THEMA

théorie économique, modélisation et applications

THEMA Working Paper n°2024-11
CY Cergy Paris Université, France

Spatial heterogeneity in vehicle license plate lottery rationing

Zhi-Chun Li, Wen-Jing Liu, André de Palma

October 2024

CERGY PARIS UNIVERSITÉ

cnrs

Spatial heterogeneity in vehicle license plate lottery rationing

Zhi-Chun Li[a,*], Wen-Jing Liu[a], André de Palma[b]

[a] *School of Management, Huazhong University of Science and Technology, Wuhan 430074, China*

[b] *CY Cergy-Paris University, Thema, France*

Abstract

This paper studies vehicle license plate lottery rationing problem with spatial heterogeneity by using a two-area modeling framework. The equilibria of residential location and vehicle ownership choices of heterogeneous households with different values of time (VOTs) are first analyzed. A social welfare maximization model is then proposed for determining the optimal number of vehicle quotas and the optimal quota proportion allocated to central and suburban areas of the city. Two alternative lottery schemes, a city-based and an area-based scheme, are explored and compared. The results show significant spatial disparities in the behavior of central and suburban residents in the license plate lottery. When the transit service level in the suburban area is relatively low, the critical VOT for participating in the lottery for suburban residents is lower than that for central residents. The area-based lottery scheme welfare-dominates the city-based lottery scheme. Compared to laissez faire, the implementation of lottery rationing schemes may cause urban sprawl or shrink, depending on the scheme adopted (area-based or city-based) and the road service level.

Keywords: License plate rationing; lottery; city-based vs. area-based; urban spatial structure; spatial heterogeneity; heterogeneous households.

JEL classification: R13, R14, R38, R41

---

[*]Corresponding author. Tel.: +86-27-8755-6490; Fax: +86-27-8755-6437.

E-mail addresses: smzcli@hust.edu.cn (Z.-C. Li), wenjing.liuro@gmail.com (W.-J. Liu), andre.de-palma@cyu.fr (A. de Palma).

1. Introduction

With rapid development of social economy and continuing promotion of urbanization, the number of motorized vehicles has dramatically been increasing in many large cities around the world. According to the report from the Beijing Municipal Government[1], the total number of registered motorized vehicles in Beijing increased from 2.1 million vehicles in 2003 to 7.6 million vehicles in 2023, meaning an average annual growth rate of 7%. However, over the past 20 years, the average annual growth rate of total road length has been only 2.5%, increasing from 14,453 km to 22,433 km. The imbalance between road supply and travel demand leads to increasingly serious traffic congestion.

In response to the exacerbating traffic congestion, various travel demand management measures have been advocated in some countries and regions, such as road pricing, road space rationing, and vehicle ownership rationing. The typical examples for road pricing include London, Singapore, Stockholm, Oslo, and Hong Kong (Small and Verhoef, 2007; de Palma and Lindsey, 2011). Driving restriction based on license plate numbers, as a road space rationing scheme, has been adopted in many cities in the world, such as Beijing China, Santiago Chile, Mexico City (Gallego et al., 2013; Gu et al., 2017). Different from the road-oriented regulatory schemes (road pricing and road space rationing), the vehicle ownership (or license plate) rationing, as a vehicle-oriented regulatory scheme, manages travel demand and mitigates traffic congestion through restricting the total number of motorized vehicles to be registered (Nie, 2017; Li et al., 2019). Such a scheme has been implemented in some cities, such as Singapore, Beijing, and Shanghai.

In 2011, Beijing introduced a city-based lottery scheme to allocate vehicle license plates, aiming to alleviate traffic congestion and improve air quality. About 240,000 new license plate quotas per year were distributed by lottery during 2011 and 2013, and the annual quota was reduced to about 100,000 after 2018. Consequently, winning the lottery has become increasingly difficult: the odds for obtaining a quota decreased from 10% in January 2011 to 0.6% in June 2024. The winners in the lottery need to purchase an auto with the obtained license plate within one year, and the license plates are non-transferable and non-tradable. In the city-based scheme, all qualified applicants from any areas of the city may participate in the lottery without spatial

[1] https://www.beijing.gov.cn.

difference. Such a city-based lottery scheme aims to ensure an equal opportunity for applicants to obtain a license plate, regardless of where they reside in the city (urban or suburban area). However, in reality, different areas have distinct characteristics, such as population size, residential density, household income, and housing prices, and thus residents' needs towards vehicle ownership change across areas. It is plausible that suburban residents may have a higher demand for private cars than central residents due to a longer travel distance, thus causing a difference in the license plate rationing for different areas. Consequently, an area-based lottery scheme with spatial heterogeneity consideration, in which license plates are allocated by area, may be a good substitute for the city-based lottery scheme that is currently being used in some cities, e.g., Beijing and Hainan, China. On the other hand, the license plate rationing may significantly affect household residential location choice and thus urban spatial structure in the long term. In general, the households with auto may prefer to live in suburban area for enjoying a large house, while the households without auto may prefer to reside in central area for the convenience of work trips. Therefore, it is important to consider the interaction between the households' preferences for residential locations and the license plate rationing scheme. This paper aims to provide a theoretical analysis of the spatial heterogeneity in the license plate lottery rationing for a city with different-income households.

We begin by developing an equilibrium model of households' residential location and auto ownership choices with spatial heterogeneity. Following the modeling framework of Brueckner and Helsley (2011), we consider a monocentric city that is discretized into two areas (or islands), i.e., a central area and a suburban area, connected by bridges. In the proposed model, residents are differentiated by their residential areas and values of time (VOTs). The critical VOTs between residents participating and non-participating in the lottery scheme are identified. Our finding reveals the spatial difference in the behavior of suburban and central residents participating in the auto quota lottery. When the transit sevice level in the suburban area is relatively low, the critical VOT for participating in the lottery for suburban residents is lower than that for central residents. Two alternative lottery schemes, namely a city-based scheme and an area-based scheme, are investigated and compared. The optimal auto quota and optimal quota allocation proportion between the central and suburban areas for each scheme are endogenously determined. We find that the vehicle quota proportion allocated to the suburban area under the area-based lottery scheme may be higher or lower than that under the city-based lottery scheme, depending on the road service level. These are significant extensions to the studies of Li et al. (2019), Yu and Li (2023), and Liu et al. (2024). They treated the whole city as a homogeneous entity and determined the optimal total auto quota on a city basis without considering the urban spatial heterogeneity. We extend their work to explicitly account for the spatial heterogeneity through developing an area-based scheme. We find that the area-based lottery scheme is more socially efficient than the city-based lottery scheme.

Based on the proposed equilibrium model, we explore the effects of the license plate lottery scheme on the urban spatial structure. The residential distribution of households between the central and suburban areas is endogenously determined. Our study extends the two-area urban models to tackle the license plate rationing issues with household relocation. To the best of our knowledge, this paper is the first study to take into account the interactions among the license plate rationing, heterogeneous households' residential distribution, and spatial heterogeneity. We show that high-VOT households tend to reside in the central area and are more likely to participate in the lottery for getting an auto quota; mid-VOT households may choose to reside in the central area and use public transit or reside in the suburban area and participate in the lottery; and low-VOT households prefer to reside in the suburban area and travel by public transit. The implementation of the license plate rationing may lead to city expand or shrink, depending on the rationing scheme adopted (i.e., city-based or area-based) and the road service level.

Regarding the license plate rationing issues, there are some related studies in the literature. They can generally be categorized into two classes according to research methods: econometric or statistical method and optimization method. The econometric or statistical method is usually used to identify key factors affecting the license plate rationing. Sample studies include Phang et al. (1996), Chin and Smith (1997), Chu et al. (2004), Chu (2002, 2012), Zhu et al. (2013), Wang and Zhao (2017), Li (2018), Yang et al. (2022), and Hu et al. (2022). In terms of the optimization method, Koh (2003, 2004) constructed a mathematical model to determine the optimal vehicle quota for Singapore's vehicle license plate system. Nie (2017) presented a theoretical framework to explore the potentials of license plate rationing and its combination with tradable permits. Li et al. (2019) presented optimization models to determine the optimal auto quota under different rationing schemes (lottery, auction, and hybrid of lottery and auction), and compared their efficiencies. Yu and Li (2023) further investigated the effects of household income distribution and implementation sequence of lottery and auction in the hybrid scheme on the optimal design of auto ownership rationing schemes. They found that the residents' income distribution has no significant effects on the model properties. Recently, Liu et al. (2024)

further extended it by developing a multi-period optimization model to determine the optimal quotas of gasoline and electric vehicles over a given time horizon.

Despite of their contributions in the field of license plate rationing, all the above-mentioned studies focused on a city-based rationing scheme, and little attention has been paid to the behavioral difference of heterogeneous households in different areas (i.e., spatial heterogeneity) and its effects on the license plate rationing strategies. Moreover, these previous studies ignored the long-term effects of the license plate rationing on the urban system, including the household residential relocation and the changes in housing market and urban spatial structure.

Building on the two-area modeling framework presented by Brueckner and Helsley (2011), further extended by Brueckner (2014), Kim (2016), and Brueckner and Franco (2018), we develop a two-area urban model to investigate the household residential relocation behavior under the license plate rationing and the long-term effects of the license plate rationing on the heterogeneous household residential distribution and urban spatial structure.

The main contributions of this paper are summarized as follows. First, we formulate the equilibria of household residential location and vehicle ownership choices under no auto purchase restriction (i.e., laissez faire) and the lottery rationing by using a two-area model. In the proposed model, residents are differentiated by their VOTs and residential areas. The critical VOTs for participating and non-participating in the lottery for central and suburban residents are identified. Second, a social welfare maximization model is proposed for determining the optimal auto quota and the optimal quota allocation share between central and suburban areas. Two alternative rationing schemes, namely a city-based lottery scheme and an area-based lottery scheme, are investigated and compared, together with laissez faire. The interactions among the license plate rationing, heterogeneous household relocation behavior, and urban spatial structure are incorporated in the model. Our study provides a useful approach for evaluating and designing the license plate rationing schemes so as to implement efficient travel demand management.

The remainder of this paper is organized as follows. Section 2 describes some basic components of the models and introduces a benchmark case without auto purchase restriction. Section 3 presents the model formulations of the city-based and area-based lottery schemes to incorporate the effects of the lottery rationing on urban spatial structure. In Section 4, a numerical example is provided to illustrate the properties and applications of the proposed models for Beijing. Section 5 concludes this paper and provides some recommendations for further studies.

2. The model

2.1. Basic setup

We adopted a two-area modeling framework, which was used in urban economic models by Brueckner and Helsley (2011), Brueckner (2014), Kim (2016), and Brueckner and Franco (2018). Consider a closed city with a total of N households. The city is monocentric and discretized into two areas: a central area and a suburban area, connected by two bridges, as illustrated in Fig. 1. The central business district (CBD) is located at the left end of the city, where all job opportunities are concentrated. Households choose to reside in the central area or the suburban area. Without loss of generality, the land area of the central area is normalized to unity. The suburban area is composed of developed land and undeveloped open space and the area of the developed land area is endogenously determined by the model.

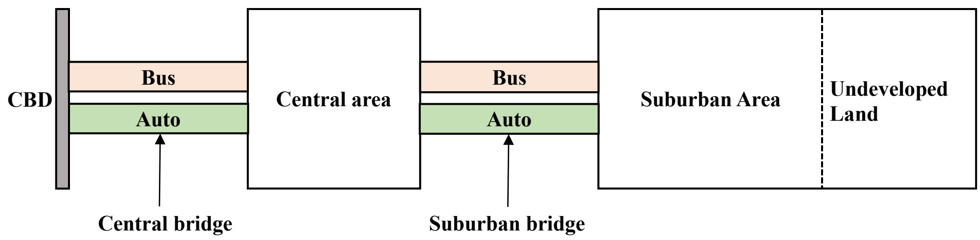


Fig. 1. Regional map of two-area city.

The residents are differentiated in terms of their VOTs and residential areas. We assume that resident's VOT τ follows a distribution over interval $\left[\underline{\theta},\overline{\theta}\right]$ with probability density function $\varphi(\tau)$ and cumulative distribution function $\Phi(\tau)$, in which $\underline{\theta}$ and $\overline{\theta}$ are the lower and upper bounds of the VOT, respectively. Central residents pass through the central bridge to reach their worksite located in the CBD, while suburban residents have to pass through both the central bridge and the suburban bridge to arrive at the workplace. For simplicity, we assume that there is no commuting cost within each area, and the travel cost only takes place on bridges, as assumed in Brueckner's model. Therefore, the suburban residents endure a longer commuting time than the central residents. There are two alternative travel modes for each of central and suburban residents, namely auto and public transit. To ensure that both travel modes are simultaneously used, we assume that auto and public transit are non-dominated both in terms of travel time cost and monetary travel cost. Specifically, auto has a shorter travel time but a more expensive monetary travel cost than transit. In the following, we in turn define the travel costs by mode for central commuters and suburban commuters, respectively. The words "commuter", "resident", and "household" can be used without difference in this paper.

2.1.1. Travel cost of central commuters

Define $G_c^{\text{au}}(\tau)$ as the travel cost of a central commuter with VOT τ by auto. It consists of monetary travel cost and travel time cost for passing through the central bridge, represented as

$$G_c^{\text{au}}(\tau) = f_c + \tau T_c^{\text{au}}, \tag{1}$$

where superscript "au" denotes the auto and subscript "*c*" denotes the central area. f_c is the monetary travel cost by auto for central commuters, which includes fuel cost, parking cost, depreciation cost etc. T_c^{au} is the travel time across the central bridge, which is assumed to be a linear function of the auto traffic volume on that bridge:

$$T_c^{\text{au}} = t_{c,f}^{\text{au}} + t_{c,v}^{\text{au}} Q = t_{c,f}^{\text{au}} + t_{c,v}^{\text{au}} \left(Q_c + Q_s \right), \tag{2}$$

where $t_{c,f}^{\text{au}}$ is the free-flow travel time and $t_{c,v}^{\text{au}}$ is the variable (or marginal) travel time on the central bridge. $t_{c,v}^{\text{au}}$ can serve as an indicator of road service level on the central bridge. A small value of $t_{c,v}^{\text{au}}$ indicates a good road service level, and vice versa. Q is the auto traffic volume on the central bridge, which is the sum of the number of auto users Q_c in the central area and Q_s in the suburban area.

The travel cost $G_c^{\text{tr}}(\tau)$ of a central commuter with VOT τ by public transit in the central area consists of fare and travel time cost, expressed as

$$G_c^{\text{tr}}(\tau) = \beta + \tau T_c^{\text{tr}}, \tag{3}$$

where superscript "tr" denotes the public transit. β is the transit fare. T_c^{tr} is the travel time of the central commuters by transit, including the in-vehicle travel time across the central bridge, walk time to/from transit stations, and waiting time for the coming transit vehicles.

2.1.2. Travel cost of suburban commuters

The suburban commuters have to pass through both the central and suburban bridges to arrive at their worksite in the CBD. The travel cost of a suburban commuter with VOT τ by auto, denoted as $G_s^{\text{au}}(\tau)$, consists of the monetary travel costs and travel time costs on both bridges, expressed as

$$G_s^{\text{au}}(\tau) = f_s + \tau\left(T_c^{\text{au}} + T_s^{\text{au}}\right), \tag{4}$$

where subscript "*s*" denotes the suburban area. f_s is the total monetary travel cost by auto for passing through both the central and suburban bridges. The travel time for the suburban commuters is the sum of the travel time across the central bridge T_c^{au} and across the suburban bridge T_s^{au}. Similar to T_c^{au}, T_s^{au} can be defined as a linear function of the auto traffic volume across the suburban bridge, represented as

$$T_s^{\text{au}} = t_{s,f}^{\text{au}} + t_{s,v}^{\text{au}} Q_s , \tag{5}$$

where $t_{s,f}^{\text{au}}$ and $t_{s,v}^{\text{au}}$ are the fixed and variable components of the travel time on the suburban bridge.

Different from central transit commuters, suburban transit commuters experience an additional travel time cost T_s^{tr} on the suburban bridge, besides on the central bridge. The travel cost, $G_s^{\text{tr}}(\tau)$, of a suburban transit commuters with VOT τ can be defined as

$$G_s^{\text{tr}}(\tau) = \beta + \tau\left(T_c^{\text{tr}} + T_s^{\text{tr}}\right). \tag{6}$$

2.2. Household residential location choice equilibrium

Traditional urban models usually assume that households are homogeneous in terms of their income levels or VOTs. However, in reality income across residents varies significantly, depending on their occupations and skills. This leads to heterogeneity in households' VOTs. In the following, we explore the difference in their residential location choices (i.e., residing in central or suburban area) due to their VOT heterogeneity.

Households obtain utility from housing consumption and non-housing goods (numeraire) consumption. We adopt a quasi-linear form of household utility function, expressed as

$$u_i^j(\tau) = z_i^j(\tau) - \frac{k}{2h_i^j(\tau)}, i = c, s \text{ and } j = \text{au, tr}, \tag{7}$$

where $u_i^j(\tau)$ is the utility of households with VOT τ residing in area *i* and traveling by mode *j*. $z_i^j(\tau)$ is the annual non-housing good consumption with a unitary price. $h_i^j(\tau)$ is the household housing size, measured in square meters of floor space. *k* is the household's preference for land, which is assumed to be a positive constant. A large value of *k* indicates a strong preference for land, and vice versa. The second term on the right-hand side (RHS) of Eq. (7) represents the household's utility derived from residential land consumption, measured in monetary units. Such a hyperbolic utility function has been adopted in some previous residential location choice models, such as Mossay and Picard (2011), Picard and Tabuchi (2013), Blanchet et al. (2016), Akamatsu et al. (2017), and Li et al. (2024b).

Household income is spent on non-housing good consumption, residential land consumption, and commuting cost:

$$w(\tau) = z_i^j(\tau) + p_i h_i^j(\tau) + \gamma G_i^j(\tau), i = c, s \text{ and } j = \text{au, tr}, \tag{8}$$

where $w(\tau)$ is the annual income of the households with VOT τ. In this paper, households' income is assumed to be proportional to their VOTs, as in Becker (1965), Small (2012), and Li et al. (2024b). The annual income of households with VOT τ can thus be calculated as $w(\tau) = \alpha\tau$, where α is the average annual working hours of commuters. p_i is the land rental price in area *i*, and thus $p_i h_i^j(\tau)$ is the annual household residential land rents. γ denotes the average annual number of round trips to the workplace. The last term on the RHS of Eq. (8) represents the annual commuting cost of households with VOT τ residing in area *i* and traveling by mode *j*.

Suppose that each household chooses its residential location, housing size, and the amount of non-housing goods (numeraire) to maximize its own utility subject to income budget constraint. From Eqs. (7) and (8), the utility maximization problem for a household with VOT τ can be expressed as

$$\max_{h_i^j} \ u_i^j(\tau) = w(\tau) - p_i h_i^j(\tau) - \gamma G_i^j(\tau) - \frac{k}{2h_i^j(\tau)}, i = c, s \text{ and } j = \text{au, tr}. \tag{9}$$

From the first-order condition $du_i^j / dh_i^j = 0$, we obtain

$$h_i = \sqrt{\frac{k}{2p_i}}. \tag{10}$$

Eq. (10) indicates that for a given area, residential land consumption per household is independent of the household income and travel mode *j* (and thus superscript "*j*" in h_i^j can be removed), and depends only on the land rental price. This implies that all the households in the same area consume the same amount of residential land.[2]

Substituting Eq. (10) into Eq. (9) yields the household indirect utility as

$$u_i^j(\tau) = \alpha\tau - \sqrt{2kp_i} - \gamma G_i^j(\tau),\ i = c, s \text{ and } j = \text{au, tr}. \tag{11}$$

Proposition 1. As for the land rental prices, housing sizes and residential distributions in the central and suburban areas, we have

(i) Compared to the suburban area, the land rental price in the central area is higher and its housing size is smaller, i.e., $p_c > p_s$ and $h_c < h_s$.

(ii) The high-income households choose to reside in the central area, whereas the low-income households choose to reside in the suburban area.

The proof of Proposition 1 is given in Appendix A. It implies a trade-off between commuting cost and land rental price: residents living in the central area have a lower commuting cost but a higher residential land rent, causing a smaller housing size, while those who reside in the suburban area have a lower residential land rent and thus a large housing size but a higher commuting cost.

At equilibrium, all the population fits inside the city. We represent N_c and N_s as the number of central residents and suburban residents, respectively. The total land consumption in the central area is $N_c h_c$. With the assumption that the central area's land area is unity, $N_c h_c = 1$ holds. Combining it and Eq. (10), housing size per household and the land rental price in the central area can be represented as

$$h_c = \frac{1}{N_c}, \text{ and } p_c = \frac{kN_c^2}{2}. \tag{12}$$

[2] The result in Eq. (10) relies on the quasi-linear utility function because of its inherent property of income-inelastic land consumption.

At the city edge, the land rental price equals the agricultural rent or opportunity cost of land, which is assumed to be a positive constant r_s. We thus have $p_s = r_s$. Housing size per household h_s and total residential land consumption in the suburban area, denoted as b_s, can then be expressed as

$$h_s = \sqrt{\frac{k}{2r_s}} \text{ and } b_s = N_s \sqrt{\frac{k}{2r_s}}. \tag{13}$$

Eq. (13) shows that the total residential land consumption b_s in the suburban area depends on the number of households there. An increase in the number of suburban residents leads to an expansion of suburban area, causing city sprawl.

2.3. Critical VOTs for participants in lottery

In this section, we investigate the behavioral difference of heterogeneous households' participation in the lottery in different residential areas. In the lottery scheme, all participants have an equal probability of obtaining an auto quota, equal to the ratio of auto quotas to be allocated to the number of participants. We represent $\hat{N}_c$ and $\hat{N}_s$ as the number of participants in the lottery in the central and suburban areas, and Q_c and Q_s as the auto quotas to be allocated to these two areas, respectively. Without loss of generality, we assume that the auto quota Q_i allocated to area *i* (where *i* is *c* for central area or *s* for suburban area) satisfies $Q_i \le \hat{N}_i$. The probability of obtaining an auto quota in area *i* is $Q_i / \hat{N}_i$, and thus the probability of failing to obtain an auto quota is $1 - Q_i / \hat{N}_i$. Therefore, the expected travel cost $\hat{G}_i(\tau)$ of a resident participating in the lottery with VOT τ in area *i* can be expressed as

$$\hat{G}_i(\tau) = \frac{Q_i}{\hat{N}_i} G_i^{\text{au}}(\tau) + \left(1 - \frac{Q_i}{\hat{N}_i}\right) G_i^{\text{tr}}(\tau), i = c, s. \tag{14}$$

The first term on the RHS of Eq. (14) represents the expected travel cost by auto if the resident obtains an auto quota, and the second term represents the expected travel cost by transit if the resident fails to acquire an auto quota.

A resident with VOT τ in area *i* would like to participate in the lottery if and only if the expected travel cost of participating in the lottery is lower than that by transit, i.e.,

$\hat{G}_i(\tau) \leq G_i^{\text{tr}}(\tau)$. Otherwise, he/she has no incentive to participate in the lottery. This means that for a given area $i = c, s$, only the residents with $\tau \geq \tau_i$ will participate in the lottery, where τ_i is the critical VOT in area *i*. From $\hat{G}_i(\tau_i) = G_i^{\text{tr}}(\tau_i)$, one can derive the critical VOT between participating and non-participating in the lottery in the central and suburban areas as follows:

$$\tau_c = \frac{f_c - \beta}{T_c^{\text{tr}} - T_c^{\text{au}}}, \text{ and } \tau_s = \frac{f_s - \beta}{T_c^{\text{tr}} + T_s^{\text{tr}} - T_c^{\text{au}} - T_s^{\text{au}}}. \tag{15}$$

As previously stated, transit mode has a lower monetary travel cost but a higher travel time cost than auto, i.e., $f_i > \beta$ and $T_i^{\text{tr}} > T_i^{\text{au}}$, and thus $\tau_i \geq \underline{\theta}$ should hold. We thus have the following property.

Proposition 2. The critical VOT for participating in the lottery for the suburban residents is lower than that for the central residents if and only if $\frac{T_s^{\text{tr}} - T_s^{\text{au}}}{f_s - f_c} \geq \frac{T_c^{\text{tr}} - T_c^{\text{au}}}{f_c - \beta}$ holds.

Proposition 2 reveals the spatial disparities in the behavior of suburban and central residents in the auto quota lottery. The condition is more likely to hold if the transit service level in the suburban area (i.e., T_s^{tr}) is relatively low, and less likely to hold otherwise.

Taking first-order partial derivatives of τ_i with regard to Q_i yields

$$\frac{\partial \tau_c}{\partial Q_c} = \frac{\partial \tau_c}{\partial Q_s} = \frac{t_{c,v}^{\text{au}} \tau_c}{t_c^{\text{tr}} - T_c^{\text{au}}} > 0, \tag{16a}$$

$$\frac{\partial \tau_s}{\partial Q_c} = \frac{t_{c,v}^{\text{au}} \tau_s}{t_c^{\text{tr}} + t_s^{\text{tr}} - T_c^{\text{au}} - T_s^{\text{au}}} > 0, \text{ and } \frac{\partial \tau_s}{\partial Q_s} = \frac{\tau_s \left(t_{c,v}^{\text{au}} + t_{s,v}^{\text{au}} \right)}{t_c^{\text{tr}} + t_s^{\text{tr}} - T_c^{\text{au}} - T_s^{\text{au}}} > 0. \tag{16b}$$

Eq. (16) shows that the critical VOTs τ_i are monotonically increasing with Q_i. This can be explained as follows. An increase in auto quota Q_i leads to an increase in the road traffic congestion level and thus in the auto travel time. As a result, the auto travel cost increases, and thus the critical VOT τ_i of participating in the lottery increases, leading to decreased numbers of participants in the lottery in the central and suburban areas.

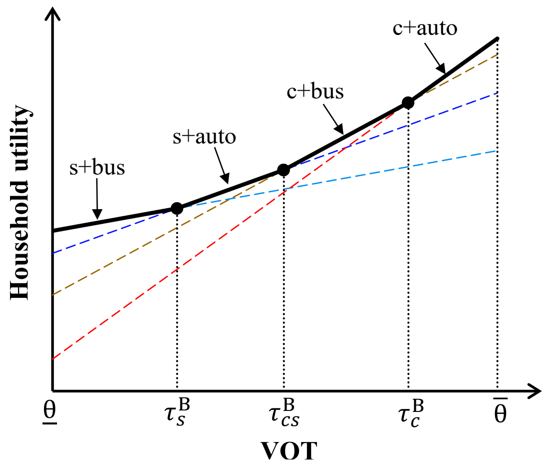
2.4. No auto purchase restriction (benchmark case)

We now consider the benchmark case with no auto purchase restriction policy (i.e., laissez faire), in which the auto purchase market can be freely entered. Households make residential location choice between central and suburban areas, and travel mode choice between auto and public transit to maximize their own utility. There are a total of four alternative location/mode choice combinations: (suburban, transit), (suburban, auto), (central, transit), and (central, auto). The households' decisions on these combinations are based on a trade-off between commuting cost and residential land rent.

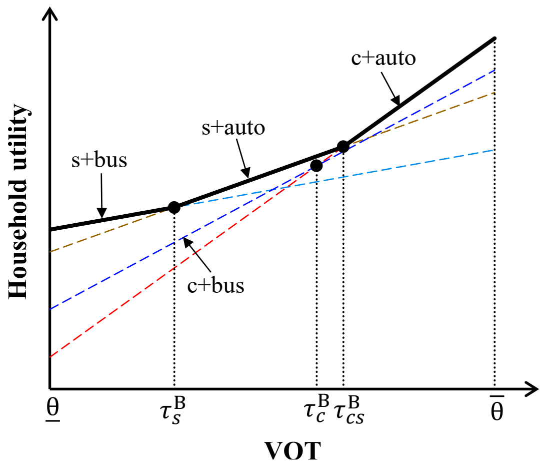
Suppose that under the laissez faire, N_c^{B} and N_s^{B} residents in the central and suburban areas decide to buy an auto, i.e., both are the numbers of auto quotas to be added in the central and suburban areas, respectively. Here, superscript "B" represents the benchmark case. Substituting N_c^{B} and N_s^{B} into Eq. (15) yields

$$\tau_c^{\mathrm{B}} = \frac{f_c - \beta}{T_c^{\mathrm{tr}} - T_c^{\mathrm{au}}(N_c^{\mathrm{B}}, N_s^{\mathrm{B}})}, \text{ and } \tau_s^{\mathrm{B}} = \frac{f_s - \beta}{T_c^{\mathrm{tr}} + T_s^{\mathrm{tr}} - T_c^{\mathrm{au}}(N_c^{\mathrm{B}}, N_s^{\mathrm{B}}) - T_s^{\mathrm{au}}(N_s^{\mathrm{B}})}. \tag{17}$$

Eq. (17) implies that the central (or suburban) residents with a VOT larger than or equal to τ_c^{B} (or τ_s^{B}) choose to purchase an auto. Otherwise, they choose to commute by transit.



(a) $\gamma\left(f_s - f_c + \tau_c^{\mathrm{B}} T_s^{\mathrm{au}}\right) \geq \sqrt{2kp_c} - \sqrt{2kp_s}$



(b) $\gamma\left(f_s - f_c + \tau_c^{\mathrm{B}} T_s^{\mathrm{au}}\right) < \sqrt{2kp_c} - \sqrt{2kp_s}$

Fig. 2. Residential location and travel mode choices of heterogeneous households under laissez faire.

According to Eq. (11), the household utility function is a linear function of VOT τ. Based on the non-dominant assumption of commuting costs between auto and transit, and the relationship between the land rental prices of the central and suburban areas, there are two possible scenarios:

$\gamma\left(f_s - f_c + \tau_c^{\mathrm{B}} T_s^{\mathrm{au}}\right) \geq \sqrt{2kp_c} - \sqrt{2kp_s}$ and $\gamma\left(f_s - f_c + \tau_c^{\mathrm{B}} T_s^{\mathrm{au}}\right) < \sqrt{2kp_c} - \sqrt{2kp_s}$, as shown in Fig. 2. Suppose that households would choose the combination of residential location and travel mode with the highest utility, and thus the utility frontier with heterogeneous households is the upper envelope of all the utility curves, as depicted by the bold line in Fig. 2. It can be seen in Fig. 2 that households with high VOTs ranging between $\max\left\{\tau_{cs}^{\mathrm{B}}, \tau_c^{\mathrm{B}}\right\}$ and $\overline{\theta}$ tend to reside in the central area and commute by auto due to shorter travel time, whereas households with low VOTs ranging between $\underline{\theta}$ and τ_s^{B} prefer to reside in the suburban area and commute by transit due to lower land rental price and low transit fare. However, households with medium VOTs may choose to reside in the central area and travel by transit or to reside in the suburban area and travel by auto, depending on the difference, $\gamma\left(f_s - f_c + \tau_c^{\mathrm{B}} T_s^{\mathrm{au}}\right)$, of the auto commuting costs of the central and suburban commuters, and the difference, $\sqrt{2kp_c} - \sqrt{2kp_s}$, of residential land rental prices between them.

Fig. 2(a) illustrates the scenario in which the additional commuting cost in the suburban area outweighs the benefit generated from its cheaper residential land rent compared to the central area, i.e., $\gamma\left(f_s - f_c + \tau_c^{\mathrm{B}} T_s^{\mathrm{au}}\right) \geq \sqrt{2kp_c} - \sqrt{2kp_s}$. It can be seen in Fig. 2a that the high-income households (i.e., $\tau \geq \tau_{cs}^{\mathrm{B}}$) will choose to reside in the central area, whereas the low-income households (i.e., $\tau < \tau_{cs}^{\mathrm{B}}$) will choose to reside in the suburban area. The critical VOT τ_{cs}^{B} satisfies

$$u_s^{\mathrm{au}}(\tau_{cs}^{\mathrm{B}}) = u_c^{\mathrm{tr}}(\tau_{cs}^{\mathrm{B}}). \tag{18}$$

As the additional commuting cost in the suburban area is lower than the benefit generated from its cheaper residential land rent, i.e., $\gamma\left(f_s - f_c + \tau_c^{\mathrm{B}} T_s^{\mathrm{au}}\right) < \sqrt{2kp_c} - \sqrt{2kp_s}$, all the households with medium VOTs would choose to reside in the suburban area and travel by auto, as shown in Fig. 2(b). For this case, at the equilibrium, the critical VOT τ_{cs}^{B} of choosing to reside in suburban and central areas becomes

$$u_s^{\mathrm{au}}(\tau_{cs}^{\mathrm{B}}) = u_c^{\mathrm{au}}(\tau_{cs}^{\mathrm{B}}). \tag{19}$$

From Eqs. (18) and (19), we can obtain the residential distribution of heterogeneous households between central and suburban areas under the laissez faire. The number of households residing in the central and suburban areas can thus be determined by

$$N_c = N\int_{\tau_{cs}^{\mathrm{B}}}^{\bar{\theta}} \varphi(\tau)\,d\tau = N\left(1-\Phi\left(\tau_{cs}^{\mathrm{B}}\right)\right), \text{ and } N_s = N\int_{\underline{\theta}}^{\tau_{cs}^{\mathrm{B}}} \varphi(\tau)\,d\tau = N\Phi\left(\tau_{cs}^{\mathrm{B}}\right), \tag{20}$$

where $\varphi(\tau)$ and $\Phi(\tau)$ are the probability density function and cumulative distribution function of households' VOTs, respectively.

Under the laissez faire, any resident who want to own an auto can obtain a license plate. At equilibrium, the numbers of auto quotas to be added in the central and suburban areas can, respectively, be given by

$$N_c^{\mathrm{B}} = N\int_{\max\left\{\tau_c,\,\tau_{cs}^{\mathrm{B}}\right\}}^{\bar{\theta}} \varphi(\tau)d\tau = N\left(1-\Phi\left(\max\left\{\tau_c,\tau_{cs}^{\mathrm{B}}\right\}\right)\right), \text{ and } N_s^{\mathrm{B}} = N\int_{\tau_s}^{\tau_{cs}^{\mathrm{B}}} \varphi(\tau)d\tau = N\left(\Phi\left(\tau_{cs}^{\mathrm{B}}\right)-\Phi\left(\tau_s\right)\right). \tag{21}$$

The social welfare can be defined as the total benefits of all parties in the system, which is the sum of the utilities of all households and the profit generated from transit operations. According to Eqs. (20) and (21), the social welfare under the laissez faire is given by

$$\begin{aligned} SW^{\mathrm{B}} = {} & \gamma\left(N - N_c^{\mathrm{B}} - N_s^{\mathrm{B}}\right)\beta + \left(N_s - N_s^{\mathrm{B}}\right)\int_{\underline{\theta}}^{\tau_s^{\mathrm{B}}} \varphi(\tau)u_s^{\mathrm{tr}}\left(\tau, N_c^{\mathrm{B}}, N_s^{\mathrm{B}}\right)d\tau + N_s^{\mathrm{B}}\int_{\tau_s^{\mathrm{B}}}^{\tau_{cs}^{\mathrm{B}}} \varphi(\tau)u_s^{\mathrm{au}}\left(\tau, N_c^{\mathrm{B}}, N_s^{\mathrm{B}}\right)d\tau \\ & + \left(N_c - N_c^{\mathrm{B}}\right)\int_{\tau_{cs}^{\mathrm{B}}}^{\max\left(\tau_c^{\mathrm{B}},\tau_{cs}^{\mathrm{B}}\right)} \varphi(\tau)u_c^{\mathrm{tr}}\left(\tau, N_c^{\mathrm{B}}, N_s^{\mathrm{B}}\right)d\tau + N_c^{\mathrm{B}}\int_{\max\left(\tau_c^{\mathrm{B}},\tau_{cs}^{\mathrm{B}}\right)}^{\bar{\theta}} \varphi(\tau)u_c^{\mathrm{au}}\left(\tau, N_c^{\mathrm{B}}, N_s^{\mathrm{B}}\right)d\tau, \end{aligned} \tag{22}$$

where SW^{B} represents the social welfare under the laissez faire. The first term on the RHS of Eq. (22) represents the profit generated by transit operations. The second term is the total utility of suburban commuters by transit, and the third term is the total utility of suburban commuters by auto. The sum of the second and third terms is thus the total utility of the suburban households. The fourth term is the total utility of central commuters by transit, and the fifth term is the total utility of central commuters by auto. The sum of the fourth and fifth terms is thus the total utility of the central households.

Under the laissez faire, N_i^{B} $(i = c, s)$ high-income households in area *i* would obtain a quota with a probability of 100%. This means that the numbers of auto quotas in the central and suburban areas equal the numbers of participants in the lottery rationing, i.e., $N_c^{\mathrm{B}} = \hat{N}_c$ and $N_s^{\mathrm{B}} = \hat{N}_s$. The lottery scheme is thus reduced to the no auto purchase restriction case. Consequently, the no auto purchase restriction case provides a lower bound of the social welfare of the lottery scheme. We thus have the following proposition.

Proposition 3. Comparing the laissez faire and the lottery scheme, we have

(i) The maximum social welfare under the lottery scheme is always higher than that under the laissez faire.

(ii) Introducing the lottery scheme leads the city to expand if and only if $G_s^{\mathrm{au}}(\tau_{cs}^{\mathrm{B}}) > \hat{G}_s^{\mathrm{au}}(\tau_{cs}^{\mathrm{B}})$ holds.

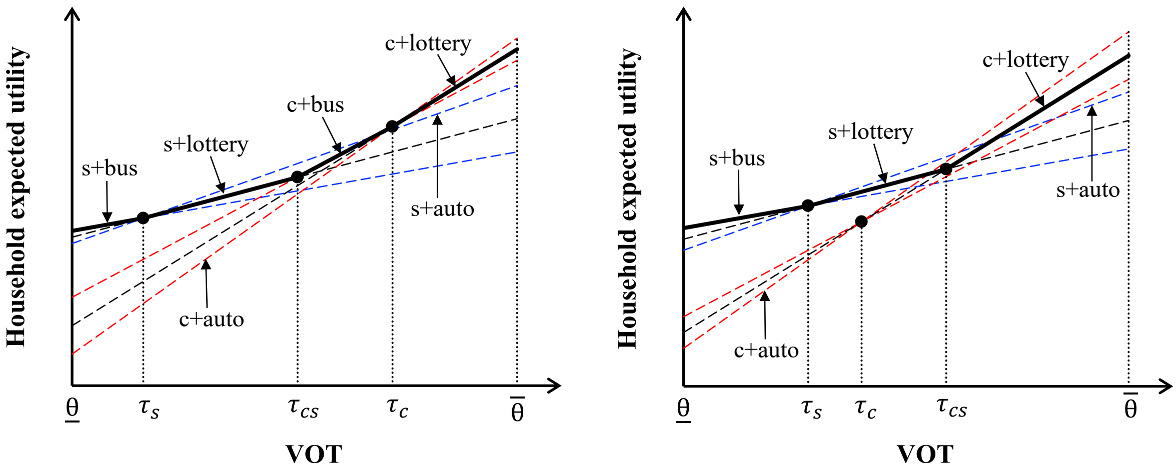
The proof of Proposition 3 is provided in Appendix B. Proposition 3 shows that the introduction of the lottery scheme increases the social welfare of the city, compared to the laissez-faire case. This justifies the implementation of a lottery scheme in practice to enhance the efficiency of the urban system. The size of the city with the lottery scheme may expand or shrink. Specifically, as the expected travel cost by auto in the suburban area decreases after introducing the lottery, the residents would migrate from the central area to the suburban area, thus leading to city expansion.

3. City-based vs. area-based lottery schemes

In this section, we discuss the effects of the lottery rationing schemes (i.e., city-based and area-based) on heterogeneous households' residential location and vehicle ownership choices between the central and suburban areas.

3.1. Critical VOT for household residential location choice

Under the license plate rationing, the households choose their residential locations and vehicle ownership (i.e., whether to attend the lottery) to maximize their utilities. There are four possible combinations: (suburban, transit), (suburban, lottery), (central, transit), and (central, lottery). According to Eqs. (11) and (14), the household utility function under each combination is a linear function of VOT τ. The utility frontier with heterogeneous households is the upper envelope of all the utility curves, as depicted by the bold line in Fig. 3. Comparing the values of critical VOTs τ_c and τ_{cs}, one obtains the following two possible relationships: $\gamma\left(\hat{G}_s(\tau_c) - \hat{G}_c(\tau_c)\right) \geq \sqrt{2kp_c} - \sqrt{2kp_s}$ and $\gamma\left(\hat{G}_s(\tau_c) - \hat{G}_c(\tau_c)\right) < \sqrt{2kp_c} - \sqrt{2kp_s}$.



(a) $\gamma\left(\hat{G}_s(\tau_c)-\hat{G}_c(\tau_c)\right)\geq\sqrt{2kp_c}-\sqrt{2kp_s}$ (b) $\gamma\left(\hat{G}_s(\tau_c)-\hat{G}_c(\tau_c)\right)<\sqrt{2kp_c}-\sqrt{2kp_s}$

Fig. 3. Residential location and travel mode choices of heterogeneous households under lottery scheme.

Fig. 3a illustrates the case where the difference in the expected travel costs of participating in lottery between central and suburban areas is greater than the difference in the residential land rents between them, i.e., $\gamma\left(\hat{G}_s(\tau_c)-\hat{G}_c(\tau_c)\right)\geq\sqrt{2kp_c}-\sqrt{2kp_s}$. It can be seen in Fig. 3a that there exist three critical VOTs (i.e., τ_s, τ_{cs}, and τ_c), dividing the households in the two-area city into four groups based on their travel mode and residential location choices. Specifically, the households with high VOTs (i.e., $\tau\in\left[\tau_c,\bar{\theta}\right]$) would choose to reside in the central area and participate in the lottery to obtain an auto quota; those with VOTs between τ_{cs} and τ_c choose to reside in the central area and commute by transit; those with VOTs between τ_s and τ_{cs} choose to reside in the suburban area and participate in the lottery for an auto quota; and those with low VOTs (i.e., $\tau\in[\underline{\theta},\tau_s)$) choose to reside in the suburban area and commute by transit.

The critical VOTs, τ_s and τ_c, between participating and non-participating in the lottery for the central and suburban areas can be calculated by Eq. (15). The critical VOT τ_{cs} of choosing to reside in the suburban and central areas satisfies:

$$\frac{Q_s}{\hat{N}_s}u_s^{\mathrm{au}}(\tau_{cs})+\left(1-\frac{Q_s}{\hat{N}_s}\right)u_s^{\mathrm{tr}}(\tau_{cs})=u_c^{\mathrm{tr}}(\tau_{cs})\,. \tag{23}$$

As the difference in the expected travel costs of participating in the lottery between the central and suburban areas is less than the difference in residential costs between them, i.e., $\gamma\left(\hat{G}_s(\tau_c)-\hat{G}_c(\tau_c)\right)<\sqrt{2kp_c}-\sqrt{2kp_s}$, all the households with medium VOTs would choose to reside in the suburban area and participate in the lottery to obtain an auto quota, as shown in Fig. 3b. For this case, at equilibrium, the critical VOT τ_{cs} of choosing to reside in the suburban and central areas satisfies

$$\frac{Q_s}{\hat{N}_s}u_s^{\text{au}}(\tau_{cs})+\left(1-\frac{Q_s}{\hat{N}_s}\right)u_s^{\text{tr}}(\tau_{cs})=\frac{Q_c}{\hat{N}_c}u_c^{\text{au}}(\tau_{cs})+\left(1-\frac{Q_c}{\hat{N}_c}\right)u_c^{\text{tr}}(\tau_{cs}). \tag{24}$$

From Eqs. (15), (23), and (24), one can obtain the residential distribution of households between the central and suburban areas. The numbers of households residing in the central and suburban areas can then be determined by

$$N_c=N\int_{\tau_{cs}}^{\bar{\theta}}\varphi(\tau)d\tau=N\left(1-\Phi\left(\tau_{cs}\right)\right),\text{ and }\ N_s=N\int_{\underline{\theta}}^{\tau_{cs}}\varphi(\tau)d\tau=N\Phi\left(\tau_{cs}\right). \tag{25}$$

The numbers of participants in the lottery in the central and suburban areas are given by

$$\hat{N}_c=N\int_{\max\{\tau_c,\tau_{cs}\}}^{\bar{\theta}}\varphi(\tau)d\tau=N\left(1-\Phi\left(\max\{\tau_c,\tau_{cs}\}\right)\right),\text{ and }\ \hat{N}_s=N\int_{\tau_s}^{\tau_{cs}}\varphi(\tau)d\tau=N\left(\Phi\left(\tau_{cs}\right)-\Phi\left(\tau_s\right)\right). \tag{26}$$

In the above, we have derived the residential location choice equilibrium of heterogeneous households under the lottery rationing, including the critical VOT of choosing to reside in the suburban and central areas and the population sizes in the central and suburban areas.

3.2. Area-based lottery scheme

In this scheme, the government determines the auto quotas for each of the central and suburban areas, aiming to maximize the social welfare of the system.

$$\begin{aligned}\max_{Q_c,Q_s}\ SW^{\text{A}}(Q_c,Q_s)=&\gamma\left(N-Q_c-Q_s\right)\beta+\left(N_s-\hat{N}_s\right)\int_{\underline{\theta}}^{\tau_s}\varphi(\tau)u_s^{\text{tr}}(\tau)d\tau+\left(\hat{N}_s-Q_s\right)\int_{\tau_s}^{\tau_{cs}}\varphi(\tau)u_s^{\text{tr}}(\tau)d\tau+Q_s\int_{\tau_s}^{\tau_{cs}}\varphi(\tau)u_s^{\text{au}}(\tau)d\tau\\&+\left(N_c-\hat{N}_c\right)\int_{\tau_{cs}}^{\max\{\tau_c,\tau_{cs}\}}\varphi(\tau)u_c^{\text{tr}}(\tau)d\tau+\left(\hat{N}_c-Q_c\right)\int_{\max\{\tau_c,\tau_{cs}\}}^{\bar{\theta}}\varphi(\tau)u_c^{\text{tr}}(\tau)d\tau+Q_c\int_{\max\{\tau_c,\tau_{cs}\}}^{\bar{\theta}}\varphi(\tau)u_c^{\text{au}}(\tau)d\tau,\end{aligned} \tag{27}$$

where SW^{A} is the social welfare under the area-based lottery scheme and superscript "A" represents the area-based scheme. The economic implications of Eq. (27) can be explained below. The first term on the RHS of Eq. (27) is the profit generated by the transit service. The second to fourth terms are the total utility of the suburban residents: the second term is the total utility of those who would not like to participate in the lottery and to travel by transit; the third term is the total utility of those who participate in the lottery but fail to get a license plate due to the quota limitation; and the fourth term is the total utility of those who obtain an auto quota through the lottery. The fifth to seventh terms are the total utility of the central residents: the fifth term is the total utility of those who choose to travel by transit; the sixth term is the total utility of those who participate in the lottery but do not get a quota; and the seventh term is the total utility of those who obtain a quota by the lottery. Note that as $\gamma\left(\hat{G}_s(\tau_c)-\hat{G}_c(\tau_c)\right)<\sqrt{2kp_c}-\sqrt{2kp_s}$, all households residing in the central area will choose to participate in the lottery to obtain an auto quota. This is because the central households are high-income ones and thus prefer to commute by auto due to its quick travel speed.

3.3. City-based lottery scheme

In the city-based lottery scheme, the government determines the total auto quota for the whole city to maximize the social welfare, represented as

$$\begin{aligned}\max_{Q}\ SW^{\mathrm{C}}(Q)=&\gamma\left(N-Q_c-Q_s\right)\beta+\left(N_s-\hat{N}_s\right)\int_{\underline{\theta}}^{\tau_s}\varphi(\tau)u_s^{\mathrm{tr}}(\tau)d\tau+\left(\hat{N}_s-Q_s\right)\int_{\tau_s}^{\tau_{cs}}\varphi(\tau)u_s^{\mathrm{tr}}(\tau)d\tau+Q_s\int_{\tau_s}^{\tau_{cs}}\varphi(\tau)u_s^{\mathrm{au}}(\tau)d\tau\\&+\left(N_c-\hat{N}_c\right)\int_{\tau_{cs}}^{\max\{\tau_c,\tau_{cs}\}}\varphi(\tau)u_c^{\mathrm{tr}}(\tau)d\tau+\left(\hat{N}_c-Q_c\right)\int_{\max\{\tau_c,\tau_{cs}\}}^{\overline{\theta}}\varphi(\tau)u_c^{\mathrm{tr}}(\tau)d\tau+Q_{\mathrm{c}}\int_{\max\{\tau_c,\tau_{cs}\}}^{\overline{\theta}}\varphi(\tau)u_c^{\mathrm{au}}(\tau)d\tau,\end{aligned}\tag{28}$$

subject to

$$\rho=\frac{\hat{N}_s}{\hat{N}_s+\hat{N}_c},\tag{29a}$$

$$Q_s=\rho Q,\ \text{and}\ \ Q_c=(1-\rho)Q,\tag{29b}$$

where SW^{C} is the social welfare under the city-based lottery scheme, and superscript "C" represents the city-based scheme. The economic implications for the components of Eq. (28) are identical to those of Eq. (27). Eq. (29a) is used to determine the proportion of auto quotas allocated to the suburban area among total quotas under the city-based lottery scheme. Eq. (29b) is used to calculate the associated auto quotas allocated to the central and suburban areas.

Compared to the social welfare maximization model under the area-based lottery scheme, the city-based lottery scheme introduces additional constraints (i.e., Eqs. (29a) and (29b)) to endogenously determine the proportion of auto quotas allocated to the central and suburban areas. Accordingly, the solution space of the area-based lottery scheme covers that of the city-based lottery scheme, and thus the solution space of the city-based lottery scheme is a subset of the solution space of the area-based lottery scheme. Therefore, the optimal solution of the area-based lottery scheme provides an upper bound for the city-based lottery scheme. We thus have the following property.

Proposition 4. Comparing the area-based and city-based lottery schemes, we have

(i) The area-based lottery scheme is more socially efficient than the city-based lottery scheme, i.e., $\max SW^{\mathrm{A}} \geq \max SW^{\mathrm{C}}$.

(ii) The proportion of auto quotas allocated to suburban area under the area-based lottery scheme is higher than that under the city-based lottery scheme if and only if $\dfrac{Q_s^{\mathrm{A}}\left(t_{c,v}^{\mathrm{au}}, t_{s,v}^{\mathrm{au}}\right)}{Q_s^{\mathrm{A}}\left(t_{c,v}^{\mathrm{au}}, t_{s,v}^{\mathrm{au}}\right) + Q_c^{\mathrm{A}}\left(t_{c,v}^{\mathrm{au}}, t_{s,v}^{\mathrm{au}}\right)} > \rho\left(t_{c,v}^{\mathrm{au}}, t_{s,v}^{\mathrm{au}}\right)$ holds.

Proposition 4 can be explained as follows. In the city-based lottery scheme, the government sets a total auto quota for the entire city, meaning that all participants in the lottery, regardless of the central or suburban area, have an equal probability of obtaining an auto quota. However, the auto demand can vary significantly between areas. For instance, suburban residents usually are more eager for a private car due to a longer commuting distance compared to the central residents. Therefore, more auto quotas should be allocated to the suburban area so as to enhance the probability of obtaining an auto quota for the suburban residents. The area-based lottery scheme can effectively address this issue by separately determining auto quotas for the central and suburban areas. More auto quotas should be allocated to the area with higher auto demand under the area-based lottery scheme, ensuring that the allocation of auto quotas can better meet the residents' car-purchasing needs at different areas. As a result, the area-based lottery scheme is a better strategy to enhance the social welfare than the city-based lottery scheme from a long-term perspective. In addition, the critical condition in Proposition 4(ii) can be directly obtained by taking a difference of the allocation proportions under the area-based and city-based lottery schemes. Proposition 4(ii) shows that the license plate proportion allocated to the suburban area under the area-based lottery scheme may be higher or lower than that under the city-based lottery scheme, depending on the road service level.

4. Model application

In order to illustrate the contributions of this paper and to gain further insights, we apply the proposed models to Beijing city of China, where a city-based lottery scheme has been implementing since 2011. We determine and compare the optimal solutions under the area-based and city-based lottery schemes, and examine the effects of lottery schemes on the urban spatial structure and social welfare of the city.

Fig 4. Geographic regions of Beijing city.

4.1. Calibrations of parameters

Thus far, the transport network of Beijing city consists of six ring roads. According to the administrative divisions of Beijing city, the area within the fifth ring road is considered as the central area, whereas the area outside it is regarded as the suburban area, as shown in Fig. 4. According to 2023 official data from the Beijing Municipal Government's statistical system[3], the total population size in Beijing is currently about 21.86 million permanent residents. Only permanent residents over 18 can participate in the lottery, as required by application of driving license. Among the permanent residents, those over 18 years old account for about 78.9%, implying that a total of 17.25 million persons are eligible for participating in the lottery, i.e., N=17,250,000. The minimum and maximum annual incomes are assumed to be \$60,000 and \$720,000 respectively, covering the majority of all residents (99.99%). Suppose that each

[3] https://www.beijing.gov.cn/renwen/bjgk/rk/rktj/202403/t20240322_3597338.html.

commuter works an average of 8 hours per day and 300 days per year, and thus the amount of the average annual work hours is 2,400 hours, i.e., $\alpha = 2,400\text{h}$. The lower and upper bounds of the residents' VOTs are, therefore, \$25 and \$300 per hour, respectively, i.e., $\underline{\theta} = \$25/\text{h}$ and $\overline{\theta} = \$300/\text{h}$. In Beijing, the average annual wage per household is \$225,000. The average VOT of residents can thus be calculated as \$93.75 per hour and the associated standard deviation is \$30 per hour. Some empirical studies showed that residents' income levels or VOTs usually follow a lognormal distribution. The residents' VOT distribution in Beijing can be calibrated as $\tau \sim LN(4.4919, 0.3122)$. Households have a residential land preference valued at 1,350,000 units. The annual agricultural rent r_s at the city boundary is \$300 per square meter.

According to 2023 Annual Report of Beijing Commuter Characteristics[4], the average daily commuting distance of residents in the central area of Beijing is about 13.3 km. The free-flow travel speeds by auto are approximately 40 km per hour, respectively. Therefore, the average free-flow travel time for central commuters by auto can be calculated as 0.33 hour, i.e., $t_{c,f}^{\text{au}} = 0.33\text{h}$. The total travel time by transit in the central area is 0.68 hour, i.e., $T_c^{\text{tr}} = 0.68\text{h}$. It is reported that there are about 300 key traffic congestion points during morning peak hours in Beijing.[5] The peak-hour factor (i.e., the ratio of peak-hour flow to daily average flow) is about 10%.[6] Hence, the average peak-hour traffic volume per key congestion point can be calculated as $21860000 \times 10\% / 300 = 7287$ vehicles. Given an average auto travel time of 0.5 hour in the central area during peak hours, then the marginal travel time by auto in this area can be calibrated as $(0.5 - 0.33)/7287 \times 3600 = 0.084\text{sec/veh}$. The monetary travel cost for a central commuter by auto is \$20 per trip, i.e., $f_c = \$20/\text{trip}$. The transit fare is \$2 per trip, i.e., $\beta = \$2/\text{trip}$.

Compared to the central residents, the suburban residents have to cross an extra suburban bridge to arrive at their workplace located in the CBD. The travel distance across the suburban bridge is approximately 14.8 km. The free-flow travel time by auto on the suburban bridge is thus 0.37 hour, i.e., $t_{s,f}^{\text{au}} = 0.37\text{h}$. The total travel time by transit crossing the suburban bridge is 0.62

---

4 https://www.beijing.gov.cn/ywdt/gzdt/202403/t20240309_3584335.html.
5 https://www.chinautc.com/templates/H_nianjian/content.aspx?nodeid=1488&page=ContentPage&contentid=97936.
6 https://www.quora.com/What-is-the-most-suitable-way-to-convert-peak-hour-traffic-volume-data-of-a-day-into-average-daily-traffic-ADT-for-the-same-day.

hour, i.e., $T_s^{\text{tr}} = 0.62\text{h}$. Given an average auto travel time of 0.75 hour in the suburban area during peak hours, the marginal travel time cost by auto in this area can be calibrated as $(0.75-0.37)/7287\times 3600 = 0.188\text{sec/veh}$. The associated monetary travel cost by auto is set as \$23 per trip, i.e., $f_s = \$23/\text{trip}$. Table 1 summarizes the baseline values of all the input parameters. Unless specifically stated otherwise, these parameter values are used as the base case in the following analysis.

Table 1 Values of input parameters for numerical study.

Parameter	Definition	Baseline value
N	Total number of permanent residents (persons)	17,250,000
$\overline{\theta}$	Upper bound of VOT (\$/h)	300
$\underline{\theta}$	Lower bound of VOT (\$/h)	25
α	Average annual working hours (h)	2,400
b_c	Area of central area (km^2)	200
r_s	Agricultural rent at city boundary (\$/ m^2)	300
k	Preferences of households for land (units)	1,350,000
γ	Average annual number of working trips per household (trips)	600
β	Transit fare (\$/trip)	2
T_c^{tr}	Average transit travel time on central bridge (h/trip)	0.68
f_c	Monetary cost by auto of central commuters (\$/trip)	20
$t_{c,f}^{\text{au}}$	Average auto free-flow travel time on central bridge (h/trip)	0.33
$t_{c,v}^{\text{au}}$	Marginal auto travel time on central bridge (sec/veh)	0.084
T_s^{tr}	Average transit travel time on suburban bridge (h/trip)	0.62
f_s	Monetary cost by auto of suburban commuters (\$/trip)	23
$t_{s,f}^{\text{au}}$	Average auto free-flow travel time on suburban bridge (h/trip)	0.37
$t_{s,v}^{\text{au}}$	Marginal auto travel time on suburban bridge (sec/veh)	0.188

4.2. Discussion of results

4.2.1. Comparison of solutions of city-based and area-based lottery schemes

Figs. 5 and 6 depict the changes of the social welfare with the auto quota under the city-based and area-based lottery schemes. It can be seen that the social welfare curves under both schemes are concave with regard to the auto quota, meaning that the optimal quota solution for each scheme with the social welfare maximization is unique. Table 2 further summarizes the optimal solutions with no auto purchase restriction, city-based lottery, and area-based lottery. It indicates that with no auto purchase restriction, the optimal total auto demand is 7.76 million vehicles (see also Fig. 5). To control excessively rapid growth in the number of motorized vehicles, Beijing city government has been implementing a city-based lottery scheme since 2011. The associated optimal total auto quota under this scheme is 3.93 million vehicles, with 53% allocated to the central area and 47% to the suburban area. This yields a social welfare of $2253.05 billion per year, leading to a welfare increase by $12.43 billion per year compared to the no auto purchase restriction case. If an area-based lottery scheme is adopted in the future, the optimal auto quotas for the central and suburban areas are, respectively, 0.82 million and 2.20 million vehicles, causing a welfare increase by $0.75 billion per year compared to the city-based lottery scheme (see also Fig. 6). This illustrates that the area-based lottery scheme is superior to the city-based lottery scheme in terms of the social welfare, as shown in Proposition 4.

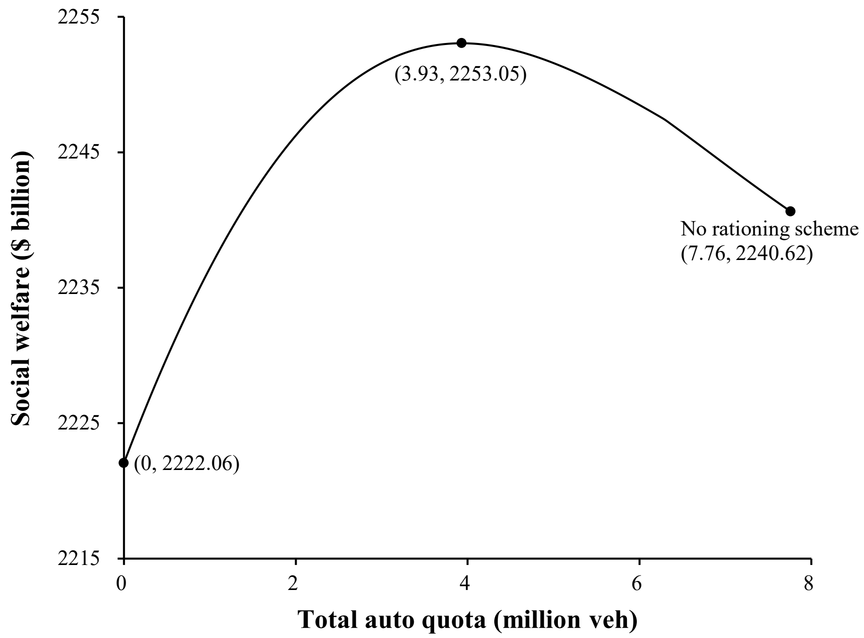


Fig. 5. Changes of social welfare with total auto quota under city-based lottery scheme.

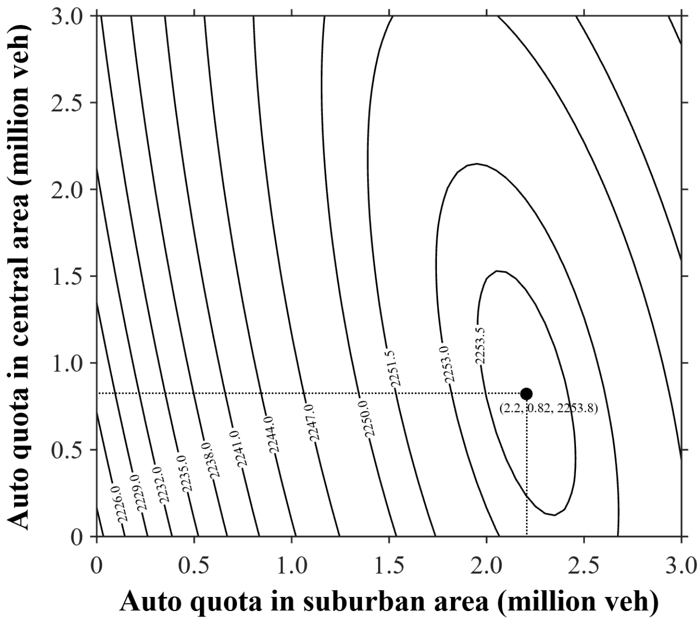


Fig. 6. Changes of social welfare with auto quota under area-based lottery scheme.

Table 2 Comparison of optimal solutions under different schemes.

Solution	No purchase restriction		Area-based lottery		City-based lottery	
	Central area	**Suburban area**	**Central area**	**Suburban area**	**Central area**	**Suburban area**
Auto quota (million veh)	4.92	2.84	0.82	2.20	2.09	1.84
Allocation proportion of auto quotas	0.63	0.37	0.27	0.73	0.53	0.47
Average housing size (m^2)	22.66	47.43	22.81	47.43	22.57	47.43
Average housing price (\$/ m^2)	1,314	300	1,298	300	1,325	300
Number of residents (million)	8.82	8.43	8.77	8.48	8.86	8.39
Area size (km^2)	200.00	399.65	200.00	402.28	200.00	397.93
Critical VOT of choosing to reside in central and suburban areas ($/h)	88.48		88.71		88.34	
Social welfare ($ billion)	2,240.62		2,253.80		2253.05	

Table 2 further shows that the changes of the urban spatial structure under the city-based and area-based lottery schemes are significantly different. Specifically, the city-based lottery scheme leads the households with VOT $\tau \in [88.34, 88.48]$ to migrate from the suburban area to the central area, causing a decrease in the number of suburban households by 40 thousand and in the city size by 1.72 km^2. The land rental price in the central area rises, and the average housing size per household decreases. Conversely, introducing the area-based lottery scheme results in a migration of 50 thousand households from the central area to the suburban area due to a lower housing rent in the suburban area, thus causing the city sprawl. As a result, the land rental price in the central area decreases, and the associated housing size per household increases.

4.2.2. Effects of lottery schemes and road service level on quota proportion allocated to central and suburban areas

In order to illustrate the effects of different rationing schemes and road service level on the quota proportion allocated to the central/suburban area and the urban spatial structure, we conduct sensitivity analysis of $t_{c,v}^{\mathrm{au}}$ (as an indicator of road service level on central bridge) under different rationing schemes, as shown in Fig. 7. It can be seen in Fig. 7 that as $t_{c,v}^{\mathrm{au}}$ increases from 0.05 to 0.09 second per vehicle, the proportion of auto demand (quotas) in the suburban area increases by 9% (from 29% to 38%) under the no rationing scheme, by 49% (from 34% to 83%) under the area-based lottery scheme, and remains nearly unchanged under the city-based lottery scheme. There exsits a critical value of $t_{c,v}^{\mathrm{au}} = 0.063$ such that at its right-hand side, the proportion of auto quotas allocated to the suburban area under the area-based lottery scheme is higher than that under the city-based lottery scheme, but lower at its left-hand side. This illustrates the result of Proposition 4.

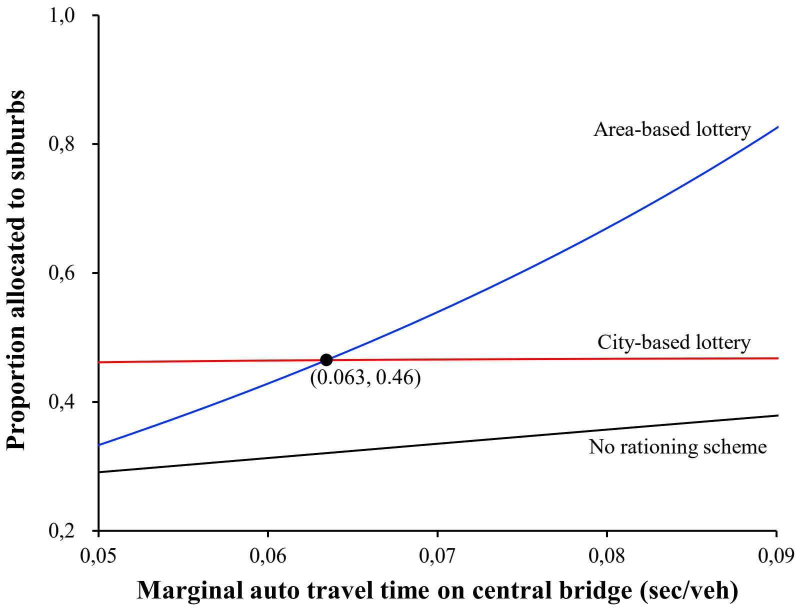


Fig. 7. Effects of marginal auto travel time $t_{c,v}^{\mathrm{au}}$ on quota proportion allocated to suburban area under different rationing schemes.

4.2.3. Effects of lottery schemes and road service level on city size

To illustrate the effects of different rationing schemes on the city size, Fig. 8 shows the changes of city size with the road service level on the central bridge under different rationing schemes. One can observe that after introducing the lottery schemes, the city may expand or shrink, depending on the scheme adopted (area-based or city-based) and the road service level. Specifically, introducing the area-based lottery scheme would cause city sprawl for a low road service level on the central bridge (i.e., $t_{c,v}^{\text{au}} > 0.063$), and city contraction for a high road service level on the central bridge (i.e., $t_{c,v}^{\text{au}} \leq 0.063$). However, the introduction of the city-based lottery scheme always leads some households to migrate from the suburban area to the central area, causing a more compact city. This result is consistent with that of Proposition 3.

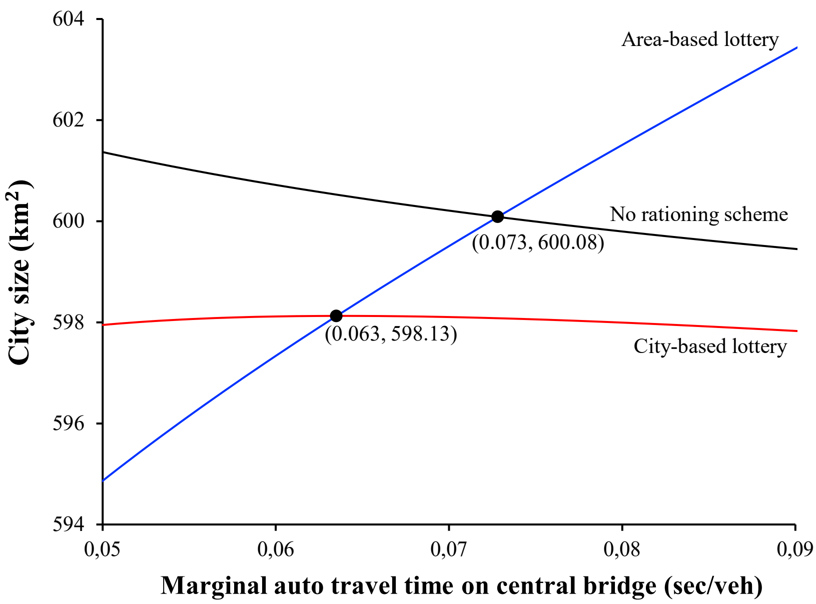


Fig. 8. Effects of marginal auto travel time $t_{c,v}^{\text{au}}$ on city size under different rationing schemes.

4.2.4. Effects of lottery schemes and road service level on social welfare

Finally, we look at the effects of different lottery schemes and the road service level on the social welfare. Fig. 9 shows the changes of the social welfare with the marginal auto travel time $t_{c,v}^{\text{au}}$ on the central bridge under the city-based and area-based lottery schemes. It can be seen in Fig. 9 that the social welfare curve under the area-based lottery scheme is always above that under the city-based lottery scheme. This means that the area-based lottery scheme is more efficient than the city-based lottery scheme in improving the social welfare of the system, which is consistent with Proportion 4. This is because that compared to the city-based lottery scheme, more auto quotas should be allocated to the area with higher auto demand under the area-based lottery scheme, so as to ensure that the allocation of auto quotas can better meet the residents' car-purchasing needs in both the central and suburban areas.

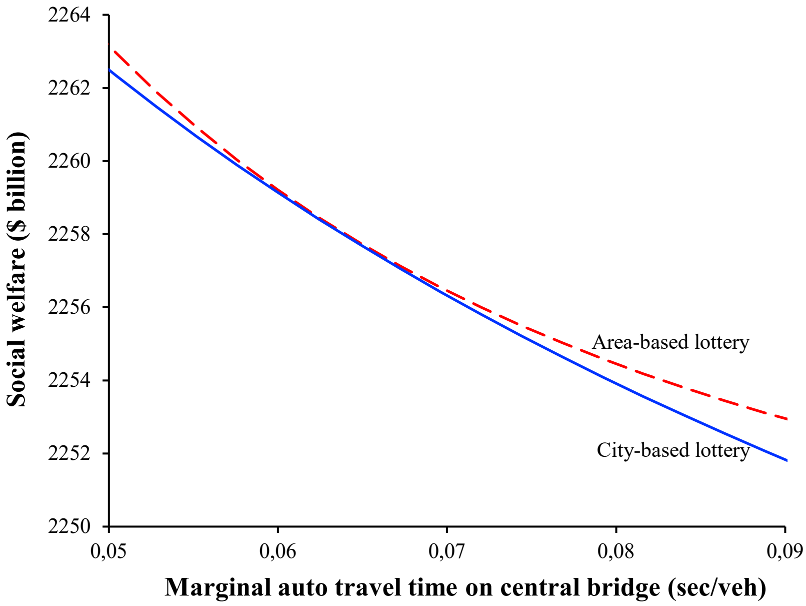


Fig. 9. Effects of marginal auto travel time $t_{c,v}^{\mathrm{au}}$ on social welfare under different lottery schemes.

5. Conclusion and further studies

In this paper, we presented a novel theory of license plate rationing that accounts for the effects of spatial heterogeneity using a two-area modeling method. Based on a formulation of the residential location and vehicle ownership choice equilibria of heterogeneous households, a social welfare maximization model was proposed for determining the optimal auto quota and optimal quota allocation proportion between central and suburban areas. The critical VOTs for participating and non-participating in the lottery for central and suburban residents are derived. Two alternative rationing schemes, i.e., a city-based and an area-based lottery scheme, were investigated and compared, together with no auto purchase restriction (i.e., laissez faire). The interactions among the license plate rationing, household heterogeneity, and urban spatial structure were incorporated in the model.

Some important and insightful findings were obtained. First, there is a significant spatial difference in the lottery behavior of suburban and central residents. When the transit service level in the suburban area is relatively low, the cirtical VOT for participating in the lottery for suburban residents is lower than that for central residents. Second, households with high VOTs choose to reside in the central area and participate in the lottery for obtaining an auto quota, while those with low VOTs choose to reside in the suburban area and commute by transit. Third, the area-based lottery scheme is more socially efficient than the city-based lottery scheme. Compared to the laissez faire, the introduction of the license plate rationing may lead to city sprawl or shrink, depending on the scheme adopted and road service level. The proposed model in this paper can serve as a useful tool for evaluating and designing the license plate rationing policies for efficient travel demand management.

Some extensions can be envisaged as follows. First, we focused on a monocentric city with two discrete areas. Extending the proposed models in this paper to a one-dimensional or two-dimensional city continuum is meaningful (Li et al., 2013, 2024a), which is left for future study. Second, this paper only discussed the auto ownership restriction policy. Other travel demand management policies, such as road pricing and driving restriction based on license plate numbers, have also been adopted in some large cities. Therefore, there is a need to compare the efficiency of the license plate rationing with these schemes. Third, license plate auction, as a market-based mechanism, has also been implemented in some cities, such as Shanghai and Singapore. It is meaningful to ascertain the long-term effects of the auction schemes on the urban system and to compare the results under the auction and lottery in a further study.

Acknowledgments

This study was supported for the first two authors by grants from the National Natural Science Foundation of China (72131008) and the Fundamental Research Funds for the Central Universities (2021GCRC014), and for the third author by AFFINITE, ANR grant (Project-ANR20-CE22-0014), France, and MAAT, France, The Netherlands and China.

References

Akamatsu, T., Fujishima, S., Takayama, Y., 2017. Discrete-space agglomeration model with social interactions: Multiplicity, stability, and continuous limit of equilibria. Journal of Mathematical Economics 69, 22-37.

Becker, G.S., 1965. A theory of the allocation of time. The Economic Journal 75, 493-517.

Blanchet, A., Mossay, P., Santambrogio, F., 2016. Existence and uniqueness of equilibrium for a spatial model of social interactions. International Economic Review 57(1), 31-60.

Brueckner, J.K., Helsley, R.W., 2011. Sprawl and blight. Journal of Urban Economics 69(2), 205-213.

Brueckner, J.K., 2014. Cordon tolling in a city with congested bridges. Economics of Transportation 3(4), 235-242.

Brueckner, J.K., Franco, S.F., 2018. Employer-paid parking, mode choice, and suburbanization. Journal of Urban Economics 104, 35-46.

Chin, A., Smith, P., 1997. Automobile ownership and government policy: The economics of Singapore's vehicle quota scheme. Transportation Research Part A 31(2), 129-140.

Chu, S., 2002. Auctioning rights to vehicle ownership: Singapore's experience with sealed-bid tenders. Transportation Research Part A 36(6), 555-561.

Chu, S., 2012. Allocation flexibility and price efficiency within Singapore's Vehicle Quota System. Transportation Research Part A 46(10), 1541-1550.

Chu, S., Koh, W.T.H., Tse, Y.K., 2004. Expectations formation and forecasting of vehicle demand: An empirical study of the vehicle quota auctions in Singapore. Transportation Research Part A 38(5), 367-381.

de Palma, A., Lindsey, R., 2011. Traffic congestion pricing methodologies and technologies. Transportation Research Part C 19(6), 1377-1399.

Gallego, F., Montero, J.P., Salas, C., 2013. The effect of transport policies on car use: Evidence from Latin American cities. Journal of Public Economics 107, 47-62.

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

Hu, M., Xiao, J., Zheng, B., 2022. The selection effect of quota rationing mechanisms on sales distribution: The convergence of auction and lottery. Journal of Economic Behavior & Organization 200, 803-819.

Kim, J., 2016. Vehicle fuel-efficiency choices, emission externalities, and urban sprawl. Economics of Transportation 5, 24-36.

Koh, W.T.H., 2003. Control of vehicle ownership and market competition: Theory and Singapore’s experience with the vehicle quota system. Transportation Research Part A 37(9), 749-770.

Koh, W.T.H., 2004. Congestion control and vehicle ownership restriction: The choice of an optimal quota policy. Journal of Transport Economics and Policy 38(3), 371-402.

Li, S., 2018. Better lucky than rich? Welfare analysis of automobile licence allocations in Beijing and Shanghai. The Review of Economic Studies 85(4), 2389-2428.

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(6), 927-937.

Li, Z.C., Cheng, L., de Palma, A., 2024a. Ring road investment, cordon tolling, and urban spatial structure. Transportation Research Part B 182, 102905.

Li, Z.C., Wu, Q.Y., Yang, H., 2019. A theory of auto ownership rationing. Transportation Research Part B 127, 125-146.

Li, Z.C., Yu, D.P., de Palma, A., 2024b. Bottleneck congestion and urban spatial structure with heterogeneous households: Equilibrium, capacity expansion and congestion tolling. Journal of Urban Economics 144, 103693.

Liu, W.J., Li, Z.C., Yang, H., 2024. Gasoline and electric vehicle ownership rationing over time: Lottery vs. first-come-first-served schemes. Transportation Research Part B 186, 102995.

Mossay, P., Picard, P.M., 2011. On spatial equilibria in a social interaction model. Journal of Economic Theory 146(6), 2455-2477.

Nie, Y., 2017. On the potential remedies for license plate rationing. Economics of Transportation 9, 37-50.

Phang, S.Y., Wong, W.K., Chia, N.C., 1996. Singapore's experience with car quotas: Issues and policy processes. Transport Policy 3(4), 145-153.

Picard, P.M., Tabuchi, T., 2013. On microfoundations of the city. Journal of Economic Theory 148(6), 2561-2582.

Small, K.A., 2012. Valuation of travel time. Economics of Transportation 1(1-2), 2-14.

Small, K.A., Verhoef, E.T., 2007. The Economics of Urban Transportation. Routledge, London.

Wang, S., Zhao, J., 2017. The distributional effects of lotteries and auctions-License plate regulations in Guangzhou. Transportation Research Part A 106, 473-483.

Yang, H., Zhai, G., Liu, X., Yang, L., Liu, Y., Yuan, Q., 2022. Determinants of city-level private car ownership: Effect of vehicle regulation policies and the relative price. Transport Policy 115, 40-48.

Yu, D.P., Li, Z.C., 2023. Income distribution, implementation sequence, and equity in auto ownership rationing. Transportation Research Part B 173, 59-89.

Zhu, S., Du, L., Zhang, L., 2013. Rationing and pricing strategies for congestion mitigation: Behavioral theory, econometric model, and application in Beijing. Transportation Research Part B 57, 210-224.

Appendix A: Proof of Proposition 1

We first consider a suburban household with VOT τ_1. Since the household's utility is maximized for a chosen location at equilibrium, changing its residential location to central area will inevitably reduce its utility, i.e., $u_s^j(\tau_1) > u_c^j(\tau_1)$. The commuting cost for a resident who lives in the suburban area is always higher than the commuting cost of the same resident who lives in the central area, i.e., $G_s^j(\tau_1) > G_c^j(\tau_1)$ ($j = \text{au, tr}$). We thus immediately have $\sqrt{2kp_c} > \sqrt{2kp_s}$ and $p_c > p_s$ according to Eq. (11).

We then analyze the residential location choices of the heterogeneous households. At equilibrium, exchanging the residential locations of any two households would reduce their utilities. Therefore, for any two households who respectively reside in the central and suburban areas with corresponding VOTs τ_1 and τ_2, $u_c^j(\tau_1) + u_s^j(\tau_2) > u_s^j(\tau_1) + u_c^j(\tau_2)$ ($j = \text{au,tr}$) should hold. From the condition, we immdiately obtain $\tau_1 > \tau_2$. This means that households residing in the central area have higher VOTs that those residing in the suburban area. This completes the proof of Proposition 1.

Appendix B: Proof of Proposition 3

Under the no vehicle purchase restriction case, the critical VOT of choosing to reside in suburban and central areas satisfies the following relationships:

$$u_s^{\text{au}}(\tau_{cs}^{\text{B}}) = \begin{cases} u_{\text{c}}^{\text{tr}}(\tau_{cs}^{\text{B}}), \text{if } \tau_{cs}^{\text{B}} < \tau_c^{\text{B}}, \\ u_c^{\text{au}}(\tau_{cs}^{\text{B}}), \text{if } \tau_{cs}^{\text{B}} \geq \tau_c^{\text{B}}. \end{cases} \tag{B.1}$$

Substituting Eq. (11) into Eq. (B.1), one obtains

$$\alpha\tau_{cs}^{\text{B}} - \sqrt{2kp_s} - \gamma G_s^{\text{au}}(\tau_{cs}^{\text{B}}) = \begin{cases} \alpha\tau_{cs}^{\text{B}} - \sqrt{2kp_c} - \gamma G_c^{\text{tr}}(\tau_{cs}^{\text{B}}), \text{if } \tau_{cs}^{\text{B}} < \tau_c^{\text{B}}, \\ \alpha\tau_{cs}^{\text{B}} - \sqrt{2kp_c} - \gamma G_c^{\text{au}}(\tau_{cs}^{\text{B}}), \text{if } \tau_{cs}^{\text{B}} \geq \tau_c^{\text{B}}. \end{cases} \tag{B.2}$$

Under the lottery scheme, the critical VOT τ_{cs} of choosing to reside in suburban and central areas becomes:

$$\frac{Q_s}{\hat{N}_s} u_s^{\text{au}}(\tau_{cs}) + \left(1 - \frac{Q_s}{\hat{N}_s}\right) u_s^{\text{tr}}(\tau_{cs}) = \begin{cases} u_c^{\text{tr}}(\tau_{cs}), \text{if } \tau_{cs} < \tau_c, \\ \dfrac{Q_c}{\hat{N}_c} u_c^{\text{au}}(\tau_{cs}) + \left(1 - \dfrac{Q_c}{\hat{N}_c}\right) u_c^{\text{tr}}(\tau_{cs}), \text{if } \tau_{cs} \geq \tau_c. \end{cases} \tag{B.3}$$

Substituting Eqs. (11) and (14) into Eq. (B.3), one obtains

$$\alpha\tau_{cs} - \sqrt{2kp_s} - \gamma \hat{G}_s^{\text{au}}(\tau_{cs}) = \begin{cases} \alpha\tau_{cs} - \sqrt{2kp_c} - \gamma G_c^{\text{tr}}(\tau_{cs}), \text{if } \tau_{cs} < \tau_c, \\ \alpha\tau_{cs} - \sqrt{2kp_c} - \gamma \hat{G}_c^{\text{au}}(\tau_{cs}), \text{if } \tau_{cs} \geq \tau_c. \end{cases} \tag{B.4}$$

After introducing the lottery scheme, the utility of the household with VOT τ_{cs}^{B} residing in the suburban area becomes

$$\hat{u}_s^{\text{au}}(\tau_{cs}) = \alpha\tau_{cs} - \sqrt{2kp_s} - \gamma \hat{G}_s^{\text{au}}(\tau_{cs}). \tag{B.5}$$

If the expected travel cost by auto in the suburban area decreases after introducing the lottery scheme, i.e., $\hat{G}_s^{\text{au}}(\tau_{cs}^{\text{B}}) < G_s^{\text{au}}(\tau_{cs}^{\text{B}})$, it follows $\hat{u}_s^{\text{au}}(\tau_{cs}^{\text{B}}) > u_s^{\text{au}}(\tau_{cs}^{\text{B}})$ by Eqs. (B.2) and (B.5). Therefore, after introducing the lottery scheme, the households with VOT $\tau \in \left[\tau_{cs}^{\text{B}}, \tau_{cs}\right]$ would like to relocate from the central area to the suburban area, causing the city expansion. Otherwise, the households with VOT $\tau \in \left[\tau_{cs}, \tau_{cs}^{\text{B}}\right]$ would like to migrate from the suburban area to the central area, leading to a more compact city. This completes the proof of Proposition 3.