

Spatial competition in a city: Hotelling revisited in an Alonso-Mills-Muth city*

Takaaki Takahashi[†]

September 2012

Abstract

The purpose of this paper is to show that a strategic interaction between firms is one of the key factors that determine the spatial structure of a city. In this regard, we expand the Hotelling model of a linear city with two retail firms to incorporate some aspects of an urban spatial structure, which are characteristic of the Alonso-Mills-Muth model. We show that in an equilibrium, the two firms are either agglomerated at a city center or dispersed at two outermost feasible locations. The outcome that arises depends on the values of parameters, especially on the relative size of commuting costs to the city center and shopping costs to a retail firm.

Keywords: agglomeration; city limits; commuting costs; land rent; linear city; maximum dispersion; shopping costs

JEL Classification Numbers: L13; R10; R32

*This research is partly supported by a Grant in Aid for Research (No. 20330045) by the Ministry of Education, Science and Culture in Japan.

[†]Center for Spatial Information Science, The University of Tokyo. E-mail address: takaaki-t@csis.u-tokyo.ac.jp

1 Introduction

One of the most marked changes in the spatial structure of cities from the beginning of the 20th century is decentralization. Both employment places and residences have been shifting from city centers to suburbs in many advanced countries. Studying this change is important not only because it is salient and omnipresent but also because it is a source of some serious economic problems that need to be solved. In many cities, for example, an increasing number of commercial activities are fleeing an old city centers toward suburban shopping centers including mega malls, which is followed by a flight of population. It is not rare that low income households are left with lesser job opportunities in an “inner city,” which is often characterized by low quality of education and a high crime rate. In contrast, in suburbs, which are represented by “edge cities,” the problem of congestion, especially road congestion, is becoming more and more intolerable.

In the field of urban economics, a classical Alonso-Mills-Muth model proves effective in explaining this change. A rigorous analysis of this model demonstrates how a city expands geographically as a result of a decline in commuting costs, which has been under way since the last century mainly because of the rise of automobiles (see review articles by Anas et al. (1998) and Glaeser and Kahn (2004), and the literature cited there).

Having said that, decentralization is accompanied with another change, namely, a transformation from a monocentric city to a polycentric city. For this change, many researchers have provided explanations. Some discuss the emergence of a subcenter at *an exogenously given location* (see White (1976), Yinger (1992), and Ross and Yinger (1995), for instance). Others examine not only when a subcenter emerges but also where it is formed (see Ogawa and Fujita (1980), Fujita and Ogawa (1982), Helsley and Sullivan (1991), and Henderson and Mitra (1996), among others).¹ These studies successfully unveil various factors that are responsible for the change toward polycentricity. However, one factor is left undiscussed; the strategic interaction between firms. The purpose of this paper is to show that the strategic interaction plays a key role in determining whether a city becomes monocentric or polycentric.

For this purpose, we construct a model of a linear city. Workers, distributed uniformly along the line, work either in a basic sector or in the retail sector. The basic sector produces goods consumed outside the city in a competitive market. It is located at the origin of the line, which is referred to as a central business district (CBD). In contrast, the retail sector consists of two firms. When a retail firm is located at a point other than the CBD, we can consider that the firm constitutes a “subcenter” of the city. Here, it is possible to interpret each retail “firm” as a group

¹First, Ogawa and Fujita (1980) and Fujita and Ogawa (1982) explore the land use pattern in a city without specifying the locations of employment and residence a priori. They show that “subcenters” where land is used only for business can develop apart from a CBD, depending on the values of parameters, particularly, commuting costs for workers and transaction costs between firms. Second, Helsley and Sullivan (1991) examine the location of a subcenter chosen by a social planner. Finally, Henderson and Mitra (1996) analyze behaviors of the developer who constructs a subcenter so as to reap capital gains from a rise in the price (rent) of nearby lands.

of firms such as those in a traditional commercial district at a city center and those in a suburban shopping center. Workers are to visit either of these retail firms to buy a variety of goods, paying a certain amount of transport costs. The two firms compete with each other by choosing their locations. When both of them decide to settle in the CBD, the city becomes monocentric: instead, when at least one firm chooses a location apart from it, the city becomes polycentric.

At this point, one would notice a resemblance between our model and the model of spatial competition by Hotelling (1929). This is true, but a linear market in the Hotelling model lacks two important features of an urban spatial structure, to which considerable attention is paid in our model, to the contrary.

The first is a variation in input prices over space. For one thing, one of the most renowned facts concerning an urban spatial structure is that land rent constantly decreases with a distance from the CBD. In contrast to the Hotelling model that does not consider land rent or, at best, assumes a fixed land rent, our model contains a process of determination of land rent in the manner of the Alonso-Mills-Muth model. That is, land rent is determined through bidding by workers who use land for residential use. At equilibrium, rent decreases as distance from the CBD increases, so as to offset the higher commuting costs to the CBD. To this extent, retail firms have an incentive to choose a location more distant from the CBD *ceteris paribus*, provided that their production activities need land as an input.

A similar remark applies to wage. The Hotelling model does not consider a possible difference in wage over space. In our model, in contrast, retail firms offer a lower wage when they are located farther away from the CBD. The reason is simple. Those who live next to a retail firm can save the commuting costs to the CBD by working for it rather than working in the basic sector. Consequently, they are willing to work for the retail firm only when the decline in wage is more than offset by the savings in commuting costs. This also gives retail firms an incentive to choose a location remote from the CBD.

The second feature missing in the Hotelling model is that the geographical limits of a city are not fixed but depend on the locations of firms. As is evident from a casual observation of the development of edge cities, the emergence of subcenters in suburbs pushes away city limits outward. Whereas city limits are fixed in the Hotelling model, they are, in our model, determined at the points where the land rent for urban (residential) use so decreases as to reach the value of the land rent for alternative (agricultural) use again in the manner of the Alonso-Mills-Muth model. Now suppose that the left retail firm relocates farther leftward. If the left city limit is fixed as in the Hotelling model, the market area of that firm will contract. If it is allowed to move, in contrast, it will move farther leftward. As a result, the market area will shrink only to a smaller degree, or might expand when the effect of the shift of the city limit is dominating. Thus, taking into account the endogenous determination of city limits results in more incentive for firms to settle in suburbs.

To sum up, our study is an attempt to reconstruct the Hotelling model in a more realistic set-

ting of a city, namely, the setting with a variation of input prices over space and the endogenous determination of city limits. Furthermore, because these two features are, as the above explanation indicates, characteristic of the Alonso-Mills-Muth model, one can say that the Hotelling model is unified with the Alonso-Mills-Muth model in this study.

Our main findings are summarized as follows. At an equilibrium, two retail firms are either agglomerated at the CBD or located dispersedly at the two extreme points that are as distant from the CBD as possible (“outermost feasible locations”): there exists no possibility of symmetric dispersion at interior points or asymmetric dispersion. A key factor that determines which pattern emerges is the relative size of the two transport costs, commuting costs to the CBD and shopping costs to a retail firm. If the former costs decline more rapidly than the latter costs over time, an equilibrium pattern may, depending on other parameters, switch from the agglomeration at the CBD to the symmetric dispersion in suburbs. In this case, a monocentric city is transformed into a polycentric city.

Two studies are closely related to this paper. First, Fujita et al. (1997) examine the location of a big firm (an entrant) in a city when all the other firms (incumbents) are concentrated in the CBD. The entrant strategically competes against the incumbents in a labor market. Although their concern is closest to ours, there are two differences. First, we examine the competition in a market of an output and not in a market of an input (labor). Moreover, in our model, both the locations of the two retail firms are variables while in theirs, only the location of the entrant is a variable (the incumbents are presumed to be located at the CBD). Second, the basic setting in our model is the same as that of Lai and Tsai (2008), who analyze the location of a monopolist in a linear city. They introduce land rent into the Hotelling model in the same way as we do. Nonetheless, they are not interested in the strategic interaction between firms.

The rest of the paper is divided into four sections. In the next section, we present a basic model. In Section 3, behaviors of firms are examined. Section 4 characterizes an equilibrium. The last section concludes.

2 Model

2.1 Basic settings

We consider a linear city surrounded by a rural area. The city extends in the area where land rent for residential use is at least as high as agricultural land rent, which is assumed to be fixed at 0. The locations of left and right city limits are denoted by $\underline{l} \leq 0$ and $\bar{l} \geq 0$, respectively.

There are two sectors. A basic sector produces export goods, which are consumed outside the city. Firms in that sector are concentrated in a central business district (CBD) located at $x_0 \equiv 0$. They pay an equal wage, denoted by w_0 , which is given for them.² The other sector is a retail

²The wage is determined by the clearing of the competitive export goods market in the entire economy.

sector, which consists of two firms selling a composite good to workers in the city at a given price. Each “firm” can be interpreted as a group of firms such as those in a traditional commercial district at a city center and those in a suburban shopping center. The locations of the retail firms are denoted by x_1 and x_2 . We assume that they are not allowed to settle outside the city, probably because of land use controls. This requires that the firms be located in a “feasible location space,” $[x, \bar{x}]$, which will be derived later. The wage rates paid by the retail firms, which depend on their locations, are denoted by w_1 and w_2 .

Workers are employed either in the basic sector or at one of the two retail firms. To go to work, they need to pay a commuting cost, which is proportional to distance. For a worker living at l , the cost in a unit period is equal to $t|l - x_j|$ with $t > 0$, where x_j is the location of her workplace ($j \in \{0, 1, 2\}$).

Workers consume a composite good and land. For simplicity, we assume that each of them consumes one unit of land irrespective of land rent. This implies that they are uniformly distributed with unit density. Furthermore, to buy a composite good, workers need to visit a retail firm. The cost for the shopping trip is also assumed to be proportional to distance. More specifically, the shopping cost in a unit period for a worker at l is equal to $s|l - x(l)|$ with $s > 0$, where $x(l)$ is the location of the retail firm that the worker visits. Here, we assume that $t > s$: the unit commuting cost per mile and per month is higher than the unit shopping cost per mile and per month, say, the former is the higher.³

It follows from these arguments that a budget constraint for a worker who lives at l and works at x_j is given by

$$pz + r(l) + t|l - x_j| + s|l - x(l)| = w_j. \quad (1)$$

Here, p and $r(l)$ are the price of the composite good and land rent at l , respectively.

The city is small and open. That is, workers enter to and exit from it until their utility levels become equal to a given level prevailing in the rest of the economy. This is one of the two standard approaches in the Alonso-Mills-Muth model. (The other is a closed city approach in which the population of the city is fixed.) By taking a unit of the composite good appropriately, we can pin down this given level of utility to the one that a worker enjoys when consuming one unit of composite good and one unit of land. In other words, we normalize the composite good so that each worker in the city consumes one unit of it in addition to one unit of land.

Now, let us turn our attention to the retail firms. They use labor as a variable input and land as a fixed input.⁴ More specifically, to sell z units of composite good, az units of labor and f units of land are necessary. Here, $a \in (0, 1]$ and $f > 0$.⁵ We make an ad hoc assumption that land rent does not depend on the amount of land used by retail firms, because that amount is supposedly

³One reason for this is that commutes to work are made much more frequently than shopping trips.

⁴This assumption is not crucial. As will be shown, wage paid by a retail firm turns out to be a linear function of the land rent at its location. Therefore, both the variable cost and the fixed cost become linear functions of land rent.

⁵Because each worker consumes one unit of composite good, her wage must be no lower than p . If $a > 1$, therefore, marginal cost would exceed p , which cannot occur at equilibrium. Hence, a should be no greater than 1.

sufficiently small compared to the amount of land for residential use at any point of the city.

In addition, the retail firms need to pay a certain amount of congestion cost, which is the cost associated with negative externalities arising from being located closer to each other. These include intensification of road congestion; increase in noise, waste and air pollution mainly caused by heavier traffic; and aggravation of social environment due to proliferation of crime. For simplicity, we assume that this cost is proportional to the distance between the two firms:

$c(\bar{x} - \underline{x} - |x_1 - x_2|)$ with $c \geq 0$, where \underline{x} and \bar{x} are, as has been mentioned earlier, the left and right limits of their feasible locations. The congestion cost decreases with an increase in the distance between the two firms. Indeed, it becomes 0 when the two firms are located at the two outermost feasible locations, \underline{x} and \bar{x} , respectively, and becomes $c(\bar{x} - \underline{x})$ when they are located at the same point.

The profit of each firm is, therefore, given by

$$\pi(x_i|x_j) = d(x_i|x_j) (p - aw_i) - fr(x_i) - c(\bar{x} - \underline{x} - |x_1 - x_2|), \quad (2)$$

where $\pi(x_i|x_j)$ and $d(x_i|x_j)$ are profit and demand for firm i located at x_i given the location of its opponent at x_j .

The retail firms play a location game by simultaneously choosing their locations that maximize their own profits, given the price of a composite good. When we interpret each “firm” as a group of firms, the decision on their locations is made by a board of the group or a developer who constructs a shopping center that houses them. Now, there are several reasons why we focus on a location game, and not a location-then-price game. First, we suspect that the role of price competition in the location decisions of actual retail firms has been more or less diminishing.⁶ Second, the location-then-price game would become much more complicated in our setting than in a standard one. This is because our model deals with two kinds of transport costs; that is, commuting costs and shopping costs. This would make firms’ decisions in the first stage more complex: in particular, the decisions would depend heavily on the functional forms of the two kinds of transport costs.

2.2 Specification of the profit function

In this subsection, we specify the profit of each retail firm given by (2) by deriving the land rent it pays, the demand it receives, and the wage it offers.

To begin with, note that workers visit the retail firm located closer to them since the two firms

⁶One reason is the rise of online shopping, which provides consumers with a standardized reservation price. Another reason is concerned with the interpretation that a retail “firm” is a group of firms in a shopping center, as is mentioned above. It is becoming more and more common that different shopping centers contain stores belonging to the same chain. In many cases, such stores arguably have a tendency to sell goods at a common price.

charge the same price. Therefore, the location of a retail firm that a worker at l visits is given by

$$x(l) = \begin{cases} \min[x_1, x_2] & \text{if } l \leq \hat{x} \\ \max[x_1, x_2] & \text{if } l > \hat{x}, \end{cases} \quad (3)$$

where $\hat{x} \equiv (x_1 + x_2)/2$ is a midpoint of the locations of the two firms.⁷

Land rent

We focus on the case where the amount of land used by each retail firm is so small that at least some land is used for residential use at any location. Then, land rent is determined by workers' bids in the manner of the Alonso-Mills-Muth model. Furthermore, we suppose that the amount of labor employed by each retail firm is also small and that at any location, there are some workers who work in the basic sector at the CBD. Then, it suffices to consider bid rents of the workers who commute to the CBD. Therefore, from (1), land rent at l is given by

$$r(l) = y - t|l| - s|l - x(l)|, \quad (4)$$

where $y \equiv w_0 - p$ is the wage in the basic sector net of the payment for the composite good and $x(l)$ is given by (3). This equation indicates that land rent is increasing in l if $l < 0$ and decreasing if $l > 0$, because $t > s$. Consequently, land rent is the highest at the CBD and declines as we go farther away from it.

Having obtained a land rent function, we can determine the city limits and a feasible location space for retail firms. First, we have $r(l) = 0$ at the city limits since the city extends in the area where the bid land rent is no lower than the agricultural land rent. Consider the left limit at \underline{l} . First, $\underline{l} < 0$ because land rent takes a maximum at the CBD. Second, a worker at the left limit buys the composite good from the firm located at $\min[x_1, x_2]$; that is, $x(\underline{l}) = \min[x_1, x_2]$. This follows from (3) and the fact that retail firms are not allowed to settle outside the city, that is, $\underline{l} \leq \min[x_1, x_2]$. Third, the last inequality implies that $\underline{l} - x(\underline{l}) \leq 0$. Taking all these observations together, (4) implies that \underline{l} solves $y - t(-\underline{l}) - s(\min[x_1, x_2] - \underline{l}) = 0$; that is,

$$\underline{l} = -\frac{y - s \min[x_1, x_2]}{t + s}. \quad (5)$$

Similarly, the right limit is given by

$$\bar{l} = \frac{y + s \max[x_1, x_2]}{t + s}. \quad (6)$$

Second, for the feasible location space, note that a point lies within the city if and only if the land rent there is no lower than the agricultural rent. Therefore, if a retail firm is located in the city, its location, x , needs to satisfy either $x < 0$ with $r(x) = y - t(-x) \geq 0$, or $x > 0$ with $r(x) = y - tx \geq 0$. This implies that the feasible location space is given by $[\underline{x}, \bar{x}]$ where

$$\underline{x} = -y/t \quad \text{and} \quad \bar{x} = y/t. \quad (7)$$

⁷If $l = \hat{x}$, a worker is indifferent between visiting firm 1 and visiting firm 2. However, it does not matter whether we suppose that she visits the firm to her left (as in (3)) or that to her right, because workers at a particular location have 0 measure.

We refer to \underline{x} and \bar{x} as the leftmost feasible location and the rightmost feasible location, respectively, and collectively as the outermost feasible locations.

Figure 1 describes a typical land rent function when $x_i < 0$, $x_j > 0$ and $\hat{x} < 0$. A worker living and working at the CBD does not have to commute; therefore, what she needs to incur is only a shopping cost. Because she visits firm j , the cost is equal to sx_j . Consequently, she can pay $y - sx_j$ for land rent. As we gradually move from the CBD to the left, *both* the amount of the commuting cost to the CBD *and* that of the shopping cost to firm j increase for the worker living there. The sum increases by $t + s$ per mile, and therefore, the bid rent declines by that amount. As we further move leftward and pass \hat{x} , the shopping cost begins to decrease. This is because workers now visit firm i , and not firm j . Since the commuting cost to the CBD is still increasing at rate t , the rate of an overall increment in costs becomes $t - s$. Therefore, the bid rent declines by $t - s$ per mile. At x_i , furthermore, a worker pays only the commuting cost to the CBD: the bid rent becomes equal to $y + tx_i$. Finally, as we go beyond x_i , *both* the commuting cost to the CBD *and* the shopping cost increase. The slope of the bid rent curve becomes $t + s$ again.

Figure 1: Bid rent

Moreover, suppose that the two firms respectively relocate to symmetric points with respect to the origin, that is, firm 1 relocates from x_1 to $-x_1$ and firm 2 does from x_2 to $-x_2$. Then, the land rent at l after the relocation will be equal to the land rent at $-l$ before the relocation. In this sense, the land rent function is *invariant with respect to the reflection through the origin*. In particular, we can apply this property to the land rent at a location of a firm. To express the result, let us denote the land rent at the location of firm i by $r(x_i|x_j)$ rather than $r(x_i)$, explicitly considering the fact that it depends on the location of the competitor. Then, the reflection invariant property implies that $r(x_i|x_j) = r(-x_i|-x_j)$ ($i = 1, 2$).

Demand

If the retail firms are located at different places (i.e., $x_1 \neq x_2$), firm i gets demand from workers at l such that $x(l) = x_i$. Instead, if the firms are located at the same place (i.e., $x_1 = x_2$), they equally share demand. Consequently, (3) implies that

$$d(x_i|x_j) = \begin{cases} \hat{x} - \underline{l} & \text{if } x_i < x_j \\ (\bar{l} - \underline{l})/2 & \text{if } x_i = x_j \\ \bar{l} - \hat{x} & \text{if } x_i > x_j \end{cases} \quad (i = 1, 2; j \neq i). \quad (8)$$

Note that the demand function is also invariant with respect to the reflection through the origin, that is, $d(x_i|x_j) \equiv d(-x_i|-x_j)$.⁸

Wage offered by retail firms

⁸When $x_i < x_j$, for instance, $d(x_i|x_j) = \hat{x} - \underline{l}$ and $d(-x_i|-x_j) = \bar{l} - \hat{x}$ since $-x_i > -x_j$. These coincide with each other.

If a worker at l works for firm i , a retail firm, she can consume $z' \equiv [w_i - r(l) - t|l - x_i| - s|l - x(l)|]/p$ units of the composite good. Instead, by working in the basic sector, she can consume 1 unit of it. Therefore, she works for firm i only if $z' \geq 1$. Firm i , trying to set wage as low as possible, offers that worker the wage that equates z' to 1, or the wage equal to $p + r(l) + t|l - x_i| + s|l - x(l)|$. Because the firm employs workers for whom this offer becomes the lowest, its wage is determined at

$$w_i = \min_l [p + r(l) + t|l - x_i| + s|l - x(l)|]. \quad (9)$$

Note that $t|l| \leq t|x_i| + t|l - x_i|$ because of a triangle inequality. This implies that $r(l) + t|l - x_i| + s|l - x(l)| \geq r(x_i) + s|x_i - x(x_i)|$ for any l . Therefore, the term in the square brackets in (9) is minimized at $l = x_i$: each retail firm employs workers who live at its location. Furthermore, because workers at x_i buy a composite good from the firm located there (i.e., $x(x_i) = x_i$; see (3)), we have

$$w_i = p + r(x_i). \quad (10)$$

By construction, workers at x_i are indifferent between working for firm i and working in the basic sector. Furthermore, workers living at the other points do not prefer working for firm i to working in the basic sector.⁹

Profit function

Using (7) and (10), we can rewrite (2) as

$$\pi(x_i|x_j) = d(x_i|x_j) [(1-a)p - ar(x_i)] - fr(x_i) - c \left(\frac{2y}{t} - |x_1 - x_2| \right), \quad (11)$$

where $r(x_i)$ and $d(x_i|x_j)$ are given by (4) and (8), respectively. It is important to note that because both the land rent function and the demand function are reflection invariant, so is the profit function:

$$\pi(x_i|x_j) \equiv \pi(-x_i|-x_j). \quad (12)$$

For a retail firm, the location affects the profit through four channels. First, it affects the amount of demand it receives, which we call the *demand effect*. For a firm i 's relocation that is sufficiently small to involve no regime change in (8), the demand effect is given by $[(1-a)p - ar(x_i)] \cdot \partial d(x_i|x_j)/\partial x_i$. The demand function, (8), and the city limit functions, (5) and (6), imply that

$$\frac{\partial d(x_i|x_j)}{\partial x_i} = \begin{cases} \frac{t-s}{2(t+s)} > 0 & \text{if } x_i < x_j \\ -\frac{t-s}{2(t+s)} < 0 & \text{if } x_i > x_j. \end{cases} \quad (13)$$

The demand effect is further decomposed into two effects. By relocation, the market boundary between the two retail firms moves. We refer to this as the *competition effect*. Simultaneously, city limits also move. As a result, each firm seizes a larger or a smaller "outer market area," which

⁹(9) implies that $[w_i - r(l) - t|l - x_i| - s|l - x(l)|]/p \leq 1$ for any l .

for the left firm, is the market area to its left, and for the right firm, is the market area to its right. This is referred to as the *city limit effect*. Second, the relocation of firm i alters the price markup by changing the land rent. This effect on the profit, given by $-ad(x_i|x_j) \cdot \partial r(x_i)/\partial x_i$, is referred to as the *marginal cost effect*. Third, the change in land rent also brings about a change in the fixed cost. This effect, which we call *fixed cost effect*, is given by $f\partial r(x_i)/\partial x_i$. Fourth and finally, the congestion cost changes as a result of the relocation of firm i , which affects its profit by $-c \cdot \partial|x_1 - x_2|/\partial x_i$. This we refer to as the *congestion effect*.

To close the section, let us impose one assumption. We focus on an interesting case with $\pi(0|0) \geq 0$: each of the two retail firms earns a nonnegative profit when they are agglomerated at the CBD. Because $r(0|0) = y$, this condition is reduced to the following inequality:

Assumption 1.

$$c \leq \bar{c} \equiv \frac{t[(1-a)p - ay - f(t+s)]}{2(t+s)}. \quad (14)$$

This assumption entails $(1-a)p - ay > 0$ because $c > 0$. Therefore, $(1-a)p - ar(x_i) \equiv (1-a)p - ay + t|x_i| > 0$, which indicates that a price markup is always positive (see (11)).

3 Firms' choice of location

In this section, we examine a choice of location by firm i when the location of firm $j \neq i$ is given. Without loss of generality, let us focus on the case where $x_j \geq 0$. We can similarly analyze the other case using the reflection invariance property, (12): a choice of firm i when $x_j < 0$ becomes a reflection of its choice when firm j is located at $-x_j > 0$.

For the choice of firm i , there are four possibilities.

Case i) $x_i \in [\underline{x}, 0]$ with $x_i \neq x_j$.

Suppose that firm i chooses $x_i \in [\underline{x}, 0]$ with $x_i \neq x_j$. For $(x_i, x_j) \in X^I \equiv \{(x_i, x_j) | x_i \in [\underline{x}, 0], x_j \in [0, \bar{x}], x_i \neq x_j\}$, the profit of firm i is given by

$$\pi^I(x_i|x_j) \equiv (\hat{x} - l) \left[(1-a)p - a(y + tx_i) \right] - f(y + tx_i) - c(2\bar{x} + x_i - x_j)$$

(see (4) and (8)). It follows that

$$\frac{\partial \pi^I(x_i|x_j)}{\partial x_i} = \frac{(1-a)p - a(y + tx_i)}{2} - \frac{s[(1-a)p - a(y + tx_i)]}{t+s} - ta \left(\frac{x_1 + x_2}{2} + \frac{y - sx_i}{t+s} \right) + tf - c. \quad (15)$$

This equation shows the four effects mentioned earlier of firm i 's slight relocation toward the CBD.

Among the five terms in the right-hand side of (15), the first two represent a demand effect. The first term represents a competition effect and the second, a city limit effect. Because the

market boundary between the two firms shifts to the right as a result of firm i 's relocation, the competition effect is positive. As firm i moves toward the CBD, on the other hand, workers who were located to the left of that firm now need to pay higher shopping costs to visit it. This implies that they cannot bid as high as earlier. Therefore, the land rents at their locations decrease; and consequently, the left city limit moves rightward. To this extent, the market area of firm i to its left shrinks: the city limit effect is negative. However, because the competition effect more than offsets the city limit effect, the total effect is positive, which is confirmed by (13) since $x_i < x_j$ in this case. In Figure 1 above, the dashed line represents the land rent after firm i relocates from x_i to x_i' . By this relocation, both the left city limit and the market boundary between the two firms move rightward, from \underline{l} to \underline{l}' , and from \hat{x} to \hat{x}' , respectively.

Moreover, the third and the fourth terms in (15) represent a marginal cost effect and a fixed cost effect, respectively. These effects are negative because firm i pays higher wage and higher land rent as it moves toward the CBD. Finally, the last term indicates a negative congestion effect: firm i needs to pay a higher congestion cost as a result of the relocation.

The sign of the total effect is ambiguous: it depends on the location of the opponent. To see this, note that $\pi^I(x_i|x_j)$ has at most one local maximum because it is strictly concave in x_i .¹⁰ We denote x_i that maximizes $\pi^I(x_i|x_j)$ as $g(x_j)$. Because it must solve $\partial\pi^I(g(x_j)|x_j)/\partial x_i = 0$, we have

$$g(x_j) = \frac{1}{2at(t-s)} \left[-at(t+s)x_j + \theta - ay(3t-s) \right], \quad (16)$$

where $\theta \equiv p(1-a)(t-s) - 2(t+s)(tf+c)$. First, if $g(x_j) \geq 0$, $\pi^I(x_i|x_j)$ is increasing for any $x_i \in [\underline{x}, 0)$. This condition can be written as $x_j \leq x^o$, where $x^o \equiv [\theta - ay(3t-s)]/[at(t+s)]$. Second, if $g(x_j) \leq \underline{x}$, $\pi^I(x_i|x_j)$ is decreasing for any $x_i \in (\underline{x}, 0]$. This occurs when $x_j \geq x^{oo}$, where $x^{oo} \equiv [\theta - ay(t+s)]/[at(t+s)]$. Note that $x^{oo} > x^o$ since $t > s$. Finally, if x_j lies in between the two critical values, $g(x_j)$ falls into interval $(\underline{x}, 0)$.

To sum up, for $x_j \in [0, \bar{x}]$,

$$\left\{ \begin{array}{l} \frac{\partial\pi^I(x_i|x_j)}{\partial x_i} > 0 \text{ for any } x_i \in [\underline{x}, 0) \\ \text{There exists } g(x_j) \in (\underline{x}, 0) \text{ such that } \frac{\partial\pi^I(x_i|x_j)}{\partial x_i} \left\{ \begin{array}{l} > \\ = \\ < \end{array} \right\} 0 \text{ for } \left\{ \begin{array}{l} \text{any } x_i \in [\underline{x}, g(x_j)) \\ x_i = g(x_j) \\ \text{any } x_i \in (g(x_j), 0] \end{array} \right\} \\ \frac{\partial\pi^I(x_i|x_j)}{\partial x_i} < 0 \text{ for any } x_i \in (\underline{x}, 0] \end{array} \right. \begin{array}{l} \text{if } x_j \leq x^o \\ \text{if } x_j \in (x^o, x^{oo}) \\ \text{if } x_j \geq x^{oo}. \end{array} \quad (17)$$

Furthermore, note that $\pi^I(x_i|x_j)$ is continuous at $x_i = 0$ and $x_i = \underline{x}$. Therefore, to any other location in $[\underline{x}, 0]$, firm i prefers the CBD if $x_j \leq x^o$, the interior location given by $g(x_j)$ if $x_j \in (x^o, x^{oo})$, and finally, the leftmost feasible location \underline{x} if $x_j \geq x^{oo}$. Whether x^o and x^{oo} are positive

¹⁰ $\partial^2\pi^I(x_i|x_j)/\partial x_i^2 = -at(t-s)/(t+s) < 0$.

or not, and whether they are no greater than \bar{x} or not depend on the size of θ : it may be the case that only one or two lines in (17) are actually applicable.

Relating to this result, two comments are in order.

First, (16) indicates that if firm j decides to settle at a location closer to the CBD, then firm i tends to do so as well. In this sense, firms are more aggressive when their opponents are aggressive.¹¹ The reason is straightforward. The location of firm j affects $\partial\pi^I(x_i|x_j)/\partial x_i$ only through the marginal cost effect (the third term in the right-hand side of (15)). When firm j is located closer to the CBD, the market area of firm i is smaller. Firm i , acquiring only a smaller amount of demand, now cares less about the marginal cost. As a result, it will suffer less harm from a rise in the marginal cost when it moves toward the CBD; that is, a smaller marginal cost effect. Therefore, the relocation toward the CBD becomes more likely to be beneficial.

Second, we can derive the counterpart of (16) for firm j . Because of the reflection invariance, $\pi(x_j|x_i) \equiv \pi(-x_j|-x_i)$. Moreover, $(x_i, x_j) \in X^I$ implies that $(-x_j, -x_i) \in X^I$. Therefore, $\pi(x_j|x_i) \equiv \pi^I(-x_j|-x_i)$. By the definition of function $g(\cdot)$, however, $\pi^I(-x_j|-x_i)$ is maximized at $-x_j = g(-x_i)$. In other words, the best response of firm j is given by $-g(-x_i)$, provided that it lies in the interior.

Case ii) $x_i \in (0, x_j)$.

Now, we turn to the possibility that firm i chooses $x_i \in (0, x_j)$. For $(x_i, x_j) \in X^{II} \equiv \{(x_i, x_j) | x_i \in (0, x_j), x_j \in [0, \bar{x}]\}$, (8) and (4) imply that the profit of firm i is given by

$$\pi^{II}(x_i|x_j) \equiv (\hat{x} - l) \left[(1-a)p - a(y - tx_i) \right] - f(y - tx_i) - c(2\bar{x} + x_i - x_j).$$

Because this is a strictly convex function of x_i , the maximum of $\pi^{II}(x_i|x_j)$ over $x_i \in [0, x_j]$ is achieved either at $x_i = 0$ or at $x_i = x_j$.¹² Consequently, firm i 's profit increases as it approaches either the CBD or the location of the competitor.

Case iii) $x_i = x_j$.

When $x_i = x_j$, demand for firm i is given by the second equation in (8). Therefore, its profit becomes equal to

$$\pi^{III}(x_i|x_j) \equiv \frac{(\bar{l} - l) \left[(1-a)p - a(y - tx_j) \right]}{2} - f(y - tx_j) - 2c\bar{x}.$$

As long as $x_j > 0$, however, $\pi^{II}(x_j|x_j) > \pi^{III}(x_j|x_j)$: a deviation to slightly left of firm j gives a higher profit to firm i (i.e., is "profitable").

¹¹True, in terms of game theory, (16) shows that the locations of the two firms are strategic substitutes, but to say so is misleading because a decrease, not an increase, in $x_i < 0$ has a meaning parallel to an increase in $x_i \geq 0$ in our setting.

¹²Note that $\frac{\partial^2 \pi^{II}(x_i|x_j)}{\partial x_i^2} = \frac{at(t-s)}{t+s} > 0$.

Case iv) $x_i \in (x_j, \bar{x}]$.

Finally, suppose that firm i chooses $x_i > x_j$. Because demand is given by the last equation in (8) for $(x_i, x_j) \in X^{IV} \equiv \{(x_i, x_j) | x_i \in (x_j, \bar{x}], x_j \in [0, \bar{x}]\}$, the profit of firm i is equal to

$$\pi^{IV}(x_i|x_j) \equiv (\bar{l} - \hat{x}) \left[(1-a)p - a(y - tx_i) \right] - f(y - tx_i) - c(2\bar{x} - x_i + x_j).$$

However, as long as $x_j > 0$, firm i does not actually choose such x_i , because it can obtain a higher profit by relocating to $-x_i$, a symmetric point to the left of the CBD, namely, in the “left part” of the city. Three observations explain this. First, by this relocation, firm i gets a larger demand by the following reason. When it is located at $x_i > x_j$, the inner market area (the area to the left of x_i) is smaller than $x_i/2$ given that $x_j > 0$. As it relocates to $-x_i < 0$, however, the inner market area (the area to the right of $-x_i$) becomes greater than $x_i/2$. Therefore, the inner market area expands by the relocation. On the other hand, the size of the outer market area remains unchanged; that is, $\bar{l} - x_i$ is equal to $-x_i - \underline{l}$. Second, the land rent firm i pays does not change by the relocation, which implies that both the marginal cost and the fixed cost remain unchanged. Third and last, the distance between the two firms increases by the relocation ($x_i - x_j < x_j - (-x_i)$) as long as $x_j > 0$. Therefore, the congestion cost declines. Because the demand effect and the congestion effect are both positive and there is no marginal cost effect nor fixed cost effect, the symmetric relocation is profitable for firm i , given that $x_j > 0$.

4 Equilibrium

In this section, we derive an equilibrium. Let us denote a pair of equilibrium locations of two firms by (x_1^*, x_2^*) . To begin with, we examine symmetric equilibria with $x_2^* = -x_1^*$. First, in Section 4.1, a symmetric equilibrium at which two firms are agglomerated at the CBD; that is, $x_1^* = x_2^* = 0$, is considered. Second, in Section 4.2, we turn to symmetric equilibria at which they are dispersed; that is, $x_i^* < 0$ and $x_j^* = -x_i^* > 0$. Next, asymmetric equilibria are examined in Section 4.3. Finally, we discuss in Section 4.4 how transport costs affect the equilibrium.

4.1 Equilibrium with agglomeration at the CBD

Suppose that firm j is located at the CBD. When does firm i decide to settle there? To answer this question, note that $\pi(x_i|0) = \pi(-x_i|0)$ because of the reflection invariance property: if firm i prefers the CBD to any point in the city’s left part, it also prefers the CBD to any point in its right part. Therefore, it suffices to consider only $x_i \leq 0$. Now, the profit of firm i is continuous at $x_i = 0$; that is, $\pi^I(0|0) = \pi^{III}(0|0)$. Consequently, (17) implies that it chooses to settle at the CBD if and only if $x_j = 0 \leq x^0$. This condition is reduced to

$$c \leq c^{AG} \equiv \frac{p(1-a)(t-s) - 2tf(t+s) - ay(3t-s)}{2(t+s)}, \quad (18)$$

which is a necessary and sufficient condition for agglomeration at the CBD.

The critical value c^{AG} depends on several parameters. In the following explanation, we will consider a relocation of firm i from $x_i < 0$ toward the CBD, and examine how parameters affect the four effects of that relocation.

First, c^{AG} decreases with the wage in the basic sector, w_0 (remember that $y \equiv w_0 - p$). When w_0 is higher, workers can afford more money for land, which results in a higher land rent. With a smaller price markup, therefore, a *given amount* of an increase in demand raises firm i 's profit only on a smaller scale: the positive demand effect of the relocation toward the CBD is smaller. However, when w_0 is higher, \underline{l} is smaller, which means that firm i receives a greater demand *ceteris paribus*. Therefore, a *given amount* of an increase in marginal cost makes a more devastating impact; that is, the negative marginal cost effect of the relocation is greater. For these reasons, the relocation toward the CBD is less beneficial for firm i and an agglomeration there is less likely to be supported as an equilibrium outcome.

Second, c^{AG} increases with the price of the composite good, p . The logic behind this result is the same as that for the change in w_0 : a lower p reduces the demand effect but enlarges the marginal cost effect.

Third, cost parameters a and f are both negatively related to c^{AG} . A higher f , requiring a larger amount of fixed inputs, increases the negative fixed cost effect of the relocation. A higher a , on the other hand, not only increases the negative marginal cost effect but also decreases the positive demand effect by reducing the price markup.

Finally, the impacts of changes in unit transport costs, t and s , are ambiguous. One interesting result is, however, that when s is sufficiently high and sufficiently close to t , c^{AG} becomes negative: (18) does not hold for any positive c . For agglomeration at the CBD to be supported as an equilibrium outcome, therefore, it is necessary that the gap between the unit commuting cost and the unit shopping cost is sizable and/or the unit congestion cost is small.

These results are summarized as follows.

Proposition 1. *Two firms are agglomerated at the CBD at an equilibrium if and only if (18) holds. This equilibrium is more likely to occur when the wage in the basic sector is lower, the price of the composite good is higher, production requires smaller amounts of inputs, and the unit congestion cost is smaller. It is not supported by an equilibrium if the unit shopping cost is too high.*

Two comments follow.

First, $c^{AG} < \bar{c}$ since $(1 - a)p - ay > 0$. If $c^{AG} > 0$, consequently, there exist *both* a low c that satisfies not only Assumption 1 but also (18), *and* a high c that satisfies Assumption 1 but does not satisfy (18). In other words, as long as $c^{AG} > 0$, there exist *both* the set of parameters for which the agglomeration at the CBD becomes an equilibrium outcome *and* the set of parameters for which it does not.

Second, let us take a look at the special case where a , t and c are equal to 0. In this case, c^{AG} becomes a negative constant, which implies that no combination of parameters satisfies (18):

in no case is agglomeration at the CBD supported by an equilibrium. This is in sharp contrast to Hotelling's result that the only equilibrium involves the agglomeration there. It might appear odd because our special case well describes the economy in the Hotelling model. The reason for this seeming discrepancy is that city limits are endogenously determined in our model whereas they are fixed in the Hotelling model. To see this, suppose that firm i , which is located in the city's left part, decides to move by a unit distance toward the CBD, where its competitor is located. Then, as long as our special case is concerned, the left city limit will move to the right by the same distance because t is assumed to be 0. At the same time, the market boundary between the two firms will move to the right by only half a unit. Therefore, the demand for firm i will diminish as a result of the relocation, which implies that the agglomeration is not supported by an equilibrium. If the city limits were fixed, to the contrary, the demand would increase and consequently, the agglomeration would become an equilibrium outcome. This indicates that *the assumption of fixed city limits is one of the indispensable building blocks for Hotelling's result.*

4.2 Equilibrium with symmetric dispersion

Next, we study a symmetric equilibrium with $x_i^* < 0$ and $x_j^* = -x_i^* > 0$.

Before examining the specific types of equilibrium, it is useful to ask one question. Does firm i not obtain a higher profit by deviating to the immediate left of firm j ? In other words, does $\pi^{II}(x_j^* - \epsilon | x_j^*) \leq \pi^I(x_i^* | x_j^*)$ hold for an arbitrarily small number $\epsilon > 0$? Because of continuity, it holds if $\pi^{II}(x_j^* | x_j^*) < \pi^I(x_i^* | x_j^*)$ and only if $\pi^{II}(x_j^* | x_j^*) \leq \pi^I(x_i^* | x_j^*)$. These conditions can be rewritten as

$$x_j^* < \tilde{x} \equiv \frac{2c(t+s) - (t-s)[(1-a)p - ay]}{at(t-s)} \quad (19)$$

and

$$x_j^* \leq \tilde{x}, \quad (20)$$

respectively. Thus, firm i 's relocation to an "almost symmetric" point is not profitable if, and only if, the locations of firms are sufficiently close to the CBD.

This result can be explained as follows. Note that land rent remains (almost) unchanged by the almost symmetric relocation. Thus, there exists no marginal cost effect nor fixed cost effect. However, the boundary between the two firms and the left city limit move to the right by a distance equal to x_j^* and a distance equal to $2sx_j^*/(t+s)$, respectively, which implies that the demand increases by $(t-s)x_j^*/(t+s)$. At the same time, the congestion cost increases by $2cx_j^*$. When the two firms are located closer to the CBD (that is, $x_j^* = -x_i^*$ is smaller), therefore, both the positive demand effect and the negative congestion cost effect are smaller. However, the former effect is much smaller and consequently, the almost symmetric relocation is less likely to be profitable.

Corner equilibrium

First, let us consider the equilibrium at which two firms are located at the outermost feasible locations, respectively: $(x_i^*, x_j^*) = (\underline{x}, \bar{x})$.

In order for this pattern of maximum dispersion to be supported by an equilibrium, two conditions need to be satisfied. First, each firm should not be able to earn a higher profit by any “local” deviation, which is a deviation that involves no change across the cases mentioned in Section 3. As far as the maximum dispersion is concerned, it is a deviation to $x_i \in (\underline{x}, 0]$ (Case i) for firm i . Equation (17) implies that the condition for no local deviation is given by

$$\bar{x} \geq x^{oo}. \quad (21)$$

This is not only a sufficient condition but also a necessary condition.

Furthermore, any “global” deviation, which accompanies a change across the cases, should not be profitable. Here, a deviation to $x_i \in (0, \bar{x}]$ is relevant for firm i . First, for firm i , the deviation to \bar{x} is dominated by the deviation to $\bar{x} - \epsilon$ according to the result obtained for Case iii in Section 3. Second, we have verified that firm i 's profit function is strictly convex in interval $(0, \bar{x})$. Third, as long as (21) is satisfied, the deviation to $x_i = 0$ is not profitable. Because firm i 's profit function is continuous at $x_i = 0$, these three observations imply that no global deviation is profitable if and only if the almost symmetric relocation to $\bar{x} - \epsilon$ is not profitable. However, we know that it is indeed not profitable if (19) is satisfied for $x_j^* = \bar{x}$, that is, if $\bar{x} < \tilde{x}$. Both the last condition and (21) are satisfied if and only if the condition

$$c > c^{MD} \equiv \frac{p(1-a)(t-s)}{2(t+s)} \quad (22)$$

holds.¹³ This is a sufficient condition for the maximum dispersion being supported by an equilibrium. A necessary condition is similarly given by

$$c \geq c^{MD}. \quad (23)$$

The critical value c^{MD} depends on four parameters: it increases with p and t but decreases with a and s . To understand these results, let us remember that the relocation of firm i to the immediate left of its opponent involves only the demand effect and the congestion effect. First, when p is higher, land rent is lower. This, along with a higher price itself, implies a greater price markup. Therefore, a given amount of expansion of market area is more beneficial, and consequently, the demand effect of the almost symmetric relocation is larger. Second, when t is higher, the left city limit moves only by a shorter distance as a result of the relocation, and thus a smaller city limit effect.¹⁴ In addition, because land rent is lower, a given amount of expansion of market area is more beneficial. By these two reasons, a higher t is associated with a greater demand effect, and, consequently, makes the relocation more likely to be profitable. Third, when a is higher, the price

¹³Note that (21) is reduced to $c \geq c^{MD} - (ay + tf)$ whereas $\bar{x} < \tilde{x}$ is to $c > c^{MD}$. Obviously, the latter is binding. A similar remark applies to the relationship between (21) and $\bar{x} \leq \tilde{x}$.

¹⁴The extent of the shift is equal to $-\frac{y - s\bar{x}}{t+s} - \left[-\frac{y - s\underline{x}}{t+s}\right] = \frac{2sy}{t(t+s)}$.

markup is smaller because the marginal cost is higher. Therefore, the demand effect is smaller, which implies that the relocation is less beneficial. Fourth and finally, when s is higher, the left city limit shifts more, and therefore, the relocation is less beneficial.

These findings are summarized as follows.

Proposition 2. *Two firms are located at the two outermost feasible locations, respectively, at an equilibrium if (22) is satisfied, and only if (23) is satisfied. This equilibrium is more likely to occur when the price of the composite good and the unit commuting cost are lower, and the unit labor requirement, the unit shopping cost, and the unit congestion cost are higher.*

Two observations follow.

First, the *relative* size of the two types of transport costs plays an important role in the determination of equilibrium. As the gap between t and s narrows, c^{MD} declines, and therefore, it becomes more likely that the maximum dispersion is supported as an equilibrium outcome. In contrast, note that $c^{MD} > \bar{c}$ if and only if

$$(1-a)p - ay - f(t+s) < \frac{p(1-a)(t-s)}{t}. \quad (24)$$

As the gap between t and s widens while their sum remains unchanged, (24) becomes more likely to be satisfied. Indeed, when s is sufficiently close to 0, (24) holds and consequently, $c^{MD} > \bar{c}$ for any t . In this case, any set of parameters that satisfies Assumption 1 does not satisfy (23): there does not exist an equilibrium with maximum dispersion.

Second, we note that $c^{MD} > c^{AG}$. This implies that there is no possibility of multiple equilibria: whenever the agglomeration at the CBD is supported by an equilibrium, maximum dispersion is not, and *vice versa*.

Interior equilibrium

Next, consider an equilibrium with $x_i^* \in (\underline{x}, 0)$ and $x_j^* = -x_i^*$. At this interior equilibrium, both $x_1^* = g(x_2^*)$ and $x_2^* = g(x_1^*)$ must hold. From (16), we derive $(x_i^*, x_j^*) = (-x^*, x^*)$ where $x^* \equiv [\theta - ay(3t-s)]/[at(3s-t)]$.

In order for such an equilibrium to exist, two conditions are necessary. First, x_i^* should indeed fall in the interval $(\underline{x}, 0)$. A necessary condition is given by

$$x_j^* = x^* \in (x^0, x^{00}) \quad (25)$$

(see (17)). Second, no *global* deviation from x_i^* should be profitable for firm i . In particular, a relocation to the immediate left of competitor (i.e., an almost symmetric relocation), should not be profitable. Now, suppose that $\pi^{II}(x_j^*|x_j^*) > \pi^I(x_i^*|x_j^*)$, or equivalently that $x_j^* > \tilde{x}$ (see the derivation of (20)). Then, there is a sufficiently small $\epsilon > 0$ such that $\pi^{II}(x_j^* - \epsilon|x_j^*) > \pi^I(x_i^*|x_j^*)$; that is, the almost symmetric relocation is profitable. Therefore, the unprofitability of global deviations requires that

$$x^* \leq \tilde{x}. \quad (26)$$

However, this contradicts (25) (for the underlying reason, see the proof of the next proposition). Hence, there exists no interior symmetric equilibrium.

Proposition 3. *There exists no symmetric equilibrium for which firms choose locations apart from each other in the interior of the feasible location space, (\underline{x}, \bar{x}) .*

Proof. The inequality $x^* < x^{00}$, which is one of the conditions in (25), is reduced to $\theta > \theta' \equiv atx^*(t+s) + ay(t+s)$. Furthermore, (26) can be rewritten as $\theta \leq \theta'' \equiv -atx^*(t-s) + ay(t-s) - 2tf(t+s)$. However, it is readily verified that $\theta' > \theta''$. Hence, there is no θ that satisfies both (25) and (26). ■

The reason for the nonexistence is straightforward. We have seen that an almost symmetric relocation is less likely to be profitable when firms are located closer to the CBD. However, the candidate locations for an interior equilibrium, given by $(-x^*, x^*)$, are too far from the CBD: at those locations, firms always have an incentive to relocate to slightly inner positions of their opponents.

4.3 Asymmetric equilibrium

Finally, we can show that there is no asymmetric equilibria with $x_2^* \neq -x_1^*$.

Proposition 4. *There exists no asymmetric equilibrium.*

The proof is tedious and thus relegated to the Appendix.

4.4 Effects of changes in transport costs

As has been mentioned in the introduction, one of the most important reasons for spatial decentralization in cities is the decline in transport costs mainly caused by increased automobile use. In this subsection, we briefly examine this assertion in light of our results obtained above.

To begin with, note that the condition for the equilibrium with agglomeration at the CBD, (18), can be rewritten as

$$s \leq \frac{t[(1-a)p - 3ay - 2(c+tf)]}{(1-a)p - ay + 2(c+tf)} \equiv s^{AG}. \quad (27)$$

If $(1-a)p - 3ay - 2c \leq 0$, s^{AG} becomes negative. In that case, no $s > 0$ satisfies (27): in no case is the agglomeration supported by an equilibrium. Thus, our attention is concentrated on interesting cases with

$$(1-a)p - 3ay - 2c > 0. \quad (28)$$

If we regard s^{AG} as a function of t , it is depicted by a s^{AG} curve in Figure 2. The curve has three properties. First, it is strictly concave in t . Second, as t goes to 0 from above, the slope of the

curve approaches a positive constant which is smaller than 1.¹⁵ Third and last, the curve passes through the origin. The agglomeration at the CBD is supported as an equilibrium outcome if and only if a combination of t and s falls into the area below or on the curve.

Next, we can rewrite the necessary condition for the equilibrium with maximum dispersion, (22), as

$$s \geq \frac{t[(1-a)p - 2c]}{(1-a)p + 2c} \equiv s^{MD}.$$

In Figure 2, s^{MD} is depicted by a ray starting from the origin, which is referred to as an s^{MD} curve. Since it is steeper than the s^{AG} curve at the origin, the s^{MD} curve always lies above the s^{AG} curve in the first quadrant; that is, $s^{MD} > s^{AG}$ for any $t > 0$. This corresponds to our earlier finding that $c^{MD} > c^{AG}$. The maximum dispersion is supported by an equilibrium if (t, s) lies in the area above the s^{MD} curve, and only if it lies in the area above or on the s^{MD} curve.

Figure 2: Effects of transport costs on the spatial structure of a city

Furthermore, when (t, s) falls into the area between the s^{AG} curve and the s^{MD} curve, there exists no equilibrium with pure strategies but there exists only an equilibrium with mixed strategies. In this paper, however, we do not discuss the mixed strategy equilibrium in more detail.

Suppose that transport costs decline. How the spatial pattern of the locations of the two firms changes depends on the nature of the decline, in particular, the *relative* change in unit commuting cost, t , and unit shopping cost, s . For instance, if s declines more rapidly than t , the spatial pattern may change from maximum dispersion to agglomeration at the CBD. The transition from point A to point B in Figure 2 depicts such a change. Conversely, if it is t that declines more rapidly, the pattern may switch from the agglomeration to the maximum dispersion, as the trajectory from point C to point D indicates.

Which type of transport cost declines more rapidly? A casual observation of our history suggests that t rather than s tends to do so. One of the immediate consequences of a rise in automobile use was congestion during a peak period on roads within a city center and roads connecting suburbs and the CBD. Because the social loss from the associated externalities, especially, the one caused by prolonged commuting time, was too heavy to overlook, governments were quick to take diverse measures to ameliorate the problem, including construction of additional highways and freeways. As a result, many cities saw an improvement in congestion, which reduced commuting time and therefore, commuting costs. This corresponds to a decline in t in our model. In suburbs, in contrast, road congestion is a relatively new problem: it arrives only after a number of new office buildings and big shopping malls were constructed to form "edge cities." Attempts to mitigate the problem, if any, were made only recently. Therefore, in suburbs, transport costs

¹⁵Note that $\frac{d^2 s^{AG}}{dt^2} = -8f \cdot \frac{[(1-a)p - 2ay][(1-a)p - ay + 2c]}{[(1-a)p - ay + 2(c+tf)]^3} < 0$, and that $\left. \frac{ds^{AG}}{dt} \right|_{t=0} = \frac{(1-a)p - 3ay - 2c}{(1-a)p - ay + 2c} \in (0, 1)$ because of (28).

declined only recently, if at all. In the context of our model, this means that s declined more slowly.

If this observation is true, our model indicates that the locational pattern is more likely to change from the agglomeration at a CBD to maximum dispersion (i.e., the second case mentioned above), than in the reverse direction. This is one explanation for the decentralization in cities: *one reason for the decentralization is that commuting costs decline more rapidly than shopping costs*. It goes without saying, however, that the questions of which type of transport costs declines more rapidly and how the two types decline over time are matters of empirical studies, which we leave for future research.

5 Concluding remarks

This paper shows that a strategic interaction between firms is one of the key factors that determine the spatial structure of a city. More specifically, we have examined two retail firms that compete against each other by choosing their locations in a linear city. This Hotelling-like setting is expanded to incorporate two most important aspects of an urban spatial structure, a variation in input prices, in particular, land rent, over space and endogenous determination of city limits, which are characteristic of the standard Alonso-Mills-Muth model.

It has been shown that at an equilibrium, the two firms are either agglomerated at the CBD or located at the two outermost feasible locations. A key factor that determines which pattern emerges is the relative size of commuting cost to the CBD and shopping cost to a retail firm. In particular, if the former cost declines more rapidly than the latter cost over time, the equilibrium pattern may switch from agglomeration at the CBD to symmetric dispersion in suburbs, which many cities around the world have witnessed since the mid-20th century.

References

- [1] Anas, A., Arnott, R., Small, K. A., 1998. "Urban spatial structure," *Journal of Economic Literature*, 36, 1426-1464.
- [2] Fujita, M., Ogawa, H., 1982. "Multiple equilibria and structural transition of non-monocentric urban configurations," *Regional Science and Urban Economics*, 12, 161-196.
- [3] Fujita, M., Thisse, J. F., Zenou, Y., 1997. "On the endogenous formation of secondary employment centers in a city," *Journal of Urban Economics*, 41, 337-357.
- [4] Glaeser, E. L., Kahn, M. E., 2004. "Sprawl and urban growth," in *Handbook of Regional and Urban Economics*, vol. 4 (eds by Henderson, J. V., Thisse, J.-F.), North Holland, Amsterdam, the Netherlands, 2482-2527.
- [5] Helsley, R. W., Sullivan, A. M., 1991. "Urban subcenter formation," *Regional Science and Urban Economics*, 21, 255-275.
- [6] Henderson, V., Mitra, A., 1996. "The new urban landscape: developers and edge cities," *Regional Science and Urban Economics*, 26, 613-643.
- [7] Hotelling, H., 1929. "Stability in competition," *The Economic Journal*, 39, 41-57.
- [8] Lai, F.-C., Tsai, J.-F., 2008. "Simplified Alonso-Mills-Muth model with a monopoly vendor," *Journal of Urban Economics*, 63, 536-543.
- [9] Ogawa, H., Fujita, M., 1980. "Equilibrium Land use patterns in a nonmonocentric city," *Journal of Regional Science*, 20, 455-475.
- [10] Ross, S., Yinger, J., 1995. "Comparative static analysis of open urban models with a full labor market and suburban employment," *Regional Science and Urban Economics*, 25, 575-605.
- [11] White, M. J., 1976. "Firm suburbanization and urban subcenters," *Journal of Urban Economics*, 3, 323-343.
- [12] Yinger, J., 1992. "City and suburb: Urban model with more than one employment center," *Journal of Urban Economics*, 31, 181-205.

Appendix

Proof of Proposition 4.

Proof. Suppose that (x_1^*, x_2^*) with $x_2^* \neq -x_1^*$ is a pair of equilibrium locations. Without loss of generality, we focus on the case where $x_2^* \geq 0$: the other case is treated similarly thanks to the reflection invariance property. We can distinguish three cases, each of which is further divided into several subcases.

i) To begin with, let us consider the case where $x_2^* = 0$.

i-a) First, suppose that $x_1^* \in [\underline{x}, 0)$. If firm 1 chooses such a location, it must be true that $\partial\pi^I(x_1|x_2^*)/\partial x_1 < 0$ for some $x_1 \in [\underline{x}, 0)$ because of the continuity of $\pi^I(x_1|\cdot)$. This requires that $0 > x^0$ (see (17)). Furthermore, in order to examine the behavior of firm 2, let us suppose that $-x_1^* > x^0$. (17) implies that $\partial\pi^I(-x_2|-x_1^*)/\partial(-x_2) < 0$ for some $-x_2 \in (\underline{x}, 0)$. Consequently, there exists $-x_2 \in (\underline{x}, 0)$ such that $\pi^I(-x_2|-x_1^*) > \pi^I(0|-x_1^*)$ because $\pi^I(-x_2|\cdot)$ is concave in $-x_2$. However, the reflection invariance property, (12), implies that $\pi(x_2|x_1^*) \equiv \pi(-x_2|-x_1^*) \equiv \pi^I(-x_2|-x_1^*)$ for any $x_2 \in (0, \bar{x})$. Thus, there exists $x_2 \in (0, \bar{x})$ such that $\pi(x_2|x_1^*) > \pi(0|x_1^*)$. This contradicts our assumption that $x_2^* = 0$. Hence, it must be true that $-x_1^* \leq x^0$. This and the previous result that $0 > x^0$ imply that $x_1^* > 0$, which is a contradiction. Hence, $x_1^* \in [\underline{x}, 0)$ cannot be the case.

i-b) Second, suppose that $x_1^* \in (0, \bar{x}]$. It follows from the reflection invariance property that $\pi(x_1^*|0) \equiv \pi(-x_1^*|0) \equiv \pi^I(-x_1^*|0)$. If firm 1 chooses x_1^* , therefore, $0 > x^0$ must hold (see (17)). Furthermore, firm 2 chooses 0 only if $\partial\pi^I(x_2|x_1^*)/\partial x_2 > 0$ for any $x_2 \in [\underline{x}, 0)$. Therefore, it must be true that $x_1^* \leq x^0$ (again see (17)). Consequently, $x_1^* \leq x^0 < 0$, which is a contradiction. Hence, $x_1^* \in (0, \bar{x}]$ cannot be the case, either.

ii) Next, we turn our attention to the case where $x_2^* \in (0, \bar{x})$.

ii-a) First, suppose that $x_1^* = \underline{x}$. If firm 1 chooses such a location, it must be true that $\partial\pi^I(\underline{x}|x_2^*)/\partial x_1 \leq 0$. This requires that $x_2^* \geq x^{00}$ (see (17)). Furthermore, for the profit of firm 2, we have $\pi(x_2|\underline{x}) \equiv \pi(-x_2|\bar{x}) \equiv \pi^I(-x_2|\bar{x})$ for any $-x_2 \in [\underline{x}, 0]$ due to the reflection invariance property. Since $-x_2^*$, which is to maximize $\pi^I(-x_2|\bar{x})$ over $-x_2 \in [\underline{x}, 0]$, lies in $(\underline{x}, 0)$, it must be true that $\partial\pi^I(-x_2^*|\bar{x})/\partial(-x_2) = 0$. This requires that $\bar{x} < x^{00}$. However, this and the previous result that $x_2^* \geq x^{00}$ imply that $x_2^* > \bar{x}$, which is a contradiction. Hence, $x_1^* = \underline{x}$ cannot be the case.

ii-b) Second, suppose that $x_1^* \in (\underline{x}, 0)$. If firm 1 chooses such a location, $\partial\pi^I(x_1^*|x_2^*)/\partial x_1 = 0$; that is, $x_1^* = g(x_2^*)$. Furthermore, for the profit of firm 2, we have $\pi(x_2|x_1^*) \equiv \pi(-x_2|-x_1^*) \equiv \pi^I(-x_2|-x_1^*)$ for any $-x_2 \in [\underline{x}, 0]$ due to the reflection invariance property. Since $-x_2^*$, which maximizes $\pi^I(-x_2|-x_1^*)$ over $-x_2 \in [\underline{x}, 0]$, lies in $(\underline{x}, 0)$, $\partial\pi^I(-x_2^*|-x_1^*)/\partial(-x_2) = 0$, or $-x_2^* = g(-x_1^*)$. By differentiating both sides of the two equations, $x_1^* = g(x_2^*)$ and $-x_2^* = g(-x_1^*)$, we have $x_2^* = -x_1^*$. This configuration is symmetric and hence $x_1^* \in (\underline{x}, 0)$ cannot occur at an asymmetric equilibrium.

ii-c) Third, $x_1^* = 0$ cannot be the case. The candidate equilibrium with $(x_1^*, x_2^*) = (0, x_2)$ for $x_2 \in (0, \bar{x})$ becomes a mirror image of that with $(x_1^*, x_2^*) = (x_1, 0)$ for $x_1 \in (\underline{x}, 0)$, which has been discussed in *i-a)*.

ii-d) Fourth, $x_1^* \in (0, x_2^*)$ cannot be the case. There is no maximum in this range because firm 1's profit function is strictly convex here.

ii-e) Fifth, $x_1^* = x_2^* > 0$ cannot be the case because firm 1 can obtain a higher profit by deviating to the immediate left of firm 2.

ii-f) Sixth and finally, $x_1^* \in (x_2^*, \bar{x}]$ cannot be the case. This is because a symmetric relocation to the left part of the city is always profitable, as has been explained.

iii) Last, let us consider the case where $x_2^* = \bar{x}$. For one thing, neither $x_1^* \in (\underline{x}, 0)$ nor $x_1^* = 0$ cannot be the case, as these are mirror images of the candidate equilibrium with $(x_1^*, x_2^*) = (\underline{x}, x_2)$ for $x_2 \in (0, \bar{x})$ and that with $(x_1^*, x_2^*) = (\underline{x}, 0)$, respectively. These have been discussed in *ii-a)* and *i-a)*, respectively. Furthermore, neither $x_1^* \in (0, x_2^*)$ nor $x_1^* = x_2^*$ can be the case by the same reasons as for *ii-d)* and *ii-e)*, respectively. ■

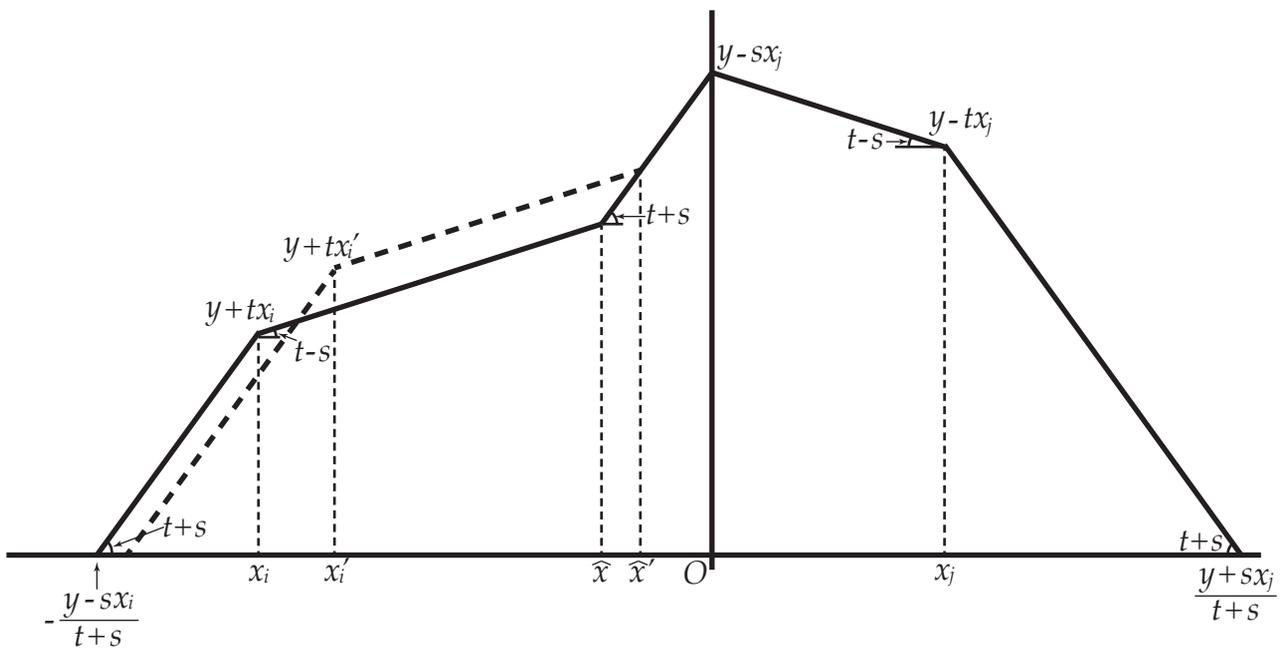


Figure 1. Bid rent

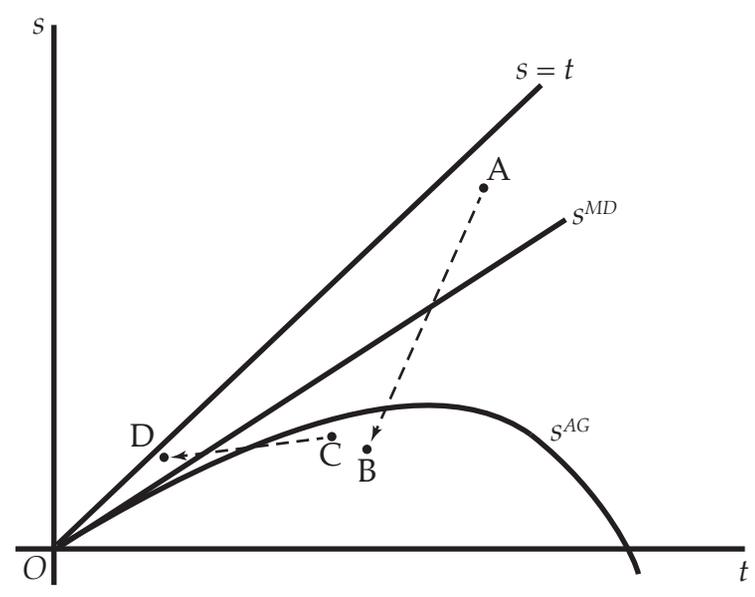


Figure 2. Effects of transport costs on the spatial structure of a city