Article:
Xu, S-X, Liu, T-L, Huang, H-J et al. (1 more author) (2018) Mode choice and railway subsidy in a congested monocentric city with endogenous population distribution. Transportation Research Part A: Policy and Practice, 116. pp. 413-433. ISSN 0965-8564

https://doi.org/10.1016/j.tra.2018.07.001

Please cite paper as:

Xu, S-X., Liu, T-L., Huang, H-J. and Liu, R. (2018) Mode choice and railway subsidy in a congested monocentric city with endogenous population distribution. **Transportation Research Part A.** 116, 413-433, https://doi.org/10.1016/j.tra.2018.07.001.

Mode choice and railway subsidy in a congested monocentric city with endogenous population distribution

Shu-Xian Xu[a,b], Tian-Liang Liu[a, 1], Hai-Jun Huang[a] and Ronghui Liu[b]

[a] MoE Key Laboratory of Complex System Analysis and Management Decision, School of Economics and Management, Beihang University, Beijing 100191, China

[b] Institute for Transport Studies, University of Leeds, Leeds LS2 9JT, United Kingdom

Abstract

The objective of this paper is to provide new insights into commuters' mode choice behavior in a monocentric closed city with endogenous population distribution, where a congested highway and a crowded railway provide commuting services for residents on a linear urban corridor. We first explore typical equilibrium mode-choice patterns with exogenous city boundary and population distribution, and then incorporate residents' mode choice into an urban spatial equilibrium model, in which residents' household consumption, residential location choice and property developers' housing production are also explicitly modeled. Using comparative static analysis, we find that the urban corridor expands with the increase of railway fare if there is no congestion in the bimodal transportation system, but it would be uncertain if highway congestion and transit crowding cannot be ignored. We provide numerical evidence to show that the urban corridor possibly shrinks with the increase of railway fare once congestion effects are considered. We also discuss the changes of urban form, utility level of residents and social welfare with different railway fare and subsidy policies. The numerical results show that the distance-based fare policy with low subsidy should be preferred because it can realize the Pareto-improved social welfare and utility level of residents.

Keywords: linear monocentric city; mode choice; residential location choice; housing market; railway subsidy

1. Introduction

In recent decades and accompanying the economic growth and technological advances, we have seen rapid expansions and complex changes in developing cities around the world, such as that

[1] Corresponding author. Tel: +86-10-8231-5060; Email: liutianliang@buaa.edu.cn

taking place in Beijing and Shanghai, China. Urban expansion results in commuters living further away from work places, which in turn dramatically increases the demand for motorized vehicles. For instance, a report by Beijing Municipal Bureau of Statistics shows that the total number of motorized vehicles reached 5.6 million at the end of 2014 from a level of 4.8 million just four years ago, even though new car registrations via a lottery system have been introduced since 2011 (BMBS, 2015). Meanwhile, rapid developments of urban subways and railway networks such as mass transit systems in these cities have broadened travel mode choices to commuters (Yang et al., 2016, Peng et al., 2017), and governments and transit agencies are putting in vast sum of investments and subsidies to provide a reasonable level of transit operations throughout the cities (Wang et al., 2015; Xu et al., 2017). These rapid and complex developments in cities raise challenging research questions, especially on the agenda of sustainable urban development.

In principle, distribution and migration of population, frequent changes in work place and residential location and so on, may all have marked effects on the travel decisions of the residents (e.g. on travel mode, time of day and route choices). Likewise, developments (and expansions) of multimodal transportation networks (such as new metro lines) and the accompanying pricing policies may lead to changes in residential location choice, land-use pattern, housing market and so on (e.g., Bravo et al., 2010; Ma and Lo, 2012; Mohammad et al., 2013; Efthymiou and Antoniou, 2013; Dubé et al., 2013; Wang and Du, 2016b; Ng and Lo, 2015, 2017). It is, therefore, of great importance to address the inter-relationship between transportation and residential location choices, and the impacts of pricing policies, land use and housing developments on these choices.

Transport planners have long since recognized the need to consider the interactions between transport and land use in making their long-term transport planning for urban areas. For cities of relatively small sizes and with stable transportation and land-use markets, traditional four-stage travel demand models have been established to analyze trip generation, trip distribution, modal split and trip assignment. However, it has long been recognized that there are inconsistencies across different levels of four-stage modeling, due in part to their sequential and independent processes and the lack of feedback loops between stages. There have been large efforts in developing combined (with feedback loops) transportation equilibrium models to overcome some of the inconsistencies in the traditional four-stage modeling (e.g. Evans, 1976; Boyce and Southworth, 1979; Safwat and Magnanti, 1988; Huang and Lam, 1992; Tam and Lam, 2000; Zhou et al., 2009). For a historical overview of combined equilibrium models, readers can refer to Boyce and Williams (2015).

Combined equilibrium models based on multi-modal discrete networks have been formulated and analyzed extensively in the past decade (e.g. Lo et al., 2004; García and Marín, 2005; Liu et al., 2015). Discrete network models are generally developed for their realism in representing the behavior of city; however such models tend to have a large number of parameters to be estimated. On the other hand, the continuum modeling approach has been shown to be able to explore general trends and patterns of commuters' behavior and their changes in response to policy changes in transportation systems at a more aggregated macroscopic level (Ho and Wong, 2006). In many continuum equilibrium traffic assignment models, densely spaced roads are treated as a continuum over which commuters are continuously distributed in a two-dimensional space (e.g., Sasaki et al., 1990; Yang et al., 1994; Wong, 1998; Jiang et al., 2011). Due to the difficulty of obtaining exact solutions and analytical properties in a two-dimensional space, a simplified one-dimensional urban corridor with a continuum of entry points and a single exit point has often been adopted. Jehiel (1993) was the first to verify the existence of the simple solution of equilibrium states under the condition that the capacities of two congested modes are constant. In a transport system with a congestible highway and a congestion-free railway, Wang et al. (2004) investigated the characteristics of equilibrium mode choice patterns before and after the introduction of a park-and-ride (P&R) service. Following the thinking of Wang et al. (2004), Liu et al. (2014) further investigated the effects of rationing and pricing on morning commuters' travel cost and modal choice behavior in each location. Taking into account the in-vehicle crowding effects of railway service and assuming a continuous P&R provision on the urban corridor, Liu et al. (2009) explored the continuum equilibrium properties by analyzing commuters' mode choice and P&R transfer decisions. Li et al. (2012) investigated the intermodal equilibrium, road toll pricing, and bus system design issues on the congested urban corridor with two alternative modes of auto and bus, which share the same roadway. These studies on the continuum equilibrium are limited in their consideration of transportation systems and rely on one key assumption that the spatial distribution of households and the length of urban corridor, i.e., the city boundary, are given exogenously. As we mentioned before, transportation systems are linked closely with urban economics. Especially in those cities with rapid spatial expansions, urban land-use and housing developments as well as residents' consumer behavior frequently interact with residents' residential location and mode choices in the long term. Therefore, it is necessary to analyze the continuum equilibrium properties of mode choice patterns in an urban spatial equilibrium modeling framework.

On the basis of the stylized monocentric city model (Alonso, 1964; Muth, 1969; Mills, 1967, 1972; Brueckner, 1987), this paper develops a bimodal urban spatial equilibrium model in which the interplays among household consumption, residential location, mode choice and housing production are explicitly modeled. Furthermore, we analyze the impacts of railway fare changes on the city boundary with the consideration of endogenous population distribution, and numerically discuss the changes of urban form, utility level of residents and social welfare with different railway fare and subsidy policies.

The remainder of this paper is organized as follows. Section 2 reviews the urban economics studies on mode choice and subsidy issues associated with monocentric cities. Section 3 describes the basic assumptions and the overall modeling framework. Section 4 explores equilibrium mode-choice patterns with exogenous city boundary and population distribution. Section 5 presents an urban spatial equilibrium model by integrating household consumption, residential location choice and housing production with mode choice. The effects of railway fare changes on the city boundary are examined in detail. Section 6 provides a numerical comparison of urban system performance with different railway fare and subsidy policies. Concluding remarks are provided in Section 7.

2. Related studies

Much urban economic analysis is made based on a particular model of urban spatial structure, the monocentric city model pioneered in the 1960s by Alonso (1964), Muth (1969) and Mills (1967). In this section, we focus on reviewing related studies on mode choice and subsidy issues associated with monocentric cities in the urban economics literature.

The earlier literature emphasized the integration of mode choice into urban economic analysis and ignored the effects of either traffic congestion or in-vehicle crowding. Capozza (1973) was the first to develop a spatial general equilibrium model of a monocentric city with two transportation modes, i.e., a land-intensive road service and a land-economizing subway service. By assuming that the subway is less expensive than roads from the Central Business District (CBD) to some location on the urban corridor, Capozza found numerically that the addition of a subway system to a city with only roads would reduce transportation costs and city size. The reason for this is such that the construction of a subway permits land to be transferred from road use to housing, thereby dominating the reduction of city size. Without the use of land in transportation, Arnott and MacKinnon (1977) used a spatial general equilibrium simulation model to study the long-run effects of transportation changes such as changing parking fees and decreasing bus travel time in a closed city. An interesting point brought out in their simulations is the welfare-interdependence of different groups resulting from their spatial interaction. Anas and Moses (1979) was the first to provide an analytical urban spatial model to examine the impact of bimodal transportation on equilibrium residential land use and urban forms. They showed that the basic urban forms can result from the relative generalized cost characteristics of competing dense and sparse radial networks. Using an extended Alonso-Muth model (Alonso, 1964; Muth, 1969) with two competing modes of commuting, LeRoy and Sonstelie (1983) explained both of why resident distribution pattern of American cities, that the rich lived on the edges while the poor lived in the centers, prevailed until the 1970s and of why it is changing.

Sasaki (1989, 1990) made a comparative static analysis of urban spatial structure in two-transport mode setting and examined the impacts of transportation system changes and income changes on users' welfare. It is found that an improvement in the public transport mode may produce a contraction in the city size and decrease the welfare of some residents. Using a monocentric city model with two transportation modes, Creutzig (2014) investigated the effect of fuel prices on public transport infrastructure, modal shares and urban form. Besides the above works with discrete mode choice, some papers developed urban spatial equilibrium models that introduces mode choice as a continuous variable by assuming residents may optimize respective travel time, speeds or costs for objective decisions (e.g., Brown, 1986; DeSalvo and Huq, 2005; Brueckner, 2005).

There are a few studies focusing on transport subsidies in a monocentric city model. Brueckner (2005) was the first to deal with transportation subsidies as a potential source of urban sprawl. They showed that transport subsidies inefficiently lead to the urban expansion if the single-mode transport system exhibits constant returns to scale. Su and DeSalvo (2008) extended the work of Brueckner (2005) to investigate the effect of transportation subsidies on urban sprawl in a two-mode urban spatial model. It is found by comparative static analysis that there are an inverse relation between transit subsidies and sprawl and a direct relation between auto subsidies and sprawl, which is different from the single-mode results obtained in Brueckner (2005). With the assumption of fixed housing consumption, Borck and Wrede (2008) made progress in addressing optimal mode choice in presence of income heterogeneities. They found subsidies toward different modes have different effects. For instance, rich automobile drivers may suffer from transit subsidies, while poor transit users may benefit from subsidies to automobiles.

To the best of our knowledge, however, few studies discussed mode choice problems in an urban spatial equilibrium setting with congestion externalities, except for Haring et al. (1976) and Buyukeren and Hiramatsu (2016). Haring et al. (1976) extended a von Thunen-type model of urban structure by Mills (1972) to include two congested modes of transportation, and then concluded by simulating representative American and European cities that one travel mode dominates transportation choice until a competing mode becomes competitive, beginning at the edge of city. This conclusion is intuitive although it is through a numerical analysis: Haring et al. (1976) did not provide any analytical proof for it, nor did they discuss subsidies for commuting in their work. Buyukeren and Hiramatsu (2016) studied how anti-congestion policies such as congestion tolls and an urban growth boundary should be designed optimally in a monocentric city with car and public transit modes. They found that modal substitution effect can limit the centralizing force of anti-congestion policies, which would make mitigating congestion cause urban sprawl. The result is obtained using a simplified two-zone monocentric model often used in the urban economics literature (e.g., Anas and Pines, 2008).

Table 1 summarizes the differences among the related studies together with this papers' contributions. In reality, the impacts of congestion and transport subsidies on residents' mode choice and residential location choice cannot be ignored. In this paper, we characterize congestion and substitution effects between transportation modes in a continuum model framework for urban spatial equilibrium, and examine the impacts of transport subsidies on urban form and utility level of residents.

Table 1. Contributions to urban economics literature.

Citation	**Transportation modes**	**Congestion effect**	**Transport subsidies**	**Urban model**	**Solutions**
Alonso (1964), Mills (1972), Muth (1969)	Highway	No	No	Continuum	Analytical
Capozza (1973), Arnott and MacKinnon (1977)	Highway & railway	No	No	Continuum	Numerical
Anas and Moses (1979), Sasaki (1989, 1990)	Highway & railway	No	No	Continuum	Analytical
Brueckner (2005)	Highway	No	Yes	Continuum	Analytical
Su and DeSalvo (2008)	Highway & railway	No	Yes	Continuum	Analytical
Haring et al. (1976)	Highway & railway	Yes	No	Continuum	Numerical
Buyukeren and Hiramatsu (2016)	Highway & railway	Yes	No	Two-zone	Analytical
This paper	Highway & railway	Yes	Yes	Continuum	Analytical & Numerical

3. Model framework for a bimodal monocentric city

In this section, we propose a continuum model framework for a linear monocentric city with two transportation modes, which is a modification of the stylized monocentric city model (Alonso, 1964; Muth, 1969; Mills, 1967, 1972; Brueckner, 1987). In the modified model, each urban resident commutes to work in the CBD along a linear urban corridor with a crowded railway (specially referred to be of either subway or light rail type with closely spaced stations) and a congested highway serving for two alternative travel modes, transit and auto (Liu et al., 2009; Du and Wang, 2014).

To facilitate the presentation of essential ideas of this paper without loss of generality, we introduce several basic assumptions as follows.

A1: The city is closed. This means that the total population is exogenously given and fixed, but the utility level of residents, city boundary and spatial population distribution are all endogenously obtained by balancing the demand and supply of housing and land markets. In the land market, the land value determines the land use patterns on the urban corridor, an urban residential area or a rural area. In the long run, the land rent at the city boundary is assumed to be equal to the exogenous agricultural rent.

A2: At the demand side of housing market, all residents are assumed to be rational and earn the same annual income at the CBD, and tastes are assumed to be identical for all individuals. After subtracting the annual commuting cost, all the remaining annual income of each resident can be used to consume two kinds of normal goods, a housing service and a composite non-housing good. The objective of each resident is to maximize his/her household utility by household consumption and residential location choice within his/her budget constraint.

A3: The annual commuting cost of each resident is endogenously determined by all residents' simultaneous mode choice decisions along the corridor. All residents can choose their preferred travel modes based on each mode's generalized travel cost in a morning rush hour, which is defined as the fixed cost plus the variable cost related to travel distance and congestion externality. For simplicity the supply of transportation, e.g., the capacity of highway and the speed of train, is assumed to be fixed and constant throughout the corridor.

A4: At the supply side of housing market, property developers determine the housing investment per unit of land on the corridor in order to maximize their respective profits. The land revenue from land rents belongs to the government, and can be partly used to subsidize the operating deficit of railway (defined as the difference between operating cost and fare income) with given fare and subsidy policies.

Based on the above assumptions, urban residents' mode choice, household consumption and residential location behavior, property developers' housing production behavior and the government's subsidy policies for railway operation are explicitly integrated in the proposed bimodal monocentric city model, and their interplays are shown schematically in Fig. 1. The proposed model framework consists of three important components:(i) mode choice equilibrium, (ii) housing demand-supply equilibrium, and (iii) railway fare and subsidy policies. For the mode choice equilibrium, each resident's annual commuting cost is generated with a given inputted city boundary and spatial population distribution from the second component and the railway fare from the third component. Taking the land and endogenous commuting cost as inputs, the second component determines the city boundary, spatial population distribution, land rents and residents' utility level in a housing demand-supply equilibrium setting. With the above land rents and residents' utility level as inputs, the third component sets the government's fare and subsidy policies for railway operating.

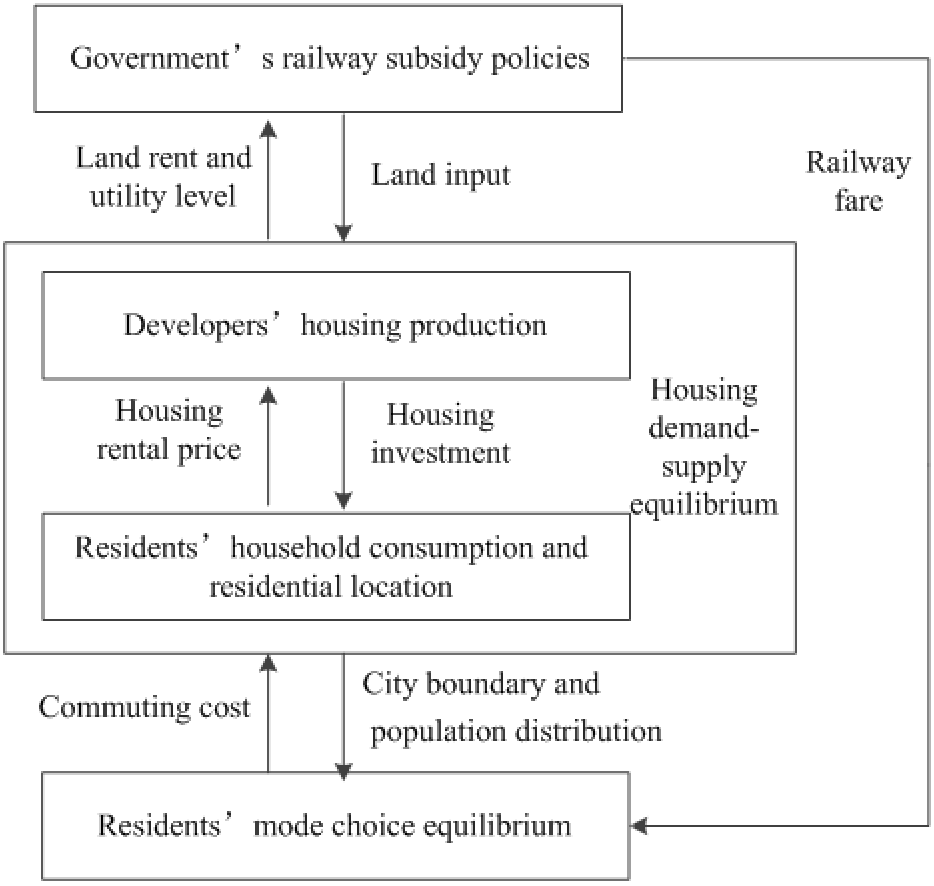

Fig.1. Model framework for a bimodal monocentric city.

4. Mode choice equilibrium

This section focuses on the first component of the proposed bimodal monocentric city model. Specifically, we will characterize typical equilibrium mode-choice patterns and the generation of endogenous annual commuting cost with exogenously given city boundary and spatial population distribution. Following Wang et al. (2004), Liu et al. (2009) and Liu et al. (2014), the commuting during a morning rush hour is modeled as a continuum of entry points and a single exit point. The exit point represents the CBD which all residents or commuters are heading for.

Let B be the city boundary or the length of urban corridor, N be the total population of city commuting to the CBD and $n(x)$ be the residential population density at location or entry point x, where x is defined as the distance from the location or entry point to the CBD. Therefore, it holds that $\int_0^B n(x)dx = N$.

4.1. Generalized travel cost functions

According to the assumption A3, all residents or commuters can choose their preferred travel modes at any entry point of the corridor based on each mode's generalized travel cost in a morning rush hour at that entry point. Before characterizing equilibrium mode-choice patterns along the corridor, we first introduce the specific components of generalized travel costs in a morning rush hour by transit mode and by auto mode, respectively.

The generalized travel cost by transit mode from location x to the CBD, $G_r(x)$, consists of three cost components: (a) the fixed cost component, a_r, which includes the access time cost using the transit mode and the fixed part of railway fare; (b) the distance-related cost component, $b_r x$, where b_r is the congestion-free variable cost per unit distance (e.g., the variable part of railway fare); and (c) the location-dependent in-vehicle crowding cost component, $c_r(x)$. It follows:

$$G_r(x) = a_r + b_r x + c_r(x). \tag{1}$$

As explained in Huang (2000), the in-vehicle crowing cost is mainly attributed to the privacy loss and body contact (uncomfortable physical proximity). The more passengers there are in the train, the larger the in-vehicle crowding cost is (Tirachini et al., 2013; Lu et al., 2015). Thus, the value of $c_r(x)$ certainly depends on the number of passengers in the train from location x to the CBD. Let $N_r(x)$ be the number of passengers in the train arriving at location x. Similar to Liu et al. (2009), we consider a function form of $c_r(x)$ as follows:

$$c_r(x) = \int_0^x g_r\left(N_r(w)\right) dw, \tag{2}$$

where $g_r\left(N_r(x)\right)$ is the in-vehicle crowing cost per unit distance at location x. Without loss of generality, it is assumed that $g_r(0) = 0$ and $g_r'\left(N_r(x)\right) > 0$.

The generalized travel cost by auto mode from location x to the CBD, $G_h(x)$, also consists of three cost components: (a) the fixed cost component, a_h, which includes the access time cost and the fixed monetary cost using the auto mode (e.g., the parking fee at the CBD); (b) the distance-related cost component, $b_h x$, where b_h is the congestion-free variable cost per unit distance including the free-flow travel time cost and the variable monetary cost for driving unit distance on the highway (e.g., fuel and insurance fees).; and (c) the location-dependent congestion time cost component, $c_h(x)$. It follows:

$$G_h(x) = a_h + b_h x + c_h(x). \tag{3}$$

Similar to Wang et al. (2004) and Liu et al. (2009), let the travel time cost $t_h\left(N_h(x)\right)$ for driving unit distance around location x be a strictly increasing function of traffic volume $N_h(x)$ at x, and t_h^0 be the free-flow travel time cost per unit distance. Thus, $t_h^0 = t_h(0)$ holds. The location-dependent congestion time cost component $c_h(x)$ can be denoted as

$$c_h(x)=\int_0^x t_h\left(N_h(w)\right)dw-t_h^0 x. \tag{4}$$

Specially, if $t_h\left(N_h(x)\right)$ takes a standard Bureau of Public Roads (BPR) function type, i.e., $t_h\left(N_h(x)\right)=t_h^0\left(1+\varepsilon\left(N_h(x)/W_h^0\right)^{\sigma}\right)$, where W_h^0 is highway capacity, ε and σ are positive parameters, then Eq. (4) can be changed as

$$c_h(x)=\int_0^x t_h^0\varepsilon\left(N_h(w)/W_h^0\right)^{\sigma}dw. \tag{5}$$

4.2. Equilibrium patterns of mode choice

Let $n_r(x)$ and $n_h(x)$ be the demand densities (the number of commuters per unit distance) of residents who choose the transit mode and auto one at location x, respectively. It follows:

$$n(x)=n_r(x)+n_h(x). \tag{6}$$

The number of passengers in the train arriving at any location x is

$$N_r(x)=\int_x^B n_r(w)dw. \tag{7}$$

And the traffic volume on the highway at any location x is

$$N_h(x)=\int_x^B n_h(w)dw. \tag{8}$$

According to Wardrop's (1952) First Principle for travel choice, a deterministic user equilibrium is achieved when no user can reduce his/her generalized travel cost by changing mode choice no matter where he/she lives.

Definition 1. Mathematically, the user equilibrium conditions for mode choice can be written as:

$$\begin{cases} n_h(x)>0\Rightarrow G_h(x)\le G_r(x) \\ n_r(x)>0\Rightarrow G_r(x)\le G_h(x) \end{cases}, x\in[0,B], \tag{9}$$

where $G_r(x)$ and $G_h(x)$ is given by Eqs. (1) and (3), respectively. This definition states that at mode choice equilibrium, residents at any location choose the mode with the minimal generalized travel cost. Therefore, the annual commuting cost for each resident located at x can be expressed as

$$C(x)=2\varphi\min\left(G_r(x),G_h(x)\right), \tag{10}$$

where the "2" denotes a daily round-trip travel between location x and the CBD (here, we assume morning and evening commuting are completely symmetrical), and φ is the annual average number of trips to the CBD per resident.

A little change in the cost components of any one mode's generalized travel cost would impact on the equilibrium mode-choice patterns. Next, we focus on the scenarios that each mode will be used at certain locations along the urban corridor, and characterize typical patterns of mode choice equilibrium under different restrictive conditions. Without much loss of generality, the following assumption is used for all the characterized equilibrium patterns:

$$a_r + b_r B > a_h + b_h B, \tag{11}$$

which assures that for all residents at the city boundary, traveling on the highway is always cheaper than that on the railway if the highway is empty (Wang et al., 2004; Liu et al., 2009). In other words, residents located closer to the city boundary always prefer to commute by auto mode. Based on this assumption, the following three cases with four typical patterns of mode choice equilibrium become possible[2] due to the gap between the fixed cost components of two modes, as shown in Fig.2.

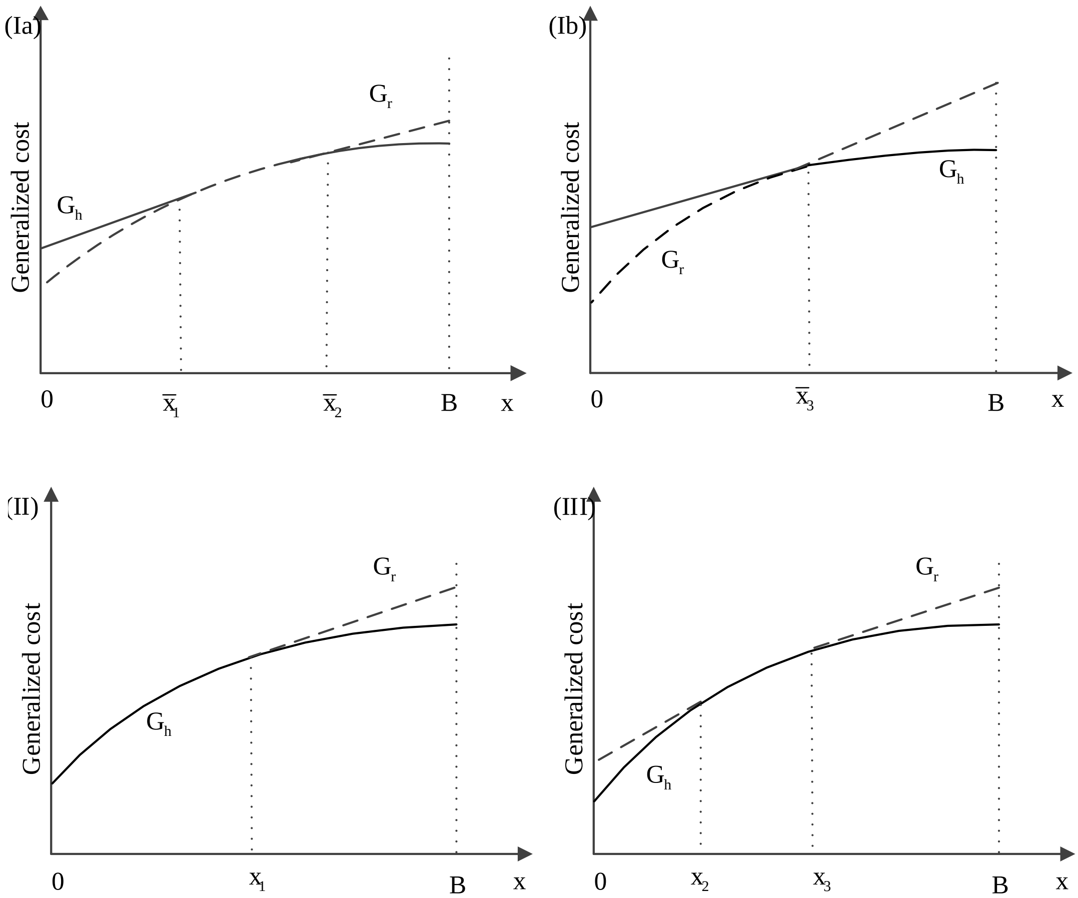

Fig. 2. Possible equilibrium mode-choice patterns

Case (I): $G_r(0) < G_h(0)$

[2] If the transit mode has no crowding effect, as analyzed in Wang et al. (2004), there exists only one possible equilibrium mode-choice pattern for the scenario that both modes will be used along the urban corridor. It is similar to that shown in Fig. 2(Ib), where a unique mode-switching point distinguish the use of transit mode and auto mode along the corridor.

This case has been discussed in Liu et al. (2009). We restate it here for comparison. In this case, the railway has the lower fixed cost than the highway, i.e.

$$G_r(0) = a_r < G_h(0) = a_h \,. \tag{12}$$

Eqs. (11) and (12) assure that

$$b_r > b_h \,, \tag{13}$$

which means that the highway has lower congestion-free variable cost per unit distance than the railway. Therefore, the railway is used by all residents living closer to the CBD, while the highway is used by those living farther out. Two typical patterns of mode choice equilibrium become possible due to the gap between the variable costs (including the distance-related and congestion cost components) of two modes, see panels (Ia) and (Ib) of Fig. 2. If there are small gaps between the fixed costs of two modes and/or between their variable costs, two mode-switching points, $\bar{x}_1$ and $\bar{x}_2$, might exist and both congested modes can be used simultaneously between them. As shown in Fig. 2(Ia), the variable costs of two modes per unit distance become equal at $\bar{x}_2$, and the two generalized travel cost curves coincide in the interval $\left[\bar{x}_1, \bar{x}_2\right]$. So that $G_r'\left(\bar{x}_2\right) = G_h'\left(\bar{x}_2\right)$ and $G_r\left(\bar{x}_1\right) = G_h\left(\bar{x}_1\right)$, i.e.

$$b_r = b_h + t_h\left(\int_{\bar{x}_2}^{B} n(w)dw\right) - t_h^0 \,, \tag{14}$$

$$a_r + b_r\bar{x}_1 + c_r(\bar{x}_1) = a_r + b_r\bar{x}_1 + \int_0^{\bar{x}_1} g_r\left(\int_x^{\bar{x}_1} n(w)dw + \int_{\bar{x}_1}^{\bar{x}_2} n_r(w)dw\right)dx$$

$$= a_h + b_h\bar{x}_1 + c_h(\bar{x}_1) = a_h + b_h\bar{x}_1 + \int_0^{\bar{x}_1} t_h\left(\int_{\bar{x}_1}^{\bar{x}_2} n_h(w)dw + \int_{\bar{x}_2}^{B} n(w)dw\right)dx - t_h^0\bar{x}_1 \,. \tag{15}$$

By solving Eqs. (14) and (15), we can get the solutions of $\bar{x}_1$ and $\bar{x}_2$ if they both exist. However, the congestion-free variable cost of transit mode per unit distance has the possibility to be large enough to exceed that of auto mode when serving all demands, i.e.

$$b_r \geq b_h + t_h(N) - t_h^0 \,. \tag{16}$$

This means Eq. (14) does not hold again. In this scenario, only one mode-switching point exists along the corridor, denoted as $\bar{x}_3$. As shown in Fig. 2(Ib), all residents living inside $\bar{x}_3$ take railway while those living beyond $\bar{x}_3$ take highway. At location $\bar{x}_3$, we have $G_h\left(\bar{x}_3\right) = G_r\left(\bar{x}_3\right)$, i.e.

$$a_r + b_r\bar{x}_3 + c_r(\bar{x}_3) = a_r + b_r\bar{x}_3 + \int_0^{\bar{x}_3} g_r\left(\int_x^{\bar{x}_3} n(w)dw\right)dx$$

$$= a_h + b_h\bar{x}_3 + c_h(\bar{x}_3) = a_h + b_h\bar{x}_3 + \int_0^{\bar{x}_3} t_h\left(\int_{\bar{x}_3}^{B} n(w)dw\right)dx - t_h^0\bar{x}_3 \,. \tag{17}$$

Case (II): $G_r(0) = G_h(0)$

As the fixed costs of two modes are equal, i.e., $G_r(0) = a_r = G_h(0) = a_h$, the solution of Eq. (15) approaches the CBD, and the "simple solution" in Jehiel (1993) emerges, which is the special case of Case (I). As shown in Fig.2(II), both congested modes are used between the CBD and location x_1, and only the auto mode is used from x_1 to the city boundary. The solution of x_1 can be obtained by resolving $G_r'(x_1) = G_h'(x_1)$, i.e.

$$b_r = b_h + t_h\left(\int_{x_1}^{B} n(w)dw\right) - t_h^0. \tag{18}$$

Case (III): $G_r(0) > G_h(0)$

In this case, the fixed cost of transit mode is larger than that of auto mode, i.e.

$$G_r(0) = a_r > G_h(0) = a_h, \tag{19}$$

which means all residents living near the CBD take highway for travel. Fig. 2(III) depicts the situation that both congested modes can be used simultaneously between location x_2 and x_3. Similar to the first panel of Case (I), the variable costs of two modes per unit distance become equal at x_3, and the two generalized travel cost curves coincide in the interval $[x_2, x_3]$. So that $G_r'(x_3) = G_h'(x_3)$ and $G_r(x_2) = G_h(x_2)$, i.e.

$$b_r = b_h + t_h\left(\int_{x_3}^{B} n(w)dw\right) - t_h^0, \tag{20}$$

$$a_r + b_r x_2 + c_r(x_2) = a_r + b_r x_2 + \int_0^{x_2} g_r\left(\int_x^{x_2} n(w)dw + \int_{x_2}^{x_3} n_r(w)dw\right)dx$$

$$= a_h + b_h x_2 + c_h(x_2) = a_h + b_h x_2 + \int_0^{x_2} t_h\left(\int_{x_2}^{x_3} n_h(w)dw + \int_{x_3}^{B} n(w)dw\right)dx - t_h^0 x_2. \tag{21}$$

By solving Eqs. (20) and (21), we can get the solutions of x_2 and x_3. Note that in this case, the solution of Eq. (20) always exists with the assumption that no mode is allowed to dominate the whole corridor.

Some properties of mode choice equilibrium can be observed for all cases from Fig. 2: (a) the variable costs of two modes per unit distance are both positive and non-increasing with x[3], $x \in [0, B]$; (b) the generalized travel costs of two modes and their lower envelope, i.e., the minimum of

[3] The proof of this property is similar to that of entry (i) of Lemma 1 in Liu et al. (2009). We omit it here in order to save space.

generalized travel costs, are both continuous and increasing with x, $x \in [0, B]$. In terms of its definition that the minimal generalized travel cost times some constants, the annual commuting costs of residents, $C(x)$, are also continuous and increasing with x, $x \in [0, B]$. In addition, it is noticed that the minimum of generalized travel costs and resultant annual commuting costs of residents are differentiable for all $x \in [0, B]$ in each panel of Fig. 2 except for $x = \bar{x}_3$ in Fig. 2(Ib).

Given the city boundary and residential population distribution, the following Proposition 1 (Proof can be found in Appendix A.1) shows how the mode-switching points $\bar{x}_2$, $\bar{x}_3$, x_1 and x_3 in Fig. 2 vary with the fixed cost component of transit travel a_r or the congestion-free variable cost per unit distance by transit mode b_r. This proposition can be used for comparison with the numerical results in Section 5.5, where the city boundary and population distribution are endogenously determined. However, due to the simultaneous use of transit mode and auto mode at certain location internals along the corridor, the variations of the mode-switching points $\bar{x}_1$ and x_2 with respect to a_r or b_r are difficultly derived, although it intuitively seems that $\bar{x}_1$ is decreasing whilst x_2 is increasing with a_r or b_r.

Proposition 1. Given the city boundary B and the residential population density $n(x)$, the mode-switching points $\bar{x}_2$, x_1 and x_3 do not vary with a_r whist $\bar{x}_3$ is decreasing with a_r. Furthermore, $\bar{x}_2$, $\bar{x}_3$, x_1 and x_3 are all decreasing with b_r.

Once the city boundary and population distribution are fixed, it is also intuitive to know that the annual commuting cost of residents $C(x)$ will increase with a_r or b_r. The following Proposition 2 (Proof can be found in Appendix A.2) only verifies this property for the scenario shown in Panel (Ib) of Fig. 2. For the other scenarios in Fig. 2, analytical derivations are more complex due to the simultaneous use of transit mode and auto mode at certain location internals along the corridor.

Proposition 2. Given the city boundary B and the residential population density $n(x)$ for the scenario shown in Panel (Ib) of Fig. 2, the annual commuting cost of residents $C(x)$ are increasing with a_r or b_r for any $x \in [0, B]$.

5. Urban spatial equilibrium

It is known in the previous section that, with a given city boundary and residential population distribution, the annual commuting cost of residents at any location along the urban corridor may be generated endogenously by modeling residents' mode choice behavior. In the long run, the city boundary and population distribution will both change with residents' household consumption and residential location choice, property developers' housing production and housing market' demand-supply equilibrium. Taking the endogenous annual commuting costs of residents as inputs, this section presents the whole urban spatial equilibrium model except for the mode choice equilibrium component.

5.1. Household consumption and residential location

This section focuses on the demand side of housing market. According to assumption A2, all residents are assumed to be identical and earn the same annual income Y at the CBD. For a rational resident at location x, his/her optimal decision on the annual consumption of two normal goods, a housing service and a composite non-housing good, is to resolve the direct utility maximization problem under his/her budget constraint. That is, for any $x \in [0, B]$,

$$U(x) = \max_{z(x),g(x)} V\left(z(x), g(x)\right), \tag{22}$$

subject to the budget constraint,

$$z(x) + p(x)g(x) = Y - C(x) . \tag{23}$$

Here, $V\left(z(x), g(x)\right)$ is a common household direct utility function, where $z(x)$ is the location-dependent consumption of a composite non-housing good and $g(x)$ is the location-dependent consumption of housing (also called the lot size), measured in square feet of floor space; $U(x)$ is the location-dependent household indirect utility function; $p(x)$ is the location-dependent housing rental price per square foot and the price of non-housing good is taken to be unity for simplicity; $C(x)$ is the annual commuting cost as defined before.

For convenience of further analysis, as assumed in Li et al. (2013) and Gubins and Verhoef (2014), the following Cobb–Douglas form of household direct utility function is adopted in this paper,

$$V\left(z(x), g(x)\right) = z(x)^{\alpha} g(x)^{\beta}, \ \alpha, \beta > 0, \ \alpha + \beta = 1, \tag{24}$$

where α and β are positive constants. Here, $\alpha + \beta = 1$ represents the household direct utility function has constant returns to scale, which is assumed in this paper for simplicity.

By solving the budget constraint in $z(x)$ and substituting it into Eq. (24), the first order condition of Eq. (24) with respect to $g(x)$ gives a unique demand for the lot size, which is implicitly defined as

$$-\frac{\partial V}{\partial z}p(x)+\frac{\partial V}{\partial g}=0. \tag{25}$$

Then we obtain

$$g(x)=\beta\left(Y-C(x)\right)/p(x), \tag{26}$$

$$U(x)=\alpha^{\alpha}\left(Y-C(x)\right)^{\alpha}g(x)^{\beta}. \tag{27}$$

Since all residents are identical, the urban spatial equilibrium must yield identical utility levels for all individuals. Let u be the utility level of residents at urban spatial equilibrium. So, we have $U\left(x\right)=u$ for all $x\in[0,B]$. Combining this with Eqs. (26) and (27), we derive $p\left(x,u\right)$ and $g\left(x,u\right)$, which are also functions of utility level u, as follows (Please refer to Appendix B):

$$p\left(x,u\right)=\alpha^{\alpha/\beta}\beta\left(Y-C(x)\right)^{1/\beta}u^{-1/\beta}, \tag{28}$$

$$g\left(x,u\right)=\alpha^{-\alpha/\beta}\left(Y-C(x)\right)^{-\alpha/\beta}u^{1/\beta}. \tag{29}$$

Eqs. (28) and (29), respectively, illustrate the housing rental price per square foot and lot size per household at equilibrium. Obviously, under a given level of utility, the housing rental price decreases and the lot size per household increases with the distance from the CBD, since the annual commuting cost is a continuous and increasing function of x, see the discussions in the previous section.

5.2. Housing production

This section focuses on the supply side of housing market. Property developers at each location along the corridor are assumed to determine the capital investment in the location-dependent housing market in order to maximize their respective profits. The following Cobb–Douglas form of housing production function is used to capture property developers' behavior (Brueckner, 1987):

$$h\left(S(x)\right)=\eta S(x)^{b}, 0<b<1,\ x\in[0,B], \tag{30}$$

where $h\left(S(x)\right)$ is the housing supply per unit of land at location x, $S(x)$ is the capital investment of housing per unit of land at location x and η and b are positive parameters.

Let $r(x)$ be the rent or value per unit of land at location x and k be the price of capital (i.e., the interest rate). Property developer's profit per unit of land at location x, $\pi(x)$, by optimizing the capital investment intensity $S(x)$, can be maximized as

$$\max_{S(x)} \pi(x) = p(x)h\left(S(x)\right) - \left(r(x) + kS(x)\right), \ x \in [0, B], \tag{31}$$

where the first term is the total housing revenue per unit of land and the second is total cost per unit of land including the land rent and production cost. The first-order optimality condition of the maximization problem (31) is

$$\frac{\partial \pi(x)}{\partial S(x)} = p(x)b\eta S(x)^{b-1} - k = 0. \tag{32}$$

Substituting $p(x,u)$ in Eq. (28) into Eq. (32) produces the capital investment intensity

$$S(x,u) = \left(\eta\alpha^{\alpha/\beta}\beta bk^{-1}\right)^{\frac{1}{1-b}} u^{-\frac{1}{\beta(1-b)}} \left(Y - C(x)\right)^{\frac{1}{\beta(1-b)}}. \tag{33}$$

Then, using Eqs. (29), (30) and (33), the residential population density at location x, $n(x,u)$, can be calculated by

$$n(x,u) = \frac{h\left(S(x,u)\right)}{g(x,u)} = \left(\eta\alpha^{\alpha/\beta}\left(\beta bk^{-1}\right)^{b}\right)^{\frac{1}{1-b}} u^{-\frac{1}{\beta(1-b)}} \left(Y - C(x)\right)^{\frac{\alpha+\beta b}{\beta(1-b)}}. \tag{34}$$

Under perfect competition (Brueckner, 1987), all property developers earn zero profit, i.e. $\pi(x) = 0$ for all $x \in [0, B]$, thus the land rent at location x is

$$r(x,u) = p(x)\eta S(x)^{b} - kS(x) = k\left(1/b - 1\right)\left(\eta\alpha^{\alpha/\beta}\beta bk^{-1}\right)^{\frac{1}{1-b}} u^{-\frac{1}{\beta(1-b)}} \left(Y - C(x)\right)^{\frac{1}{\beta(1-b)}}. \tag{35}$$

From Eqs. (33) – (35), we can easily obtain $\frac{\partial S(x,u)}{\partial C(x)} < 0$, $\frac{\partial n(x,u)}{\partial C(x)} < 0$ and $\frac{\partial r(x,u)}{\partial C(x)} < 0$. These inequalities state that the capital investment intensity, residential population density and land value all decrease with the distance from the CBD under a given level of utility, since $C(x)$ is increasing with x, $x \in [0, B]$, which is observed from Fig. 2.

5.3. Housing demand-supply equilibrium

Balancing the housing supply and demand requires two conditions that characterize the overall spatial equilibrium of the closed city (Brueckner, 1987). The first equilibrium condition requires that property developers outbid agricultural users for all lands used in housing production. Since the land rent decreases with distance from the CBD, the land rent for urban area reaches the minimum at the city boundary, at least equal to the exogenous agricultural rent r_a. Therefore, it follows:

$$r(B) = r_a. \tag{36}$$

The second equilibrium condition requires all residents live inside the urban areas. Since the total population of the closed city is fixed as N, it holds that

$$\int_0^B n(x,u)dx = N . \qquad (37)$$

Eqs. (36) and (37) are used to solve for the city boundary B and the utility level of residents at equilibrium u.

5.4. Solution procedure

In the presented closed city model, the total population is exogenously fixed whilst the city boundary and the equilibrium utility level of residents are both endogenous. The step-by-step procedure for calculating the equilibrium solutions of the model is presented as follows:

Step 1: Give the initial values of city boundary $B^{(0)}$, and residential density $n^{(0)}(x)$ for all $x \in \left[0, B^{(0)}\right]$. Residents is assumed to be uniformly distributed on the urban corridor at beginning.

Step 2: Evaluate the generalized travel cost $C^{(0)}(x)$ according to Eq. (10). Set $l = 1$.

Step 3: Use an iterative process to yield the values of utility level $u^{(l)}$ and city boundary $B^{(l)}$. Specifically, keeping the values of other variables in Eq. (36) and (37) fixed and using the value of $B^{(l-1)}$, first solve Eq. (37) to obtain the value of u, and then update the value of B by solving Eq. (36) based on the Bisection algorithm. Repeat the above process until the values of B and u both satisfy Eqs. (36) and (37).

Step 4: Calculate the values of $p^{(l)}(x)$, $g^{(l)}(x)$, $S^{(l)}(x)$, $n^{(l)}(x)$ and $r^{(l)}(x)$ by solving the Eqs (28), (29), (33), (34) and (35) using the values of $u^{(l)}$ and $B^{(l)}$ obtained in Step 3.

Step 5: Obtain the auxiliary travel cost $\overline{C}^{(l)}(x)$ by Eq. (10). Then, set

$$C^{(l+1)}(x) = C^{(l)}(x) + \left(\overline{C}^{(l)}(x) - C^{(l)}(x)\right) \Big/ l .$$

Step 6: If the relative error $\left\| C^{(l+1)}(x) - C^{(l)}(x) \right\| \Big/ \left\| C^{(l)}(x) \right\|$ is less than an acceptable level, then terminate; Otherwise, replace $C^{(l)}(x)$ with $C^{(l+1)}(x)$. Let $l = l + 1$. Go to Step 3.

5.5. Effects of railway fare changes

It is known in Section 4 that, the equilibrium mode-choice patterns along the corridor may vary with the relative cost differences between using public transit and using private automobile, which are measured by comparing the fixed or variable components of generalized travel costs by both modes. The switching among possible mode choice patterns will bring a significant change in the annual commuting costs of residents, which leads to different household consumption, residential location choice and housing production in a closed city. Accordingly, an urban expansion or contraction might occur. Besides those in the time costs and fuel price, the changes in the railway fare determine the cost differences between two modes, which may finally lead to different urban forms. This section focuses on the effects of railway fare changes on the equilibrium mode choice patterns and urban forms. Without loss of much generality, the railway fare at location x, $f(x)$, is assumed to be distance-based and linear with x, i.e., $f(x) = f_r^0 + f_r x$. Here, f_r^0 is the fixed part of railway fare and f_r is the variable part of railway fare per unit distance. Next, we discuss the effects of parameters f_r^0 and f_r on the city boundary, utility level and equilibrium mode choice patterns, respectively.

(1) Effect of parameter f_r^0

As mentioned in Section 4.1, the fixed part of railway fare is included into the fixed cost component of transit travel a_r. Hence, with the other factors fixed, a change in f_r^0 is exactly equivalent to that in a_r. Notice that the city boundary B and the utility level u are endogenously determined in Eqs. (36) and (37). Totally differentiating Eqs. (36) and (37) with respect to f_r^0 produces:

$$\left.\frac{\partial r}{\partial u}\right|_{x=B}\frac{du}{df_r^0}+\left.\frac{\partial r}{\partial x}\right|_{x=B}\frac{dB}{df_r^0}+\left.\frac{\partial r}{\partial f_r^0}\right|_{x=B}=0, \quad (38)$$

$$n(B)\frac{dB}{df_r^0}+\frac{du}{df_r^0}\int_0^B\frac{\partial n}{\partial u}dx+\int_0^B\frac{\partial n}{\partial f_r^0}dx=0. \quad (39)$$

Combing Eq. (38) and Eq. (39), we have

$$\frac{dB}{df_r^0}=\frac{\left.\frac{\partial r}{\partial u}\right|_{x=B}\int_0^B\frac{\partial n}{\partial f_r^0}dx-\left.\frac{\partial r}{\partial f_r^0}\right|_{x=B}\int_0^B\frac{\partial n}{\partial u}dx}{\left.\frac{\partial r}{\partial x}\right|_{x=B}\int_0^B\frac{\partial n}{\partial u}dx-n(B)\left.\frac{\partial r}{\partial u}\right|_{x=B}}, \quad (40)$$

$$\frac{du}{df_r^0}=\frac{n(B)\left.\frac{\partial r}{\partial f_r^0}\right|_{x=B}-\left.\frac{\partial r}{\partial x}\right|_{x=B}\int_0^B\frac{\partial n}{\partial f_r^0}dx}{\left.\frac{\partial r}{\partial x}\right|_{x=B}\int_0^B\frac{\partial n}{\partial u}dx-n(B)\left.\frac{\partial r}{\partial u}\right|_{x=B}}. \quad (41)$$

Since $C(x)$ is increasing with x from Fig. 2, According to Eqs. (34) and (35), we easily get $\partial n(x,u)/\partial u<0$, $\partial r(x,u)/\partial u<0$ and $\left.\partial r(x,u)/\partial x\right|_{x=B}<0$. Thus, the denominators of Eq. (40) and Eq. (41) are both positive, and the signs of dB/df_r^0 and du/df_r^0 are determined by that of the numerators of Eq. (40) and Eq. (41), respectively, which truly depend on the degree of highway congestion and transit crowding.

When there are no highway congestion and transit crowding on the studied urban corridor, $G_r(x)$ and $G_h(x)$ are both linear with x, and only one mode-switching point $\bar{x}_3$ exists along the corridor, similar to that shown in Fig. 2(Ib). Obviously, according to Eq. (10) and the assumption (11), $\partial C(B)/\partial f_r^0 = 0$, $\partial C(x)/\partial f_r^0 \geq 0$ for all $x \in [0, B)$, and there exists a corridor interval where $\partial C(x)/\partial f_r^0 > 0$, which lead to $\int_0^B \partial n(x,u)/\partial f_r^0 dx < 0$ and $\partial r(x,u)/\partial f_r^0\big|_{x=B} = 0$ from Eqs. (34) and (35). As a result, we have

$$\frac{dB}{df_r^0} = \frac{\left.\frac{\partial r}{\partial u}\right|_{x=B} \int_0^B \frac{\partial n}{\partial f_r^0} dx}{\left.\frac{\partial r}{\partial x}\right|_{x=B} \int_0^B \frac{\partial n}{\partial u} dx - n(B) \left.\frac{\partial r}{\partial u}\right|_{x=B}} > 0, \tag{42}$$

$$\frac{du}{df_r^0} = \frac{-\left.\frac{\partial r}{\partial x}\right|_{x=B} \int_0^B \frac{\partial n}{\partial f_r^0} dx}{\left.\frac{\partial r}{\partial x}\right|_{x=B} \int_0^B \frac{\partial n}{\partial u} dx - n(B) \left.\frac{\partial r}{\partial u}\right|_{x=B}} < 0, \tag{43}$$

which are consistent with the results analyzed in Sasaki (1989, 1990) and Su and DeSalvo (2008). However, when there exist highway congestion and transit crowding, it is difficult to judge on the signs of $\int_0^B \partial n(x,u)/\partial f_r^0 dx$ and $\partial r(x,u)/\partial f_r^0\big|_{x=B}$ due to the complex nested relationships between $C(x)$ and $n(x,u)$. Hence, the signs of dB/df_r^0 and du/df_r^0 are un-determinate in this situation and may be different case by case. Taking the values of model parameters in Table 2 as inputs, Table 3 and Fig. 3 show some numerical examples with the consideration of highway congestion and transit crowding, where $dB/df_r^0 < 0$ and $du/df_r^0 < 0$ hold. To summarize, we have the following proposition.

Proposition 3. Without highway congestion and transit crowding, the city boundary will expand and the utility level of residents will reduce as the fixed part of railway fare increases. However, there are possibilities that an increase in the fixed part of railway fare results in a shrink in the city boundary if highway congestion and transit crowding are considered.

With the given values of model parameters in Table 2, Table 3 shows the changes of some endogenous variables in the studied city model with different values of f_r^0, such as the city boundary, the utility level of residents, the mode-switching points and the total number of transit passengers.

We can see clearly that, as f_r^0 takes values from 2 to 10, the city boundary, the utility level of residents and the total number of transit passengers gradually decrease. This confirms the latter part of Proposition 3. Furthermore, the mode-switching point farther from the CBD always decreases with f_r^0, which is similar to the conclusion drawn in Proposition 1, where both of the city boundary and population distribution are exogenously given. But, different from one farther from the CBD, the mode-switching point closer to the CBD first decreases till being zero when $f_r^0 = 6$, and then increases. This is because all residents living close to the CBD in fact use different travel modes when the value of f_r^0 is smaller or larger than 6.

Table 2. Values of model parameters.

Symbol	Definition	Value
Parameters associated with city model		
N	Total number of residents in the city	90000
Y	Annual income (RMB)	150000
r_a	Agricultural rent at the city boundary (RMB)	300000
φ	Annual average number of trips to the CBD per resident	350
α, β	Parameters in utility function	$\alpha = 0.75$, $\beta = 0.25$
b, η	Parameters in housing production function	$b = 0.7$, $\eta = 0.8 \times 10^{-8}$
k	Interest rate	5%
ξ	parameter that converts utility level into equivalent monetary units	80
Parameters for auto travel		
W_h^0	Highway capacity (veh/h)	5400
a_h	Fixed cost component of auto travel (RMB)	11
t_h^0	Free flow travel time cost per unit distance on highway (RMB/km)	1/3
$b_h - t_h^0$	Congestion-free variable cost per unit distance except for t_h^0 (RMB)	0.2
ε, σ	Parameters in BPR function	$\varepsilon = 0.5$, $\sigma = 1$
Parameters for transit travel		
f_r^0	Fixed part of railway fare (RMB)	2, 4, 6, 8, or 10
$a_r - f_r^0$	Fixed cost component of transit travel except for f_r^0 (RMB)	5
f_r	Variable part of railway fare (RMB)	0.4, 0.8, or 1.8
b_r	Congestion-free variable cost per unit distance except for f_r (RMB)	0.6
c_o	Fixed operating cost of railway per year (RMB)	2×10^8
γ, δ	Parameters in in-vehicle crowding cost $g_r\left(N_r(x)\right) = \gamma\left(N_r(x)/W_r^0\right)^{\delta}$	$W_r^0 = 8000$, $\gamma = 0.5$, $\delta = 1$

Note: In all numerical examples, the city corridor is uniformly discretized into 100 sections for approximately solving the model.

Fig. 3(g) – (i) depict the equilibrium mode-choice patterns along the urban corridor associated with $f_r^0 = 2$, 6 and 10, respectively, which are similar to that shown in panels (Ia), (II) and (III) of

Fig. 2. When $f_r^0 = 2$, the railway has the smaller fixed cost than the highway, i.e., $a_r < a_h$, thus all residents living near the CBD choose the railway for travel. When $f_r^0 = 6$, the fixed costs by railway and by highway are equal, i.e., $a_r = a_h$, thus residents living in location $x = 23.23$ to the CBD will use highway and railway simultaneously. When $f_r^0 = 10$, the railway has the larger fixed cost than the highway, i.e., $a_r > a_h$, leading all residents living near the CBD choose the highway for travel.

For comparison, besides that with highway congestion and transit crowding shown in panels (g) – (i), Fig. 3 also gives the equilibrium mode-choice patterns for cases without highway congestion and transit crowding and with only highway congestion, which correspond to panels (a) – (c) and panels (d) – (f), respectively. Clearly, when both of highway congestion and transit crowding are ignored as depicted in Fig. 3(a) – (c), the city boundary is increasing with f_r^0, which is consistent with the former part of Proposition 3. However, even if only highway congestion is considered, the opposite change of the city boundary possibly occurs with the increase of f_r^0. Fig. 3(d) – (f) provide such numerical examples. Further, when transit crowding is considered together with highway congestion, the city boundary always is smaller than that with only highway congestion for different values of f_r^0, by comparing Fig. 3(d) – (f) with Fig. 3(g) – (i).

Table 3. Changes of endogenous variables with different values of f_r^0

f_r^0	City boundary	Utility level	Mode-switching point closer to the CBD	Mode-switching point farther from the CBD	Total number of transit passengers
2	75.57	285.93	4.53	23.43	44552
4	75.18	284.77	3.00	23.30	38346
6	74.92	283.78	0	23.23	24753
8	74.61	283.06	4.48	23.13	15038
10	74.29	282.48	6.69	23.03	11066

Note: These results are calculated based on $f_r = 0.4$ and other parameter values in Table 2.

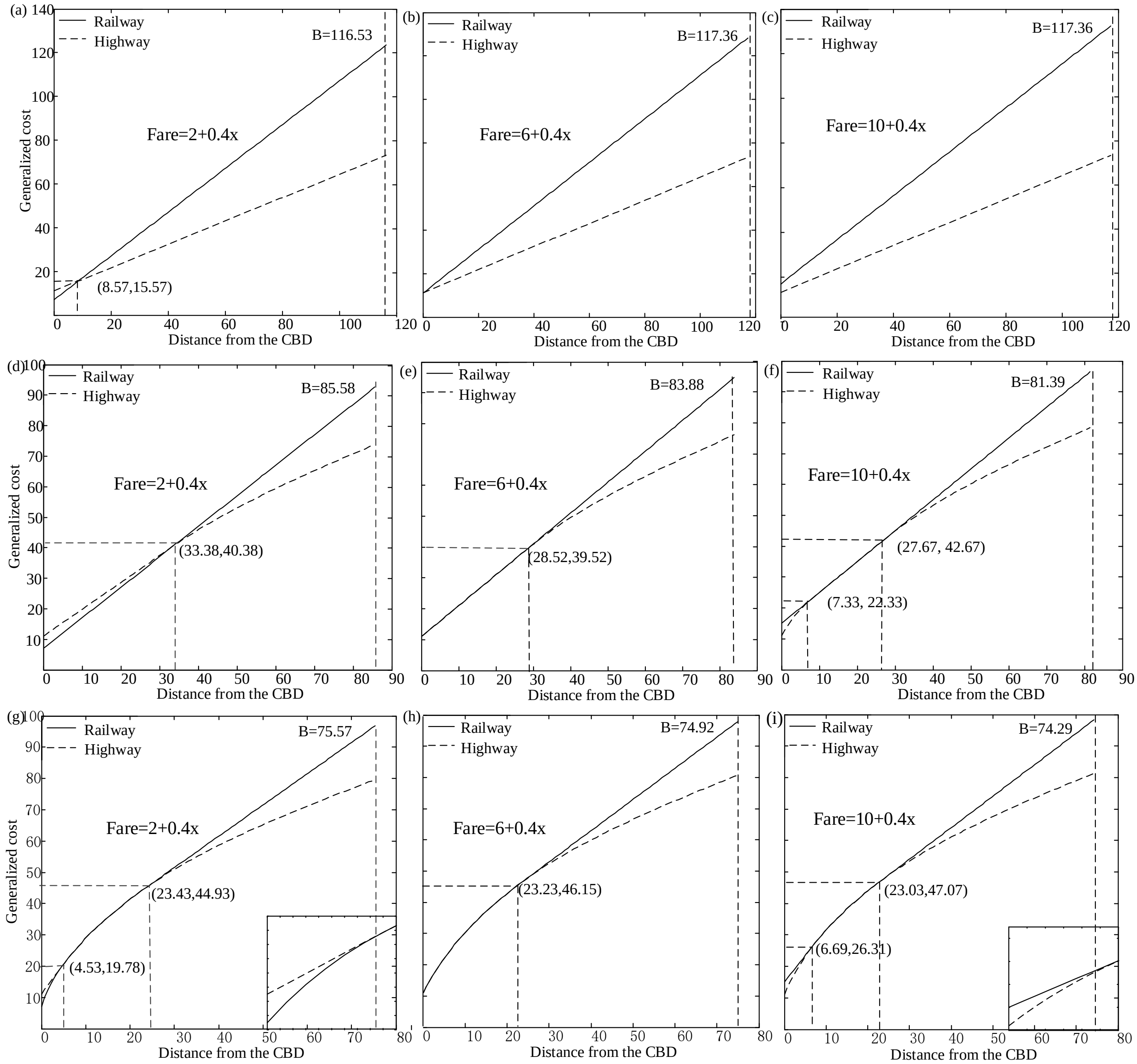

Fig. 3. Mode-choice patterns with different f_r^0 for cases without highway congestion and transit crowding ((a) – (c)), with only highway congestion ((d) – (f)) & with highway congestion and transit crowding ((g) – (i)).

Next, we examine what changes would result if both of the city boundary and population distribution are exogenously given. We first fix the city boundary and population distribution as those endogenously generated by the studied city model with $f_r^0 = 10$ or 2, and then observe the changes of equilibrium mode- switching points by adjusting f_r^0, as shown in Table 4. Clearly, with the increase of f_r^0, the mode-switching point closer to the CBD first decreases till being zero and then increases whilst the mode-switching point farther from the CBD always decreases, which is consistent with the results with endogenous city boundary and population distribution. Furthermore, it can be seen by comparing the results in Table 4 and Table 3 that, when the city boundary is fixed as

75.57, corresponding to the endogenous model with $f_r^0 = 2$, all mode-switching points with different values of f_r^0 are larger than that obtained using the endogenous model. In addition, when the city boundary is fixed as 74.29, corresponding to the endogenous model with $f_r^0 = 10$, the opposite trend comes true. This means, the equilibrium mode-choice patterns would be inaccurately predicted if endogenous properties of city boundary and population distribution are ignored.

Table 4. Equilibrium mode-choice points with different f_r^0 when the city boundary and population density are exogenously given.

Exogenous examples	f_r^0	Mode-switching point closer to the CBD	Mode-switching point farther from the CBD
$B = 74.29$, corresponding to the endogenous model with $f_r^0 = 10$	2	4.46	23.10
	4	2.97	23.07
	6	0	23.06
	8	4.46	23.03
	10	6.69	23.03
$B = 75.57$, corresponding to the endogenous model with $f_r^0 = 2$	2	4.53	23.43
	4	3.02	23.35
	6	0	23.27
	8	4.63	23.20
	10	6.80	23.14

Note: The results are calculated based on $f_r = 0.4$ and the parameter values in Table 2.

(2) Effect of parameter f_r

As mentioned in Section 4.1, the variable part of railway fare is included into the congestion-free variable cost component of transit travel b_r. Hence, with the other factors fixed, a change in f_r is exactly equivalent to that in b_r. Similar to the analysis on the effect of parameter f_r^0, differentiating Eqs. (36) and (37) with respect to f_r, respectively, and rearranging them, we have

$$\frac{dB}{df_r} = \frac{\left.\frac{\partial r}{\partial u}\right|_{x=B} \int_0^B \frac{\partial n}{\partial f_r} dx - \left.\frac{\partial r}{\partial f_r}\right|_{x=B} \int_0^B \frac{\partial n}{\partial u} dx}{\left.\frac{\partial r}{\partial x}\right|_{x=B} \int_0^B \frac{\partial n}{\partial u} dx - n(B) \left.\frac{\partial r}{\partial u}\right|_{x=B}}, \tag{44}$$

$$\frac{du}{df_r} = \frac{n(B) \left.\frac{\partial r}{\partial f_r}\right|_{x=B} - \left.\frac{\partial r}{\partial x}\right|_{x=B} \int_0^B \frac{\partial n}{\partial f_r} dx}{\left.\frac{\partial r}{\partial x}\right|_{x=B} \int_0^B \frac{\partial n}{\partial u} dx - n(B) \left.\frac{\partial r}{\partial u}\right|_{x=B}}. \tag{45}$$

As done before, it is easily to verify that the denominators of Eqs. (44) – (45) are both positive, thus the signs of their numerators determine that of dB/df_r and du/df_r. When there exist highway congestion and transit crowding on the urban corridor, it is difficult to judge on the signs of $\int_0^B \partial n(x,u)/\partial f_r dx$ and $\partial r(x,u)/\partial f_r|_{x=B}$ due to the complex nested relationships between $C(x)$ and $n(x,u)$. Thus, the signs of dB/df_r^0 and du/df_r^0 are unknown in this situation. With the given values of model parameters in Table 2, Table 5 and Fig. 4 show some numerical examples with the consideration of highway congestion and transit crowding, where $dB/df_r < 0$ and $du/df_r < 0$ hold.

When there are no highway congestion and transit crowding, similar to the analysis on the effect of parameter f_r^0, it is easy to verify that $\partial C(B)/\partial f_r = 0$, $\partial C(x)/\partial f_r \geq 0$ for all $x \in [0, B)$, and there exists a corridor interval where $\partial C(x)/\partial f_r > 0$, which lead to $\int_0^B \partial n(x,u)/\partial f_r dx < 0$ and $\partial r(x,u)/\partial f_r|_{x=B} = 0$ from Eqs. (34) and (35). As a result, we have

$$\frac{dB}{df_r} = \frac{\left.\frac{\partial r}{\partial u}\right|_{x=B} \int_0^B \frac{\partial n}{\partial f_r} dx}{\left.\frac{\partial r}{\partial x}\right|_{x=B} \int_0^B \frac{\partial n}{\partial u} dx - n(B) \left.\frac{\partial r}{\partial u}\right|_{x=B}} > 0, \tag{46}$$

$$\frac{du}{df_r} = \frac{-\left.\frac{\partial r}{\partial x}\right|_{x=B} \int_0^B \frac{\partial n}{\partial f_r} dx}{\left.\frac{\partial r}{\partial x}\right|_{x=B} \int_0^B \frac{\partial n}{\partial u} dx - n(B) \left.\frac{\partial r}{\partial u}\right|_{x=B}} < 0, \tag{47}$$

which are also consistent with the results analyzed in Sasaki (1989, 1990) and Su and DeSalvo (2008). To summarize, we have the following proposition.

Proposition 4. Without highway congestion and transit crowding, the city boundary will expand and the utility level of residents will reduce as the variable part of railway fare increases. However, , there are possibilities that an increase in the variable part of railway fare results in a shrink in the city boundary if highway congestion and transit crowding are considered.

With the given values of model parameters in Table 2, Table 6 shows the changes of some endogenous variables in the studied city model with f_r for different cases with $f_r^0 = 2$, 6 or 8, such as the city boundary, the utility level of residents, the mode-switching points and the total number of transit passengers. Obviously, no matter what value of f_r^0 is fixed, as f_r increases from 0.4 to 1.8, the city boundary, the utility level of residents and the total number of transit passengers gradually decrease. This confirms the latter part of Proposition 4. Furthermore, different from the changes varying with f_r^0 as discussed before, both the mode-switching points are decreasing with f_r. Fig. 4(e) – (f) depict the equilibrium mode-choice patterns along the urban corridor associated with $f_r^0 = 2$ and $f_r = 0.8$ or 1.8, which are similar to that shown in panels (II) and (Ib) of Fig. 2. In these scenarios, the railway has the smaller fixed cost than the highway, thus all residents living between the CBD and the mode-switching point closer to the CBD choose the railway for travel. However, when $f_r^0 > 6$, the railway has the larger fixed cost than the highway, leading all residents living near the CBD choose the highway for travel. The equilibrium mode-choice patterns corresponding to this situation are not depicted here due to space limitations.

Table 5. Changes of endogenous variables with different values of f_r

f_r^0	f_r	City boundary	Utility level	Mode-switching point closer to the CBD	Mode-switching point farther from the CBD	Total number of transit passengers
2	0.4	75.57	285.93	4.53	24.18	44552
	0.8	74.01	285.04	4.44	13.32	40813
	1.8	72.69	283.30	4.36	×	30872
6	0.4	74.92	283.78	×	23.23	24753
	0.8	73.30	282.90	×	13.19	20468
	1.8	72.03	281.15	×	3.60	9760
10	0.4	74.29	282.48	6.69	23.03	11066
	0.8	72.87	281.58	6.56	13.12	6419
	1.8	71.80	280.51	×	×	0

For comparison, besides that with highway congestion and transit crowding shown in panels (e) – (f), Fig. 4 also gives the equilibrium mode-choice patterns with $f_r^0 = 2$ and $f_r = 0.8$ or 1.8 for cases without highway congestion and transit crowding and with only highway congestion, which correspond to panels (a) – (b) and panels (c) – (d), respectively. Clearly, when both highway congestion and transit crowding are ignored as depicted in Fig. 4(a) – (b), the city boundary is increasing with f_r, which is consistent with the former part of Proposition 4. However, even if only highway congestion is considered, the opposite change of the city boundary possibly occurs with the increase of f_r. Fig. 4(c) – (d) provide such numerical examples. Further, when transit crowding is considered together with highway congestion, the city boundary is always smaller than that with only highway congestion for different values of f_r, by comparing Fig. 4(c) – (d) with Fig. 4(e) – (f).

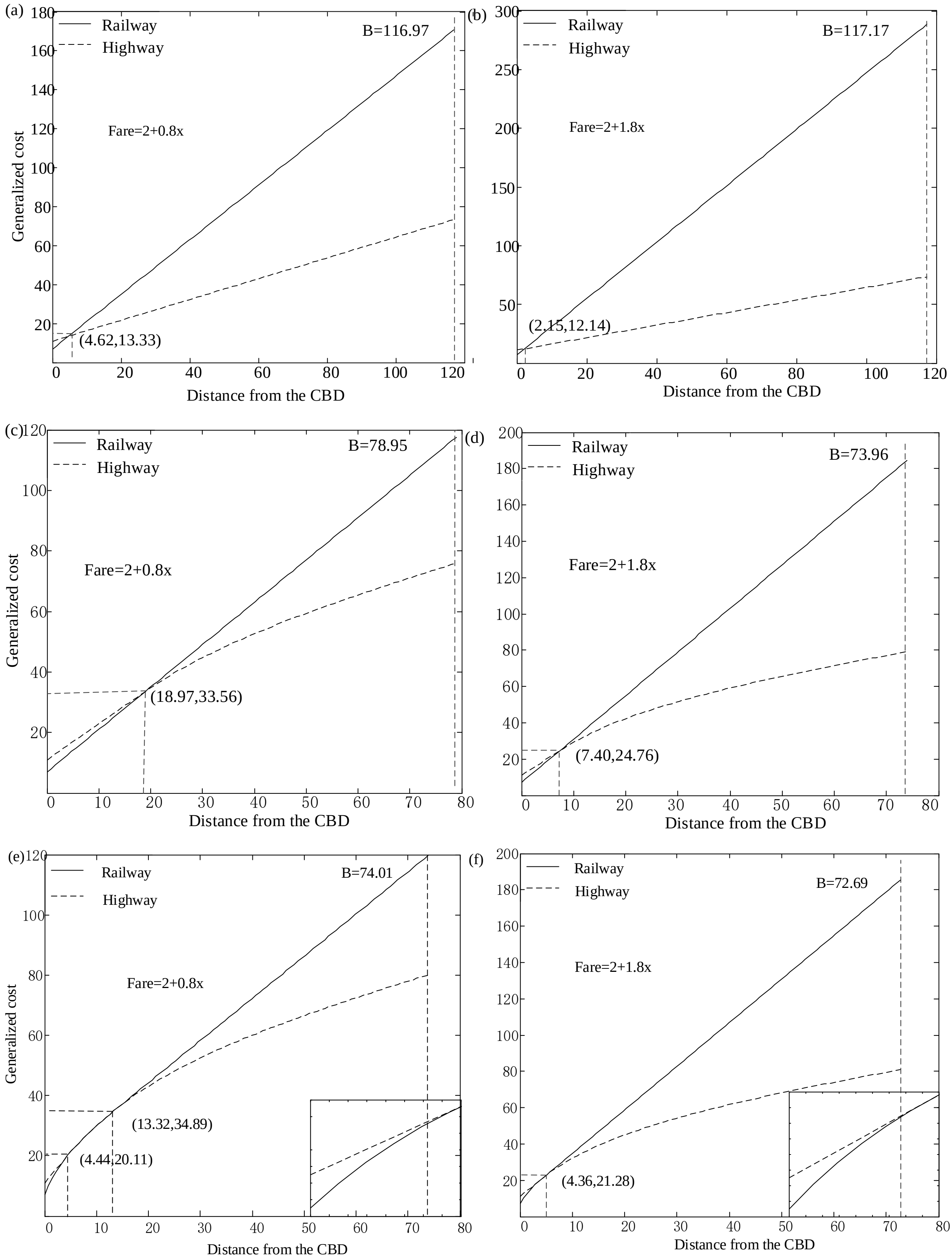

Fig. 4. Mode-choice patterns with different f_r for cases without highway congestion and transit crowding ((a) – (b)), ,with only highway congestion ((c) – (d)) & with highway congestion and transit crowding ((e) – (f)).

Next, we examine the changes when both of the city boundary and population distribution are exogenously given. We first fix the city boundary and population distribution as those endogenously generated by the studied city model with the four combination of f_r^0 and f_r, i.e., (2, 1.8), (2, 0.4), (10, 1.8) and (10, 0.4), and then observe the changes of equilibrium mode-choice points by adjusting f_r, as shown in Table 6. The results show that both the mode-switching points are decreasing with f_r, which is consistent with the results with endogenous city boundary and population distribution. Furthermore, it can be seen by comparing the results in Table 6 with that in Table 5 that, when the city boundary is fixed as 72.69 or 71.80, corresponding to the endogenous model with $f_r = 1.8$, all mode-switching points with different values of f_r are smaller than that obtained using the endogenous model. In additon, when the city boundary is fixed as 75.57 or 74.29, corresponding to the endogenous model with $f_r = 0.4$, the opposite trend comes true. This again implies that the equilibrium mode-choice patterns would be inaccurately predicted if endogenous properties of city boundary and population distribution are ignored.

Table 6. Equilibrium mode-choice points with different f_r when the city boundary and population density are exogenously given.

Exogenous examples	f_r^0	f_r	Mode-switching point closer to the CBD	Mode-switching point farther from the CBD
$B = 72.69$, corresponding to the endogenous model with $f_r = 1.8$	2	0.4	4.43	22.53
		0.8	4.42	13.08
		1.8	4.36	×
$B = 75.57$, corresponding to the endogenous model with $f_r = 0.4$	2	0.4	4.53	24.18
		0.8	4.52	14.35
		1.8	4.45	×
$B = 71.80$, corresponding to the endogenous model with $f_r = 1.8$	10	0.4	6.46	21.54
		0.8	6.45	12.92
		1.8	×	×
$B = 74.29$, corresponding to the endogenous model with $f_r = 0.4$	10	0.4	6.69	23.03
		0.8	6.68	14.11
		1.8	×	×

6. Railway fare and subsidy policies

In the previous sections, both the fixed and variable parts of railway fare are taken as exogenous parameters when we explore possible equilibrium mode-choice patterns with or without endogenous city boundary and population distribution. In this section, we focus on the comparison of different railway fare and subsidy policies, and investigate the influence of them on the population distribution, city boundary, utility level of residents, which are essential to develop a sustainable city.

6.1. Model setting

In most cities around the world, fare incomes are not high enough to cover the investment and the operating costs of transit system. Thus, direct financial subsidies are often provided by local governments to ensure suitable coverage of transit service (Gwilliam 2008; Tscharaktschiew and Hirte, 2012; Drevs et al., 2014). The source of transit subsidies mainly comes from local land revenue, property taxes, gasoline taxes, road tolls or others (Frankena, 1973; Creutzig, 2014; Xu et al., 2017). In this paper, the land revenue from land rents belongs to the government. To reveal the nature of the city model developed, we only consider part of the land revenue as the unique source of transit subsidies. Next, we first introduce two benchmark models for railway pricing without explicit transit subsidy, and then give the definition of high or low transit subsidy policies against them.

The first benchmark model is called the profit maximization model, in which the railway operator determines the fixed and variable parts of railway fare to maximize the profit, π_r. That is,

$$\max_{f_r^0, f_r} \pi_r = 2\varphi \int_0^B (f_r^0 + f_r x) n_r(x) dx - c_0, \tag{48}$$

where the first term is the annual fare income and c_0 is the fixed operating cost of railway. Hereafter, $(\tilde{f}_r^0, \tilde{f}_r)$ denotes the profit maximization solution.

The social welfare maximization model is the second benchmark model, in which the government aims to maximize the social welfare of urban system by optimizing the fixed and variable parts of railway fare. It can be formulated as

$$\max_{f_r^0, f_r} SW = \xi u N + \int_0^B \left(r(x) - r_a\right) dx + \pi_r, \tag{49}$$

where ξ is a parameter that converts the utility level of residents into the equivalent monetary units, the first term is associated with the total utility of residents, the second term is the government's land revenue from land rents after deducting agricultural rents, and the third term is the railway operator's profit. In the following text, $(\hat{f}_r^0, \hat{f}_r)$ denotes the social welfare maximization solution.

Given a specific fare policy without transit subsidy, the profit of the railway operator π_r might be either positive or negative since it depends on the relative values of fare income and fixed operating cost of railway. The operating of railway would be unsustainable in reality if $\pi_r < 0$. Thus, it is necessary to subsidy the railway to be operated at least at breakeven point in this situation. That is, the following expression must hold:

$$\pi_r + \theta_e \int_0^B \left(r(x) - r_a\right) dx \geq 0, \tag{50}$$

where θ_e is the minimum expenditure proportion of land revenue used for railway subsidy. Without loss of much generality, we only consider the case that Eq. (50) takes equality. Consequently, it holds that

$$\theta_e = \max\left(-\pi_r \Big/ \int_0^B \left(r(x) - r_a\right) dx, 0\right). \tag{51}$$

Further, letting θ_s denote the subsidy ratio of fixed operating cost of railway, we have

$$\theta_s = \theta_e \int_0^B \left(r(x) - r_a\right) dx \Big/ c_0 = \max\left(1 - 2\varphi \int_0^B (f_r^0 + f_r x) n_r(x) dx \Big/ c_0, 0\right). \tag{52}$$

It can be easily observed from Eqs. (51) – (52) that, the subsidy ratio θ_s decreases proportionally with the fare income $2\varphi \int_0^B (f_r^0 + f_r x) n_r(x) dx$ whilst the expenditure proportion θ_e might not necessarily. Consider a special case with $f_r = 0$, which means that all transit users will be charged the same flat fare, i.e., $f(x) = f_r^0$, for any $x \in [0, B]$. With the given parameter values in Table 2, where $c_0 = 2 \times 10^8$, Fig. 5 shows the changes of fare income, θ_s and θ_e varying with flat fare f_r^0. Obviously, when $f_r^0 = 5.3$, the fare income reaches the maximum and θ_s takes the minimum. In contrast, θ_e is minimal at $f_r^0 = 5.27$.

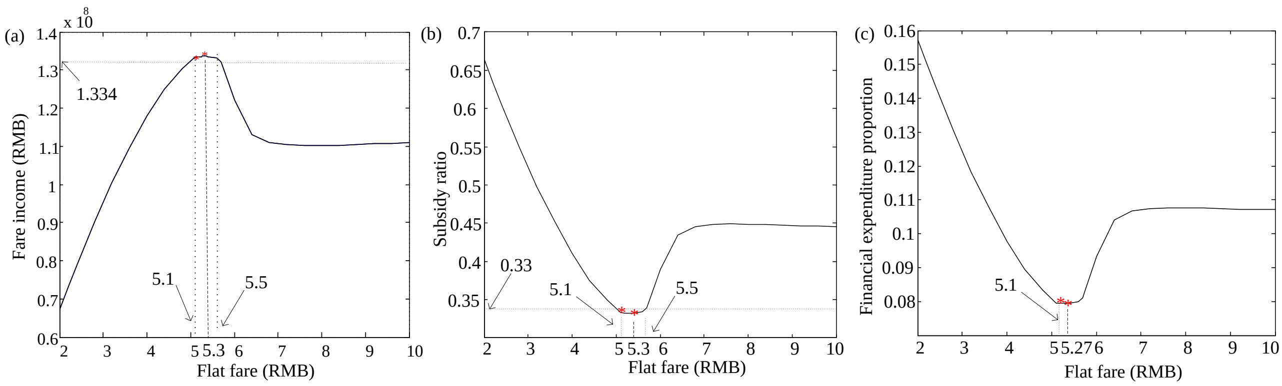

Fig.5. Changes of fare income, θ_s and θ_e with flat fare

Since the fare income under social welfare maximization is not higher than that under profit maximization, we give the definition of high or low transit subsidy policies as follows.

Definition 2. It is a high subsidy level if the fare income of railway under a specific transit fare and subsidy policy is lower than that under social welfare maximization. On the contrary, it is a low subsidy level if the fare income of railway is higher than that under social welfare maximization, but is lower than that under profit maximization.

Take the case with flat fare as an illustrative example. It is shown in Fig. 5 that, the fare income under social welfare maximization is 1.334×10^{8}, corresponding to $f_r^0 = 5.1$. In this situation, the operating deficit of railway may be fully subsidized if the subsidy ratio θ_s is at least 0.33. Thus, according to Definition 2, if the fare income of railway under a specific transit fare and subsidy policy is lower than 1.334×10^{8}, it is called as a fare policy with high subsidy. Otherwise, it is called as a fare policy with low subsidy.

Table 7. Influence of fixed operating cost on fare income and profit of railway under social welfare maximization, profit maximization and breakeven without transit subsidy

Policy		Flat fare			Distance-based fare with $f_r^0 = 2$			Distance-based fare with $f_r^0 = 8$		
$c_0\,(10^8)$		1.3	1.335	2.0	1.3	1.335	2.0	1.3	1.335	2.0
Social welfare maximization	(f_r^0, f_r)	(5.1, 0)	(5.1, 0)	(5.1, 0)	(2, 0.6)	(2, 0.6)	(2, 0.6)	(8, 0.25)	(8, 0.25)	(8, 0.25)
	Income (10^7)	13.34	13.34	13.34	13.69	13.69	13.69	13.17	13.17	13.17
	$\pi_r\,(10^4)$	340	–10	–6660	690	340	–6310	170	–180	–6830
Profit maximization	(f_r^0, f_r)	(5.3, 0)	(5.3, 0)	(5.3, 0)	(2, 0.8)	(2, 0.8)	(2, 0.8)	(8, 0.3)	(8, 0.3)	(8, 0.3)
	Income (10^7)	13.36	13.36	13.36	14	14	14	13.25	13.25	13.25
	$\pi_r\,(10^4)$	360	10	–6640	1000	650	–6000	250	–100	–6750
Breakeven	(f_r^0, f_r)	(4.8, 0), (5.75, 0)	(5.2, 0), (5.4, 0)	×	(2, 0.45), (2, 1.4)	(2, 0.5), (2, 1.28)	×	(8, 0.2), (8, 0.4)	×	×
	Income (10^7)	13	13.35	×	13	13.35	×	13	×	×
	$\pi_r\,(10^4)$	0	0	×	0	0	×	0	×	×

6.2. Numerical comparison

With the given values of parameters of the city model in Table 2, Table 7 shows the influence of fixed operating cost on the fare income and profit of railway under social welfare maximization, profit maximization and breakeven without transit subsidy. In this table, three specific fare policies, i.e., flat fare and distance-based fares with low or high fixed component ($f_r^0 = 2$ or 8), are also examined for comparison. Clearly, the profit of railway is decreasing with the fixed operating cost. Furthermore, when the railway has a larger fixed operating cost, e.g., $c_0 = 2\times10^{8}$, the profit of railway is always negative even under profit maximization. This renders the no-existence of breakeven solutions. Considering these, we next make a numerical comparison of urban system performance with different fare and subsidy policies in the case with $c_0 = 2\times10^8$, which is summarized in Table 8.

Table 8. Urban system performance with different fare and subsidy policies.

Performance index	Flat fare		Distance-based fare with $f_r^0 = 2$		Distance-based fare with $f_r^0 = 8$	
	Low subsidy	High subsidy	Low subsidy	High subsidy	Low subsidy	High subsidy
(f_r^0, f_r)	(5.3,0)	(2,0)	(2,0.8)	(2,0.2)	(8,0.3)	(8,0)
Subsidy ratio θ_s	33.21%	66.33%	30.03%	46.51%	33.77%	44.89%
Expenditure proportion θ_e	7.67%	15.19%	6.88%	10.64%	7.82%	10.40%
B	79.08	79.63	74.01	76.91	75.19	78.74
Average population density	1138	1130	1216	1170	1197	1143
Standard deviation of population density	1706	1727	1955	1805	1829	1719
Land revenue (10^8)	10.66	10.67	11.50	11.07	11.19	10.667
Residual land revenue after subsidizing (10^8)	9.04	9.05	9.75	9.39	9.49	9.046
Average land value (10^7)	1.096	1.097	1.18	1.14	1.15	1.10
Standard deviation of land value (10^7)	2.08	2.11	2.36	2.20	2.21	2.09
Utility level	284.62	286.45	285.04	286.34	283.28	283.56
Social welfare (10^9)	2.82	2.77	2.84	2.81	2.806	2.79

Note: Average population density = N/B, Standard deviation of population density = $\sqrt{\int_0^B \left(n(x) - N/B\right)^2 dx \Big/ B}$, Land revenue = $\int_0^B \left(r(x) - r_a\right) dx$, Residual land revenue after subsidizing = $(1-\theta_e)\int_0^B \left(r(x) - r_a\right) dx$, Average land value = $\int_0^B r(x)dx \Big/ B$, and Standard deviation of land value = $\sqrt{\int_0^B \left(r(x) - \int_0^B r(x)dx \Big/ B\right)^2 dx \Big/ B}$.

It can be seen from Table 8 that, it does not matter whether it is flat fare or distance-based fare with different fixed components, the policies with low subsidy always cause a decrease in the city boundary and utility level of residents and an increase in the average population density and social welfare compared to those with high subsidy. However, it is different if the other performance indexes of urban system, such as the standard deviation of population density, land revenue, residual land revenue after subsidizing, average land value and standard deviation of land value, are examined. Under the flat fare policy, low subsidy leads to more even population distribution and land value along the corridor and to lower land revenue and average land value. However, under both distance-based fare policies, low subsidy leads to the opposite results. This is because residents tend to live closer to the CBD when facing with the non-identical fares along the corridor. Furthermore, it can be found that, the distance-based fare policy with (2, 0.8) should be preferred among all three fare ones with low subsidy since the subsidy ratio is minimal, and the social welfare and utility level of residents are both maximal. If further lowering f_r on the basis of the fare policy, the social welfare would be worse although the utility level of residents becomes better, please see the fare policy with (2, 0) or (2, 0.2) for comparison.

7. Concluding remarks

We presented an urban spatial equilibrium model by integrating residents' household consumption, residential location choice and property developers' housing production with residents' mode choice. In this model, all residents are assumed to commute from their home to work at the CBD on a linear urban corridor, where a highway and a railway together form a competitive bimodal transportation system. The city boundary and population distribution become endogenous determinants in response to residents' consumption of housing and one composite non-housing good, and their residential location and mode choice decisions. Different from the existing bi-modal urban economics analysis (e.g., Capozza, 1973; Arnott and MacKinnon, 1977; Anas and Moses, 1979; LeRoy and Sonstelie, 1983; Sasaki, 1989, 1990; Su and DeSalvo, 2008; Creutzig, 2014), residents' transportation costs are also endogenously generated due to highway congestion and transit crowding in the proposed model.

The main findings and highlights of this paper are summarized as follows. Firstly, with exogenously given city boundary and population distribution, we derived the four possible equilibrium mode-choice patterns along the urban corridor by comparing the relative fixed cost of using transit mode with that of using auto mode. Comparably, only the case of smaller distance-free fixed cost by transit mode than that by auto mode was discussed in Liu et al. (2009). It is found that for any possible mode-choice pattern along the corridor, the mode-switching point farther from the CBD always decreases with the fixed cost component and the congestion-free variable cost per unit distance of railway travel. Secondly, we examined the effects of railway fare changes on the mode choice patterns and urban forms and found that a decrease of railway fare, whether in the fixed or variable components, would result in a spatial expansion of urban corridor if congested effects in the bimodal transportation system cannot be ignored. This result is different to the conclusion in the congestion-free case drawn in the urban economics literature (e.g., Sasaki, 1989, 1990; Su and DeSalvo, 2008). Finally, with the assumption that railway operation is subsidized from land rent revenue by the government to reach the breakeven point, we numerically compared the urban system performance under different railway fare and subsidy policies. We found that high railway subsidy (or low fare policy) will induce the spatial expansion of urban corridor and reduce the social welfare, no matter the fare is flat or distance-based. Furthermore, the distance-based fare policy with low subsidy should be preferred, under which the social welfare and utility level of residents can be Pareto improved.

Our work can be extended in several ways to investigate the in-depth interactions between transportation systems and land use patterns. Firstly, all residents were assumed to be homogenous in this paper. However, income levels of residents obviously determine their household consumption, auto vehicle ownership, and then residential location choice (Sasaki, 1990; Borck and Wrede, 2008). Therefore, residents' income heterogeneity should be incorporated to the model. Secondly, in reality, morning peak-hour congestion is generally dynamic, and commuters may choose to use the less congested mode to travel and/or to depart early or late in order to reduce congestion (Gubins and Verhoef, 2014; Wang and Du, 2016a; Xu et al., 2017). It is of interest to model residents' departure time choice, mode choice and residential location choice in an integrated urban framework. Thirdly, land rent revenue is only used to subsidize public transport, and only railway fare policies are compared in this paper. In a fast growing city, it is of importance to investigate the issues of fiscal subsidy for highway or railway construction.

Acknowledgments

The research described in this paper was jointly supported by grants from the National Natural Science Foundation of China (71771007), the Fundamental Research Funds for the Central Universities, the UK Rail Safety and Standard Board Project RSSB-T1071 and the Royal Academy of Engineering Newton Fund project UK-CIAPP\286. The work is carried out while the first author is on a one-year study visit to the University of Leeds funded by the China Scholarship Council.

Appendixes

Appendix A: Mode choice equilibrium

A.1. Proof of Proposition 1

Proof. Since the city boundary B and residential population density $n(x)$ are fixed here, it is easy to know from Eqs. (14), (18) and (20) that the mode-switching points $\bar{x}_2$, x_1 and x_3 are independent of a_r, but depends on the value of b_r. Since $t_h' > 0$ according to the assumption, take the first-order derivatives of both sides of Eqs. (14), (18) and (20) with respect to b_r, respectively, and we have

$$d\bar{x}_2/db_r = -1\Big/\left(t_h' n(\bar{x}_2)\right) < 0, \tag{A.1}$$

$$dx_1/db_r = -1\Big/\left(t_h' n(x_1)\right) < 0, \tag{A.2}$$

$$dx_3/db_r = -1\Big/\left(t_h' n(x_3)\right) < 0. \tag{A.3}$$

Next, we analyze the variation of $\bar{x}_3$ with respect to a_r or b_r. Taking the first-order derivative of both sides of Eq. (17) with respect to a_r or b_r leads to:

$$\frac{d\bar{x}_3}{da_r} = \frac{-1}{b_r - b_h + t_h^0 - t_h\left(\int_{\bar{x}_3}^{B} n(w)dw\right) + n(\bar{x}_3)\left(\int_0^{\bar{x}_3} g_r'\left(\int_x^{\bar{x}_3} n(w)dw\right)dx + \bar{x}_3 t_h'\left(\int_{\bar{x}_3}^{B} n(w)dw\right)\right)} < 0, \tag{A.4}$$

$$\frac{d\bar{x}_3}{db_r} = \frac{-\bar{x}_3}{b_r - b_h + t_h^0 - t_h\left(\int_{\bar{x}_3}^{B} n(w)dw\right) + n(\bar{x}_3)\left(\int_0^{\bar{x}_3} g_r'\left(\int_x^{\bar{x}_3} n(w)dw\right)dx + \bar{x}_3 t_h'\left(\int_{\bar{x}_3}^{B} n(w)dw\right)\right)} < 0, \tag{A.5}$$

where the inequalities hold due to $g_r' > 0$, $t_h' > 0$, and $b_r - b_h + t_h^0 - t_h\left(\int_{\bar{x}_3}^{B} n(w)dw\right) > 0$ from the condition (16). This completes the proof. □

A.2. Proof of Proposition 2

Proof. According to the assumption, the city boundary B and residential population density $n(x)$ are fixed as constants for the scenario shown in Panel (Ib) of Fig. 2. Without loss of generality, suppose that the fixed cost component of transit travel increases from a_r^o to a_r^n, or the congestion-free variable cost per unit distance by transit mode increases from b_r^o to b_r^n. Here, the variables with superscripts "o" and "n" denote the "original" and "new" ones, respectively. Obviously, $\bar{x}_3^o > \bar{x}_3^n$ holds from Proposition 1, which means residents located at $x \in \left(\bar{x}_3^n, \bar{x}_3^o\right]$ change their travel mode from transit to auto at new mode-choice equilibrium. As a result, it holds that $N_r^n(x) \le N_r^o(x)$, $c_r^{n\prime}(x) = g_r\left(N_r^n(x)\right) \le c_r^{o\prime}(x) = g_r\left(N_r^o(x)\right)$, $N_h^n(x) \ge N_h^o(x)$, and $G_h^n(x) > G_h^o(x)$ hold for any $x \in (0, B]$. Next, we analyze the variation of annual commuting cost $C(x)$ by dividing the whole corridor into three parts, i.e., $x \in \left[0, \bar{x}_3^n\right]$, $\left(\bar{x}_3^n, \bar{x}_3^o\right]$ and $\left(\bar{x}_3^o, B\right]$, respectively.

(1) For residents located at $x \in \left(\bar{x}_3^{o}, B\right]$, they always drive to the destination regardless of the variation of a_r or b_r. Hence, $C^n(x) = 2\varphi G_h^n(x) > 2\varphi G_h^o(x) = C^o(x)$ for any $x \in \left(\bar{x}_3^{o}, B\right]$.

(2) For residents located at $x \in \left(\bar{x}_3^{n}, \bar{x}_3^{o}\right]$, they change their travel mode from transit to auto at new mode-choice equilibrium. Hence, $C^n(x) = 2\varphi G_h^n(x) > 2\varphi G_h^o(x) \geq C^o(x) = 2\varphi G_r^o(x)$ for any $x \in \left(\bar{x}_3^{n}, \bar{x}_3^{o}\right]$, where the second inequality is due to Eq. (10).

(3) For residents located at $x \in \left[0, \bar{x}_3^{n}\right]$, they always travel by transit mode regardless of the variation of a_r or b_r. For proving $C^n(x) \geq C^o(x)$, it is sufficient and necessary to verify $G_r^n(x) \geq G_r^o(x)$. Since $G_h^n(x) > G_h^o(x)$ holds for any $x \in (0, B]$, $G_r^n(\bar{x}_3^{n}) = G_h^n(\bar{x}_3^{n}) > G_h^o(\bar{x}_3^{n}) > G_r^o(\bar{x}_3^{n})$ according to the user equilibrium conditions (9). Note that $G_r(0) = a_r$ and $G_r{}'(x) = b_r + c_r'(x)$ from Eq. (1). Next, we discuss different cases with the increase of a_r or b_r.

When the fixed cost component of transit travel increases from a_r^o to a_r^n, we have $G_h^n(0) = a_r^n > G_h^o(0) = a_r^o$, and $G_r^{n\prime}(x) = b_r + c_r^{n\prime}(x) < G_r^{o\prime}(x) = b_r + c_r^{o\prime}(x)$ for any $x \in \left(0, \bar{x}_3^{n}\right]$. Hence, considering the continuity of $G_r(x)$, $G_r^n(x) > G_r^o(x)$ must hold for any $x \in \left[0, \bar{x}_3^{n}\right]$.

When the congestion-free variable cost per unit distance by transit mode increases from b_r^o to b_r^n, we have $G_h^n(0) = G_h^o(0) = a_r$ and $G_r^{n\prime}(0) = b_r^n > G_r^{o\prime}(0) = b_r^o$. Since $c_r^{n\prime}(x) \leq c_r^{o\prime}(x)$ for any $x \in (0, B]$, there are at most a point $y \in \left[0, \bar{x}_3^{n}\right]$ such that $G_r^{n\prime}(y) = b_r^n + c_r^{n\prime}(y) = G_r^{o\prime}(y) = b_r^o + c_r^{o\prime}(y)$. Hence, considering the continuity of $G_r(x)$, $G_r^n(x) \geq G_r^o(x)$ must hold for any $x \in \left[0, \bar{x}_3^{n}\right]$.

This completes the proof. □

Appendix B: Derivations of rental price and lot size

Since all residents are identical, we have $U(x) = u$ for all x. Accordingly, combining it with Eq. (27), we have

$$u = \alpha^{\alpha} \left(Y - C(x)\right)^{\alpha} g(x)^{\beta}. \qquad \text{(B1)}$$

This leads to

$$g\left(x,u\right)=\alpha^{-\alpha/\beta}\left(Y-C(x)\right)^{-\alpha/\beta}u^{1/\beta}. \quad \text{(B2)}$$

Since $g(x)=\beta\left(Y-C(x)\right)/p(x)$ according to Eq. (26), we easily get

$$p\left(x,u\right)=\alpha^{\alpha/\beta}\beta\left(Y-C(x)\right)^{1/\beta}u^{-1/\beta}. \quad \text{(B3)}$$

This completes the derivations of Eqs. (28) and (29). □

References

Alonso, W., 1964. Location and Land Use: Toward a General Theory of Land Rent. Harvard University Press, Cambridge, Massachusetts.

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

Anas, A., Pines, D., 2008. Anti-sprawl policies in a system of congested cities. Regional Science and Urban Economics 38(5), 408–423.

Arnott, R.J., MacKinnon, J.G., 1977. The effects of urban transportation changes: a general equilibrium simulation. Journal of Public Economics 8(1), 19–36.

Beijing Municipal Bureau of Statistics, 2015. Beijing Statistical Yearbook 2014. available from http://www.bjstats.gov.cn/tjsj/tjgb/ndgb/201511/t20151124_327764.html.

Borck, R., Wrede, M., 2008. Commuting subsidies with two transport modes. Journal of Urban Economics 63, 841–848.

Boyce, D.E., Southworth, F., 1979. Quasi-dynamic urban-location models with endogenously determined travel costs. Environment and Planning A 11, 575–584.

Boyce, D.E., Williams, H.C., 2015. Forecasting Urban Travel: Past, Present and Future. Edward Elgar Publishing, Glos, UK.

Bravo, M., Briceno, L., Cortes, C.E., Martinez, F.J., 2010. An integrated behavioral model of the land-use and transport systems with network congestion and location externalities. Transport Research Part B 44(4), 584–596.

Brown, B., 1986. Modal choice, location demand, and income. Journal of Urban Economics 20(2), 128–139.

Brueckner, J.K., 1987. The structure of urban equilibria: A unified treatment of the Muth-Mills model. Handbook of Regional and Urban Economics 2, 821–845.

Brueckner, J.K., 2005. Transport subsidies, system choice, and urban sprawl. Regional Science and Urban Economics 35(6), 715–733.

Buyukeren, A.C., Hiramatsu, T., 2016. Anti-congestion policies in cities with public transportation. Journal of Economic Geography 16 (2), 395–421.

Capozza, D.R., 1973. Subways and land use. Environment Planning 5, 555–577.

Creutzig, F., 2014. How fuel prices determine public transport infrastructure, modal shares and urban form. Urban Climate 10, 63–76.

DeSalvo, J.S., Huq, M., 2005. Mode choice, commuting cost, and urban household behavior. Journal of Regional Science 45(3), 493–517.

Drevs, F., Tscheulin, D.K., Lindenmeier, J., Renner, S., 2014. Crowding-in or crowding out: An empirical analysis on the effect of subsidies on individual willingness-to-pay for public transportation. Transportation Research Part A 59, 250–261.

Du, B., Wang, D.Z.W., 2014. Continuum modeling of park-and-ride services considering travel time reliability and heterogeneous commuters – A linear complementarity system approach. Transportation Research Part E 71, 51–81.

Dubé, J., Thériault, M., Des Rosiers, F., 2013. Commuter rail accessibility and house values: The

case of the Montreal South Shore, Canada, 1992–2009. Transportation Research Part A, 54, 49–66.
Efthymiou, D., Antoniou, C., 2013. How do transport infrastructure and policies affect house prices and rents? Evidence from Athens, Greece. Transportation Research Part A 52, 1–22.
Evans, S.P., 1976. Derivation and analysis of some models for combining trip distribution and assignment. Transportation Research 10(1), 37–57.
Frankena, M.W., 1973. Income distributional effects of urban transit subsidies. Journal of Transport Economics and Policy 7(3), 215–230.
García, R., Marín, A., 2005. Network equilibrium with combined modes: models and solution algorithms. Transportation Research Part B 39(3), 223–254.
Gubins, S., Verhoef, E.T., 2014. Dynamic bottleneck congestion and residential land use in the monocentric city. Journal of Urban Economics 80, 51–61.
Gwilliam, K., 2008. A review of issues in transit economics. Research in Transportation Economics 23, 4-22.
Haring, J.E., Slobko, T., Chapman, J., 1976. The impact of alternative transportation systems on urban structure. Journal of Urban Economics 3(1), 14–20.
Ho, H.W., Wong, S.C., 2006. A review of the two-dimensional continuum modeling approach to transportation problems. Journal of Transportation Systems Engineering and Information Technology 6(6), 53–72.
Huang, H.J., Lam, W.H.K., 1992. Modified Evans' algorithms for solving the combined trip distribution and assignment problem. Transportation Research Part B 26, 325–337.
Huang, H.J., 2000. Fares and tolls in a competitive system with transit and highway: the case with two groups of commuters. Transportation Research Part E 36(4), 267–284.
Jehiel, P., 1993. Equilibrium on a traffic corridor with several congested modes. Transportation Science 27 (1), 16–24.
Jiang, Y., Wong, S.C., Ho, H.W., Zhang, P., Liu, R., Sumalee, A., 2011. A dynamic traffic assignment model for a continuum transportation system. Transportation Research Part B 45(2), 343–363.
LeRoy, S.F., Sonstelie, J., 1983. Paradise lost and regained: Transportation innovation, income, and residential location. Journal of Urban Economics 13(1), 67–89.
Li, Z.C., Lam, W.H.K., Wong, S.C., 2012. Modeling intermodal equilibrium for bimodal transportation system design problems in a linear monocentric city. Transportation Research Part B 46, 30–49.
Li, Z.C., Chen, Y.J., Wang, Y.D., Lam, W.H.K., Wong, S.C., 2013. Optimal density of radial major roads in a two-dimensional monocentric city with endogenous residential distribution and housing prices. Regional Science and Urban Economics 43, 927–937.
Liu, P., Liao, F., Huang, H. J., Timmermans, H., 2015. Dynamic activity-travel assignment in multi-state supernetworks. Transportation Research Part B 81, 656–671.
Liu, T.L., Huang, H.J., Yang, H., Zhang, X.N., 2009. Continuum modeling of park-and-ride services in a linear monocentric city with deterministic mode choice. Transportation Research Part B 43, 692–707.
Liu, W., Yang, H., Yin, Y.F., 2014. Traffic rationing and pricing in a linear monocentric city.Journal of Advanced Transportation 48, 655–672.
Lo, H.K., Yip, C.W., Wan, Q.K., 2004. Modeling competitive multi-modal transit services: a nested logit approach. Transportation Research Part C 12(3-4), 251–272.
Lu, X.S., Liu, T.L., Huang, H.J., 2015. Pricing and mode choice based on nested logit model with trip-chain costs. Transport Policy 44, 76–88.
Ma, X.S., Lo, H.K., 2012. Modeling transport management and land use over time. Transportation Research Part B 46, 687–709.
Mohammad, S.I., Graham, D.J., Melo, P.C., Anderson, R.J., 2013. A meta-analysis of the impact of rail projects on land and property values. Transportation Research Part A 50, 158–170.

Mills, E.S., 1967. An aggregative model of resource allocation in a metropolitan area. American Economic Review 57(2), 197–210.
Mills, E.S., 1972. Urban Economics. Scott Foresman, Glenview, Illinois.
Muth, R.F., 1969. Cities and Housing. University of Chicago Press, Chicago.
Ng, K.F., Lo, H.K., 2015. Optimal housing supply in a bid-rent equilibrium framework. Transportation Research Part B 74, 62–78.
Ng, K.F., Lo, H.K., 2017. On joint railway and housing development: Housing-led versus railway-led schemes. Transportation Research Part B 106, 464–488.
Peng, Y.T., Li, Z.C., Choi, K., 2017. Transit-oriented development in an urban rail transportation corridor. Transportation Research Part B 103, 269–290.
Safwat, K.N., Magnanti, T.L., 1988. A combined trip generation, trip distribution, modal split, and trip assignment model. Transportation Science 18, 14–30.
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(3), 322–343.
Sasaki, T., Iida, Y., Yang, H., 1990. User equilibrium traffic assignment by continuum approximation of network flow. In: Proceedings of the 11th International Symposium on Transportation and Traffic Theory, Japan, Yokohama, pp. 233–252.
Su, Q., DeSalvo, J.S., 2008. The effect of transportation subsidies on urban sprawl. Journal of Regional Science 48 (3), 567–594.
Tam, M.L., Lam, W.H.K., 2000. Maximum car ownership under constraints of road capacity and parking space. Transportation Research Part A 34(3) 145–170.
Tirachini, A., Hensher, D.A., Rose, J.M., 2013. Crowding in public transport systems: effects on users, operation and implications for the estimation of demand. Transportation research part A 53, 36–52.
Tscharaktschiew, S., Hirte, G., 2012. Should subsidies to urban passenger transport be increased? A spatial CGE analysis for a German metropolitan area. Transportation Research Part A 46 (2), 285–309.
Wardrop, J.G., 1952. Some theoretical aspects of road traffic research. In: Proceedings of Institution of Civil Engineers, Part II, vol. 1, pp. 325–378.
Wang, D.Z.W., Du, B., 2016a. Continuum modelling of spatial and dynamic equilibrium in a travel corridor with heterogeneous commuters—A partial differential complementarity system approach. Transportation Research Part B, 85, 1–18.
Wang, D.Z.W., Du, B., 2016b. Optimal urban development density along a multimodal linear travel corridor with time–distance toll scheme. Transportmetrica A, 12 (10), 853–877.
Wang, Z.J., Li, X.H., Chen, F. 2015. Impact evaluation of a mass transit fare change on demand and revenue utilizing smart card data. Transportation Research Part A 77, 213–224.
Wang, J.Y.T., Yang, H., Lindsey, R., 2004. Locating and pricing park-and-ride facilities in a linear monocentric city with deterministic mode choice. Transportation Research Part B 38 (8), 709–731.
Wong, S.C., 1998. Multicommodity traffic assignment by continuum approximation of network flow with variable demand. Transportation Research Part B 32 (8), 567–581.
Xu, S.X., Liu R., Liu, T.L., Huang, H.J., 2017. Pareto-improving policies for an idealized two-zone city served by two congestible modes. Transportation Research Part B, http://dx.doi.org/10.1016/j.trb.2017.08.010.
Yang, J., Quan, J., Yan, B., He, C., 2016. Urban rail investment and transit-oriented development in Beijing: Can it reach a higher potential?. Transportation Research Part A 89, 140–150.
Yang, H., Yagar, S., Iida, Y., 1994. Traffic assignment in a congested discrete/continuous transportation system. Transportation Research Part B 28 (2), 161–174.

Zhou, Z., Chen, A., Wong, S.C., 2009. Alternative formulations of a combined trip generation, trip distribution, modal split, and traffic assignment model. European Journal of Operational Research 198, 129–138.