frac{\partial \pi}{\partial P}$ in Eq. (7) yields the new first-order maximization condition:

(13) $$\left(\frac{1}{\eta-1}\right) \cdot [P^{1-\eta} - (P + th)^{1-\eta}] = (P - c)P^{-\eta}.$$

If we again assume that *th* is small compared to *P* (and *c*), Eq. (13) simplifies using a first-order Taylor approximation for the term involving *P+th* to

(14) $$P \approx c + th.$$

The markup ratio is then

(15) $$\frac{P-c}{c} \approx \frac{th}{c}.$$

The second-order condition for profit maximization is satisfied if $\eta \geq 0$.

There is a sense in which the result for the markup ratio in Eq. (15) accords with the standard formula based on the elasticity of demand in Eq. (10). The standard result holds if η is interpreted not as the parameter in the individual demand curve in Eq. (3) but rather as the full magnitude of the elasticity of *Q* with respect to *P*—including the extensive margin whereby a store loses all of the business at its borders by raising its price. In this setting, the micro-level elasticity of demand, η—effectively the intensive margin for quantity demanded—does not affect the markup ratio (as an approximation when $th<<c$ holds).[9] Note also that the solution for the price markup in Eqs. (14) and (15) holds even when $\eta \leq 1$.[10]

In the equilibrium, where all stores price according to Eq. (14), the value *th* on the right-hand side is the maximum markup on marginal cost that an individual store can employ without losing all of its business at its borders. Moreover, when $th<<c$, the borders are close enough to a store's location to be effective constraints on its pricing strategy. That is, while each store might want to employ a higher markup (related to the elasticity of demand, η), the threat of losing border customers to neighboring stores makes this option unattractive. If stores could collude to carve out customer territories (so that $\frac{\partial h}{\partial P} = 0$ holds for each store), the pricing strategies and the equilibrium markups would be different (and would correspond to the standard results in Eqs. [9] and [10]).

Return now to the assumption that stores are evenly spaced around the circle in Figure 1. Suppose that store 1 takes as given the positions of stores 2 and *N*. Suppose, starting from a position equidistant between stores *N* and 2, that store 1 considers moving its location, say by a distance *X* to the right. The market border with store *N* would then shift from a distance *h* to a distance $h+X/2$ from store 1's location. The border with store 2 would shift from a distance *h* to a distance $h-X/2$. Therefore, the total distance covered by store 1 would remain at $2h$. However, there is a shift toward customers who are relatively far from the store (in the region toward store *N*) and away from those who are relatively close (in the region toward store 2). Because of

---

[9]A second-order Taylor approximation for the term involving $P+th$ in Eq. (13) results in the condition $P \approx c + th \cdot (1 - \frac{1}{2}\eta th/P)$. In this case, as in Eq. (10), a higher η results in a smaller markup ratio, $(P-c)/c$. However, this effect is negligible if $th<<P$.

[10]The key mechanism here is that a rise in *P* results in a contraction of the border, *h*, with *h* reaching zero for sufficiently high *P*. When $\eta \leq 1$, this mechanism prevents a store from achieving infinite profit by charging an infinite price.

the downward-sloping demand curve in Eq. (3), the more distant customers buy a smaller quantity and are less profitable for store 1. Hence, on net, store 1's profit declines. It follows that store 1 would not move and is best off remaining equidistant between stores 2 and N. That is, the equal-spacing pattern is an equilibrium outcome.[11] Note that this result depends on the downward-sloping demand curve—the result would not apply under the common assumption that each household buys exactly one unit of the good (for example, in the version of the Salop model described in Tirole [1988, Section 7.1.2]).

The results in Eqs. (14) and (15) are reminiscent of models of contestable markets, as described by Baumol, Panzar, and Willig (1982). In those models, an individual firm's markup is limited by the cost of production for potential entrants into a market, and marginal-cost pricing prevails under assumptions that include zero costs of entry and exit and equal access of all firms to all technologies. A difference in the present, Hotelling context is that the relevant competition depends on actual participants in the market—neighboring stores—rather than hypothetical entry by potential participants. Moreover, the Hotelling-type model does not predict marginal-cost pricing but rather price markups that depend on the distances between active stores in the market.

III. Free-Entry Condition

Given store 1's price, $P = c + th$ from Eq. (14), the total quantity sold, Q, by store 1 (or any store) can be determined from Eq. (4) to be:

(16) $$Q \approx 2hA(c + th)^{-\eta},$$

where the 2 reflects store 1's sales on both sides of its location in Figure 1. The associated profit is given from Eq. (6) by

---

[11]This conclusion accords with Vickrey (1964, p. 330).

(17) $$\pi \approx 2tAh^2 \cdot (c+th)^{-\eta} - f.$$

If $\pi > 0$, there is an incentive for new stores to enter the market, thereby raising *N* and lowering $h = \frac{1}{2}\frac{H}{N}$.[12] The decrease in *h* lowers π in Eq (17) (for sure if *th*<<*c*, the condition that was used in deriving Eq. [17]). This process continues as long as $\pi > 0$. If the integer constraint on *N* can be neglected, as will be satisfactory if *N* is large,[13] the free-entry condition will be $\pi \approx 0$. Equation (17) then implies

$$h^2 \approx \frac{f}{2tA(c+th)^{-\eta}}.$$

The derivation has already assumed *th*<<*c*, which implies the further approximation:

(18) $$h \approx \sqrt{\frac{f}{2tAc^{-\eta}}}.$$

Therefore, under free entry, the spacing between stores, which equals *2h*, follows a square-root rule, whereby *h* is larger the lower the transport cost per unit, *t*, the higher the fixed cost f of operating a store, and the smaller the quantity that would be sold to each buyer, $Ac^{-\eta}$, if buyers paid the marginal cost, *c*, for their goods.[14] The last two effects arise because the fixed cost is effectively scaled by the quantity sold to each buyer. Note that *h* does not depend on the overall size of the city, represented by *H* in Figure 1.

The solution for market size, *h,* in Eq. (18) implies that the number of firms is given from Eq. (1) by

---

[12]The assumption here is that firms can immediately relocate to preserve their equal spacing around the circle in Figure 1. More realistically, the model applies to a long-run, steady-state situation, rather than to the dynamics of entry and exit.

[13]For example, for New York City, with a central city population of 8 million in 2019, the New York State Comptroller reports that the number of retail businesses at that time was about 33,000, each servicing on average 242 domestic residents. In this context, where *N* is *33,000*, the integer constraint would be of no consequence. Mankiw and Whinston (1986) discuss integer constraints for a related model.

[14]The condition *th*<<*c* requires $c^2 \gg \frac{tf}{2Ac^{-\eta}}$.

(19) $$N = \frac{H}{2h} \approx H \cdot \sqrt{\frac{tAc^{-\eta}}{2f}}.$$

Therefore, N is proportional to city size, H, and rises with the square root of the transport cost per unit, t, and the quantity sold to each buyer, $Ac^{-\eta}$, if buyers paid the marginal cost, c, for their goods. The number N varies inversely with the square root of the fixed cost, f, of operating a store.

The free-entry result for market size, h, in Eq. (18) implies from Eq. (14) that the price charged will be

(20) $$P \approx c + \sqrt{\frac{tf}{2Ac^{-\eta}}}.$$

That is, the markup on marginal cost, c, is increasing with the transport-cost parameter, t, and the fixed cost f of operating a store and decreasing with the quantity sold to each buyer, $Ac^{-\eta}$, if buyers paid the marginal cost, c, for their goods. The markup does not depend on H, the overall size of the city. The markup also does not depend on the elasticity of demand, η, except for a positive effect that involves the quantity sold, $Ac^{-\eta}$. Moreover, the results hold for any $\eta \geq 0$, including Hotelling's case where $\eta = 0$.

The theoretical results can be related to the empirical analysis of Chevalier, Kashyap, and Rossi (2003) concerning pricing patterns of a major supermarket chain. Their key finding is that markups for affected categories of products are relatively low at times of peak demand, notably during major holidays and events such as Christmas, Thanksgiving, and Lent. Their empirical results rule out an explanation for this pricing pattern based on the elasticity of demand being unusually high at times of peak demand. This finding accords with the present model in that the price markup is not predicted to fall when the demand elasticity, η, rises in Eq.(14) for a given market size, h, or in Eq. (20), where h is determined by a free-entry condition. In the model, peak demand would be represented by a temporarily high value of the parameter *A*, which enters into the demand function in Eq. (3). High *A* has no effect on the price markup in Eq. (14) but reduces this markup in Eq. (20). An application of this last result to the setting of Chevalier, Kashyap, and Rossi (2003) depends on there being significant entry during peak demand periods. Possibly this entry can take the form not of new stores but of longer store hours or increased advertising (a margin emphasized by Chevalier, Kashyap, and Rossi).

IV. Socially-Optimal Entry

Heuristically, there are two elements in the model that may cause the free-entry choices of spacing, $2h$, and number of stores, N, from Eqs. (18) and (19) to deviate from socially-optimal values. The first "distortion" is the markup, th, which corresponds to an excess of the effective price, P^*, over social marginal cost, $c+tz$, in Eq. (3). This "excessive" price leads to quantities of goods consumed that fall short of socially-optimal values. That is, at the existing quantity sold, each household's willingness to pay, P^*, for an additional unit—corresponding to the inverse-demand function implied by Eq. (3)—exceeds the social marginal cost, $c+tz$. Therefore, the profit flow in Eq. (6), which is the signal for entry, is too low in the sense of falling short of the amount that corresponds to the socially-optimal quantity of goods produced. On this ground, one would expect the free-entry choice of the number of stores, N, to be too low, corresponding to the choice of spacing, $2h$, being too high.

The second "distortion" is that an entering firm's profit includes revenue that is transferred from incumbent firms. That is, there is a "business-stealing effect,"[15] which is a

[15]This term appears in Mankiw and Whinston (1986, p. 49) and also in Tirole (1988, p. 284). The business-stealing effect plays a large role in growth models that feature creative destruction, as in Aghion and Howitt (1992). The term creative destruction was used by Schumpeter (1942, p. 83): "This process of Creative Destruction is the essential fact about capitalism. It is what capitalism consists in and what every capitalist concern has got to live in."

private reward for entry that has no social benefit. (The distortion here can be viewed as generated by the lack of property rights for incumbent firms in their profit flows.) On this ground, one would expect the free-entry choice of the number of stores, *N*, to be too high, corresponding to a choice of spacing, *2h,* that is too low.

Tirole's (1988, Ch. 7, p. 284) analysis of the Salop-Hotelling model finds that entry is excessive in that model: "… we compare the free-entry equilibrium with the allocation selected by a social planner. We already know that the price charged by the firms is greater than the marginal cost. However, in this case, where consumers all receive the same utility from the good and each consumes only one unit, this price introduces no distortion." Tirole's conclusion that the free-entry equilibrium features too much entry applies because only the distortion from business-stealing was considered.

Suppose in the present model that the social planner can dictate the spacing between stores, *2h,* and the quantity of goods purchased, *q(z)*, by each household. However, the planner is assumed to accept the uniform distribution of consumers around the circle in Figure 1. Optimality entails the equation of each customer's willingness to pay, corresponding to the demand price from the inversion of Eq. (3), to the social marginal cost, $c+tz$. The socially-optimal value of *h* can then be derived, following the approach of Mankiw and Whinston (1986, p. 50), by maximizing an expression for aggregate consumer surplus (netting out costs of producing and transporting goods and the fixed costs of operating stores). The result, derived in the appendix, is that the social planner's choices of store spacing, *2h,* and number of stores, *N*, correspond to the free-entry choices given in Eqs. (18) and (19). (The

---

Schumpeter's analysis of creative destruction was heavily influenced in Schumpeter (1942, part I) by his reading of Karl Marx. It is unclear whether Schumpeter's vision was preceded by Picasso's famous quote: "Every act of creation begins with an act of destruction."

social-planner and free-entry solutions are each approximations based on the condition $th \ll c$.) The key point is that the two distortions cancel out as an approximation when $th \ll c$ holds, and there is no tendency in this model for entry of stores to be too high or too low from a social perspective.

A difference between the social-planner and free-entry outcomes is that the quantity $q(z)$ consumed in each location in the social-planner's solution exceeds that in the free-entry case. In the social-planner's problem, $q(z)$ is the quantity given by the demand curve in Eq. (3) when the effective price, P^*, equals the social marginal cost, $c+tz$, which is inclusive of costs of production, c, and transport costs, tz. In the free-entry equilibrium, $q(z)$ equals the quantity given by the demand curve in Eq. (3) when $P^*=c+tz+th$, which also includes the markup, th.

In a decentralized setting, the social optimum could be achieved by implementing the appropriate subsidy on purchases of goods, financed by lump-sum taxes. No subsidy/tax would apply to entry of stores. In practice, the goods subsidy would be difficult to implement when there is heterogeneity in markups. Moreover, lump-sum taxes are unlikely to be available.

Free entry and socially-optimal entry have also been studied in non-spatial models of monopolistic competition, which include Spence (1976), Dixit and Stiglitz (1977), and Mankiw and Whinston (1986). In this literature, free entry sometimes results in the socially optimal amount of entry, but this result is not general. However, an important limitation of these models is that the various products are treated symmetrically, with each good equally substitutable with each alternative good. In contrast, the key to Hotelling-type models and related Lancaster-type models is the strong substitution with "neighboring" producers and products. This feature underlies the markup-pricing result in Eq. (14), which is driven by the threat of losing business at the borders with adjacent stores. This formula for markup pricing is, in turn, a key ingredient in the analyses of free entry and socially-optimal entry.

V. Summary

In a circular Hotelling model, customers are uniformly distributed around a circle with circumference H and number of stores N. Equal spacing of stores applies in equilibrium. A customer's transportation cost per unit is tz, where z is the distance from the nearest store. The effective price paid is $P^*=P+tz$, where P is the price at the store's location. The marginal cost of production is constant, and the elasticity of each customer's quantity demanded with respect to P^* is constant.

Stores serve customers out a distance $h=H/2N$. If production costs dominate transport costs, each store's equilibrium price approximates the markup th over marginal cost, independent of the elasticity of demand. The key element in a store's markup is the threat of losing all of the business at its borders. In a free-entry equilibrium, h is larger the lower t, the higher the fixed cost of operating a store, and the smaller the scale of quantity demanded by each customer. Given these factors, the number of stores, N, is proportional to city size, H, and the markup is independent of H.

Two distortions affect the equilibrium: the excess of price over marginal cost implies inefficiently low quantities of goods produced, and a business-stealing effect implies that a new entrant's profit includes revenue transferred from incumbent stores. These distortions have opposing effects on entry and approximately cancel out; hence, the number of stores and their spacing approximate socially-optimal values. However, quantities consumed are inefficiently low.

References

Aghion, Philippe and Peter Howitt (1992). "A Model of Growth through Creative Destruction," *Econometrica*, 60, 323-351.

Barro, Robert J. and Rachel M. McCleary (2005). "Which Countries Have State Religions?" *Quarterly Journal of Economics*, 130, 1331-1370.

Baumol, William J. (1967). "Calculation of Optimal Product and Retailer Characteristics: The Abstract Product Approach," *Journal of Political Economy*, 75, 674-685.

Baumol, William J., John C. Panzar, and Robert D. Willig (1982). *Contestable Markets and the Theory of Industrial Structure*, New York, Harcourt Brace Jovanovich.

Brueckner, Jan K. (1987). "A Unified Treatment of the Muth-Mills Model," in Edwin S. Mills, ed., *Handbook of Regional and Urban Economics*, Amsterdam, Elsevier.

Chevalier, Judith A., Anil K. Kashyap, and Peter E. Rossi (2003). "Why Don't Prices Rise during Periods of Peak Demand? Evidence from Scanner Data," *American Economic Review*, 93, 15-37.

Dixit, Avinash K. and Joseph E. Stiglitz (1977). "Monopolistic Competition and Optimum Product Diversity," *American Economic Review*, 67, 297-308.

Downs, Anthony (1957). "An Economic Theory of Political Action in a Democracy," *Journal of Political Economy*, 65, 135-150.

Hay, Donald A. (1976). "Sequential Entry and Entry-Deterring Strategies in Spatial Competition," O*xford Economic Papers*, 28, 240-257.

Hotelling, Harold (1929). "Stability in Competition," *Economic Journal*, 39, 41-57.

Lancaster, Kelvin (1975). "Socially Optimal Product Differentiation," *American Economic Review*, 65, 567-585.

Lerner, Abba P. (1934). "The Concept of Monopoly and the Measurement of Monopoly Power," *Review of Economic Studies*, 1, 157-175.

Mankiw, N. Gregory and Michael D. Whinston (1986). "Free Entry and Social Inefficiency," *Rand Journal of Economics*, 17, 48-58.

Salop, Steven C. (1979). "Monopolistic Competition with Outside Goods," *Bell Journal of Economics*, 10, 141-156.

Schmalensee, Richard (1978). "Entry Deterrence in the Ready-to-Eat Cereal Industry," *Bell Journal of Economics*, 9, 305-327.

Schumpeter, Joseph A. (1942). *Capitalism, Socialism, and Democracy*, New York, Harper & Brothers Publishers.

Smithies, Arthur A. (1941). "Optimum Location in Spatial Competition," *Journal of Political Economy*, 49, 423-439.

Spence, Michael (1976). "Product Selection, Fixed Costs, and Monopolistic Competition," *Review of Economic Studies*, 43, 217-235.

Tirole, Jean (1988). *The Theory of Industrial Organization*, Cambridge MA, MIT Press.

Vickrey, William S. (1964). *Microstatics*, New York, Harcourt, Brace and World.

Appendix

Social-Planner Problem

The social planner effectively prices goods at the social marginal cost, which is $c+tz$ for a household located at distance z from its nearest store. The associated socially optimal quantity of goods is $q(z) = A \cdot (c + tz)^{-\eta}$ from Eq. (3). The consumer surplus, $\Omega(z)$, for an agent at z is then

$$\Omega(z) = \int_0^{q(z)} \left[\left(\frac{A}{q}\right)^{1/\eta} - c - tz \right] dq, \tag{A1}$$

where $\left(\frac{A}{q}\right)^{1/\eta}$ is the demand price associated with q. The integral can be evaluated to get

$$\Omega(z) = \frac{A(c+tz)^{1-\eta}}{\eta-1}, \tag{A2}$$

where $\eta>1$ is assumed at this stage (but is ultimately an unnecessary restriction).

The consumer surplus given in Eq. (A2) applies for $z=(0,h)$ to the right of store 1 in Figure 1. Including also the analogous term to the left of store 1, the total consumer surplus for all customers of store 1 is

$$\Omega = 2\int_0^h \frac{A(c+tz)^{1-\eta}}{\eta-1} dz. \tag{A3}$$

Using a change of variable from z to $c+tz$, the integral can be evaluated to get

$$\Omega = \frac{2A}{t(\eta-1)(\eta-2)} [c^{2-\eta} - (c + th)^{2-\eta}]. \tag{A4}$$

The analysis at this stage assumes $\eta>2$, but this restriction (or $\eta>1$) is ultimately not needed for the results.

The consumer surplus in Eq. (A4) applies from Eq. (1) to $N=H/2h$ markets. The fixed cost of operations is Nf. The social planner chooses h to maximize the overall net surplus:

$$\text{(A5)} \quad Net\ consumer\ surplus = \frac{H}{h}\left\{\frac{A}{t(\eta-1)(\eta-2)}\left[c^{2-\eta}-(c+th)^{2-\eta}\right]-\frac{f}{2}\right\}.$$

Setting the derivative of Eq. (A5) with respect to h to zero to get the first-order maximization condition leads to

$$\frac{f}{2}=\frac{A}{t(\eta-1)(\eta-2)}\left[c^{2-\eta}-(c+th)^{2-\eta}\right]-\frac{Ah}{\eta-1}(c+th)^{1-\eta}.$$

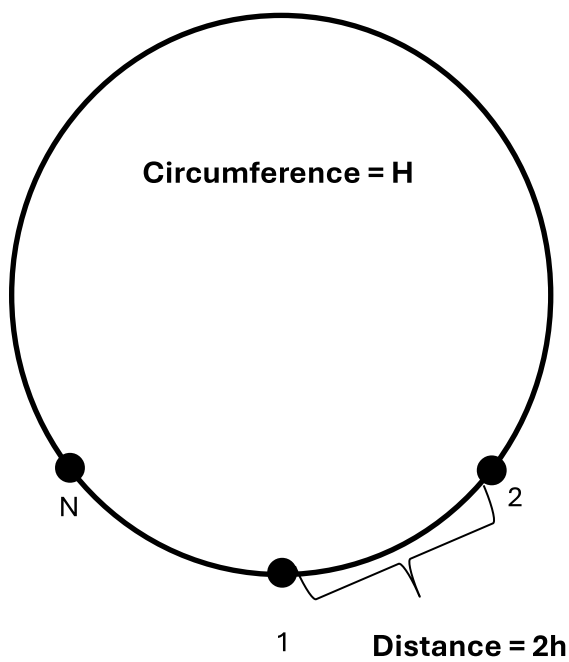
Using the condition from before, $th<<c$, the right-hand side simplifies as an approximation to $Ath^2c^{-\eta}$. Therefore, the approximate solution for h is

$$\text{(A6)} \qquad h \approx \sqrt{\frac{f}{2tAc^{-\eta}}},$$

which is the same as Eq. (18). Therefore, subject to the condition $th<<c$, the free-entry solution generates the socially optimal amount of entry.

Figure 1

Salop-Hotelling Circular City



Note: Consumers are uniformly distributed around a circle with circumference *H*. Stores are located at positions 1, 2, …, *N*, with a spacing of *2h*. Store 1's market extends out a distance *h* to the right and left of its location.