CSIS Discussion Paper No. 158

Transportation mode choice and spatial structure of a city

Takaaki Takahashi

Center for Spatial Information Science, The University of Tokyo,
5-1-5, Kashiwa-no-ha, Kashiwa, Chiba 277-8568, Japan

March 2019

Abstract

This study analyzes urban spatial structure, paying special attention to the interplay between choice of transportation mode and urban spatial structure of a city. To this end, we extend the Alonso—Mills–Muth monocentric city framework by incorporating a transit firm. In the determination of transportation costs, workers' choice of transportation mode between rapid transit and automobiles as well as the firm's decision on prices play an important role. We show that the emergence of a rapid transit system in a city depends on a few key parameters, such as the degree of patience of the transit firm, the wage rate, monetary costs of automobiles, travel times of the two modes, and efficiency in the production of transit services.

Keywords: monetary transportation costs; population density; rapid transit; time transportation costs; transportation mode; transit firm

JEL Classification Numbers: R14; R40

E-mail address: takaaki-t@csis.u-tokyo.ac.jp

I would like to thank Yasuhiro Sato and Jacques-F. Thisse for insightful discussions. I also appreciate comments from seminar participants at various institutions. This research was partly supported by Grants in Aid for Research by the Ministry of Education, Science and Culture in Japan (No. 25285070 and No. 18H00841) and a research grant from the Obayashi Foundation in 2017.

1 Introduction

Since the seminal works of Alonso (1964), Mills (1967), and Muth (1969), a number of theoretical studies have explored the effects of transportation costs upon the spatial organization of a monocentric city (for reviews, see Anas et al. 1998; Brueckner 2001; Glaeser and Kahn 2004; Duranton and Puga 2015).[1] In recent years, there has been a resurgence of interest in the topic, and several empirical works that use the more sophisticated econometric tools have come out. For instance, Baum-Snow (2007b) conducts rigorous econometric analysis to verify that the construction of new freeways causes population decline in a central city. Mayer and Trevien (2017) reveal that the Regional Express Rail in the greater Paris region expands the employment in municipalities connected to the city, while Gonzalez-Navarro and Turner (2018) obtain evidence that the construction of subways promotes the decentralization of a city.

One of the most important factors that determine transportation costs is the transportation mode used in a city.[2] It is often argued, for example, that Los Angeles has become a vastly spread city with its population sparsely distributed owing to the transportation cost characteristics associated with the use of automobiles, whereas Paris is a quite concentrated city owing to characteristics associated with the use of horse-drawn wagons, trams, and subways. In this line, quite a few researchers study how the transportation mode used in a city shapes its spatial structure (see e.g., Anas and Moses 1979; Sasaki 1989; Baum-Snow 2007a). In particular, there is a considerable amount of research conducted on the locations of different income groups within a city when each consumer chooses a transportation mode to use from more than one alternative (LeRoy and Sonstelie 1983; Sasaki 1990; Gin and Sonstelie 1992; DeSalvo and Huq 1996; Glaeser et al. 2008).[3]

At the same time, the reverse causality from urban spatial structure to the choice of transportation mode is also salient. For instance, automobiles are widely used in Los Angeles, because the population and employment are widespread and sparsely distributed, whereas public transport plays a key role in Paris, because population and employment are concentrated in a quite limited geographical area. However, this causality has attracted the interest of only a small number of

[1]In the literature of new economic geography too, transportation costs are considered as one of the most important determinants of the spatial distribution of economic activities. In the most fundamental model, for instance, lower transportation costs work as a centripetal force (see e.g., Fujita et al. 1999; Combes et al. 2008). However, few studies in the new economic geography literature deal with the spatial distribution *within* a city.

[2]Most theoretical studies assume that transportation costs for a particular mode are given parameters. Exceptions are works by Takahashi (2006, 2011), Behrens et al. (2009), and Behrens and Picard (2011), which study how transportation costs are endogenously determined.

[3]A basic observation is that the poor, whose opportunity cost of time is lower than that of the rich, tend to use less expensive but more time-consuming public transport whereas the rich tend to use automobiles, which have the opposite characteristics. This results in a spatial structure in which the poor live near a city center while the latter live in suburbs.

researchers. First, several empirical studies examine the effect of urban spatial structure on city residents' choice of transportation mode from a given set of modes. For example, Bento et al. (2005) show that households in more compact cities are less likely to own automobiles. Second, no study deals with the effect of spatial structure on the availability of transportation modes, except Takahashi (2006), which explores the interdependence of the economic geography and the adopted transportation technology. However, Takahashi (2006) does not deal with the spatial organization of a city but rather that of two distinct regions. The purpose of this study is to shed light on this reverse causality and to explain urban spatial structure both as a consequence of the choices among different transportation modes and as affecting the set of available transportation modes.

One reason for studying this interplay between urban spatial structure and a mode choice is that it helps us to appraise the benefits of public transportation investments correctly. First, as documented in the above-mentioned studies, transportation investments drastically transform urban spatial structure in the long run because they induce city residents to use a different transportation mode. Second, the change in spatial structure alters the profitability of public transport. These processes are important because, for one thing, they determine whether a private company enters the public transportation business. Furthermore, when a local government runs the business instead, it affects how much tax revenue is needed to maintain public transportation services, or, in other words, the extent to which public transport is "sustainable" in terms of the tax burden on residents. Nevertheless, it is rare for the impacts of public transportation investments on urban spatial structure and the impacts of changes in urban spatial structure on the profitability of public transport to be considered in practical cost–benefit analyses. Even if these impacts are considered, furthermore, they are usually estimated without an adequate scientific basis.[4] At the same time, there has been recent movement in urban policies toward concentrating residences and workplaces in a narrower area around a traditional city center, which is expected to help us take advantage of agglomeration economies further and reduce environmental burdens. This often accompanies brand-new construction and extensive renovation of public transportation systems, such as subways and trams (LRT's). For this Transit Oriented Development (TOD), there is a growing need to appraise the benefits of such investments correctly.

We extend the monocentric city framework á la Alonso–Mills–Muth to include a transportation firm, which constructs and operates a rapid transit system if and only if doing so yields

[4]Recently, not a few researchers have been asserting the importance of considering "wider economic impacts," which are wider than the impacts considered in the conventional cost–benefit analysis. However, with regard to the impacts of the changes in spatial structure, emphasis is placed on the workings of agglomeration economies and changes in rent revenue. The impact on profitability of public transport is not discussed. See, for example, Venables (2007, 2016), Hensher et al. (2012), and Wangsness et al. (2017).

enough profit. In other words, whether a city witnesses a rapid transit system is equivalent to whether the (potential) transit firm can earn nonnegative profit. Furthermore, workers who commute from their residence to a city center choose from two transportation mode alternatives: automobiles and rapid transit. There are linkages between transportation modes and urban spatial structure in both directions. First, consumers' decisions on a location and the amount of land consumption depend on the mode used, which in turn depends on the availability of a rapid transit system and the prices of transit services. At the same time, the firm's decision making on entry and prices depend on the population density in the areas where it provides transit services. Specifically, prices are determined so as to balance contradicting two incentives. On the one hand, the firm has an incentive to raise prices to earn larger price margins. On the other hand, the Alonso—Mills–Muth framework implies that in a small open city, population density is negatively related to the level of transportation cost at each location.[5] Therefore, the firm has an incentive to hold prices down so as to take advantage of denser population distribution.

We show that several key parameters determine whether a rapid transit system is introduced in a city. Most important is the degree of "patience" of the transit firm. A firm that attaches sufficiently heavy weight to future profits enters the rapid transit business. This is because the introduction of a rapid transit system brings about denser population distribution in the future, which enhances future profits rather than current profit. Furthermore, for a special class of utility function, it is shown that a rapid transit system is more likely to be introduced in a city when workers there earn a higher wage, they enjoy lower utility, the monetary and time costs of automobiles are lower, the rapid transit system takes a shorter time, and the transit firm has more efficient technology. In addition, we derive parallel results for the length of a rapid transit line that the firm chooses to construct and to operate.

The rest of the paper consists of five sections. In Section 2, the model is introduced. We first present the basic framework, then introduce the transportation modes, and finally explain the structure of the game. The next section explores the properties of equilibrium. In Section 4, we derive several welfare implications. In Section 5, we specify the functional form of a utility function to obtain further properties of equilibrium. Section 6 concludes.

[5]In this study, we limit our analysis to a small open city, for which workers are allowed to migrate across cities. In a closed city, whose population is fixed, the direction of the relationship between population density and the levels of transportation costs becomes the opposite.

2 Model

2.1 Spatial structure of a city

We consider a linear city with unit width extending in a homogeneous plane. The city is the set of points within 0.5 miles of a straight line, which we interpret as the area served by a rapid transit line running along the straight line with a maximum walking distance for commuters of 0.5 miles and "stations" placed continuously. In this interpretation, it is more natural to assume a linear city than a disk-shaped city, provided that only one transit line is concerned.

The city is monocentric, that is, all production activities are conducted in the city center. In what follows, we study only the half of the city, which is spatially symmetric with respect to the center. Locations within the city are identified by the distance from the center, $d > 0$.

In the city, workers are endowed with 1 unit of time. Spending some portion of their time, they work to earn wages, whose rate is fixed at w. Their income consists entirely of the wage. First, land is owned by absentee landlords, which is a conventional assumption in the urban land use theory. Second, no profit of a transit firm, if any, is taken by workers. The owner of the transit firm either lives outside the city, or is a government that spends the earned profit for implementing policies other than a monetary transfer, such as the construction and improvement of urban infrastructure and amenities.

Workers consume land for housing, a composite good, and leisure, whose amounts are denoted by x, z, and e, respectively. Here, we take the approach of Train and McFadden (1978), in which workers, facing a trade-off between consuming a greater amount of goods and enjoying more time for leisure, freely choose the length of leisure time and thereby work hours.[6]

Furthermore, in order to work, workers need to commute to the city center, incurring monetary and time costs. We assume that the number of workdays is fixed, although the work hours per workday are freely chosen by workers. Then, the cost of commuting does not depend on the amounts of leisure consumption, and thus, we assume that it depends only on the distance of commute. The monetary and time costs for a worker living at d are denoted by $m(d)$ and $t(d)$, respectively, which are increasing functions.

Naturally, we assume that $m(d)$ does not exceed $w\left[1 - t(d)\right]$, which is the maximum amount of wage income that would be earned if a worker worked for all the available time. We call this a *feasibility condition*. Now, it is convenient to define $c(d) \equiv wt(d) + m(d)$ as a *generalized cost* of commuting inclusive of both the monetary and the time costs. It is an increasing function. Then,

[6] Another approach involves fixed work hours. The length of leisure time is automatically determined as a residual of the total available time after subtracting work hours and commuting time. The reality probably lies somewhere between the depictions of the two approaches. For the differences in their implications, see, for example, Jara-Diaz (2007).

the feasibility condition can be rewritten as

$$y(d) \equiv w - c(d) \geq 0. \tag{1}$$

Here, $y(d)$ is *potential disposable income*, which is equal to the maximum income, $w\big[1 - t(d)\big]$ less the monetary cost of commuting, $m(d)$. It is a decreasing function. Furthermore, it is noteworthy that the (1) necessitates $t(d) < 1$ for positive $m(d)$: the commute cannot use up the available time, which has been normalized to unity.

Let us consider a representative worker living at $d > 0$. Taking the composite good as a numeraire, we obtain his or her budget constraint as $r(d)x(d) + z(d) + m(d) = w\big[1 - e(d) - t(d)\big]$, or equivalently,

$$r(d)x(d) + z(d) + we(d) = y(d), \tag{2}$$

where $r(d)$ is the land rent at d.

Workers have the same preference, represented by utility function $U\big(x(d), z(d), e(d)\big)$, which they maximize subject to (2). Because the division of spending between the two goods and leisure is not a main concern in this study, we specify the utility function as $U\big(x(d), z(d), e(d)\big) = \beta \ln u\big(x(d), z(d)\big) + (1-\beta)\ln e(d)$. Then, the amount of leisure consumption for a worker at d is given by $e(d) = (1-\beta)y(d)/w$. Note that β becomes equal to the ratio of actually earned disposable income to potential disposable income, $\big\{w[1 - e(d) - t(d)] - m(d)\big\}/y(d)$. In this sense, β represents a relative measure of work hours.

Choosing x and z, a worker at d maximizes the sub-utility function, $u\big(x(d), z(d)\big)$, subject to

$$r(d)x(d) + z(d) = \beta y(d). \tag{3}$$

The first-order necessary condition is given by

$$u_1\Big(x(d),\ \beta y(d) - r(d)x(d)\Big) - r(d)u_2\Big(x(d),\ \beta y(d) - r(d)x(d)\Big) = 0, \tag{4}$$

where $u_i(\cdot)$ denotes a partial derivative of $u(\cdot)$ with respect to the i-th argument.

We suppose that the city is small and open, that is, through migration, the utility level of the workers in this city becomes, over a sufficiently long period of time, equal to that of workers outside the city, which is given and denoted by $\bar{u}$. In other words,

$$U\left(x(d),\ \beta y(d) - r(d)x(d),\ \frac{(1-\beta)y(d)}{w}\right) = \bar{u} \tag{5}$$

for any $d > 0$ at a *location equilibrium*.

Solving (4) and (5) yields an equilibrium land lot size and an equilibrium land rent. This land rent is workers' bid rent. However, since all the workers are homogeneous, this rent actually becomes a market rent. Furthermore, it is important to note that the derived size of a land lot is the demand *constrained* to yield a fixed level of utility.

Now, note that the variables associated with a particular value of d are determined independently of those associated with a different value of d. In other words, variables do not depend on macro variables, such as population distribution and geographical size of the city. This *independence property* is a consequence of the small open city approach. One of the property's implications is that the equilibrium variables at each location depend not on the transportation cost schedule over the entire city but only on the transportation cost from that particular location to the city center. In addition, we can readily verify from (4) and (5) that d affects variables only through $y(d) \equiv w - c(d)$. It follows from these observations that the land lot size and land rent at an equilibrium for each d are obtained as functions of only the location-specific potential disposable income. Thus, we can denote them as $\widetilde{x}\big(y(d)\big)$ and $\widetilde{r}\big(y(d)\big)$.

Moreover, agricultural land extends outside the city, and the land is rented out at a fixed rent, r_a. Then, the location of the city's boundary, b, is given by the solution to

$$\widetilde{r}\big(y(b)\big) = r_a. \tag{6}$$

One condition needs to be satisfied. The sum of the spending on land and the composite good must be nonnegative. This condition is expressed as $c(d) \leq w$ because of (3). We concentrate on the case in which r_a is so high and therefore, the city is so small that this condition is satisfied for any $d \in [0, b]$. This is the case when

$$c(b) \leq w. \tag{7}$$

In addition, the total population of the city becomes equal to $\int_0^b 1/\widetilde{x}\big(y(d)\big)\, \mathrm{d}d$.

It is useful to review the impacts of changes in transportation costs upon the land lot size and land rent, for they play a key role in the firm's decision making about the prices of rapid transit services. Let us define $\Phi \equiv u_{11} - 2u_{12}\widetilde{r} + u_{22}\widetilde{r}^2$ and $\Psi \equiv u_{12} - u_{22}\widetilde{r}$, where u_{ij} denotes the second-order partial derivative of $u(\cdot)$ with respect to the i-th and j-th arguments. Here, the arguments of functions are omitted for brevity. Note that $\Phi < 0$ as long as the sub-utility function is strictly quasi-concave and the first-order condition, (4), is satisfied. Furthermore, we assume that the land is a normal good, that is, $\partial\widetilde{x}/\partial\,[\beta y(d)] \geq 0$. Because differentiating (4) yields $\partial\widetilde{x}/\partial\,[\beta y(d)] = -\Psi/\Phi$, this assumption is equivalent to

$$\Psi \geq 0. \tag{8}$$

Furthermore, since the marginal rates of substitution become equal to relative prices, $U_1(x,z,e)/U_2(x,z,e) = \widetilde{r}$ and $U_1(x,z,e)/U_3(x,z,e) = \widetilde{r}/w$. Using these results and (8) after totally differentiating (4) and

(5), we obtain

$$\widetilde{x}'\big(y(d)\big) = \frac{1}{\Phi\widetilde{x}}\big[u_2 + (1-\beta)\Psi\widetilde{x}\big] < 0 \tag{9}$$

and

$$\widetilde{r}'\big(y(d)\big) = \frac{1}{\widetilde{x}\big(y(d)\big)} > 0. \tag{10}$$

Since $y(d)$ is a decreasing function of $c(d)$, *the constrained demand for land of each worker increases while the land rent declines as the transportation costs increase*. The reason is simple: as the transportation cost rises, the potential disposable income decreases. Therefore, to accomplish the same utility level, the land rent needs to decline; otherwise, the utility level would diminish. As a result, workers increase their consumption of land.

2.2 Transportation modes

In the economy, a transit firm, which constructs and operates a rapid transit system, enters the rapid transit business when it expects to earn a profit. The "firm" may be a private company or a government sector. To simplify the analysis, we confine our discussion to the case in which the firm can have only one rapid transit line running from the city center to a terminal station, whose location, denoted by l, is chosen by the firm. The cost to construct the system, which increases with the length of the line, is represented by an *investment cost function*, $\Gamma(l)$.

The firm can provide a transportation service between any point on the line and the city center. In other words, stations are placed continuously along the line. The firm can price discriminate based on station location, that is, it decides a price schedule $P \equiv \{p(d) \mid d \in (0, l]\}$, where $p(d)$ is the price of the service between d and the city center. To avoid an unnecessary complication, we concentrate on the case in which $p(d)$ is continuous. Furthermore, the marginal cost to carry passengers between d and the city center is constant and equals $a(d)$. It is assumed that $a(d)$ increases with the distance of a commute: $a'(d) > 0$.

Workers can choose a transportation mode for commuting from two alternatives. First, every worker can drive an automobile. We refer to this mode as automobile or mode A. The time and monetary costs for commuting from point $d \in (0, b]$ by this mode are denoted by $t_A(d)$ and $m_A(d)$, respectively. The generalized cost and potential disposable income equal $c_A(d) \equiv wt_A(d) + m_A(d)$ and $y_A(d) \equiv w\big[1 - t_A(d)\big] - m_A(d)$, respectively. Moreover, we assume that this mode satisfies the feasibility condition, (1), that is,

$$m_A(d) \leq w\big[1 - t_A(d)\big] \text{ for any } d \in (0, b]. \tag{11}$$

Instead, workers may have access to rapid transit mode, mode T. The transportation cost for this mode is equal to the price charged by the transit firm, that is, $m(d) \equiv p(d)$ for mode T. Furthermore, for this mode, $t(d)$, $c(d)$ and $y(d)$ are denoted by $t_T(d)$, $c_T(d)$ and $y_T(d)$, respectively, where $c_T(d) \equiv wt_T(d) + p(d)$ and $y_T(d) \equiv w\big[1 - t_T(d)\big] - p(d)$.

2.3 Structure of the game

Before presenting a formal structure of the game to be analyzed, we explain how rapid transit services are launched in a city and how its spatial structure changes.

Initially, all the workers in the city use automobiles. Therefore, its initial spatial structure is prescribed by the transportation cost schedule associated with mode A, which is given by $\Big\{c_A(d) \mid d \in (0, b_A]\Big\}$. Here, b_A is the location of a boundary of this city, that is, the solution to $\widetilde{r}\big(y_A(b_A)\big) = r_a$ (see (6)). We say that a city with such a spatial structure is *A type*.

Then, a transit firm makes a decision on whether to enter the rapid transit business. When deciding to do so, it chooses the length of a rapid transit line, l, for which the city becomes an *l-mile transit city*. In addition, the firm decides the prices of its services. When both modes of transport are available, workers decide which mode to use. We denote the set of locations of the workers who use mode T by $D(P, l) \subset (0, l]$ when the price schedule is P. After the inauguration of a transit line, the firm earns revenue.

At the same time, workers change their amount of land consumption in response to the introduction of a new transportation mode, which alters the urban spatial structure. However, the change proceeds gradually, because housing is more or less durable and its economic value does not decay instantaneously. After a sufficiently long period of time, the spatial structure reaches a stationary state corresponding to the transportation cost schedule in an l-mile transit city, which is given by $\Big\{c_T(d) \mid d \in D(P, l)\Big\} \cup \Big\{c_A(d) \mid d \in (0, b_T(l)] - D(P, l)\Big\}$. Here, $b_T(l)$ denotes the location of a boundary of this city. We say that a city with such a spatial structure is $T(P, l)$ *type*.

Because the spatial structure has the characteristic that it is not adjusted instantaneously, there is ambiguity about when the firm makes its decision on prices. We can assume, for instance, that the firm sets prices only once and to retain them during and after the adjustment period in the spatial structure. Alternatively, it may be the case that the firm chooses a sequence of prices over a certain period of time to maximize its inter-temporal sum of profits. Nevertheless, our interest does not lie in the process of decision making itself, but rather in the profitability of the rapid transit investment.[7] Therefore, we take one of the simplest approaches: two periods. In one period, which we refer to as the "short run," the spatial structure of the city does not change but the city remains an A type. The firm chooses prices *taking the structure at the initial state as given* and receives revenue. The short-run prices and the short-run price schedule are denoted by $p^S(d)$

[7] Further complicating the matter is that the *expectation* of a future spatial structure can play a significant role in the firm's decision making.

and $P^S \equiv \{p^S(d) \mid d \in (0,l]\}$, respectively. By contrast, the spatial structure changes in the other period, the "long run." First, the firm sets new prices. Then, workers move one after another and the spatial structure is gradually adjusted. The firm, without changing prices, receives revenue after all the adjustments are completed. An important point is that in the long run, the firm chooses prices *based on the profit earned when the spatial structure reaches a stationary state*. The long-run prices and the long-run price schedule are denoted by $p^L(d)$'s and $P^L \equiv \{p^L(d) \mid d \in (0,l]\}$, respectively.

Because the population density at d equals $1/x(d)$, the operating profits raised from a worker there in the two periods are given by

$$\pi^S\big(d, p^S(d)\big) = \frac{p^S(d) - a(d)}{\widetilde{x}\big(y_A(d)\big)} \quad \text{and} \quad \pi^L\big(d, p^L(d)\big) = \frac{p^L(d) - a(d)}{\widetilde{x}\big(y^L(d)\big)},$$

respectively. Here, $y^L(d)$ is the potential disposable income of the worker who pays $p^L(d)$ for the transit service, that is, $y^L(d) \equiv w\big[1 - t_T(d)\big] - p^L(d)$ (the relevant subscript T is dropped from $y^L(d)$ for a simpler exposition). The aggregates of the operating profits received in the two periods are given by

$$\Pi^S(P^S, l) = \int_{d \in D(P^S, l)} \pi^S\big(d, p^S(d)\big)\, \mathrm{d}d \quad \text{and} \quad \Pi^L(P^L, l) = \int_{d \in D(P^L, l)} \pi^L\big(d, p^L(d)\big)\, \mathrm{d}d.$$

The firm enters business if the *gross profit*, $\Pi(P^S, P^L; l) \equiv \rho\Pi^S(P^S, l) + (1-\rho)\Pi^L(P^L, l)$, is at least as high as the investment cost, or equivalently, if the *net profit*, $\Omega(P^S, P^L; l) \equiv \Pi(P^S, P^L; l) - \Gamma(l)$, is nonnegative:

$$\Omega(P^S, P^L; l) \geq 0. \tag{12}$$

Here, ρ denotes the weight of the operating profit in the short run ($1-\rho$ in the long run). This parameter can be interpreted in several ways. First, ρ can represent the degree of the firm's time preference: a more patient firm has lower ρ. Second, it is related to the adjustment speed of the spatial structure. When workers relocate more quickly, the spatial structure reaches a stationary state more rapidly. In this case, the firm places more weight on long-run profit, that is, ρ is lower.

Now, we are equipped to present our three-stage game. We consider a city that is initially an A type. In the first stage, a transit firm decides whether to enter the rapid transit business. It decides to do so if and only if the net profit is nonnegative, that is, if (12) holds. In addition, the firm chooses l so as to maximize $\Omega(P^S, P^L; l)$. In the second stage, which corresponds to the short run, the firm sets P^S that maximizes $\Pi^S(P^S, l)$. The city remains an A type regardless of its chosen price schedule. At the same time, workers decide the amounts of a composite good and leisure to consume, and the transportation mode to use. Furthermore, the firm receives $\Pi^S(P^S, l)$ at the end of this stage. In the third and final stage, which corresponds to the long run, the firm charges P^L that maximizes $\Pi^L(P^L, l)$. The definition of $\Pi^L(P^L, l)$ implies that the firm takes into account the effect of its decision on the spatial structure. In other words, the firm maximizes the operating profit that is earned *after all the adjustments in the spatial structure in response to P^L are carried out*. Workers decide the amounts of consumption and the transportation mode to use, which transforms the city into a $T(P^L, l)$ type. Finally, the firm receives $\Pi^L(P^L, l)$.[8]

3 Equilibrium

We begin the analysis by examining workers' decision on the consumptions of a composite good and leisure in the second stage. Next, we study their choice of transportation mode. Then, the firm's decision-making processes are analyzed. Specifically, we discuss the decisions on price schedules, the length of a transit line, and finally, whether to enter the rapid transit business.

3.1 Consumption of a composite good and leisure in the second stage

In the second stage, the city remains an A type, that is, the amount of land consumed by each worker is fixed at $\tilde{x}\big(y_A(d)\big)$. Therefore, workers consume a composite good and leisure by the amounts dictated by their budget constraint, $\beta y_A(d) - \tilde{r}\big(y_A(d)\big)\tilde{x}\big(y_A(d)\big)$ units and $(1-\beta)y_A(d)/w$ units, respectively.

3.2 Transportation mode to use

Consider a worker at d who has access to mode T in addition to mode A. The worker uses the mode that gives a higher utility. Let us assume that workers use mode T when indifferent between the two modes.[9] If $y_T(d) \geq y_A(d)$,

$$U\left(x(d),\ \beta y_T(d) - r(d)x(d),\ \frac{(1-\beta)y_T(d)}{w}\right) \geq U\left(x(d),\ \beta y_A(d) - r(d)x(d),\ \frac{(1-\beta)y_A(d)}{w}\right). \tag{13}$$

for any $x(d)$ and any $r(d)$, since the utility function is increasing in each argument. Therefore, the worker uses mode T. Instead, if $y_T(d) < y_A(d)$, the direction of the inequality in (13) becomes the opposite and consequently, the worker uses mode A regardless of $x(d)$ and $r(d)$. Because no workers living outside the city boundary commute, we conclude that the range of the locations

[8]One may doubt the time consistency, as after workers finish relocating in the last stage, the transit firm may have an incentive to change prices. This would hold if the spatial structure does not change after the firm changes prices. In fact, if prices change, workers relocate and the spatial structure changes again. Because the firm earns higher profit in the original stationary state than in the new stationary state, it is discouraged from changing prices.

[9]This assumption is made only for a technical reason. We can obtain similar results assuming otherwise.

of mode-T users is given by

$$D(P,l) = \left\{ d \;\middle|\; y_T(d) \geq y_A(d),\ d \in \Big(0, \min\{b_T(l), l\}\Big] \right\}. \tag{14}$$

This is equivalent to the area where the generalized cost of mode T is lower than that of mode A, namely, $c_T(d) \leq c_A(d)$.

3.3 Prices of transit services

In the second and third stages, the transit firm chooses the price schedules that maximize its gross profit. Note that profit maximization in the short run is independent of that in the long run. Furthermore, in each of the two periods, the firm price discriminates based on station locations. Therefore, for each d, the firm maximizes $\pi^S\big(d, p^S(d)\big)$ and $\pi^L\big(d, p^L(d)\big)$ in the two periods, respectively.

The firm faces two upper bounds for prices. First, if a price does not satisfy the feasibility condition for mode T, no worker uses that mode. Therefore, the firm does not charge a price higher than a *feasibility price*, $\bar{p}(d) \equiv w\big[1 - t_T(d)\big]$. Second, it does not charge a price higher than a *reservation price* for workers, $p^0(d)$, which is the highest price for which they use mode T rather than mode A. Since they use mode T if and only if $c_T(d) \leq c_A(d)$, the reservation price is derived from $c_T(d) = c_A(d)$: $p^0(d) \equiv m_A(d) + w\big[t_A(d) - t_T(d)\big]$. Now, we can verify that (11) implies that $\bar{p}(d) \geq p^0(d)$. Provided that mode A satisfies the feasibility condition, the binding upper bound for prices is $p^0(d)$.

To begin, we examine decision making in the second stage. Since the city remains an A type regardless of the price the firm chooses, it can increase $\pi^S\big(d, p^S(d)\big)$ by charging a higher price provided that it does not exceed $p^0(d)$. Therefore, the firm chooses a limit price equal to $p^0(d)$:

$$p^{S*}(d) = p^0(d) \equiv m_A(d) + w\big[t_A(d) - t_T(d)\big], \tag{15}$$

where the asterisk denotes an equilibrium value. Two observations follow. First, the potential disposable income of a worker who uses mode T becomes equal to $y_A(d)$ for this price. Second, (15) implies that the equilibrium price is higher when the monetary and time costs of mode A are higher, the time cost of mode T is lower, and the wage rate is higher.

One additional requirement is that $p^0(d)$ should be higher than the marginal cost, $a(d)$. If this were not the case, then charging $p^0(d)$ would cause the firm to suffer a loss or break even. In the first case, it would rather choose a price higher than $p^0(d)$ to discourage workers from using mode T. In the second case, the firm might choose such a price, because doing so yields the same amount of profit (i.e., 0). To exclude these possibilities, we assume that $p^0(d) > a(d)$, or equivalently, $c_A(d) - wt_T(d) > a(d)$; that is, the generalized cost of mode A is sufficiently high and the time cost of mode T is sufficiently low compared to the marginal cost.

It follows from (14) that the equilibrium range of the locations of mode-T users is given by

$$D\big(P^{S*}, l\big) = \big(0,\ \min\{l, b_A\}\big], \tag{16}$$

where P^{S*} is the price schedule that consists of $p^{S*}(d)$'s. Thus, all the workers in this A-type city who have access to mode T use it.[10]

Next, we study the long-run prices that the firm chooses in the second stage. Consider a price of the service from a particular location, d. Note that $p^L(d)$ affects $\pi^L\big(d, p^L(d)\big)$ not only directly but also indirectly through $\tilde{x}\big(y^L(d)\big)$. On the one hand, as the firm raises $p^L(d)$, its long-run profit increases to the extent that a price mark-up rate rises. On the other hand, the constrained demand for land by each worker increases, as shown in (9). This leads to a less dense population distribution, which reduces the firm's profit. The price is determined at the level that balances these two effects, the positive *mark-up rate effect* and the negative *density effect*.

Which effect is dominant depends on the elasticity of the constrained demand for land with respect to the price of a transit service, $\varepsilon\big(p^L(d)\big) \equiv \mathrm{d}\ln\tilde{x}\big(y^L(d)\big)/\mathrm{d}\ln p^L(d) > 0$. To ensure the existence and uniqueness of an equilibrium, we impose three regularity assumptions:

$$\varepsilon'\big(p^L(d)\big) > 0, \qquad \lim_{p^L(d)\to 0+} \varepsilon\big(p^L(d)\big) = 0, \quad \text{and} \quad \lim_{p^L(d)\to \bar{p}(d)} \varepsilon\big(p^L(d)\big) > \zeta \text{ for any } \zeta. \tag{17}$$

These assumptions seem to be innocuous, as they are indeed satisfied, as we show in Subsection 5.1, when the sub-utility function, $u(x,z)$, is log-linear. The $1/\varepsilon$ *curve* in Fig. 1 shows how the inverse of the elasticity changes with the price. By (17), the curve is downward sloping, goes to positive infinity as $p^L(d)$ approaches 0 from above, and finally, approaches 0 as $p^L(d)$ diverges to the feasibility price, one of the upper bounds for prices.

Insert Fig. 1 around here.

Furthermore, it is readily verified that $\partial\pi^L\big(d, p^L(d)\big)/\partial p^L(d) \lesseqqgtr 0$ if $1/\varepsilon\big(p^L(d)\big) \lesseqqgtr \mu\big(p^L(d)\big)$ for a given value of d where $\mu\big(p^L(d)\big) \equiv \big[p^L(d) - a(d)\big]/p^L(d)$ is a mark-up rate. Therefore, the profit is maximized at $p^L(d) = \hat{p}(d)$ that solves

$$\frac{1}{\varepsilon\big(\hat{p}(d)\big)} = \mu\big(\hat{p}(d)\big), \quad \text{or equivalently,} \quad \frac{\hat{p}(d) - a(d)}{\tilde{x}\big(\hat{y}(d)\big)} = -\frac{1}{\tilde{x}'\big(\hat{y}(d)\big)}. \tag{18}$$

Here, $\hat{y}(d)$ is potential disposable income associated with $\hat{p}(d)$ and t_T, that is, $\hat{y}(d) \equiv w\big[1 - t_T(d)\big] - \hat{p}(d)$. This result corresponds to the well-known result that a monopolist maximizes

[10] We can say that the city is not only an A type but also a $T(P^{S*}, l)$ type, because the $T(P^{S*}, l)$-type city becomes isomorphic to the A-type city.

its profit by charging the price that equalizes the inverse of demand elasticity with a mark-up rate. Note that (18) indicates that the equilibrium price mark-up is smaller when the constrained demand is more elastic. This is because the negative density effect is stronger, which discourages the firm from setting a high price.

Fig. 1 is useful for illustrating the solution to (18). The figure describes $\mu\big(p^L(d)\big)$ by an upward-sloping μ *curve*, which passes through $\big(a(d), 0\big)$ and asymptotically approaches a horizontal line, $1/\varepsilon\big(p^L(d)\big) = 1$. The profit-maximizing price is given at the intersection of this curve and the $1/\varepsilon$ curve. Since the former is upward sloping while the latter is downward sloping, they intersect each other only once. It immediately follows from the above properties of the μ curve that $\hat{p}(d) > a(d)$.

For $\hat{p}(d)$ to be the actual profit-maximizing price, it must be true that $\hat{p}(d) \leq p^0(d)$ for each d. Otherwise, no workers at d would use mode T and $p^0(d)$ would yield higher profit than $\hat{p}(d)$. Furthermore, if $\hat{p}(d) = p^0(d)$ for *all* the relevant values of d, the introduction of a transit system would have no impact on the spatial structure. To rule out such an uninteresting circumstance, we focus on the case in which $\hat{p}(d) \neq p^0(d)$ for some d. These two conditions are combined into one: $\hat{p}(d) \leq p^0(d)$ for any $d \in (0, l]$ with the inequality sign for some d. Because $\hat{p}(d)$ does not depend on $p^0(d)$, we can safely consider such a high value of the given parameter $p^0(d)$. Since $\hat{p}(d) \left\{ \begin{smallmatrix} \leqq \\ > \end{smallmatrix} \right\} p^0(d)$ if $\hat{y}(d) \left\{ \begin{smallmatrix} \geqq \\ < \end{smallmatrix} \right\} y_A(d)$, this condition is rewritten as

$$\hat{y}(d) \geq y_A(d) \text{ for any } d \in (0, l] \quad \text{and} \quad \hat{y}(d) > y_A(d) \text{ for some } d \in (0, l]. \tag{19}$$

Because this prescribes that mode T be affordable enough, we refer to it as an *affordability condition*.

We have established that

$$p^{L*}(d) = \hat{p}(d), \tag{20}$$

as long as the regularity condition, (17), and the affordability condition, (19), are satisfied.

Let $\hat{b}$ be the solution to $\widetilde{r}\big(\hat{y}(\hat{b})\big) = r_a$: it is the location of a boundary of a hypothetical city in which the monetary and time costs of commuting are equal to $\hat{p}(d)$ and $t_T(d)$, respectively, not only in the range where transit services are available but *in the entire range* of the city. It immediately follows from (19) that all the workers in a $T(P^{L*}, l)$-type city who have access to mode T use that mode. Therefore,

$$D(P^{L*}, l) = \big(0,\ \min\{l, \hat{b}\}\big]. \tag{21}$$

Putting all the results together, we express the gross profit evaluated at the equilibrium prices as a function of the length of a transit line. Specifically, this *gross profit function* is derived by putting (15), (18), and (20) into $\Pi(P^S, P^L; l)$ and using (16) and (21):[11]

$$\Pi(P^{S*}, P^{L*}; l) = \rho \int_0^{\min\{l, b_A\}} \pi^S\big(d, p^{S*}(d)\big)\, \mathrm{d}d + (1-\rho) \int_0^{\min\{l, \hat{b}\}} \pi^L\big(d, p^{L*}(d)\big)\, \mathrm{d}d, \tag{22}$$

where

$$\pi^S\big(d, p^{S*}(d)\big) = \frac{p^0(d) - a(d)}{\widetilde{x}\big(y_A(d)\big)} \quad \text{and} \quad \pi^L\big(d, p^{L*}(d)\big) = -\frac{1}{\widetilde{x}'\big(\hat{y}(d)\big)}.$$

Before closing this subsection, it is useful to study the location of the boundary of a $T(P^{L*}, l)$-type city, denoted by $b^L(l)$. The affordability condition, (19), implies that $\hat{b} \geq b_A$ since $\widetilde{r}'\big(y(d)\big) > 0$ (see (10)). Thus, we can identify three cases depending on the size of l:

$$\begin{cases} b^L(l) = b_A & \text{if} \quad l < b_A, \\ b^L(l) = l & \text{if} \quad l \in [b_A, \hat{b}], \\ b^L(l) = \hat{b} & \text{if} \quad l > \hat{b}. \end{cases} \tag{23}$$

In the first case, the workers living in $(0, l]$ use mode T while those in $(l, b_A]$ use mode A. The land rent discretely changes at $d = l$. In the second case, those living in $(0, l]$ use mode T and no workers use mode A. In the last case, those living in $(0, \hat{b}]$ use mode T and no workers use mode A. These results are summarized in Fig. 2.

Insert Fig. 2 around here.

3.4 Length of transit line

In the first stage, the firm determines the length of a transit line so as to maximize $\Omega(P^{S*}, P^{L*}; l) = \Pi(P^{S*}, P^{L*}; l) - \Gamma(l)$. We denote such l by l^*.

We begin the analysis with the observation that the firm does not extend a line beyond $\hat{b}$: $l^* \leq \hat{b}$. This is because no workers beyond this location use mode T although the firm needs to pay additional investment costs for that section of the line. This, together with (23), implies that

$$b^L(l^*) = \max\{l^*, b_A\}. \tag{24}$$

Since $b_A \leq \hat{b}$, the entire space is divided into five sub-spaces: $\{0\}$, $\underline{D} \equiv (0, b_A)$, $\{b_A\}$, $\overline{D} \equiv (b_A, \hat{b}]$ and $\{d | d > \hat{b}\}$. $\underline{D}$ is a “built-up area” within the boundary of an A-type city. Instead, $\overline{D}$ is an “undeveloped area” beyond that boundary but within the boundary of the hypothetical city, where the monetary and time costs are $\hat{p}(d)$ and $t_T(d)$, respectively, for any d. The terminal

[11]The integrals in (22) go from $d = 0$. However, the prices, in fact, are not defined for $d = 0$. Nonetheless, the expression is still correct, because $d = 0$ occurs with measure 0.

station is placed in one of the three sub-spaces, $\underline{D}$, $\{b_A\}$, or $\overline{D}$. It is noteworthy that a transit firm may build a terminal station in the undeveloped area. In that case, the section of a transit line beyond b_A is a waste in the short run because that section carries no workers. However, in the long run, the city may expand, and the terminal station may fall within a new city boundary. If the construction cost is low enough, the firm may make such a preemptive investment decision.

To examine the maximization problem, it is convenient to derive the "marginal profit" with respect to the length of a transit line:

$$\frac{\partial \Pi(P^{S*}, P^{L*}; l)}{\partial l} = \begin{cases} \underline{\Delta}(l) \equiv \rho \pi^S\big(l, p^{S*}(l)\big) + (1-\rho)\pi^L\big(l, p^{L*}(l)\big) > 0 & \text{if} \quad l \in \underline{D} \\ \overline{\Delta}(l) \equiv (1-\rho)\pi^L\big(l, p^{L*}(l)\big) > 0 & \text{if} \quad l \in \overline{D}. \end{cases} \tag{25}$$

The two terms in $\underline{\Delta}(l)$ represent the short-run and long-run effects of an infinitesimal extension of a transit line in the built-up area, respectively. However, when construction of the rapid transit system extends into the undeveloped area, the short-run effect disappears, which results in the second line in (25). Several properties of the gross profit function follow from (25). First, it is an increasing function of l. Second, it is continuous but has a kink at $l = b_A$: it is steeper in the left neighborhood of b_A than in its right neighborhood, because $\underline{\Delta}(b_A) > \overline{\Delta}(b_A)$. Third, the function might or might not be concave in l.[12] In what follows, however, we limit our attention to a concave gross profit function. One reason is that it is indeed concave when the sub-utility function is log-linear, as will be shown in Subsection 5.2. Fig. 3 depicts $\Pi(P^{S*}, P^{L*}; l)$ as a function of l. The line, which we refer to as a *gross profit curve*, is upward sloping and concave with a kink at $l = b_A$.

Insert Fig. 3 around here.

To derive further insights, we focus on a linear or convex investment cost function. In other words, we assume that $\Gamma''(l) \geq 0$. The function is represented in Fig. 3 by an *investment cost curve*.

As long as the gross profit curve is concave, l^* is shorter than b_A if the investment cost curve is steeper than the gross profit curve at $l = b_A$. Conversely, l^* is longer than b_A if the gross profit curve at $l = b_A$ is steeper than the investment cost curve. Thus, we establish the following result.

[12] The sign of

$$\frac{\partial^2 \Pi(P^{S*}, P^{L*}; l)}{\partial l^2} = \begin{cases} \dfrac{\rho}{\widetilde{x}\big(y_A(l)\big)} \left[p^{0\,\prime}(l) - a'(l) - \dfrac{\{p^0(l) - a(l)\}\, \widetilde{x}'\big(y_A(l)\big)\, y_A{}'(l)}{\widetilde{x}\big(y_A(l)\big)} \right] + \dfrac{(1-\rho)\, \widetilde{x}''\big(\hat{y}(l)\big)\, \hat{y}'(l)}{\widetilde{x}'\big(\hat{y}(l)\big)^2} & \text{if} \quad l \in \underline{D}, \\ \dfrac{(1-\rho)\, \widetilde{x}''\big(\hat{y}(l)\big)\, \hat{y}'(l)}{\widetilde{x}'\big(\hat{c}(l)\big)^2} & \text{if} \quad l \in \overline{D} \end{cases}$$

is ambiguous.

Proposition 1

Suppose that the gross profit function is concave and that the investment cost function is linear or convex. Then,

$$\begin{cases} l^* \in \underline{D} & \textit{if} \quad \Gamma'(b_A) > \underline{\Delta}(b_A) \\ l^* = b_A & \textit{if} \quad \Gamma'(b_A) \in \left[\overline{\Delta}(b_A), \underline{\Delta}(b_A)\right] \\ l^* \in \overline{D} & \textit{if} \quad \Gamma'(b_A) < \overline{\Delta}(b_A). \end{cases}$$

According to this proposition, whether a transit line goes beyond b_A depends on the relative size of the slope of the gross profit curve to that of the investment cost curve at b_A. When the former is smaller than the latter (the first case), the transit line is limited within the boundary, that is, in the built-up area. If $\lim_{l \to 0^+} \underline{\Delta}(l) > \Gamma'(l)$, the optimal length is not 0 but positive and l^* is uniquely given by the solution to $\underline{\Delta}(l^*) = \Gamma'(l^*)$. On the contrary, if the gross profit curve is steeper than the investment cost curve at b_A (the third case), the transit line extends to the undeveloped area, transcending the boundary. In this case, our problem has an inner solution unless the gross profit curve is still steeper at $\hat{b}$. In other words, l^* is given by the solution to $\overline{\Delta}(l^*) = \Gamma'(l^*)$ if $\overline{\Delta}(\hat{b}) < \Gamma'(\hat{b})$. We assume that the two conditions for an inner solution are satisfied so that either $\underline{\Delta}(l^*) = \Gamma'(l^*)$ or $\overline{\Delta}(l^*) = \Gamma'(l^*)$ holds. In other words, we focus on the equilibrium length that satisfies

$$\frac{\partial \Pi\left(P^{S*}, P^{L*}; l^*\right)}{\partial l} = \Gamma'(l^*). \tag{26}$$

It is straightforward to examine the effects of changes in parameters on the optimal length. Suppose that a certain parameter, λ, changes by such a small amount that the associated l^* stays within the same sub-space, either the built-up area or the undeveloped area. Totally differentiating (26) yields

$$\frac{\mathrm{d}l^*}{\mathrm{d}\lambda} = \frac{\partial^2 \Pi\left(P^{S*}, P^{L*}; l^*\right)}{\partial l \partial \lambda} \Bigg/ \left[\Gamma''(l^*) - \frac{\partial^2 \Pi\left(P^{S*}, P^{L*}; l^*\right)}{\partial l^2}\right]. \tag{27}$$

The denominator on the right-hand side is positive, since we consider a concave gross profit function and a linear or convex investment cost function. If the gross profit curve becomes steeper (i.e., marginal profit declines) as a result of the increase in λ, the numerator on the right-hand side of (27) is positive, and therefore, the line becomes longer. On the other hand, if the curve becomes less steep, the line becomes shorter.

Unfortunately, most of the parameters affect $\partial \Pi\left(P^{S*}, P^{L*}; l^*\right)/\partial l$ through more than one channel and the signs of the overall effects are ambiguous. One exception is ρ. Note that

$$\pi^S\left(d, p^{S*}(d)\right) = \frac{p^0(d) - a(d)}{\widetilde{x}\left(y_A(d)\right)} \leq \max_{p^L(d)} \frac{p^L(d) - a(d)}{\widetilde{x}\left(y^L(d)\right)} = \pi^L\left(d, p^{L*}(d)\right) \tag{28}$$

for any $d \in [0, l]$. This is because in the long run, the firm can choose the price of $p^0(d)$ if it wants to, in which case the size of a land lot becomes equal to $\tilde{x}\big(y_A(d)\big)$. The next proposition follows from this result.

Proposition 2

A more patient firm chooses a longer length of transit line.

Proof

Equation (28) implies that $\partial\underline{\Delta}(l)/\partial\rho = \pi^S\big(l, p^{S*}(l)\big) - \pi^L\big(l, p^{L*}(l)\big) \leq 0$ (see (25)). In addition, we obtain $\partial\overline{\Delta}(l)/\partial\rho \leq 0$. Consequently, $\partial^2\Pi\big(P^{S*}, P^{L*}; l^*\big)/\partial l \partial\rho \leq 0$ and it follows from (27) that $\mathrm{d}l^*/\mathrm{d}\rho \leq 0$ for any $l \in \underline{D} \cup \overline{D}$. ■

The reason for this result is that a patient firm places greater value on relatively higher long-run profit, which boosts the profitability of investment.

3.5 Entry of a transit firm

In this subsection, we examine how various factors affect the conditions for the firm's entry, namely, $\Omega(P^{S*}, P^{L*}; l^*) \geq 0$.

First, it is obvious that the equilibrium profit is higher when the investment cost is lower, and thus, the entry is more likely to occur.

Second, (22) implies that when the absolute value of $\tilde{x}'\left(\hat{y}(d)\right)$ is lower at each location, $\Omega(P^{S*}, P^{L*}; l^*)$ is larger and therefore, a rapid transit system is more likely to be introduced in a city. As explained, population distribution becomes sparser as the price of the transit service rises. This density effect is weaker when the constrained demand is less elastic. To this extent, the firm is encouraged to raise the price. Thus, the less elastic constrained demand is, the greater is profit.

Third, it follows from the envelope theorem that

$$\frac{\Pi\big(P^{S*}, P^{L*}; l^*\big)}{\partial\rho} = \begin{cases} \displaystyle\int_0^{b_A} \pi^S\big(d, p^{S*}(d)\big) - \pi^L\big(d, p^{L*}(d)\big)\,\mathrm{d}d - \int_{b_A}^{l^*} \pi^L\big(d, p^{L*}(d)\big)\,\mathrm{d}d \leq 0 \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{if} \quad l^* > b_A, \\ \displaystyle\int_0^{l^*} \pi^S\big(d, p^{S*}(d)\big) - \pi^L\big(d, p^{L*}(d)\big)\,\mathrm{d}d \leq 0 \quad \text{otherwise,} \end{cases}$$

where the inequalities follow from (28). The gross profit is larger for a more patient firm.

The effects of changes in the other parameters are ambiguous. The following proposition summarizes the results in this subsection.

Proposition 3

It is more likely that a rapid transit system is introduced in a city when

i) the investment cost is lower,

ii) the constrained demand for land is less elastic, and

iii) the transit firm is more patient.

4 Welfare implications

This section discusses the effects of the introduction of a rapid transit system on the welfare of a city.

First, as in the conventional Alonso—Mills–Muth model, the utility level for each worker in a city ends up being equal to a level that is pervasive outside the city, namely, $\bar{u}$, owing to the assumption of a small open city. Caution is necessary. In the short run, it is possible that this utility equalization does not hold for the following reason. In that period, workers in a city first decide the amounts of land to consume assuming that they use mode A. Then, migration occurs until their utility levels are equalized with $\bar{u}$. However, thereafter, a rapid transit system may be introduced in the city. In that case, the potential disposable income of a worker who switches a mode from A to T changes and therefore, his or her utility level diverges from $\bar{u}$. This argument is true *ex ante* but not *ex post*, as the transportation costs of mode T become equal to those of mode A, that is, $c_T(d) = c_A(d)$ for any $d \in (0, l]$, *at the equilibrium*. Therefore, the worker actually enjoys the same utility level after he or she changes the transportation mode of use.[13]

Because the utility level remains unchanged, the introduction of a rapid transit system does not affect social welfare of workers living in the city. However, a different story emerges if we consider the welfare of absentee landlords, who receive land rents, and absentee owners of the transit firm, who receive profit shares, in addition to the workers living in the city (when a local government runs the rapid transit system, the welfare of absentee owners is replaced by the welfare of taxpayers).

To derive the aggregate land rent, we add up the rents in the short and long runs attaching the same weights, ρ and $1-\rho$, as the transit firm does when it sums up the short-run and long-run operating profits to determine the gross profit. Then, the aggregate land rent in a city without a rapid transit system equals $\int_0^{b_A} \widetilde{r}\left(y_A(d)\right) \mathrm{d}d$ while that in a city with a rapid transit system

[13] This possibility of equality not holding does not arise in the long run, even ex ante. In the long run, the transportation mode that workers plan to use when they make decisions on consumption coincides with the mode that they actually use. This is because workers know the availability of mode T at the time of their decision making, and the transit firm keeps committing to the original price it offered.

equals $\rho \int_0^{b_A} \widetilde{r}\big(y_A(d)\big)\, \mathrm{d}d + (1-\rho)\left[\int_0^{l^*} \widetilde{r}\big(\hat{y}(d)\big)\, \mathrm{d}d + \int_{l^*}^{b_A} \widetilde{r}\big(y_A(d)\big)\, \mathrm{d}d\right]$ (recall that the boundary of a $T(P^{L*}, l^*)$-type city is located at $\max\{b_A, l^*\}$, see (24)). Therefore, the increase in the aggregate land rent due to the introduction of a rapid transit system equals

$$\Delta R = (1-\rho)\int_0^{l^*} \widetilde{r}\big(\hat{y}(d)\big) - \widetilde{r}\big(y_A(d)\big)\mathrm{d}d.$$

Since $\widetilde{r}\,(y(d))$ is an increasing function (see (10)), the affordability condition, (19), implies that ΔR is positive. The aggregate land rent increases as a result of the introduction of a rapid transit system.

The size of the increase in aggregate land rent depends on how far $\hat{y}(d)$ exceeds $y_A(d)$. Since $\hat{y}(d)$ decreases with $\hat{p}(d)$, the increase in aggregate land rent is greater when $\hat{p}(d)$ is lower, other things being equal.

Proposition 4

The aggregate land rent increases as a result of the introduction of a rapid transit system. The aggregate land rent increases more when the transit firm chooses lower long-run prices.

The last approach is to define social welfare as the aggregate of the welfare of city residents, absentee landlords, and absentee owners of a transit firm. Since ΔR is positive, social welfare improves as a result of the introduction of a rapid transit system even if the transit firm suffers a loss, as long as the loss is sufficiently small. In that case, with some redistribution schemes, the city can successfully establish a rapid transit system. For instance, a government may be able to collect part of the increase in the aggregate land rent by some policy instruments, such as taxation, and spends the raised money on making up the loss of the transit firm. The same scenario is valid when the transit system is run by a local government: it can absorb some of the rent increase to cover its deficit in the transportation sector.

5 Special case of a log linear utility function

One way to examine equilibrium further is to specify the functional form of a utility function. Thus, in this section, we consider a log linear sub-utility function and explore the properties of the equilibrium.

Suppose that $u(x,z) \equiv \alpha \ln x + (1-\alpha)\ln z$ for $\alpha \in (0,1)$. Let $\theta \equiv \alpha\beta$ and $K \equiv \big[(1-\alpha)\beta\big]^{\frac{1-\alpha}{\alpha}} (1-\beta)^{\frac{1-\beta}{\theta}} \bar{u}^{-\frac{1}{\theta}} w^{-\frac{1-\beta}{\theta}}$, both of which are positive constants. The constrained demand for land and the land rent that simultaneously solve (4) and (5) become

$$\widetilde{x}\big(y(d)\big) = K^{-1}y(d)^{-\frac{1-\theta}{\theta}} \quad \text{and} \quad \widetilde{r}\big(y(d)\big) = K\theta y(d)^{\frac{1}{\theta}},$$

respectively. The location of a city boundary is derived from (6) as a solution to $c(b) = w - \left[r_a/(K\theta)\right]^{\theta}$, which implies that the boundary qualification (7) is always satisfied for our preference.

5.1 Long-run prices of transit services

Next, we examine the long-run prices of transit services.

The long-run operating profit is reduced to

$$\pi^L\big(d, p^L(d)\big) = K\big[\bar{p}(d) - p^L(d)\big]^{\frac{1-\theta}{\theta}}\big[p^L(d) - a(d)\big].$$

Because the elasticity of the constrained demand, $\varepsilon\big(p^L(d)\big) = (1-\theta)p^L(d)/\big[\theta y(d)\big]$, satisfies the three regularity conditions in (17), the maximum is indeed attained at $\hat{p}(d)$, which solves (18):

$$\hat{p}(d) = \theta\bar{p} + (1-\theta)a(d) = \theta w\big[1 - t_T(d)\big] + (1-\theta)a(d).$$

This shows that the equilibrium long-run price, which is unique, equals the weighted average of the feasibility price and the marginal cost. Moreover, we can readily verify that

$$\frac{\partial\hat{p}(d)}{\partial a(d)} = 1-\theta > 0, \quad \frac{\partial\hat{p}(d)}{\partial t_T(d)} = -\theta w < 0, \quad \text{and} \quad \frac{\partial\hat{p}(d)}{\partial w} = \theta(1-t_T) > 0: \tag{29}$$

The long-run price is higher when the transit firm uses a less efficient technology, its service is superior in terms of travel time, and the wage rate is higher.

Because

$$\hat{y}(d) = (1-\theta)\Big[w\big\{1 - t_T(d)\big\} - a(d)\Big], \tag{30}$$

the affordability condition, (19), is more likely to be satisfied when the transit firm uses a more efficient technology ($a(d)$ is lower), mode A is more time-consuming ($t_A(d)$ is larger), mode T is less time-consuming ($t_T(d)$ is smaller), and/or the monetary cost of mode A is higher ($m_A(d)$ is larger). First, as (29) shows, a smaller $a(d)$ implies a lower $\hat{p}(d)$, which is responsible for the first result. Second, although a smaller $t_T(d)$ implies a higher $\hat{p}(d)$, it results in a higher $p^0(d)$ at the same time. Because the latter effect outweighs the former effect, the price differential $p^0(d) - \hat{p}(d)$ widens and, therefore, the income differential $\hat{y}(d) - y_A(d)$ also widens. Finally, a higher $t_A(d)$ and a higher $m_A(d)$ magnify the disadvantage of mode A, which the transit firm tries to exploit by setting a higher price in the short run.

5.2 Length of transit line

Next, we turn our attention to the optimal length of a transit line.

Equation (30) and the result that

$$p^0(d) - a(d) = \frac{\hat{y}(d)}{1-\theta} - y_A(d)$$

enable us to rewrite the gross profit function as

$$\Pi(P^{S*}, P^{L*}; l) = \frac{K}{1-\theta}\left[\rho \int_0^{\min\{l, b_A\}} y_A(d)^{\frac{1-\theta}{\theta}} \left\{\hat{y}(d) - (1-\theta) y_A(d)\right\} \mathrm{d}d + \theta(1-\rho) \int_0^{\min\{l, \hat{b}\}} \hat{y}(d)^{\frac{1}{\theta}} \mathrm{d}d\right].$$

Furthermore, note that $\underline{\Delta}(l)$ and $\overline{\Delta}(l)$ are given by

$$\begin{cases} \underline{\Delta}(l) = \dfrac{K}{1-\theta}\left[\rho y_A(l)^{\frac{1-\theta}{\theta}}\left\{\hat{y}(l) - (1-\theta) y_A(l)\right\} + \theta(1-\rho)\hat{y}(l)^{\frac{1}{\theta}}\right] & \text{for} \quad l \in \underline{D}, \quad \text{and} \\ \overline{\Delta}(l) = \dfrac{K\theta(1-\rho)\hat{y}(l)^{\frac{1}{\theta}}}{1-\theta} & \text{for} \quad l \in \overline{D}, \end{cases}$$

respectively. Consequently, as long as the affordability condition holds, the gross profit curve is actually upward sloping. Furthermore, it is obvious that $\underline{\Delta}(l) > \overline{\Delta}(l)$ for any $l \in \underline{D} \cup \overline{D}$, which implies that the slope of the curve discontinuously decreases at $l = b_A$. Finally, it turns out that the gross profit function is indeed concave in l, as assumed in Subsection 3.4.[14]

We are now ready to examine the effects of changes in parameters on the optimal length of a transit line.

To begin with, we have already demonstrated that $\partial l^*/\partial \rho \leq 0$ for the general form of a utility function. When the firm is more patient, it provides a longer transit lines.

The impacts of the other parameters are examined by (27). For a particular parameter λ ($\lambda \neq \theta$, $\lambda \neq \rho$), we obtain

$$\begin{cases} \dfrac{\partial \underline{\Delta}(l^*)}{\partial \lambda} = \dfrac{\underline{\Delta}(l^*)}{K}\dfrac{\partial K}{\partial \lambda} + \dfrac{K\rho}{\theta} y_A(l^*)^{\frac{1-\theta}{\theta}}\left[\dfrac{\hat{y}(l^*)}{y_A(l^*)} - 1\right]\dfrac{\partial y_A(l^*)}{\partial \lambda} & \\ \qquad\qquad + \dfrac{K}{1-\theta}\left[\rho y_A(l^*)^{\frac{1-\theta}{\theta}} + (1-\rho)\hat{y}(l^*)^{\frac{1-\theta}{\theta}}\right]\dfrac{\partial \hat{y}(l^*)}{\partial \lambda} & \text{for} \quad l^* \in \underline{D}, \\ \dfrac{\partial \overline{\Delta}(l^*)}{\partial \lambda} = \dfrac{\underline{\Delta}(l^*)}{K}\dfrac{\partial K}{\partial \lambda} + \dfrac{K(1-\rho)}{1-\theta}\hat{y}(l^*)^{\frac{1-\theta}{\theta}}\dfrac{\partial \hat{y}(l^*)}{\partial \lambda} & \text{for} \quad l^* \in \overline{D}. \end{cases} \tag{31}$$

[14] It follows from $\hat{y}(l) \geq y_A(l)$, $y_A'(l) < 0$ and $\hat{y}'(l) < 0$ that

$$\frac{\partial^2 \Pi(P^{S*}, P^{L*}; l)}{\partial l^2} = \begin{cases} \dfrac{K}{\theta(1-\theta)}\left[\theta\hat{y}'(l)\left\{\rho y_A(l)^{\frac{1-\theta}{\theta}} + (1-\rho)\hat{y}(l)^{\frac{1-\theta}{\theta}}\right\} + \rho y_A'(l) y_A(l)^{\frac{1-\theta}{\theta}}\left\{\dfrac{\hat{y}(l)}{y_A(l)} - 1\right\}\right] < 0 & \text{if} \quad l \in \overline{D}, \\ \dfrac{K(1-\rho)\hat{y}'(l) y_A(l)^{\frac{1-\theta}{\theta}}}{1-\theta} < 0 & \text{if} \quad l \in \overline{D}. \end{cases}$$

First, consider the prevailing utility level, $\bar{u}$. Since $\partial K/\partial\bar{u} < 0$, $\partial y_A(l^*)/\partial\bar{u} = 0$ and $\partial\hat{y}(l^*)/\partial\bar{u} = 0$, both $\partial\underline{\Delta}(l^*)/\partial\bar{u}$ and $\partial\overline{\Delta}(l^*)/\partial\bar{u}$ are negative, which implies that $\partial^2\Pi\left(P^{S*}, P^{L*}; l^*\right)/\partial l\partial\bar{u}$ is always negative. Therefore, we conclude from (27) that a lower $\bar{u}$ results in a longer transit line. This is because a lower utility level is associated with a denser population distribution.

Second, the impact of w is ambiguous. It follows from the affordability condition, (19), that $\hat{y}(l^*)/y_A(l^*) \geq 1$. Therefore, (31) implies that both $\partial y_A(l^*)/\partial w > 0$ and $\partial\hat{y}(l^*)/\partial w > 0$ work to increase the length of a transit line. However, this effect is counteracted by $\partial K/\partial w < 0$: although a higher wage rate enables the firm to charge a higher price both in the short and long runs (positive price effect), it also brings about a less dense population distribution (negative density effect). The sign of the overall effect is ambiguous, as it depends on the relative sizes of these two effects.

Nevertheless, we can show that the sign is necessarily positive when w is sufficiently small. To observe this, let us define $\underline{w} \equiv m_A(d)/\left[1 - t_A(d)\right]$, for which $y_A(d) = 0$: $\underline{w}$ is the lowest value of w's that satisfy both $y_A(d) \geq 0$ and $\hat{y}(d) \geq 0$.[15] Since

$$\lim_{w\to\underline{w}} y_A(d)^{\frac{1-\theta}{\theta}}\left[\frac{\hat{y}(d)}{y_A(d)} - 1\right] = \begin{cases} 0 & \text{if} \quad \theta < \frac{1}{2} \\ \hat{y}(d) & \text{if} \quad \theta = \frac{1}{2} \\ \infty & \text{if} \quad \theta > \frac{1}{2} \end{cases} \tag{32}$$

for any d, we have

$$\lim_{w\to\underline{w}} \frac{\partial\underline{\Delta}(l^*)}{\partial w} = \begin{cases} \Upsilon(l^*) & \text{if} \quad \theta < \frac{1}{2} \\ \Upsilon(l^*) + \dfrac{K\rho m_A(l^*)\hat{y}(l^*)}{\underline{w}\theta} & \text{if} \quad \theta = \frac{1}{2} \\ \infty & \text{if} \quad \theta > \frac{1}{2}, \end{cases}$$

where $\Upsilon(d) \equiv \dfrac{K(1-\rho)}{\underline{w}(1-\theta)}\hat{y}(d)^{\frac{1-\theta}{\theta}}\left[\beta\hat{y}(d) + (1-\theta)a(d)\right] > 0$. Therefore, $\lim_{w\to\underline{w}} \partial\underline{\Delta}(l^*)/\partial w > 0$. Furthermore, it is straightforward to verify that $\lim_{w\to\underline{w}} \dfrac{\partial\overline{\Delta}(l^*)}{\partial w} = \Upsilon(l^*) > 0$. These observations imply that $\lim_{w\to\underline{w}} \partial^2\Pi\left(P^{S*}, P^{L*}; l^*\right)/\partial l\partial w > 0$. Hence, $\lim_{w\to\underline{w}} \mathrm{d}l^*/\mathrm{d}w > 0$: as long as the wage rate is not too high, the price effect outweighs the density effect and consequently, the transit line becomes longer when the wage rate is higher.

Third, to examine the impact of $m_A(d)$, let us regard it as a function of a shift parameter λ with $\partial m_A(d;\lambda)/\partial\lambda > 0$ without loss of generality; that is, a higher λ is associated with a higher $m_A(d)$. Note that $\partial K/\partial\lambda = 0$, $\partial y_A(l^*)/\partial\lambda < 0$ and $\partial\hat{y}(l^*)/\partial\lambda = 0$. These results and (31) imply that $\partial\underline{\Delta}(l^*)/\partial\lambda < 0$ and $\partial\overline{\Delta}(l)/\partial\lambda = 0$. Therefore, given that the transit line is limited in the built-up area, the transit firm provides a shorter line when the monetary costs of mode A are lower. The

[15] Because of the affordability condition, $y_A(d) \geq 0$ implies that $\hat{y}(d) > 0$. Therefore, the necessary and sufficient condition for both $y_A(d) \geq 0$ and $\hat{y}(d) \geq 0$ being satisfied is that w is so high that $y_A(d) \geq 0$.

reason is as follows. On the one hand, in an A-type city with a lower m_A, workers are distributed more densely, and therefore, the marginal profit is higher in the short run. On the other hand, the firm is, in the short run, induced to charge a lower price in that city (see (15)), which adversely affects the marginal profit. Because it turns out that the former effect dominates the latter effect, the marginal profit is higher in the city with a lower m_A. If, instead, the transit line penetrates the undeveloped area, $m_A(d)$ does not affect the optimal length.

Fourth, the impact of $t_A(d)$ is similar to that of $m_A(d)$.

Fifth, the impact of $t_T(d)$ is examined in a parallel manner. λ is now a shift parameter for $t_T(d)$ with $\partial t_T(d;\lambda)/\partial\lambda > 0$. Because $\partial K/\partial\lambda = 0$, $\partial y_A(l^*)/\partial\lambda = 0$ and $\partial\hat{y}(l^*)/\partial\lambda < 0$, both $\partial\underline{\Delta}(l^*)/\partial\lambda$ and $\partial\overline{\Delta}(l^*)/\partial\lambda$ are negative. Therefore, a lower $t_T(d)$ is associated with a longer transit line regardless of whether l^* is in the built-up or undeveloped area. This is because a lower $t_T(d)$ brings about a denser population distribution and a higher price in the long run, both of which lead to higher marginal profit.

Finally, we consider the impact of $a(d)$. For a shift parameter λ with $\partial a(d;\lambda)/\partial\lambda > 0$, we have $\partial K/\partial\lambda = 0$, $\partial\underline{\Delta}(l^*)/\partial\lambda < 0$ and $\partial\overline{\Delta}(l^*)/\partial\lambda < 0$. Therefore, a lower $a(d)$ results in a longer transit line. The reason is explained as follows. First, a lower $a(d)$ implies a higher short-run profit, because the price mark-up is greater (short-run price is constant). Second, a lower $a(d)$ brings about a higher long-run profit for the following two reasons. The price mark-up becomes greater also in the long run although the long-run price becomes lower (see (29)); and population distribution in the long run is denser owing to a lower long-run price. Thus, the marginal profit is higher in both the short and long runs.

These findings are summarized as follows.

Proposition 5

Suppose that the sub-utility function is log linear and that the investment cost curve is linear or convex. Then, the firm provides a longer transit line when

i) it is more patient,

ii) the prevailing level of utility is lower,

iii) the wage rate is higher, provided that it is not too high,

iv) the monetary cost of mode A, the time cost of the mode, or both are lower if the terminal station is located in the built-up area (those costs do not affect the length of the line if the terminal station is located in the undeveloped area), v) the time cost of mode T is lower (regardless of whether the terminal station is in the built-up or undeveloped area), and

vi) the marginal cost for the transit firm is lower.

5.3 Entry of transit firm

Finally, we examine the impacts of parameters on the entry decision of a transit firm.

We already demonstrated that a more patient firm with a lower ρ is more likely to enter the rapid transit business irrespective of the form of utility function. For the other parameters, the following result is useful:

$$\frac{\partial\Omega\left(P^{S*},P^{L*},l^*\right)}{\partial\lambda} = \frac{\Pi\left(P^{S*},P^{L*},l^*\right)}{K}\frac{\partial K}{\partial\lambda} + \frac{K\rho}{\theta}\int_0^{\min\{l^*,b_A\}} y_A(d)^{\frac{1-\theta}{\theta}}\left[\frac{\hat{y}(d)}{y_A(d)}-1\right]\frac{\partial y_A(d)}{\partial\lambda}\,\mathrm{d}d$$
$$+\frac{K}{1-\theta}\left[\rho\int_0^{\min\{l^*,b_A\}} y_A(d)^{\frac{1-\theta}{\theta}}\frac{\partial\hat{y}(d)}{\partial\lambda}\,\mathrm{d}d + (1-\rho)\int_0^{l^*}\hat{y}(d)^{\frac{1-\theta}{\theta}}\frac{\partial\hat{y}(d)}{\partial\lambda}\,\mathrm{d}d\right] \tag{33}$$

for parameter λ ($\lambda \neq \theta, \lambda \neq \rho$).

First, consider $\bar{u}$. As mentioned in Subsection 5.2, $\partial K/\partial\bar{u} < 0$, $\partial y_A(d)/\partial\bar{u} = 0$ and $\partial\hat{y}(d)/\partial\bar{u} = 0$. Therefore, $\partial\Omega\left(P^{S*},P^{L*},l^*\right)/\partial\bar{u} < 0$ by (33) for $\lambda = \bar{u}$; it is more likely that a rapid transit system is introduced in a city when $\bar{u}$ is lower. The reason is the same as before: a lower utility level is associated with a denser population distribution.

Second, the sign of the effect of a change in w is ambiguous. The logic is again the same as before: although a higher wage rate enables the firm to charge a higher price in both the short and long runs, it brings about a less dense population distribution. However, the sign of the effect can be unambiguously determined when w is sufficiently low. To observe this, we obtain the following from (32) and (33):

$$\lim_{w\to\underline{w}}\frac{\partial\Omega\left(P^{S*},P^{L*},l^*\right)}{\partial w} = \begin{cases} \int_0^{l^*} \mathrm{Y}(d)\,\mathrm{d}d & \text{if} \quad \theta < \frac{1}{2} \\ \int_0^{l^*} \mathrm{Y}(d)\,\mathrm{d}d + \frac{K\rho}{\theta}\int_0^{\min\{l^*,b_A\}}\hat{y}(d)\left[1-t_A(d)\right]\mathrm{d}d & \text{if} \quad \theta = \frac{1}{2} \\ \infty & \text{if} \quad \theta > \frac{1}{2}. \end{cases}$$

Therefore, $\lim_{w\to\underline{w}}\partial\Omega\left(P^{S*},P^{L*},l^*\right)/\partial w > 0$. The positive price effect outweighs the negative density effect when w is sufficiently low. Therefore, it is more likely that a rapid transit system is introduced in a city when the wage rate is higher but not too high.

Third, we turn our attention to $m_A(d)$. For a shift parameter λ for which $\partial m_A(d;\lambda)/\partial\lambda > 0$, we can show that $\partial\Omega\left(P^{S*},P^{L*},l^*\right)/\partial\lambda \leq 0$: when $m_A(d)$ is smaller, it is more likely that a rapid transit system is introduced in a city. We can obtain a similar result for $t_A(d)$. The same explanation as that for the optimal length of a line applies. A lower cost of mode A brings about thick population density, on the one hand, but deters the firm from charging a high price, on the other hand. The former effect dominates the latter, which results in higher profit.

Fourth and finally, consider $t_T(d)$. For a shift parameter λ for which $\partial t_T(d;\lambda)/\partial\lambda > 0$, we have $\partial\Omega\left(P^{S*}, P^{L*}, l^*\right)/\partial\lambda < 0$: the lower $t_T(d)$ is, the more likely it is that a rapid transit system is introduced in a city. A similar result applies to $a(d)$.

Proposition 6

Suppose that the utility function is log linear and that the investment cost curve is linear or convex. Then, it is more likely that a rapid transit system is introduced in a city when
i) the transit firm is more patient,
ii) the prevailing level of utility is lower,
iii) the wage rate is higher, provided that it is not too high,
iv) the monetary cost of mode A, the time cost of the mode, or both are lower,
v) the time cost of mode T is lower, and
vi) the marginal cost for the transit firm is lower.

6 Concluding remarks

In this study, we examined the introduction of public transport represented by a rapid transit system into a city, paying special attention to the interplay between urban spatial structure and choice of transportation mode. To this end, we extended the Alonso–Mills–Muth monocentric city framework by incorporating a transit firm which may construct a rapid transit system and provide transportation services. In the determination of transportation costs, workers' choice of a transportation mode between rapid transit and automobiles and the firm's choice of prices play an important role. We showed that the development of a rapid transit system in a city depends on a few key parameters, most importantly, the degree of "patience" of the transit firm. The firm is more likely to begin business when it places a heavier weight on future profits. Assuming a specific form for the utility function, we obtained some additional results. A rapid transit system is more likely to be introduced in a city when workers there earn a higher wage, enjoy lower utility, need to spend less money and time on automobile travels, and enjoy shorter transit time, as well as when the transit firm has more efficient technology. In addition, we derived parallel results for the effects of those parameters on the length of the transit line that the firm operates.

A limitation of this study is that we made several strong assumptions to focus on the most important factors. It seems important to reconsider particularly the following assumptions in future studies. First, we assumed that the demand for commutes was inelastic: the frequency of commutes for each worker in a certain period was fixed regardless of their monetary and time costs. Second, the problem of congestion was not considered. In reality, however, congestion plays an important role in workers' decisions on what transportation mode to use. Third, workers were assumed to be homogeneous. As discussed in the introduction, a stream of literature has investigated the relationship between workers' income classes and urban spatial structure. The difference in workers' income should be incorporated into our research.

Our findings are intended to aid the correct evaluation of actual economic policies for the Transit Oriented Development. Specifically, we showed how the development of a rapid transit system changes urban spatial structure. We should take into account the benefit of the development through this change.

References

Alonso, W. (1964) *Location and Land Use*, Harvard University Press, Cambridge, MA, US.

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

Anas, A., and L. N. Moses. (1979) "Mode choice, transport structure and urban land use," *Journal of Urban Economics*, 6, 228–246.

Baum-Snow, N. (2007a) "Suburbanization and transportation in the monocentric model," *Journal of Urban Economics*, 62, 405–423.

Baum-Snow, N. (2007b) "Did highways cause suburbanization?" *The Quarterly Journal of Economics*, 122, 775–805.

Behrens, K., C. Gaigné, and J.-F. Thisse. (2009) "Industry location and welfare when transport costs are endogenous," *Journal of Urban Economics*, 65, 195–208.

Behrens, K., and P. Picard. (2011) "Transportation, freight rates, and economic geography," *Journal of International Economics*, 85, 280–291.

Bento, A. M., M. L. Cropper, A. M. Mobarak, and K. Vinha. (2005) "The effects of urban spatial structure on travel demand in the United States," *Review of Economics and Statistics*, 87, 466–478.

Brueckner, J. K. (2001) "Urban sprawl: Lessons from urban economics," in *Brookings-Wharton Papers on Urban Affairs* ed. by J. K. Brueckner, E. Mills, and M. Kremer, Brookings Institution Press, 65–97.

Combes, P.-P., T. Mayer, and J.-F. Thisse. (2008) *Economic Geography*, Princeton University Press, Princeton, NJ, US.

DeSalvo, J. S., and M. Huq. (1996) "Income, residential location, and mode choice," *Journal of Urban Economics*, 40, 84–99.

Duranton, G., and D. Puga. (2015) "Urban land use," in *Handbook of Regional and Urban Economics* vol. 5 ed. by G. Duranton, and W. C. Strange, Elsevier, Amsterdam, Netherlands, 467–560.

Fujita, M., P. Krugman, and A. J. Venables. (1999) *The Spatial Economy: Cities, Regions, and International Trade*, MIT Press, Cambridge, MA, US.

Gin, A., and J. Sonstelie. (1992) “The streetcar and residential location in nineteenth century Philadelphia,” *Journal of Urban Economics*, 32, 92–107.

Glaeser, E. L., and M. E. Kahn. (2004) “Sprawl and urban growth,” in *Handbook of Regional and Urban Economics* vol. 4 ed. by J. H. Henderson, and J.-F. Thisse, Elsevier, Amsterdam, Netherlands, 2481–2527.

Glaeser, E. L., M. E. Kahn, and J. Rappaport. (2008) “Why do the poor live in cities? The role of public transportation,” *Journal of Urban Economics*, 63, 1–24.

Gonzalez-Navarro, M., and M. A. Turner. (2018) “Subways and urban growth: Evidence from earth,” *Journal of Urban Economics*, 108, 85–106.

Hensher, D. A., T. P. Truong, C. Mulley, and R. Ellison. (2012) “Assessing the wider economy impacts of transport infrastructure investment with an illustrative application to the North-West Rail Link project in Sydney, Australia,” *Journal of Transport Geography*, 24, 292–305.

Jara-Diaz, S. (2007) *Transport Economic Theory*, Elsevier, Amsterdam, Netherlands.

LeRoy S. F., and J. Sonstelie. (1983) “Paradise lost and regained: Transportation innovation, income, and residential location,” *Journal of Urban Economics*, 13, 67–89.

Mayer, T., and C. Trevien. (2017) “The impact of urban public transportation evidence from the Paris region,” *Journal of Urban Economics*, 102, 1–21.

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

Muth, R. F. (1969) *Cities and Housing*, University of Chicago Press, Chicago, IL, US.

Sasaki, K. (1989) “Transportation system change and urban structure in two-transport mode setting,” *Journal of Urban Economics*, 25, 346–367.

Sasaki, K. (1990) “Income class, modal choice, and urban spatial structure,” *Journal of Urban Economics*, 27, 322–343.

Takahashi, T. (2006) “Economic geography and endogenous determination of transport technology,” *Journal of Urban Economics*, 60, 498–518.

Takahashi, T. (2011) “Directional imbalance in transport prices and economic geography,” *Journal of Urban Economics*, 69, 92–102.

Train, K., and D. McFadden. (1978) "The goods/leisure trade-off and disaggregate work trip mode choice models," *Transportation Research*, 12, 349–353.

Venables, A. J. (2007) "Evaluating urban transport improvements: Cost-benefit analysis in the presence of agglomeration and income taxation," *Journal of Transport Economics and Policy*, 41, 173–188.

Venables, A. J. (2016) "Incorporating wider economic impacts within cost-benefit appraisal," *International Transport Forum Discussion Paper*, OECD, 2016-05.

Wangsness, P. B., K. L. Rødseth, and W. Hansen. (2017) "A review of guidelines for including wider economic impacts in transport appraisal," *Transport Reviews*, 37, 94–115.

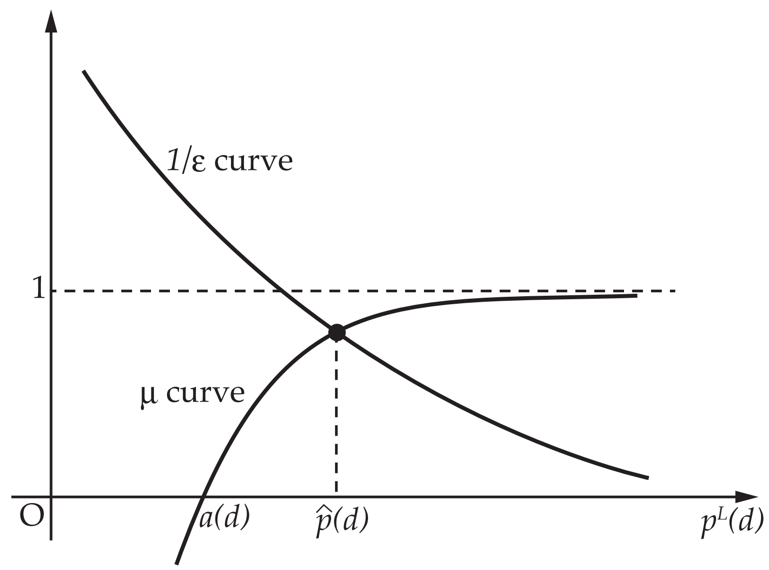

Fig. 1 Determination of a long-run price

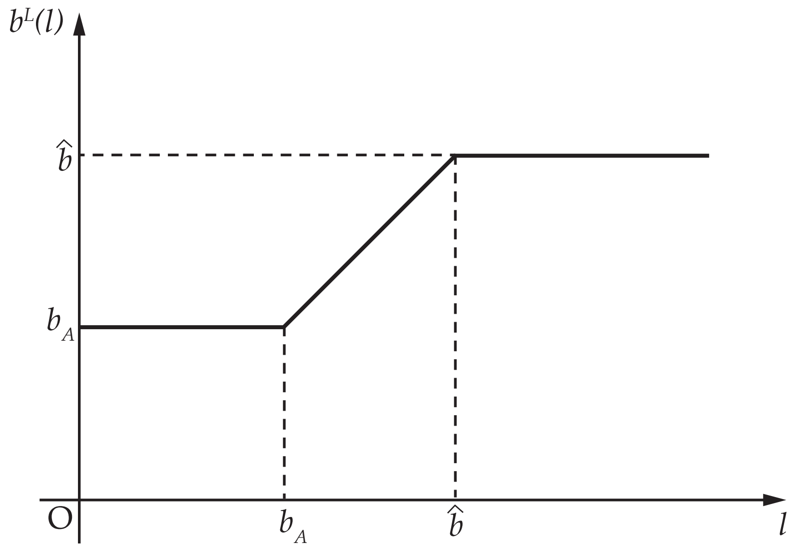

Fig. 2 Determination of a long-run boundary

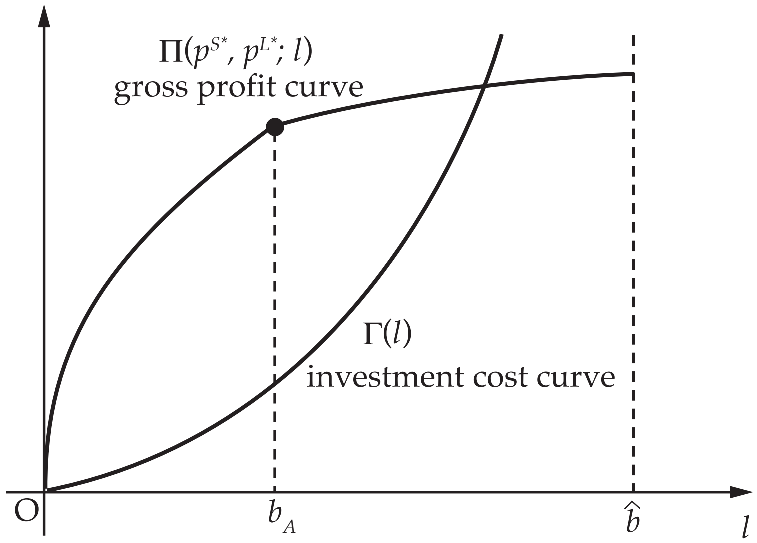

Fig. 3 Gross profit and investment cost