A social optimum model for impacts of traffic congestion on inter- and intra-city commuting and social welfare

Xiangyu Li[a], Zhiwei Yin[b], Sida Luo[b], Hansen Mark[a,*]

[a]*Department of Civil and Environmental Engineering, University of California, Berkeley, United States*
[b]*Key Laboratory of Transport Industry of Big Data Application Technologies for Comprehensive Transport, Beijing Jiaotong University, Beijing, China*

Abstract

Owing to rapid urbanization, traffic congestion has become increasingly severe, thereby affecting the selection of employment and residence and commuting behaviors of urban residents. Recently, the focus of studies on urban commuting has shifted from single-center models to two-city models. However, most of the two-city models were linear single-center city models, where the effects of congestion are not considered. As a result, these studies are not ideal reflections of the practical cases. In this study, a social optimum model considering the urban congestion effect and the corresponding congestion function under both population endogenesis and exogenesis is proposed. The proposed model comprises two circular single-center city models considering the heterogeneity of congestion distribution within the city and one high-speed rail (HSR). Under the essence of social optimum conditions, non-convex planning is developed and solved to clarify optimized employment/residence distribution and commuting population in the two-city system. It is hypothesized that the congestion level at a specific location is inversely proportional to its population and distance to the downtown. Based on the proposed model, the gap between the current status of the two-city system and its optimized status can be evaluated and the key factors limiting further improvement of social welfare can be identified.

Keywords: traffic congestion; two-city system; employment/residence distribution; social optimum; nonlinear planning

1. Introduction

The rapid development of high-speed rail (HSR) networks in China has significantly transformed inter-city commuting. By the end of 2023, the total HSR mileage in China exceeded 45,000 km, connecting 95% of cities with populations over 500,000. With speeds reaching up to 400 km/h, HSR has strengthened inter-city connections, fostering the growth of urban agglomerations and metropolitan areas (Liu et al., 2020). This expansion has played a pivotal role in reshaping commuting patterns, particularly in regions like the Yangtze River Delta, where high-frequency HSR services facilitate daily travel between cities such as Shanghai and Suzhou. For instance, inter-city commuting in this region has increased by over 60% since 2018, with more than 155,000 individuals regularly traveling between Shanghai and nearby satellite cities, such as Nantong, Suzhou, Jiaxing, and Wuxi, in 2023 (Liu et al., 2024). The accessibility of HSR has not only promoted regional integration but also addressed high housing prices in metropolitan centers by enabling individuals

*Corresponding author

to reside in satellite cities while maintaining employment opportunities in core urban areas (Gao and Pan, 2023).

However, urban traffic congestion has emerged as a critical issue, hindering national economic growth and urban functionality, particularly in metropolises. In economic terms, urban traffic congestion represents a negative externality that remains insufficiently addressed despite solutions such as congestion tolls and traffic restrictions (Arnott, 2007; Brinkman, 2016; Gibson and Carnovale, 2015; Gu and Deakin, 2017; Ho, 2021). Traffic congestion also exacerbates environmental degradation, energy consumption (Mandhare and Kharat et al., 2018; Beagle and Belmont, 2019), and societal challenges (Ke and Chen et al., 2022). For example, congestion in Beijing alone cost the city significant productivity losses, as reflected in its persistent traffic index of 2.08—a 5.6% increase in commuter peak congestion in 2023 (Global Business Journalism, 2023). In regions like Beijing, residents of satellite cities, such as Yanjiao, frequently experience extended commuting times, spending 50–80% more time in traffic during peak hours (Zhao and Hu, 2019). Consequently, housing prices, living environments, and traffic congestion increasingly influence individuals' decisions regarding residence and employment locations, which, in turn, shape inter- and intra-city commuting behaviors.

The Guidance on Cultivation and Development of Modern Metropolitan Areas issued by China's National Development and Reform Commission emphasizes the need for comprehensive research into metropolitan commuting, defining a metropolitan area as "a spatial pattern of urbanization within an urban agglomeration, characterized by metropolises with strong radiating capabilities and a one-hour commuting circle" (Hu et al., 2020). Against this backdrop, this study develops a novel social optimum model to evaluate the interplay between urban agglomeration effects (positive externalities) and crowding effects (negative externalities) in a two-city system. This model investigates the interactions among employment location, residential distribution, and commuting behaviors while incorporating the effects of urban traffic congestion.

The remainder of this paper is structured as follows: Section 2 develops the theoretical model, including assumptions, congestion functions, and social optima formulations. Section 3 conducts an empirical study on the Shanghai-Jiaxing system, analyzing social welfare outcomes. Section 4 concludes with key findings, policy implications, and future research directions.

2. Literature Review

The investigation of occupation, residential choices, and commuting patterns in urban and metropolitan settings can be traced back to seminal studies in the 1960s that explored the interplay between housing prices and commuting costs within single-city systems (Alonso, 1964). Over the decades, an extensive body of research has emerged focusing on single-city dynamics, with significant contributions addressing single-center city models, which conceptualize urban economies around a central business district (Mills, 1967; Li et al., 2013; Buyukeren et al., 2016), as well as multi-center city models that account for the presence of multiple urban cores (Fujita and Ogawa, 1982; Zhang et al., 2016). Single-center city models typically assume that all employment opportunities are concentrated in the downtown area, with suburban areas serving primarily as residential zones. These models are further categorized into continuous city models, which consider a seamless spatial distribution of urban activities (Alonso, 1964), and urban-suburban city models, which explicitly differentiate between residential and commercial zones based on spatial location (Xu et al., 2018).

While these studies have provided valuable insights into the spatial structure and internal commuting costs of isolated cities, they often fail to offer detailed characterizations of the fundamental features of urban systems, such as the interactions between multiple cities. Early work by Henderson (1974) extended the scope of urban modeling to multi-city systems, examining the relationship between occupational choices, residential distribution, and commuting behaviors at the scale of urban agglomerations. Henderson's work introduced the concept of social optimum and competitive equilibrium in multi-city systems. However, the limited number of urban agglomeration models that followed (Henderson, 1974; Anas, 2013; Albouy, 2019) often relied on idealized assumptions regarding urban patterns. For example, many studies investigated residential choices using one-dimensional, linear single-center city models, applying these frameworks to policy evaluations, such as road pricing, public transportation systems, and urban rail transit design (Li et al., 2021; Yang et al., 2022). Although these linear models provide valuable theoretical frameworks, they deviate significantly from real-world urban structures, rendering their conclusions less applicable to policy-making in complex urban agglomerations.

In addition, urban agglomeration studies that rely on discrete choice models calibrated with field survey data (Li et al., 2012) often fail to generalize the relationship between residential and occupational choices and commuting behaviors across varying contexts. A particularly critical limitation in this body of work is the insufficient incorporation of traffic congestion, a significant factor that directly affects urban mobility and economic productivity. Recently, a growing number of studies have incorporated congestion into single-center city models (McConnell, 1982; McDonald, 2009; Franco, 2017), offering new perspectives on urban dynamics. For instance, Xu and Liu (2018) analyzed traffic congestion by focusing on bottleneck sections along highways connecting two densely populated zones. However, such simplifications inadequately capture the spatial heterogeneity of congestion, as traffic patterns and their associated costs vary significantly across different urban locations.

Furthermore, the external costs associated with traffic congestion, such as increased travel time, energy consumption, and environmental impacts, remain underexplored in many models (Quinet, 2004; Strotz, 2016). Although numerous studies have examined inter-city commuting, few have explicitly addressed the critical interplay between inter-city commuting and traffic congestion. This gap in the literature underscores the need for more comprehensive models that integrate congestion as a key variable in analyzing the dynamics of urban and inter-city systems.

To address these shortcomings, this study introduces a novel circular two-city social optimum model that incorporates the negative externalities of urban traffic congestion. The model investigates the impact of congestion on population distribution and commuter dynamics within two-city systems, offering a comprehensive framework for analyzing the selection of employment and residential locations, as well as commuting behaviors under varying congestion scenarios. To capture the spatial heterogeneity of traffic congestion, it is incorporated as a variable in the inter-city commuting model, reflecting its differential effects across urban areas.

The proposed model is formulated as a nonlinear optimization problem with six decision variables, encompassing occupation and residential choices along with commuting patterns. The travel cost per resident is expressed as a function of the total population and commuting population, serving as a critical optimization variable to represent the congestion effects induced by population agglomeration. Through numerical experiments, the model is applied to the Shanghai-Jiaxing region, using a linear congestion function as an illustrative example. Results demonstrate that inter-city railway commuting significantly enhances the spatial allocation of labor and increases net social income. However, the findings also reveal that the current volume of inter-city commuters remains substantially below the optimal level, highlighting the potential for improved system efficiency.

References	City model	Intracity commute	Intercity migration	Intercity commute	Congestion effect
Alonso (1964); Muth (1969); Mills (1967)	Monocentric and one-dimensional	Yes	No	No	No
Murata and Thisse (2005); Tabuchi, et al. (2005)	Monocentric and one-dimensional	Yes	Yes	No	No
Vandyck and Proost (2012)	Two cities and one-dimensional	No	No	Yes	Yes
Sorek (2009)	Two cities and one-dimensional	No	Yes	Yes	No
Borck and Wrede (2009); Dong, et al. (2022)	Two cities and one-dimensional	Yes	Yes	Yes	No
This study	Two cites and two dimensional	Yes	Yes	Yes	Yes

Table 1: Literature example

3. Model

3.1. Two-city system

In this study, an economy system composed of two cities (City 1 and City 2, see Fig. 1) were established and the basic characteristics of the two cities were reflected by a single-center model. Herein, locations of downtowns of City 1 and City 2 were fixed and urban areas are dependent on the distribution of the population. Meanwhile, the two cities are connected by HSR and both HSR stations are located in suburban areas. Residents in both cities need to work and both working places are available to them. Two options are available for employment location. The first option is the downtown of the respective city and the second option is the downtown of the other city. Therefore, residents in the two cities would be exposed to commuting expenses in different forms. Additionally, time delay caused by traffic congestion is considered as a large portion of commuting expenses.

In this study, the corresponding model was established based on six hypotheses.

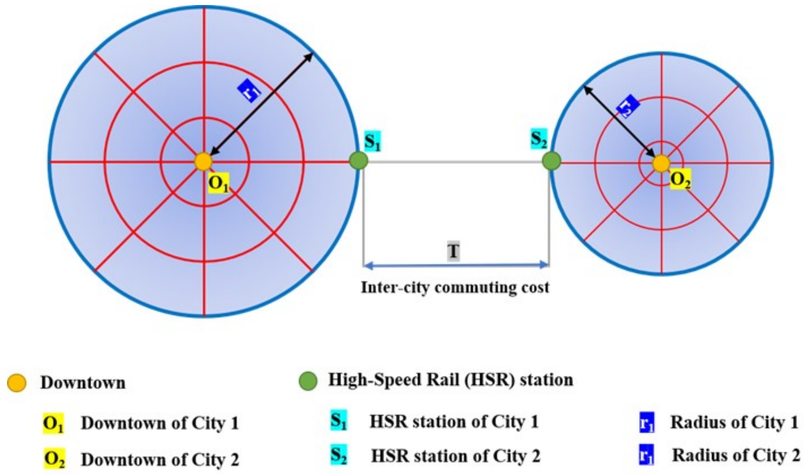


Figure 1: Schematic of a two-city system

Hypothesis 1: All cities in the system are circular and single-centered. In this study, r_i denotes the radius of City i. Inside the city, all road networks covering the city are circular and radial.

Hypothesis 2: There are differences in resource endowment and infrastructure between the two cities and these differences are reflected by the total factor productivity. A_i denotes the total factor productivity of City i. $A_1 > A_2$. City 1 and City 2 are called metropolis and satellite city, respectively.

Hypothesis 3: All production activities take place in a production agglomeration area that does not occupy land in downtown. All enterprises are in a completely competitive market and labor force is the sole production factor. Meanwhile, enterprises are not held by the residents in the system, meaning that the enterprise profits are not regarded as the income of the two-city system.

Hypothesis 4: Production output is proportional to input and the growth rate of output is greater than that of input. $P_i(W_i)$ denotes the overall output of City i and W_i denotes the number of employees in City i. It is assumed that $P_i = A_i W_i^{1+\gamma}\,(\gamma > 0)$. Housing consumption by all residents is set as 1. Additionally, the land rent at boundary of City i is determined by the agricultural rent R_i and $R_1 > R_2$.

Hypothesis 5: The system has a definite and constant population (N) and all people are homogeneous and evenly distributed in the city. For simplicity, the population density is assumed to be "$1/km^2$". Residents can choose to work in their respective city or the other city. Residents who live and work in the same city and those who live and work in different cities are denoted as intracity and intercity commuters, respectively.

Hypothesis 6: People are rational, that is, people always choose the path with minimum cost. Herein, the cost includes the commuting expense of free flow and additional cost caused by congestion.

Based on hypotheses mentioned above, intracity commuters travel directly from residence to downtown; intercity commuters shall travel to the HSR station of the living city at minimum expense and then to the other city by train. Hence, traffic volumes of different roads in the city may vary. Herein, a macroscopic model of traffic congestion was established. Local residents travel to the downtown along different central angles, resulting in uniform traffic congestion in different directions of central angles in the city; intercity commuters also generate a congestion effect in the sector with the axis of downtown - HSR station. In this study, it is assumed that the congestion cost of this city can be obtained by linear superposition of the two congestion effects. In the proposed model, the congestion level decreases with the increase of the distance from the downtown. The commuting cost consists of intercity and intracity commuting costs. Since HSR pricing and speed are not likely to fluctuate drastically in practical, it is assumed that intercity commuting cost is a definite constant denoted as T. For intracity commuting costs, h_i denotes the unit travel cost of a single person in City i ($0 \leq h_i < +\infty$). In order to investigate the effects of traffic congestion on intracity and intercity commuting, the congestion effect will be thoroughly discussed in the following section.

3.2. Congestion effect

The intracity commuting cost (h_i) can be divided into two parts ($h_i = t_i + f_i$). Herein, t_i refers to unit distance cost per capita in City i under free flow ; f_i refers to agglomeration-induced unit congestion cost per capita in City i. The units of both parameters are 10,000 yuan / (capita×meter).

$$f_i = f_{ia} + f_{ib} \tag{1}$$

where f_{ia} refers to the congestion cost induced by intracity commuters. As mentioned above, f_{ia} increases as the distance from downtown decreases; f_{ib} refers to the congestion cost induced by intercity commuters and f_{ib} induces additional congestion cost in the zone affected by the HSR station only. As shown in Fig. 2, $\Phi \geq 0$ reflects the range of additional congestion cost induced by intercity commuters. Specifically, intercity commuters travel to the downtown via a radial path within Sector $O_iS_iT_i$ upon arrival to the HSR station. Especially, $\Phi = 0$ suggests that all intercity commuters select roads that are directly connected to downtown and HSR station. Hence, the value of Φ is related to the traffic network near the HSR station of City i and generally accepted detour distances.

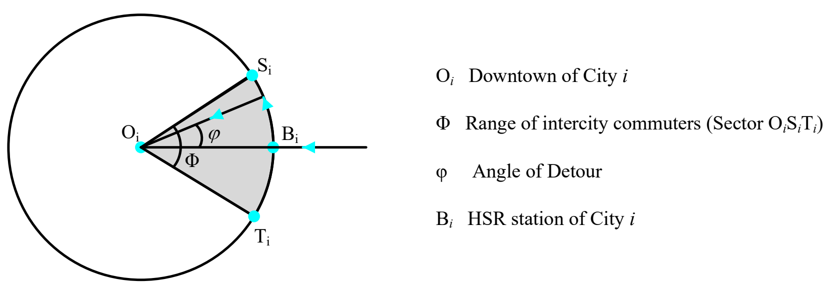


Figure 2: Scope of influence of intercity commuters on commuting near the HSR station

With downtown as the pole and the HSR station direction as the positive direction, a polar coordinate system is established. As hypothesized, the system has a known and constant population (N) and all residents are homogeneous and uniformly distributed in the city. The population density is 1 and the unit is /km2. As a result, for a point in City i, f_{ia} is a function of the radius of City i (r_i) and the population of City i (N_i). In other words, $f_{ia} = f_{ia}(r_i, r)$. Without losing generality, it can be considered that $f_{ia}(r_i, r_i) = 0$; N_{ji} refers to the population who live in City j but work in City i; within Φ, f_{ib} is a function of N_{ji}, namely $f_{ib} = f_{ib}(N_{ji})$. Herein, φ indicates the angle at which the commuters enter the downtown. The angle of the sector area affected by congestion in large cities, caused by the inter-city commuting population, increases with the rise in the number of inter-city commuters.

f_{ib} is within the range of Φ and can be expressed as

$$f_{ib} = \begin{cases} f_{ib}\left(N_{ji}\right) & -\frac{\Phi}{2} \leq \varphi \leq \frac{\Phi}{2} \\ \\ 0 & Other \end{cases} \tag{2}$$

3.3. Establishment of social optimum model

Social optimum refers to the case where total welfare of the system is maximized. In social optimum, the central government has the absolute power to determine the residential populations of all cities in the urban agglomeration and the intercity commuting populations, thus maximizing social welfare. Hence, social welfare = aggregate output - (aggregate commuting cost + aggregate agricultural rent). Consider the population endogenesis case, that is, the residential populations of the two cities are employed as the optimization variables in the model.

3.4. Theorem 1

Theorem 1: In case of population endogenesis, residents in City 1 do not commute to City 2 in social optimum.

Proof: See Appendix A. Contradiction method. If some residents commute from City 1 to City 2 in social optimum, define the optimized population distribution as (N_1^*, N_2^*, N_{12}^*), N_{12}^* refers to the total quantity of commuting from City 1 to City 2. Based on that, a new population distribution ($N_1^{'}, N_2^{'}$) is established under the premise of population endogenesis. Meanwhile, $N_1^{'} = N_1^* - N_{12}^*$, $N_2^{'} = N_2^* + N_{12}^*$ and no intercity commuting is observed. These two population distributions have no influences on the overall output and commuting cost of the two-city system, while they lead to reduced aggregate agricultural rent ($N_{12}^*\left(R_1 - R_2\right)$) as $R_1 > R_2$, demonstrating that new population distributions lead to enhanced social welfare. In other words, the original population distribution is not optimal. This conflicts with the hypothesis.

Theorem 1 demonstrated that residents in City 1 will not commute to City 2, suggesting that $N_{12} = 0$ and $N_{21} \geq 0$. Residents in City 1 always prefer radial paths as the way to downtown. Therefore, social optimum can be expressed as:

$$\max_{N_1,N_2,N_{21}} SW = P_1 + P_2 - (ACC_1 + ACC_2 + N_{21}T) - (R_1N_1 + R_2N_2) \tag{3}$$

s.t.

$$N_1 + N_2 = N \tag{3.1}$$

$$N_{21} \leq N_2 \tag{3.2}$$

$$N_1 \geq 0, N_2 \geq 0, N_{21} \geq 0 \tag{3.3}$$

where SW refers to social welfare, ACC_1 and ACC_2 refer to the intracity commuting costs (traffic expenses of residents from both cities in a specific city, excluding the expenses on HSR) of City 1 and City 2, respectively. Herein, the objective function (3) includes aggregate output, aggregate commuting cost and aggregate agricultural rent. The decision variables include population of City 1 (N_1), population of City 2 (N_2) and quantity of intercity commuting (N_{21}). Constraint (3.1) indicates that the total population of the system is determined and remains constant, Constraint (3.2) indicates that the quantity of intercity commuting does not exceed the population of City 2 and Constraint (3.3) is a non-negative constraint.

In order to solve the planning problem mentioned above, the three parts of objective function shall be expressed as functions of decision variables.

For the aggregate output,

$$\begin{aligned} P_1 &= A_1 \left(N_1 + N_{21}\right)^{1+\gamma} \\ P_2 &= A_2 \left(N_1 - N_{21}\right)^{1+\gamma} \end{aligned} \tag{4}$$

The house price (including rent) is transferred between subjects in the two-city system and aggregate agricultural rent reflects the cost of land reclamation by the central government. Since housing consumptions by all residents are 1, the aggregate agricultural rent is$R_1 N_1 + R_2 N_2$ (sum of the products of population and agricultural rent of the two cities).

For the aggregate commuting cost, ACC_1 and ACC_2 shall be denoted as functions of decision variables. Herein, ACC_1 is generated by both residents in City 1 and intercity commuters and ACC_2 includes the commuting costs to the HSR station and to the downtown of City 2. In order to reflect ACC_1 and ACC_2, identity and travel path of residents in City 2 must be solved. Identity refers to the role of intracity commuter or intercity commuter, while travel path refers to the selection of travel path ("circumferential + radial" or "radial + radial") to the HSR station by intercity commuters. As shown in Fig. 3, residents at Points M and Q travel to the HSR station via "circumferential + radial" and "radial + radial" paths, respectively. In both cases, congestion has key influences on identity and travel path. If City 1 is over-crowded, it is possible that negligible or no intercity commuters from City 2 are observed; if City 2, especially its downtown, is over-crowded, it is possible that negligible or no commuters select the "radial +radial" path.

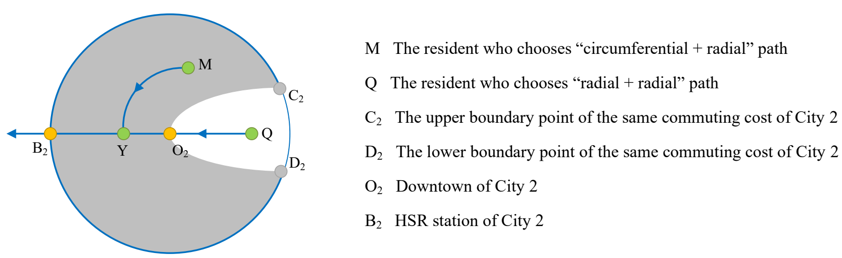


Figure 3: Different paths to HSR Station

First, the traveling paths of intercity commuters in City 2 are determined. Indeed, selection of "circumferential + radial" or "radial + radial" paths traveling to HSR station is dependent on the starting point of a commuter. If the commuting cost of "circumferential + radial" is lower than that of "radial + radial", the area consisting of the starting points of corresponding commuters is denoted as Area L (shown in Fig. 4). Otherwise, the area is denoted as Area S.

Herein, Area L refers to the gray area corresponding to major Arc C_2D_2. The boundaries of these two areas are Curves C_2O_2 and D_2O_2. Notably, Areas L and S are only present in cities with intercity commuters, meaning that Areas L and S are not present in City 1 under population endogenesis. Then, the specific forms of Curves C_2O_2 and D_2O_2 in City 2 are deduced. As shown in Fig. 4, only points on Curve D_2O_2 in City 2 shall be considered owing to the symmetry. Its polar coordinates are set as $\left(r, \bar{\theta}_2\right)$ and the costs traveling to the HSR station via "circumferential + radial" and "radial + radial" paths are equivalent.

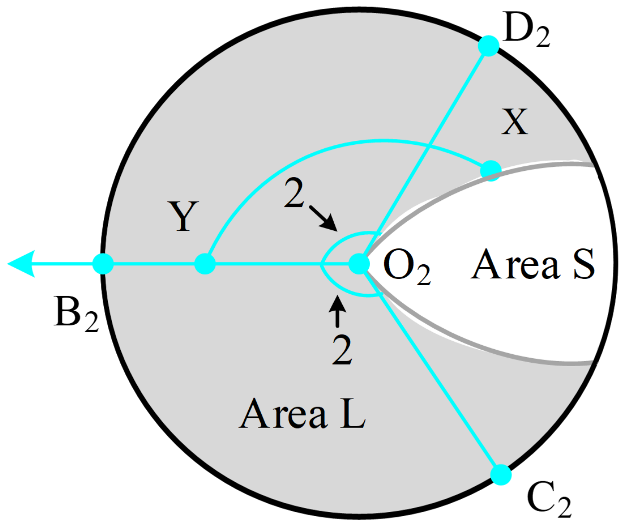


Figure 4: Geographic division of City 2

Hence,

$$h_2\bar{\theta}r + \int_r^{r_2} h_2 dr = \int_0^r h_2 dr + \int_0^{r_2} h_2 dr \tag{5}$$

where $h_2\bar{\theta}r$ refers to the commuting cost traveling to Point Y via circumferential paths; $\int_r^{r_2} h_2 dr$ refers to the commuting cost traveling from Point Y to Point B_2; $\int_0^r h_2 dr$ refers to the commuting cost traveling to Point O_2via radial paths; $\int_0^{r_2} h_2 dr$ refers to the commuting cost traveling from Point O_2 to Point B_2.

$$\bar{\theta}_2(r) = \frac{2\int_0^r h_2 dr}{h_2 r} \tag{6}$$

The conclusions about equations of boundary curves of Areas L and S are as follows:

Conclusion 1: if $r = 0$, $\bar{\theta}_2$ (polar coordinate of the point on the boundary of Areas L and S of City 2) is minimized and $\bar{\theta}_2^{min} = 2$.

Proof : See Appendix B.

Inference 1: both Areas L and S shall be present in City 2 regardless of the congestion situation.

Proof : See Appendix C.

Theorem 2: In social optimum, the differences of practical commuting cost and opportunity commuting cost of all points on the boundary of intercity commuting area AC_1 are consistent.

The following inference can be obtained based on **Theorem 2:**

Inference 2: If two residents with different identities have equal differences between practical costs and opportunity commuting costs, switching of their identities do not affect the overall commuting cost.

Proof: See Appendix D.

Theorem 2 and Inference 2 may be employed to deduce the boundary of intercity commuters. Owing to symmetry, only a random point E (see Fig. 5), whose polar coordinates are $(r, \widetilde{\theta}_2)$, on Boundary AC_1, shall be considered. In this way, equations for intercity commuting boundaries under population endogenesis in

Type-I and Type-II social optima can be deduced.

3.4.1. Type-I social optimum

In order to deduce the equation of boundary of intercity commuting area in Type-I social optimum, the intersection point of boundary curve and B_2O_2 in intercity commuting area (i.e., Point A in Fig. 5) is described using η. Herein, $\eta \in [0,1]$, $|O_2A| = \eta r_2$, η reflects the relative position of Point A to the downtown. Nevertheless, η would be employed as optimization variable in model optimization in the following section and η reflects the attractiveness of City 1 to City 2. Specifically, the attractiveness of City 1 to residents in City 2 increases as η decreases.

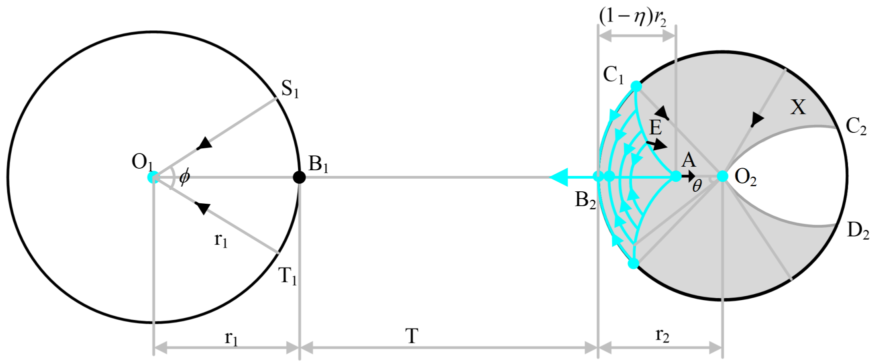


Figure 5: Intercity commuting area in Type-I social optimum

According to Theorem 2, differences of commuting costs of Points E and A are equivalent to each other. At Type-I social optimum, all intercity commuters are from Area L:

$$h_2\tilde{\theta}r + \int_r^{r_2} h_2 dr + T + \int_0^{r_1} h_1 dr - \int_0^{r} h_2 dr = \int_{\eta r_2}^{r_2} h_2 dr + T + \int_0^{r_1} h_1 dr - \int_0^{\eta r_2} h_2 dr \tag{7}$$

Terms on the left side of the equation reflect the difference between practical and opportunity commuting costs of Point E, while terms on the right side of the equation reflect the difference between practical and opportunity commuting costs of Point A.

Further simplification leads to:

$$\tilde{\theta}_2(r) = \frac{2\int_{\eta r_2}^{r} h_2 dr}{h_2 r} \tag{8}$$

If $\eta > 0$, $\widetilde{\theta}_2$ is minimized at $r = \eta r_2$ and $\widetilde{\theta}_2^{min} = 0$.

In this way, expression of boundary curves of intercity commuters in the distribution area of City 2 is obtained. N_{21} can be easily obtained based on boundary equation andN_{21} will be uniquely identified by η in this case. Hence, N_{21} can be substituted by η as decision variable. The area of intercity commuters in Area L increases as η decreases. Hence, η denotes the attractiveness of City 1 on City 2.

The boundary curve equation of intercity commuters has a determining effect on the identity (intercity or intracity commuters) of residents in City 2, while the boundary curve equation of Areas L and S ($\bar{\theta}_2$) has a determining effect on intercity commuters' selection of the path to HSR station. The correlation of $\bar{\theta}_2$and $\widetilde{\theta}_2$ reflects identity and path selection of residents in City 2.

Conclusion 2: At Type-I social optimum, regional boundary curves include boundary curves of commuters, namely $\widetilde{\theta}_2 \leq \bar{\theta}_2$.

Conclusion 3: If Points A and O_2coincide (see Fig. 5), regional boundary curves coincide with boundary curves of commuters. In other words, $\widetilde{\theta}_2 \leq \bar{\theta}_2$ if $\eta = 0$.

Since maximum angles of the two boundary curves ($\bar{\theta}_2^{max}$ and $\widetilde{\theta}_2^{max}$) cannot be determined, Type-I social optimum may be present in various patterns (see Table 1) and this is directly related to the traffic congestion level. As shown in Table 1, blue area denotes intercity commuters, while grey and white areas denote intracity commuters. If forced to commute to City 1, residents in grey and white areas would travel to HSR station via the "circumferential + radial" path and the "radial + radial" path, respectively.

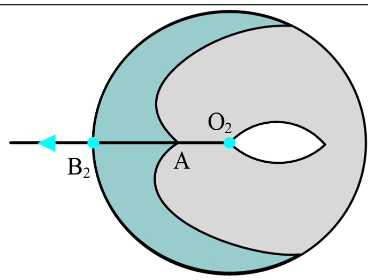
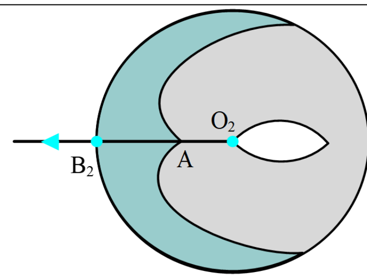
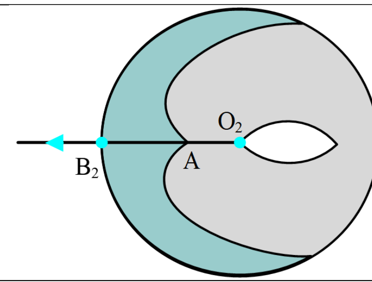
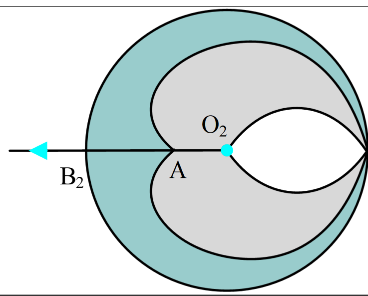
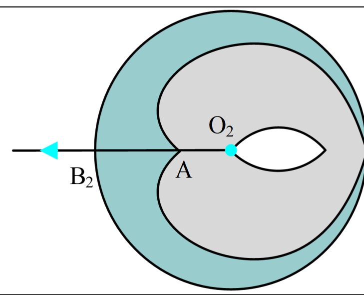
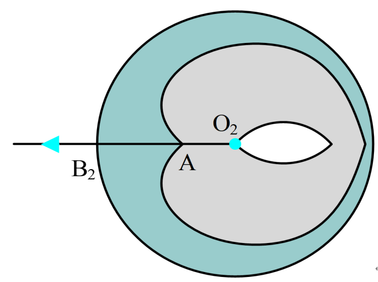
$\widetilde{\theta}_2 < \bar{\theta}_2,\ \widetilde{\theta}_2^{max} < \bar{\theta}_2^{max} < \pi$	$\widetilde{\theta}_2 < \bar{\theta}_2,\ \widetilde{\theta}_2^{max} < \bar{\theta}_2^{max} = \pi$	$\widetilde{\theta}_2 < \bar{\theta}_2,\ \widetilde{\theta}_2^{max} < \pi < \bar{\theta}_2^{max}$
		
$\widetilde{\theta}_2 < \bar{\theta}_2,\ \widetilde{\theta}_2^{max} = \bar{\theta}_2^{max} = \pi$	$\widetilde{\theta}_2 < \bar{\theta}_2,\ \widetilde{\theta}_2^{max} = \pi < \bar{\theta}_2^{max}$	$\widetilde{\theta}_2 < \bar{\theta}_2,\ \pi < \widetilde{\theta}_2^{max} < \bar{\theta}_2^{max}$
		

Table 2: Possible cases of boundary curves

According to Inference 2 and Table 1, $\widetilde{\theta}_2^{max} > \pi$is possible if $\bar{\theta}_2^{max} > \pi$, which interferes determination of $N_{21}$and ACC_2. For consistent expression, it is assumed that $\widetilde{\theta}_2 = \widetilde{\theta}_2^{max}$ when $r = \widetilde{r}_2$. Based on that, $\widetilde{r}_2$ and $\widetilde{\theta}_2^{max}$ are re-defined: $\widetilde{r}_2 = min\{r_2,\ \widetilde{r}_2\}$, $\widetilde{\theta}_2^{max} = min\left\{\widetilde{\theta}_2^{max},\ \pi\right\}$. In this case, all inferences mentioned above are still valid. Additionally, the expressions of N_{21} and ACC_2 using N_1, $N_2$and η were deduced. N_{21}, which equals to the area of intercity commuting zone, can be obtained by double integral:

$$N_{21} = 2\int_{\eta r_2}^{\widetilde{r}_2}\int_0^{\theta_2(r)} r dr d\theta + \pi\left(r_2^2 - \widetilde{r}_2^2\right) \tag{9}$$

ACC_2 consists of the costs of intercity and intracity commuters:

$$ACC_2 = 2\left[\int_{\eta r_2}^{\widetilde{r}_2} rdr \int_0^{\widetilde{\theta}(r)} \left(h_2\theta r + \int_r^{r_2} h_2 dr\right) d\theta + \int_{\widetilde{r}_2}^{r_2} rdr \int_0^{\pi} \left(h_2\theta r + \int_r^{r_2} h_2\right) d\theta + \int_0^{\widetilde{r}_2} \left(\pi r \int_0^r h_2 dr\right) dr - \int_{\eta r_2}^{r_2} \left(\tilde{\theta} r \int_0^r h_2 dr\right) dr\right] \quad (10)$$

Herein, the first part refers to the cost of intercity commuters traveling to the HSR station via "circumferential + radial" paths, while the second part refers to intracity commuters traveling to the downtown via "radial + radial" paths. Additionally, the proof that $\widetilde{\theta}_2^{max} > \pi$ is not possible when $\eta \to 1$ is presented.

Proof:See Appendix E.

3.4.2. Type-II social optimum

Before deduction of equation of boundary curve of intercity commuting area in Type-II social optimum, the correlation of the two social optima is clarified based on Theorem 3.

Theorem 3: In social optimum, residents in Area S shall not commute to City 1 for work unless all residents in Area L are required to commute to City 1 for work.

Theorem 3 indicates that residents in Area S shall commute to City 1 for work only if $\eta = 0$.

Proof: See Appendix F.

In Type-II social optimum, intercity commuters are from Areas L and S, while all residents in Area L are intercity commuters. Similar to the case of Type-I social optimum, the distribution area of intercity computers in Area S is determined first. Indeed, the difference of practical and opportunity commuting costs of a random Point (r,θ) in Area $S(\int_0^r h_2dr + \int_0^{r_2} h_2dr + T + \int_0^{r_1} h_1dr - \int_0^r h_2dr = \int_0^{r_2} h_2dr + T + \int_0^{r_1} h_1dr)$ is a constant. According to Inference 2, the aggregate commuting cost is independent from the residence of intercity commuters, suggesting that the residence of intercity commuters can be arranged randomly if the overall quantity of intercity commuters in Area S remains constant.

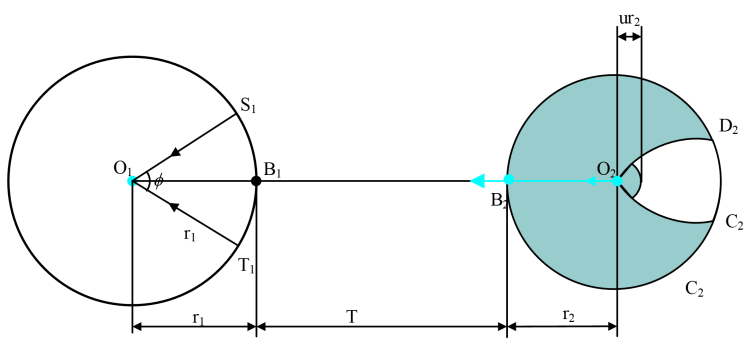


Figure 6: Intercity commuting area in Type-II social optimum

As shown in Fig. 6, the intercity commuting area of Area S is reflected by a sector with radius of $\mu r_2\,(0 \le \mu \le 1)$. Similar to ηin Type-I social optimum, μ would be employed as optimization variable in model optimization, μ reflects the attractiveness of City 1 to City 2 in Type-II social optimum. Specifically, the attractiveness of City 1 to residents in City 2 increases asμincreases.

In Type-II social optimum, all residents in Area L are intercity commuters and the two boundaries coincide. Table 3 lists all possible cases of boundary curves in Type-II social optimum.

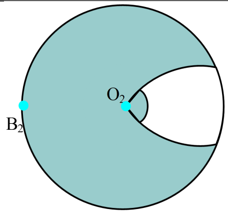
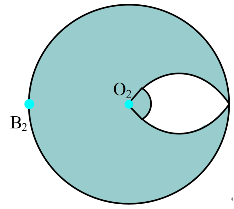
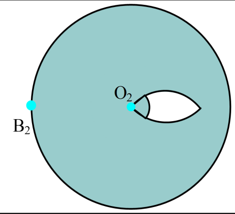
$\bar{\theta}_2^{max} < \pi$	$\bar{\theta}_2^{max} = \pi$	$\bar{\theta}_2^{max} > \pi$
		

Table 3: Possible cases of boundary curves in Type-II social optimum

Similar to the case of Type-I social optimum, N_{21} and ACC_2 in Type-II social optimum can be expressed as:

$$N_{21} = 2\left(\int_0^{\widetilde{\Gamma_2}} \int_0^{\theta(r)} r dr d\theta + \int_0^{\mu\widetilde{r_2}} (\pi - \tilde{\theta}) r dr\right) + \pi\left(r_2^2 - \widetilde{r}_2^2\right) \tag{11}$$

$$\begin{aligned} A^{2CC_2} = 2 & \left(\int_0^{\Gamma_2} r dr \int_0^{\theta^\theta(r)} \left(h_2 \theta r + \int_r^{r_2} h_2 dr\right) d\theta\right. \\ & + \int_0^{\mu\widetilde{2}_2} \left((\pi - \tilde{\theta}) r \left(\int_0^r h_2 dr + \int_0^{r_2} h_2 dr\right)\right) dr \\ & + \int_{\mu\widetilde{r_2}}^{\widetilde{r}_2} \left((\pi - \tilde{\theta}) r \int_0^r h_2 dr\right) dr + \int_{\widetilde{r}_2}^{r_2} r dr \int_0^{\pi} \left(h_2 \theta r + \int_r^{r_2} h_2\right) d\theta \end{aligned} \tag{12}$$

3.4.3. Social optimum model

In this section, ACC_1 is deduced. Residents in City 1 always travel to the downtown via radial paths, while intercity commuters tend to travel along circumferential paths of the boundary of City 1 before moving to the downtown via radial paths. We simply think that intercity commuters are evenly distributed on the boundary, so the average distance they detour is $\frac{\phi r_1}{4}$. With h_1^* denoting the commuting cost per unit distance of intercity commuters on the boundary of City 1, ACC_1can be expressed as:

$$\mathrm{ACC}_1 = \int_0^{r_1} \left(2\pi r \int_0^r h_1 dr\right) dr + N_{21}\left(\frac{\phi r_1}{4} h_1^* + \int_0^{r_1} h_1 dr\right) \tag{13}$$

where the first term denotes the commuting cost of residents in City 1 traveling to the downtown via radial paths and the second term denotes the commuting cost of intercity commuters traveling to the downtown via "circumferential + radial" paths.

N_{21} and ACC_2 can be expressed as:

$$N_{21} = 2\left(\int_{\eta r_2}^{\widetilde{r}_2} r dr \int_0^{\vartheta(r)} d\theta + \int_0^{\mu\widetilde{r}_2} (\pi - \tilde{\theta}) r dr\right) + \pi\left(r_2^2 - \widetilde{r}_2^2\right) \tag{14}$$

$$
\begin{aligned}
ACC_2 = 2 \Bigg(& \int_{\eta r_2}^{\Gamma_2} r dr \int_0^{\vartheta\theta(r)} \left(h_2 \theta r + \int_r^{r_2} h_2 dr \right) d\theta + \int_0^{\eta r_2} \left(\pi r \int_0^r h_2 dr \right) dr \\
& + \int_0^{\mu \widetilde{2}_2} \left((\pi - \tilde{\theta}) r \left(\int_0^r h_2 dr + \int_0^{r_2} h_2 dr \right) \right) dr \\
& + \int_{\eta r_2 + \mu \widetilde{r_2}}^{\widetilde{r}_2} \left((\pi - \tilde{\theta}) r \int_0^r h_2 dr \right) dr + \int_{\widetilde{r}_2}^{r_2} r dr \int_0^{\pi} \left(h_2 \theta r + \int_r^{r_2} h_2 \right) d\theta \Bigg)
\end{aligned} \tag{15}
$$

According to Theorem 3, $\mu > 0$ is possible only if $\eta = 0$, otherwise $\mu = 0$. Hence, with 0-1 variable δ introduced, social optimum (4) can be equivalent to:

$$
\max_{N_1, N_2, \eta, \mu, \delta} SW = P_1 + P_2 - (ACC_1 + ACC_2 + N_{21} T) - (R_1 N_1 + R_2 N_2) \tag{16}
$$

$$
\begin{aligned}
\text{s.t. } N_1 + N_2 &= N && (16.1) \\
\eta \delta &= 0 && (16.2) \\
N_1 &\geq 0 && (16.3) \\
N_2 &\geq 0 && (16.4) \\
0 \leq \mu &\leq \delta && (16.5) \\
0 \leq \eta &\leq 1 && (16.6) \\
\delta &\in \{0, 1\} && (16.7)
\end{aligned}
$$

As the proposed social optimum model is essentially nonlinear planning and the subject of this study is relatively small, three solving algorithms in MATLAB were employed and tested. Additionally, parameters in these algorithms were determined.

The three solving algorithms exhibit the following features:

1. fmincon and default interior point algorithm in Matlab are used as the model solver and the initial values before iterations are selected by random simulation.

2. The genetic algorithm (GA) kit in Matlab is used as the model solver.

3. The advantage of fmincon in local optimization and the advantage of GA in global searching are combined. In other words, iteration results by GA are used as input of fmincon for local optimization and the results are used as input of next iteration cycle. This is denoted as GA+f.

3.4.4. Population exogenesis

Section 2.3.3 introduces the social optimum model under population endogenesis. Under population exogenesis induced by household registration (i.e., both N_1 and N_2 are constants), residents in City 1 may commute to City 2 in social optimum due to the congestion effect. This may be attributed to the fact that social welfare can be enhanced by residents in City 1 commuting to City 2 for work as congestion-induced additional costs are high if City 1 is over-crowded. As a result, Theorem 1 is not valid and the problem would evolve. Therefore, the following theorem shall be proved first:

Theorem 4: Simultaneous commuting between metropolis and satellite city is not possible. **Proof:** See Appendix F.

Theorem 4 suggests that intercity commuting is either from City 2 to City 1 or from City 1 to City 2. Therefore, population exogenesis can be divided as two population endogenesis by the following method to determine the social optimum.

City A and City B constitute a two-city system and the total factor productivity of City A is higher than that of City B. Under population endogenesis, City A is denoted as City 1 and City B is denoted as City 2. Apparently, residents commute from City 2 to City 1. Under population exogenesis, City A and B are defined as City 1 and 2, respectively;$N_1$and N_2 denoted populations of City A and B, respectively. the optimized solution (SW^1) can be obtained; by defining City B as City 1 and City A as City 2. The optimized solution (SW^2) can be obtained by solving the social optimum model mentioned in Section 2.2.3. Additionally, the larger of them is taken as the optimized solution under population exogenesis, namely $SW = \max\{SW^1, SW^2\}$.

In this way, a social optimum model of two-city system can be established and the corresponding solution method is given.

Additionally, the possibility of intercity commuting and that residents in City 2 commute to City 1 for work when N_1 and N_2 remain constant is investigated based on the proposed social optimum model.

The presence of intercity commuting guarantees that:

$$\lim_{N_{21}\to 0} \frac{\partial SW}{\partial N_{21}} > 0 \tag{17}$$

Since $N_{21} \to 0$ is equivalent to $\eta \to 1$, it has been demonstrated that $\widetilde{\theta}_2^{max} > \pi$ is not possible in this case. Hence,

$$\begin{aligned}
\frac{\partial SW}{\partial N_{21}} &= (1+\gamma)\left(A_1\left(N_1+N_{21}\right)^{\gamma} - A_2\left(N_2-N_{21}\right)^{\gamma}\right) - \frac{\partial ACC_1}{\partial N_{21}} - \frac{\partial ACC_2}{\partial N_{21}} - T \\
&= (1+\gamma)\left(A_1\left(N_1+N_{21}\right)^{\gamma} - A_2\left(N_2-N_{21}\right)^{\gamma}\right) \\
&\quad - \int_0^{r_1}\left(\phi r \int_0^r \frac{\partial h_1}{\partial N_{21}} dr\right) dr - \frac{\phi r_1}{4} h_1^* - \int_0^{r_1} h_1 dr - N_{21}\int_0^{r_1} \frac{\partial h_1}{\partial N_{21}} dr \\
&\quad - 2\left(\int_{\eta r_2}^{r_2} h_2 dr - \frac{\int_{\eta r_2}^{r_2} \frac{\int_0^r h_2 dr}{h_2} dr}{\int_{\eta r_2}^{r_2} \frac{1}{h_2} dr}\right) - T
\end{aligned} \tag{18}$$

It can be obtained from Eq (18):

$$\begin{aligned}
\lim_{N_{21}\to 0} \frac{\partial SW}{\partial N_{21}} &= (1+\gamma)\left(A_1 N_1^{\gamma} - A_2 N_2^{\gamma}\right) + 2\int_0^{r_2} h_2 - \int_0^{r_1}\left(\phi r \int_0^r \frac{\partial h_1}{\partial N_{21}} dr\right) dr \\
&\quad - \int_0^{r_1}\left(t_1 + f_{1a}\right) dr - \frac{t_1 \phi r_1}{4} - T
\end{aligned} \tag{19}$$

As ϕ is typically small, In this case, the presence of intercity commuting leads to:

$$\lim_{N_{21}\to 0}\frac{\partial SW}{\partial N_{21}}=(1+\gamma)\left(A_1{N_1}^{\gamma}-A_2{N_2}^{\gamma}\right)+2\int\limits_0^{r_2}h_2-T-\int\limits_0^{r_1}\left(t_1+f_{1a}\right)dr>0 \tag{20}$$

Namely

$$T<(1+\gamma)\left(A_1N_1^{\gamma}-A_2N_2\gamma\right)+2\int_0^{r_2}h_2-\int_0^{r_1}\left(t_1+f_{1a}\right)dr \tag{21}$$

When the population difference between City 1 and City 2 is huge, Eq (20) tend to be valid. Generally, intercity commuting will always be observed if the overall population of the system exceeds a critical level.

The case where all residents in City 2 commute to City 1 for work is equivalent to the case where $\eta=0$ and $\mu=1$. This situation is indeed Type-II social optimum. The population of Area L of City 2 is denoted as N_2^L, which can be determined by Eq (15). In this case, $N_{21}\in[N_2^L,N_2]$ and $\eta=0$, $\mu=1$ is equivalent to:

$$(\frac{\partial SW}{\partial N_{21}})\geq 0 \tag{22}$$

Additionally, $\frac{\partial SW}{\partial N_{21}}\geq 0$ is always valid if $N_{21}=N_2^L$ as the situation is Type-II social optimum. According to Eq (15):

$$\begin{aligned}\frac{\partial SW}{\partial N_{21}}&=(1+\gamma)\left(A_1\left(N_1+N_{21}\right)^{\gamma}-A_2\left(N_2-N_{21}\right)^{\gamma}\right)-\frac{\partial ACC_1}{\partial N_{21}}-\frac{\partial ACC_2}{\partial N_{21}}-T\\ &=(1+\gamma)\left(A_1\left(N_1+N_{21}\right)^{\gamma}-A_2\left(N_2-N_{21}\right)^{\gamma}\right)-\int_0^{r_1}\left(\phi r\int_0^{r}\frac{\partial h_1}{\partial N_{21}}dr\right)dr-\frac{\phi r_1}{4}h_1^*\\ &\quad-\int_0^{r_1}h_1dr-N_{21}\int_0^{r_1}\frac{\partial h_1}{\partial N_{21}}dr-\int_0^{r_2}h_2dr-T\end{aligned} \tag{23}$$

If $-1<\gamma<0$, $\lim\limits_{N_{21}\to N_2}\frac{\partial SW}{\partial N_{21}}=-\infty$ as suggested by Eq (19), demonstrate that $(\frac{\partial SW}{\partial N_{21}})_{min}<0$. Therefore, not all residents in City 2 would choose to commute to City 1 for work. If $\gamma>0$, only cases where congestion function is constant (i.e., h_1, h_2 are constants) are investigated. According to Eq (19):

$$\frac{\partial SW}{\partial N_{21}}=(1+\gamma)\left(A_1\left(N_1+N_{21}\right)^{\gamma}-A_2\left(N_2-N_{21}\right)^{\gamma}\right)-h_1r_1-\frac{h_1\phi r_1}{4}-h_2r_2-T \tag{24}$$

Further evolution leads to:

$$\frac{\partial^2 SW}{\partial N_{21}^2}=\gamma(1+\gamma)\left(A_1\left(N_1+N_{21}\right)^{\gamma-1}+A_2\left(N_2-N_{21}\right)^{\gamma-1}\right)>0 \tag{25}$$

According to Eq (21), $\frac{\partial SW}{\partial N_{21}}$ increases monotonically with N_{21}. Hence, $(\frac{\partial SW}{\partial N_{21}})_{min}\geq 0$, demonstrating that all residents in City 2 would commute to City 1 for work.

4. Empirical study

4.1. Target area

Shanghai and Jiaxing cities in China, both of which are in the Yangtze River Delta region, are selected as the object of empirical study. As a regional and global financial center, Shanghai is the core of the Yangtze River Delta region. Jiaxing is a second-tier city adjacent to Shanghai. Before the operation of HSR, a train trip from Jiaxing to Shanghai took 1.5 hours. After the operation of HSR, such a trip takes 0.45 h. Therefore, various residents in Jiaxing commute to Shanghai for work, resulting in huge intercity commuting. Despite their large areas, considerable parts of both Shanghai and Jiaxing are still under- or undeveloped. In 2022, the GDP of the central area of Shanghai was approximately 3.82 trillion, which is 2/3 of the total GDP of Shanghai, indicating that economic activities in Shanghai are concentrated in this region. Hence, the central areas of Shanghai and Jiaxing (red line in Fig. 7) are defined as urban boundaries of Shanghai and Jiaxing and all parameters are extracted from the central area. Additionally, both Shanghai Hongqiao Station and Jiaxing South Station are located at the urban boundaries and multiple HSR train departures are available to satisfy commuting needs. The practical situation is highly consistent with the model.

Figure 7: Downtowns of Shanghai and Jiaxing

4.2. Parameter calibration

Other parameters are calibrated based on practical data listed in Table 4 (mainly extracted from government statistics). Herein, social welfare and resident utility are calculated in years. Hence, both resident income and commuting cost refers to annual values per capita.

Table 4: Definitions of input symbols

Input	Description	Unit	Value	Data source
$\overline{N}_i$	Population of City i		Shanghai:12,436,800 Jiaxing: 1,288,700	Shanghai Bureau of Statistics; Jiaxing Bureau of Statistics
$\overline{N}_{21}$	Current intercity commuting population		3100	*Annual Report onIntercity Commuting in the Yangtze River Delta in 2019*
s_i	Housing space per capita in City i	m^2	Shanghai:37 Jiaxing: 41.26	Interface news; 2018 Statistical Bulletin of National Economic and Social Development of Jiaxing City
v_i	Average vehicle travel speed in City	km/h	Shanghai:37 Jiaxing: 41.26	
λ_i	Money cost of commuting per unit km in City	¥10,000	Shanghai:0.00015 Jiaxing: 0.0001	Estimated according to taxi pricing.
tc	Time value of intercity commuting	Hour	0.65	Official website of China Railway 12306
mc	Money cost of intercity commuting	¥ 10,000	0.00385	Official website of China Railway 12306
GDP_i	GDP of City	¥ 10,000	Shanghai:213,388,800 Jiaxing: 12,464,200	Shanghai Bureau of Statistics; Jiaxing Bureau of Statistics
tw	Total annual working days	Day	232 Notice of the General Office of the State Council on the Arrangement of Some Holidays in 2018	

The total output of City i can be measured by its GDP (GDP_i) and population $\bar{N}_i$:

$$A_i = \frac{GDP_i}{N_i^{1+\lambda}} \tag{26}$$

In this study, A_1 and A_2 were 7.58 and 4.79, respectively.

In 2018, average annual incomes of residents in Shanghai and Jiaxing were 140270 and 89311 yuan, respectively. With 10 working hours per day, the average hourly wage of residents in Shanghai was about 74 yuan. Based on that, the time value of residents involved is set to be 74 yuan/h (denoted as vot, vot = 74 yuan/h). The intercity commuting cost per year can be determined using working days per year (tw), time value (vot) and monetary cost of intercity commuting per trip (mc):

$$T = tw \times (2 \times mc + 2 \times vot \times tc) = 40182\, yuan \tag{27}$$

The annual commuting cost under intracity free flow can be determined using tw, vot, monetary cost per km of City i (λ_i), housing space per capita (s_i) and vehicle speed under free flow (v_i):

$$t_i = tw \times (2 \times \lambda_i \times \frac{\sqrt{s_i}}{1000} + 2 \times vot \times \frac{\sqrt{s_i}}{1000 v_i}) \tag{28}$$

In this study, $t_1$andt_2 were 7.7 yuan and 8.5 yuan, respectively.

With agricultural rent R_i substituted by the average housing price of suburban area of City i, R_1 and R_2 are 75200 and 23100 yuan, respectively.

4.3. Determination of congestion cost

As discussed above, congestion costs of City 1 and City 2 can be expressed as:

$$h_1 = t_1 + f_{1a} + f_{1b},\ h_2 = t_2 + f_{2a} \tag{29}$$

Then, f_{1a}, f_{2a} and f_{1b} are determined.

4.3.1. Determination of f_{1a} and f_{2a}

It is assumed that $f_{1a}$and f_{2a} are linear:

$$f_{1a} = k_{1a}(r_1 - r) \tag{30}$$

$$f_{2a} = k_{2a}(r_2 - r) \tag{31}$$

where k_{1a} and k_{2a} are two undetermined constants that can be determined based on current data:

$$\int_0^{\bar{r}_i} f_{ia} dr = m_{ia} t_i \bar{r}_i \quad i = 1, 2 \tag{32}$$

$$k_{ia} = \frac{2 m_{ia} t_i}{\bar{r}_i} \tag{33}$$

where $\bar{r}_i = \sqrt{\frac{\bar{N}_i}{\pi}}$.

In 2022, the congestion delay coefficients of Shanghai and Jiaxing are 1.74 and 1.35, respectively. With travel cost considered, $m_{1a} = 0.74$ and $m_{2a} = 0.35$. The calibration results of t_i and N_i mentioned in Section 3.2 are substituted into Eq. (27): k_{1a} and k_{2a} are 7.5×10^{-3} and $k_{2a} = 1.13 \times 10^{-2}$, respectively.

In 2018, the congestion delay coefficients of Shanghai and Jiaxing are 1.82 and 1.25, respectively. With travel cost considered,$m_{1a} = 0.82$ and $m_{2a} = 0.25$. The calibration results of $t_i$and N_i mentioned in Section 3.2 were substituted into Eq (24): k_{1a} and k_{2a} were 6.3588×10^{-7} and $k_{2a} = 6.3029 \times 10^{-7}$, respectively.

4.3.2. Determination of f_{1b}

According to the analysis in Section 2, the congestion cost caused by intercity commuters is expressed as a linear function of intercity commuters:

$$f_{1b} = k_{1b} N_{21} \tag{34}$$

Similar to determination of k_{1a},based on current data, k_{1b} can be determined by:

$$k_{1b} N_{21} \bar{r}_1 = m_{1b} t_1 \bar{r}_1 \tag{35}$$

where $\bar{r}_1 = \sqrt{\frac{\bar{N}_1}{\pi}}$. In other words, intercity commuters induced congestion cost near the HSR station of City 1 is m_{1b} times of that under free flow.

Concerning commuters' incentive to avoid the congestion, we calibrate phi such that the number of commuters coincide with the number of residences of the region induced by phi in big city.

Considering that the congestion near the high-speed railway station of city 1 is caused by the population of city 1 and the commuter population, in this paper, we use the following formula to determine m_{1b}

$$m_{1b} = \frac{\bar{N}_{21}}{\bar{N}_{21} + \frac{1}{2}\Phi r^2} m_{1a} \tag{36}$$

Then, it can be obtained that $k_{1b} = \frac{2 m_{1b} t_1}{\bar{r}_1}$.

4.4. Analysis of social optimum

The intercity commuting in Shanghai and Jiaxing under both population exogenesis and population endogenesis are analyzed. Analysis of population exogenesis revealed the difference between the current and optimized commuting status of Jiaxing and Shanghai, thus providing references to short-term and medium-term commuting policies by the government. Likewise, analysis of population endogenesis also provides references to development strategy and long-term planning by the government. This can be attributed to the fact that in population exogenesis, changes in the household registration of the residents in the two cities are not allowed (i.e., residents in Jiaxing are not allowed to move to Shanghai) and the populations of the system and the two cities are constant; in population endogenesis, changes in the household registration of the residents are allowed if the total population of the system remains constant. Current social welfare can be obtained by substituting the current residential and commuting populations of Jiaxing and Shanghai into the model. Additionally, social welfare with and without congestion taken into consideration are 9.7296×10^{11} and 1.4394×10^{12} yuan, respectively.

4.4.1. Social optimum under population exogenesis

Fig. 8 shows the trend of social welfare as a function of η (μ) in Type-I and Type-II social optimum without congestion taken into consideration. As observed, social optimum can only be reached if all residents in Jiaxing commute to Shanghai for work (mu = 1 under Type-II social optimum). In this case, the Type-II social welfare is maximized (1.4394×10^{12} yuan). According to statistics, the population commuting from Jiaxing to Shanghai is currently 12400. In summary, results obtained without congestion taken into consideration reveal that this system is far away from social optimum, while it is apparently unrealistic for all residents in Jiaxing to commute to Shanghai for work.

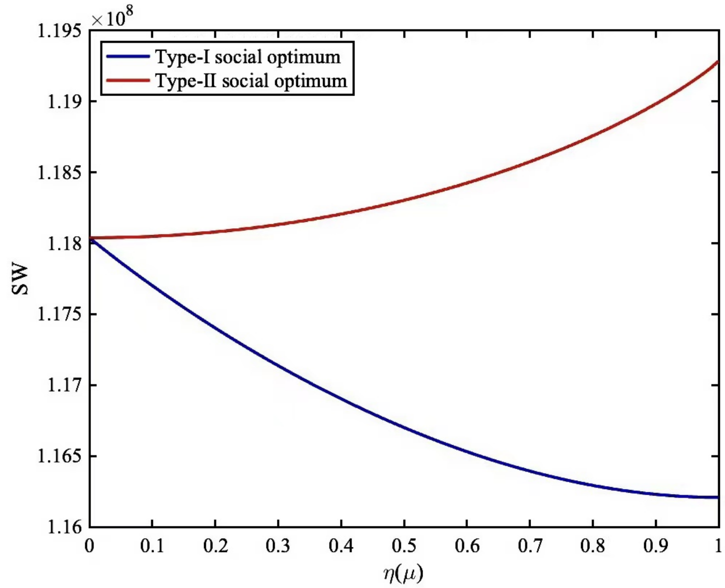


Figure 8: Social optimum without congestion taken into consideration

With congestion taken into consideration, trigonometric function is employed to verify the rationality of linear f_{1a} and f_{2a}. The trigonometric function can be expressed as:

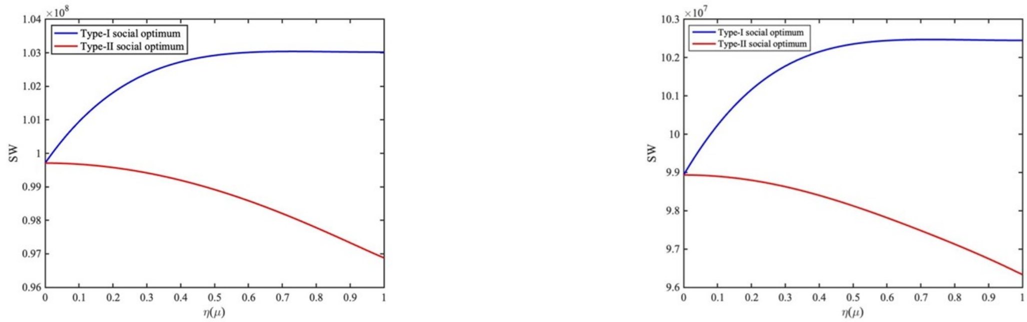
$$f_{1a} = k_{1a}\frac{r_1}{2}\left(1 + \cos\frac{\pi r}{r_1}\right) \tag{37}$$

$$f_{2a} = k_{2a}\frac{r_2}{2}\left(1 + \cos\frac{\pi r}{r_2}\right) \tag{38}$$

As shown in Fig. 9, the results of trigonometric function and linear function are consistent, and their SW* were 9.8574×10^{11} *and* 9.7296×10^{11}, respectively, demonstrating that the linear congestion function is rational. According to Fig. 11a, social optimum is achieved at $\eta = 0.78$, $\mu = 0$, demonstrating that if the congestion level of Shanghai remains unchanged, social optimum shall encourage some residents in Jiaxing people to commute to Shanghai. Therefore, under the condition of exogenous population, the systems of Jiaxing and Shanghai still has much room to improve.

4.4.2. Social optimum under population endogenesis

Without congestion taken into consideration, maximum population distribution and social welfare of Shanghai and Jiaxing obtained are: N_1, N_2 and N_{21} are 18.50, 3.38 and 3.38 million, respectively. SW^* =



(a) linear function

(b) trigonometric function

Figure 9: Trend of social welfare as a function of n() in Type-I and Type- II social optimum with traffic congestion taken into consideration

$1.4444 \times 10^{12}\ yuan$

In this case, social optimum can only be achieved if everyone is a resident in Shanghai, which is similar to the population exogenesis case.

Optimized population distribution and social welfare under different scenarios are investigated. Table 5 shows optimized population distributions in different situations.

Scenario		N_1	N_2	N_{21}
Current population		19.09 million	2.33 million	12400
Without congestion	Exogenous population	19.09 million	2.33 million	2.33 million
	Endogenous population	18.05 million	3.38 million	3.38 million
With congestion	Exogenous population	19.09 million	2.33 million	75431
	Endogenous population	16.98 million	4.44 million	3997087

Table 5: Optimized population distribution of the Shanghai-Jiaxing two-city system

Investigation of optimized population distribution and social welfare under different scenarios can provide references to policies by government. As observed, the results obtained with congestion taken into consideration are more consistent with the practical situation compared to those obtained without congestion taken into consideration. Specifically, exogenous population requires 322 times increase of the intercity commuting population compared with the case of current population, indicating that the positive externalities brought by population agglomeration remain significant for Shanghai if the residential population remains unchanged. Meanwhile, population agglomeration induces large congestion cost. As a result, social welfare under population endogenesis can only be enhanced by reducing the population of the metropolis and increasing the population of the satellite city and the intercity commuting population. Additionally, population distributions in social optimum under current and exogenesis populations are significantly different from that in ideal cases.

Table 6 summarizes optimized social welfare under different scenarios compared with the current situation. This reflects the effects of policies implemented.

<table>
<tr><th colspan="2">Scenario</th><th>Social optimum</th><th>Increase (compared with current situation)</th><th>Social welfare loss induced by congestion</th></tr>
<tr><td rowspan="3">Without congestion</td><td>Current Population</td><td>1.1621×10^{12}</td><td></td><td></td></tr>
<tr><td>Exogenous Population</td><td>1.1928×10^{12}</td><td>2.64%</td><td></td></tr>
<tr><td>Endogenous Population</td><td>1.1962×10^{12}</td><td>2.93%</td><td></td></tr>
<tr><td rowspan="3">With congestion</td><td>Current Population</td><td>1.0302×10^{12}</td><td></td><td>11.35%</td></tr>
<tr><td>Exogenous Population</td><td>1.0304×10^{12}</td><td>0.02%</td><td>13.61%</td></tr>
<tr><td>Endogenous Population</td><td>1.0316×10^{12}</td><td>0.14%</td><td>13.76%</td></tr>
</table>

Figure 10: Optimized social welfare of the Shanghai-Jiaxing two-city system

The results reveal that improvements of social welfare with congestion taken into consideration are significant (up to 1.22%) although population distribution and intercity commuting population in this case remain significantly different from those in ideal cases, suggesting that the Shanghai-Jiaxing two-city system is far from optimized social welfares. Additionally, congestion causes severe loss of social welfare (about 32%) practically, indicating that the commuting model in urban areas shall consider the influences by congestion.

4.4.3. impacts of regional congestion on residents

The influences of congestion on residents are investigated in this section. Geographic division and intercity commuting area division of City 2 can be obtained by determining the specific function patterns of $\bar{\theta}_2(r)$ and $\tilde{\theta}_2(r)$, as shown in Fig. 10.

$$\bar{\theta}_2(r) = \frac{2t_2 + k_2(2r_2 - r)}{t_2 + k_2(r_2 - r)} \tag{39}$$

$$\tilde{\theta}_2(r) = \frac{2t_2(r - \eta r_2) + k_2(2r_2 r - r^2 - 2\eta r_2^2 + \eta^2 r_2^2)}{(t_2 + k_2(r_2 - r))\, r} \tag{40}$$

As shown in Fig. 10, the blue area refers to the intercity commuting area of Jiaxing, the grey and white areas refer to Area L and S, respectively. Herein, $\tilde{\theta}_2 < \bar{\theta}_2$,$\tilde{\theta}_2^{max} < \bar{\theta}_2^{max} < \pi$, which belongs to the first case in Table 2.

4.5. Sensitivity

The effects of parameters such as congestion on SW^* (optimized social welfare) and social welfare per capita (SW per capita) are discussed in this section. Sensitivity analysis aims to explore the influences of

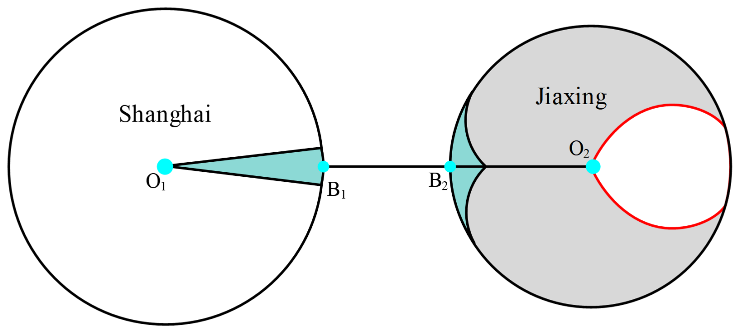


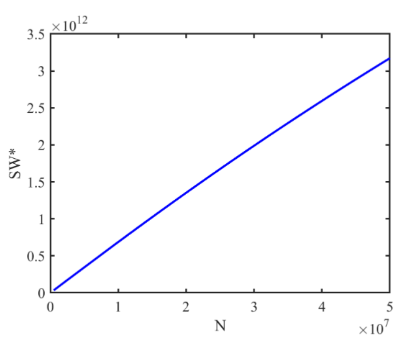
Figure 11: Social optimum without congestion taken into consideration

parameters such as congestion on social optimum are investigated. Eight parameters and their corresponding ranges (see Table 7) are selected and the control variable method is employed for analysis.

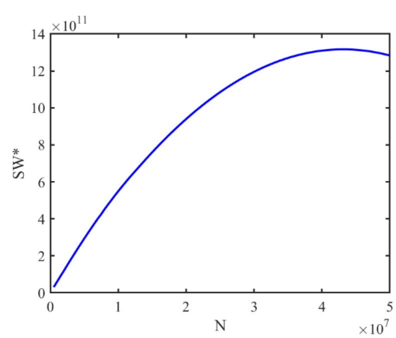
Parameter	Definition	Amplitude
N	Population scale of the system	$[0,\ 50\ million]$
λ	Production function parameters	$[\frac{\gamma}{2}, 2\gamma]$
tc	Time value of intercity commuting	$[\frac{tc}{2}, 2tc]$
Φ	Scope of influence of intercity commuters	$[\frac{\phi}{2}, 2\phi]$
R_i	agricultural rent of City	$[\frac{R_i}{2}, 2R_i]$
m_{1i}	Congestion-induced delay index of City	$[\frac{m_{1i}}{2}, 2m_{1i}]$

Table 6: Parameters selected

4.5.1. *Effects of population scale on social optimum*



(a) Results obtained without congestion taken into consideration

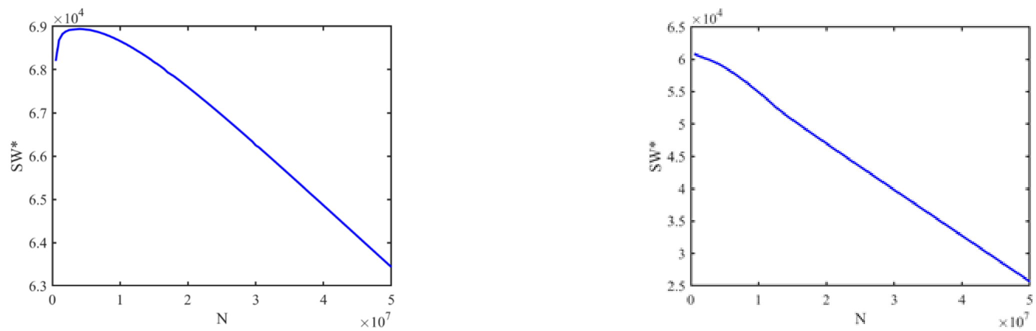


(b) Results obtained with congestion taken into consideration

Figure 12

The effects of the population scale of the two-city system on social optimum (under population endogenesis) are discussed in this section. As shown in Figs. 12 (a), the optimized overall social welfare increases with the total population without considering congestion, suggesting that increased investment of labor force leads to enhanced production capacity and increased social fortune. With congestion taken into consideration (Fig. 11b), social welfare increases gradually with the population before 43 million and decreases after 43 million, indicating that loss induced by congestion exceeds gain in social welfare induced by population agglomeration once the population exceeds 43 million. Therefore, the population of a city shall be controlled at a rational level (within 40 million). Additionally, the effects of the population scale of the system on SW per capita are investigated. Figs. 12(a) and 12(b) shows the results without and with congestion taken into consideration.

4.5.2. Effects of population scale on social optimum



(a) SW per capita vs. population scale without congestion

(b) SW per capita vs. population scale with congestion

Figure 13

The optimized social welfares per capita decreased as the total population increased as the commuting costs caused by congestion and increased city size counteract the increase in personal income induced by clustering effects despite that SW^* increased as the total population increased. With congestion taken into consideration, The social welfare per capital is decreasing along with the increasing total population in this system. Also, the decreasing rate of optimized social welfares per capita with increasing population would be significantly higher than that without congestion taken into consideration. Hence, optimization of overall social welfare shall be combined with rational personal welfare as the cost of long-distance commuting and the additional costs induced by congestion have a significant impact on personal welfare. Instead of pursuing increases in social welfares at the cost of reduced personal welfare, appropriate allocation of residents in an urban agglomeration system and improvement of social welfare per capita by relieving congestion shall be achieved.

Macroscopically, despite the increases of overall social welfare induced by increasing population, traffic congestion has a negative influence on the improvement of social welfare, resulting in limited improvement. Microscopically, personal welfare is sensitive to population scale; increasing population leads to rapidly increasing SW^* per capita when the population scale is small. With traffic congestion taken into consideration, SW^* per capita starts to decrease in a short time as the total population increases further and the correlation is linearly negative. The results demonstrate that over-large population scale may lead to severe loss of personal welfare although the overall social welfare is enhanced. Additionally, increasing total population has negligible effects on social welfare under congestion if the population scale is large. In order to significantly enhance both social and personal welfare, government shall pay great attention to congestion management.

4.5.3. Effects of other parameters on social optimum results

Table 8 shows the effects of other parameters on commuting population and social welfare results in social optimum with congestion taken into consideration. The sensitivity analysis revealed that tc, R_i and m_{1a} significantly affected N_{21}^{*}; R_i m_{1a} and m_{2a} affected SW^{*}. Herein, the influence level follows the sequence of R_1 ¿ m_{1a} ¿ tc, demonstrating that the housing price and the congestion are the main factors hindering improvement of social welfare.

Since agricultural rent R_i does not fluctuate in short term, government shall pay great attention to congestion management, especially in metropolis. Meanwhile, increasing social welfare basically corresponds to increasing intercity commuting population, suggesting that the infrastructure of intercity traffic shall also be enhanced to provide sufficient supply of intercity traffic. The time for intercity commuting is a key factor influencing the selection of occupation in Shanghai by intercity commuters from Jiaxing to Shanghai. Especially, few residents in Jiaxing would commute to Shanghai when the time for daily intercity commuting exceeds 0.87 h, indicating that the HSR would induce intercity commuting and facilitate intercity communications by improving the social welfare of intercity commuters. In terms of the angle of population detour (Φ), intercity commuters can avoid severely congested areas by bypassing; a complete road network shall be constructed near the HSR station to provide more paths for intercity commuters so that they can travel to the CBD with no significant congestion. Meanwhile, the housing price in Jiaxing has significant impacts on the selection by commuters. For instance, all residents in Jiaxing shall commute to Shanghai for optimized social welfare if the housing price in Jiaxing exceeds 41,000 yuan/m2. Indeed, the housing price has significant impacts on intercity commuting and selection of occupation and residence: more residents choose to live in Jiaxing as the housing price in Shanghai increase, while more residents choose to live and work in Shanghai as the housing price in Jiaxing increases. Additionally, increasing housing prices in both cities, especially Shanghai, would lead to degraded social welfare. For this reason, policies shall be released to maintain the housing prices in a rational ranges. Excessive intercity commuting has negative impacts on transportation and the happiness of residents. Specifically, the congestion in Shanghai has severe impacts on the population of intercity commuters in Jiaxing. When the congestion delay index of Shanghai increases from 1.6 to 2.4, the population of intercity commuters in Jiaxing decreases by 50%. Congestions in both Shanghai and Jiaxing would have significant impacts on the two-city system of social welfare and a 25% reduction in social welfare can be induced as the congestion delay index of Shanghai increases from 1.6 to 2.4. Therefore, congestion in Shanghai shall be relieved and the congestion delay index shall be kept at a low lever so that the overall social welfare of the two-city system is maintained at a good level.

5. Conclusions and outlook

5.1. Main conclusions

The congestion effect is introduced to accurately describe the social optimum of two-city system. An optimization model was established to clarify most rational residential and commuting populations of the two-city system under both location conditions and population migration. Additionally, an empirical study of the Shanghai-Jiaxing two-city system was conducted. Social optimum results obtained under population exogenesis and endogenesis were investigated, the Shanghai-Jiaxing two-city system was evaluated, measures by government to enhance social welfare of two-city system were discussed. This study provides references

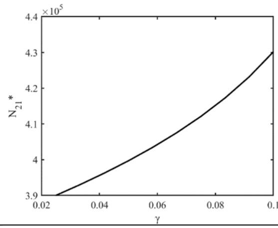
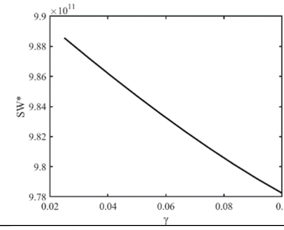
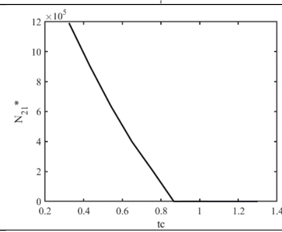
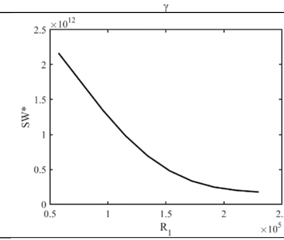
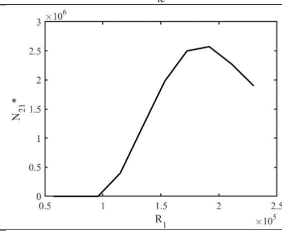
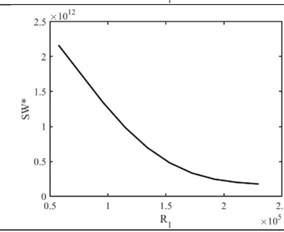
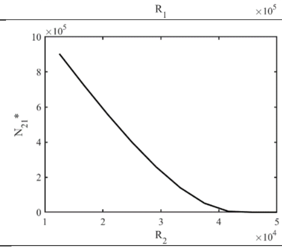
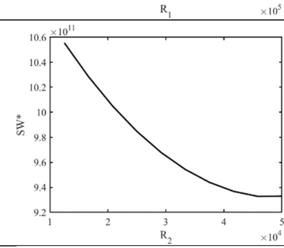
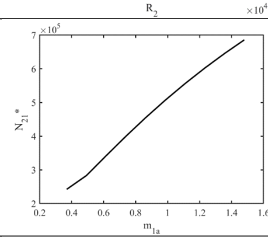
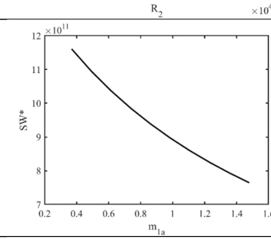
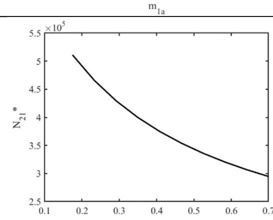
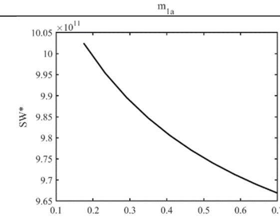
Variable \ Parameter	N_{21}^{*}	SW^{*}
γ		
tc		
R_1		
R_2		
m_{1a}		
m_{2a}		

Table 7: Sensitivity analysis

and suggestions to social development and population issues based on two-city model considering congestion. The main conclusions are as follows:

1. Excessive population agglomeration leads to high congestion cost in the city, thereby severe loss of

social welfare. Specifically, traffic congestion is a significant negative externality.

2. Intercity commuting can enhance social welfare via re-allocation of labor force. Nevertheless, cost induced by traffic congestion is the main factor hindering further improvement of social welfare if the population scale is large. Additionally, severe time delay of individuals may be caused.

3. The population scale shall be controlled according to the objectives of traffic management. Nevertheless, attractiveness of metropolis to satellite cities would be enhanced as a result of congestion management. Therefore, the infrastructure of intercity traffic shall be further improved and large volume transportation shall be encouraged.

5.2. Outlook

First, it is assumed in this study that all residents in the system are homogeneous (e.g., consistent housing consumption and time value). Future studies may assume that residents in different cities have different time values. Then, it is assumed in this study that the urban system has one single industry only and residents have one single consumption only (housing consumption). Future studies may involve multiple industries in urban system and residents have multiple consumptions to establish an improved social optimum model. Additionally, social optimum of two-city system was investigated in this study. Future studies may extend social optimum to urban agglomeration based on two-city systems.

References Afrin, T. and N. Yodo (2020). "A survey of road traffic congestion measures towards a sustainable and resilient transportation system." Sustainability 12 (11): 4660.

Akbar, Prottoy, Victor Couture, Gilles Duranton, and Adam Storeygard. 2023. "Mobility and Congestion in Urban India." American Economic Review, 113 (4): 1083-1111.

Albouy, D., Behrens, K., Robert-Nicoud, F., Seegert, N. (2019). The optimal distribution of population across cities. Journal of Urban Economics, 110, 102-113.

Alonso, W. (1964). Location and Land Use: Toward a General. East-West Center Press.

Anas, A., Pines, D. (2013). Public goods and congestion in a system of cities: how do fiscal and zoning policies improve efficiency?. Journal of Economic Geography, 13(4), 649-676.

Arnott R. Congestion tolling with agglomeration externalities[J]. Journal of Urban Economics, 2007, 62(2): 187-203. Beagle E, Belmont E. Comparative life cycle assessment of biomass utilization for electricity generation in the European Union and the United States[J]. Energy Policy, 2019, 128: 267-275.

Borck, R., Wrede, M. (2009). Subsidies for intracity and intercity commuting. Journal of Urban Economics, 66(1), 25-32.

Brinkman J C. Congestion, agglomeration, and the structure of cities[J]. Journal of Urban Economics, 2016, 94: 13-31.

Brueckner, J. K. (1987). "The structure of urban equilibria: A unified treatment of the Muth-Mills model." Handbook of regional and urban economics 2 (20): 821-845.

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

Dong, T., Jia, N., Ma, S., Xu, S. X., Ong, G. P., Liu, P., Huang, H. J. (2022). Impacts of intercity commuting on travel characteristics and urban performances in a two-city system. Transportation research part E: logistics and transportation review, 164, 102792.

Franco, S. F. (2017). "Downtown parking supply, work-trip mode choice and urban spatial structure." Transportation Research Part B: Methodological 101: 107-122.

Fujita, M., Ogawa, H. (1982). Multiple equilibria and structural transition of non-monocentric urban configurations. Regional science and urban economics, 12(2), 161-196.

Gao, Ya, and Haixiao Pan. ”Understanding the Role of High-Speed Rail on Intercity Commuting: Evidence from the Shanghai Metropolitan Area.” In International Workshop on HSR Socioeconomic Impacts, pp. 205-224. Cham: Springer Nature Switzerland, 2023.

Gibson M, Carnovale M. The effects of road pricing on driver behavior and air pollution[J]. Journal of Urban Economics, 2015, 89: 62-73.

Global Business Journalism (2023). ”Why Beijing traffic is so congested and what can be done about it.” Global Business Journalism. Accessed November 2024.

Gu Y, Deakin E, Long Y. The effects of driving restrictions on travel behavior evidence from Beijing[J]. Journal of Urban Economics, 2017, 102: 106-122.

Hall J D. Can tolling help everyone? estimating the aggregate and distributional consequences of congestion pricing[J]. Journal of the European Economic Association, 2021, 19(1): 441-474.

Henderson, J. V. (1974). The Sizes and Types of Cities. American Economic Review, 64(4), 640-656.

Hensher, D. A. (2018). ”Tackling road congestion–What might it look like in the future under a collaborative and connected mobility model?” Transport Policy 66: A1-A8.

Hu, Y., Deng, T., Zhang, J. (2020). Can commuting facilitation relieve spatial misallocation of labor?. Habitat International, 106, 102136.

Ke, J. and X. M. Chen, et al. (2022). ”Coordinating supply and demand in ride-sourcing markets with pre-assigned pooling service and traffic congestion externality.” Transportation Research Part E: Logistics and Transportation Review 166: 102887.

Li, Z. and Y. Chen, et al. (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 (6): 927-937.

Li, Z. C., Lam, W. H., Wong, S. C., Sumalee, A. (2012). Design of a rail transit line for profit maximization in a linear transportation corridor. Transportation Research Part E: Logistics and Transportation Review, 48(1), 50-70.

Li, Z. C., Lam, W. H., Wong, S. C. (2012). Modeling intermodal equilibrium for bimodal transportation system design problems in a linear monocentric city. Transportation Research Part B: Methodological, 46(1), 30-49

Li, Z. C., Ma, J. C. (2021). Investing in inter-city and/or intra-city rail lines? A general equilibrium analysis for a two-city system. Transport Policy, 108, 59-82.

Niu, X. Annual Report on Cross City Commuting in Yangtze River Delta in 2023. 2024.

Liu, Y., Zhang, X., Pan, X., Ma, X., Tang, M. (2020). The spatial integration and coordinated industrial development of urban agglomerations in the Yangtze River Economic Belt, China. Cities, 104, 102801.

Mandhare, P. A. and V. Kharat, et al. (2018). ”Intelligent road traffic control system for traffic congestion: a perspective.” International Journal of Computer Sciences and Engineering 6 (07): 2018.

McConnell, V. D., Straszheim, M. (1982). Auto pollution and congestion in an urban model: An analysis of alternative strategies. Journal of Urban Economics, 11(1), 11-31.

McDonald, J. F. (2009). ”Calibration of a monocentric city model with mixed land use and congestion.” Regional Science and Urban Economics 39 (1): 90-96.

Mills, E. S. (1967). ”An Aggregate Model of Resource Allocation in a Metropolitan Area.” American Economic Review 57 (2): 197-210.

Murata, Y., Thisse, J. F. (2005). A simple model of economic geography à la Helpman–Tabuchi. Journal of Urban Economics, 58(1), 137-155.

Quinet, E. (2004). ”A meta-analysis of Western European external costs estimates.” Transportation Research Part D Transport Environment 9 (6): 465-476.

Sorek, G. (2009). Migration costs, commuting costs and intercity population sorting. Regional Science and Urban Economics, 39(4), 377-385.

Strotz, R. H. (2016). Urban transportation parables. In The public economy of urban communities (pp. 127-169). Routledge.

Tabuchi, T., Thisse, J. F., Zeng, D. Z. (2005). On the number and size of cities. Journal of Economic Geography, 5(4), 423-448.

Vandyck, T., Proost, S. (2012). Inefficiencies in regional commuting policy. Papers in Regional Science, 91(3), 659-689.

Xu, S. X., Liu, R., Liu, T. L., Huang, H. J. (2018). Pareto-improving policies for an idealized two-zone city served by two congestible modes. Transportation Research Part B: Methodological, 117, 876-891.

Xu, S. X., Liu, T. L., Huang, H. J., Liu, R. (2018). Mode choice and railway subsidy in a congested monocentric city with endogenous population distribution. Transportation Research Part A: Policy and Practice, 116, 413-433.

Yang, X. Q., Huang, H. J. (2022). Effects of HSR station location on urban spatial structure: A spatial equilibrium analysis for a two-city system. Transportation Research Part E: Logistics and Transportation Review, 166, 102888.

Zhao, P. and H. Hu (2019). ”Geographical patterns of traffic congestion in growing megacities: Big data analytics from Beijing.” Cities 92: 164-174.

Appendix A. Proof of Theorem 1

Contradiction method. If some residents commute from City 1 to City 2 in social optimum, define the optimized population distribution as $(N_1^*, N_2^*,\ N_{12}^*)$, N_{12}^* refers to the total quantity of commuting from City 1 to City 2. Based on that, a new population distribution (N_1', N_2') is established under the premise of population endogenesis. Meanwhile, $N_1' = N_1^* - N_{12}^*$, $N_2' = N_2^* + N_{12}^*$ and no intercity commuting is observed. Apparently, these two population distributions have no influences on the overall output and commuting cost of the two-city system, while they lead to reduced aggregate agricultural rent $(N_{12}^*(R_1 - R_2))$ as $R_1 > R_2$, demonstrating that new population distributions lead to enhanced social welfare. In other words, the original population distribution is not optimal. This conflicts with the hypothesis.

Appendix B. Proof of Conclusion 1

According to the mean value theorem of integrals, $0 < \xi < r$ so that the following equation holds:

$$\bar{\theta}_2(r) = \frac{2\int_0^r h_2 dr}{h_2 r} = \frac{2h_2(\xi)r}{h_2 r} \tag{B.1}$$

Since h_2 decreases monotonically with the distance (i.e., $\frac{\partial h_2}{\partial r} < 0$), then

$$\bar{\theta}_2(r) > \frac{2h_2(r)r}{h_2 r} = 2 \tag{B.2}$$

Since $\bar{\theta}_2$ has no definition at r=0, the ultimate value of $\bar{\theta}_2$ at r=0 $(lim_{r\to 0}\frac{\partial h_2}{\partial r} > -\infty)$ is investigated:

$$\lim_{r\to 0}\bar{\theta}_2(r) = \lim_{r\to 0}\frac{2\int_0^r h_2 dr}{h_2 r} = 2\lim_{r\to 0}\frac{h_2}{\frac{\partial h_2}{\partial r}r + h_2} = 2 \tag{B.3}$$

According to Eq (B.3), it is defined that $\bar{\theta}_2\,(0) = 2$ so that $\bar{\theta}_2^{min} = 2$.

Appendix C. Proof of Inference 1

According to Conclusion 1, the minimum angle of boundary curves of Areas L and S satisfies $\bar{\theta}_2^{min} = 2 < \pi$. Hence, Area S is definitely present in the city.

Social optima can be divided as Type-I and Type-II social optimum according to Inference 1. In Type-I social optimum, all intercity commuters are from Area L in City 2; in Type-II social optimum, intercity commuters are from Areas L and S in City 2.

Then, identities of residents (i.e., spatial distribution of intercity commuters) in City 2 are determined based on Theorem 2 to determine expressions of ACC_1 and ACC_2. The distribution areas of intercity commuters should be continuous and symmetrical about B_2O_2, while residents outside this area work locally.

Appendix D. Proof of Theorem 2

Contradiction method. As shown in Fig. 5, Point E is defined as any point different from Point A on Curve AC_1, a_1 and e_1 denote the practical commuting costs of A and E, respectively, while a_2 and e_2 denote the opportunity commuting costs of A and E, respectively. ACC_2^* denotes aggregate commuting cost in City 2 in social optimum. Apparently, ACC_2^* is the minimized in social optimum. Assuming that $a_1 - a_2 \neq e_1 - e_2$ if the commuting cost of City 2 reaches ACC_2^*, it can be deduced that $a_1 - a_2 > e_1 - e_2$. Then, Points A' and E' are identified near Points A and E so that practical and opportunity commuting costs of Point A' are $a_1 - \Delta$ and $a_2 + \Delta$, respectively, while practical and opportunity commuting costs of Point E' are $e_1 + \Delta$ and $e_2 - \Delta$, respectively. Herein, $\Delta > 0$ refers to minimal increment. Owing to the continuity of the Area, residents of Point A' and E' are intercity and intracity commuters, respectively. If identities of residents of Point A' and E' are switched (the overall quantity of intercity commuters remains constant, while the aggregate commuting cost of City 2 increases by $e_1 - e_2 - (a_1 - a_2) + 4\Delta)$, $\Delta^* > 0$ must be possible so that $e_1 - e_2 - (a_1 - a_2) + 4\Delta^* < 0$ as $\lim \Delta \to 0(e_1 - e_2 - (a_1 - a_2) + 4\Delta) = e_1 - e_2 - (a_1 - a_2) < 0$. Therefore, $ACC_2^*$is not the optimal solution, which is conflict with the hypothesis.

Appendix E. Proof of Theorem 3

The commuting costs of working in this city and the other city are defined as opportunity and practical commuting costs, respectively. Assuming that, some residents in Area S commute to City 1 for work in social optimum when $\eta > 0$, residents at Point Q (in Area S), whose distance from the downtown is r, are intercity commuters and $r < \eta r_2$. As a result, there must be Point Q', whose distance from the downtown is r, on the connection of HSR station and downtown in Area L. Residents at Point Q' are intracity commuters as it is on Segment OA. Apparently, opportunity commuting costs of Points Q and Q' are equivalent, while the practical commuting cost of Point Q' is lower than that of Point Q. Hence, social welfare can be enhanced by switching identities of residents at Points Q and Q', which conflicts to the hypothesis.

Appendix F. Proof of Conclusion 3

Proof: If $\widetilde{\theta}_2^{max} > \pi$, $\widetilde{r_2}$ that is strictly less than $r_2$can be found so that $\widetilde{\theta}_2$ is maximized at $\widetilde{r_2}$. Hence,

$$2\left(\frac{1}{r} - \frac{\left(h_2 + \frac{\partial h_2}{\partial r}r\right)\int_{\eta r_2}^{r_2} h_2 dr}{h_2^2 r^2}\right) < \frac{\partial \tilde{\theta}_2}{\partial r} = 2\left(\frac{1}{r} - \frac{\left(h_2 + \frac{\partial h_2}{\partial r}r\right)\int_{\eta r_2}^{r} h_2 dr}{h_2^2 r^2}\right) < \frac{2}{r}$$

Therefore,

$$\lim_{\eta \to 1} \frac{\partial \tilde{\theta}_2}{\partial r} = \frac{2}{r} > 0$$

When$\eta \to 1$, $\widetilde{\theta}_2$increases monotonically with r. In this case, $\widetilde{r_2}$ that is strictly less than r_2 cannot be found so that $\widetilde{\theta}_2$ is maximized at$\widetilde{r_2}$.

Appendix G. Proof of Theorem 4

Assuming that simultaneous commuting between metropolis and satellite city is present in social optimum, there must be Point A near the HSR station of metropolis so that residents at this point commute from metropolis to satellite city; likewise, there must be Point E near the HSR station of satellite city so that residents at this point commute from satellite city to metropolis. a_1 and e_1 refer to practical commuting costs of Points A and E, respectively; a_2 and e_2 refer to opportunity commuting costs of Points A and E; $a_1 < a_2$, $e_1 < e_2$. In this case,$a_1 > e_2$, $a_2 < e_1$. Since it is assumed that $a_1 < a_2$, then $e_1 > e_2$. This is contradictory to the hypothesis that $e_1 < e_2$.

Symbol	Description	Unit
A_i	Total factor productivity of City i	
P_i	Total output of City i	Yuan
W_i	Number of employees working in City i	
R_i	Agricultural rent at the boundary of City i	Yuan
N	The definite population of the system	
N_i	Population of City i	
N_{ji}	Number of intercity commuters traveling from City j to City i	
N_1^*	Population of City 1 at social optimum	
N_2^*	Population of City 2 at social optimum	
N_{12}^*	Number of employees traveling from City 1 to City 2 at social optimum	
T	Intercity commuting cost	Yuan
h_i	Commuting cost per capita per unit distance in City i	Yuan
h_1^*	Commuting cost per capita per unit distance of intercity commuters from City 2 at the boundary of City 1	Yuan
t_i	Commuting cost per capita per unit distance in City i under free flow	Yuan/capita·km
f_i	Congestion cost per capita per unit distance induced by agglomeration in City i	Yuan/capita·km
f_{ji}	Congestion cost of City i induced by residents in City j	Yuan/capita·km
Φ	Range of additional congestion cost induced by intercity commuters	
φ	The angle between the connection of the location of an intercity commuter and the downtown of his/her working city and the connection of the HSR station and the downtown of the working city	
SW	System social welfare	Yuan
ACC_1	Commuting cost induced by intra- and intercity commuters in City 1	Yuan
ACC_2	Commuting cost induced by intra- and intercity commuters in City 2	Yuan
r_i	Radius of City i	km
r	Distance from the resident location to the downtown of the city	km
$\tilde{r}_2$	Maximum distance from the boundary of intercity commuters to the downtown of City 2	km
θ	The angle between the connection of the resident's location and the downtown of the city and the connection of the HSR station and the downtown of the city	
$\bar{\theta}_i$	The angle between the connection of the location of an intercity commuter of City i and the downtown of City i and the connection of the HSR station of City i and the downtown of City i when the costs of the intercity commuter of City i traveling to the HSR station via "circumferential + radial" and "radial + radial"	
$\tilde{\theta}_i$	Practical boundary of the intercity commuter of City i (The net income of working in one city is equivalent to that of working in the other city)	
η, μ	Attractiveness of City 1 to City 2 (η or μ substitutes N_{21} as the decision variable)	
δ	0-1 variable	
k_{ji}	Congestion constant of residents in City j to City i	
m_{ji}	Times of congestion cost of City i induced by residents of City j under current conditions of congestion cost of City i under free flow	
C_{ij}	congestion coefficient of residents in City i commuting to City j	
C_{21}^*	congestion coefficient of intercity commuters from City 2 in City 1	
ACC_{ij}	Commuting cost of residents in City i to City j under congestion conditions	Yuan
$FACC_{ij}$	Commuting cost of residents in City i to City j under free flow	Yuan

Figure 14: List of Key Notations