Article:

Zhang, F orcid.org/0000-0003-3584-4679, Yang, H and Liu, W (2014) The Downs–Thomson Paradox with responsive transit service. Transportation Research Part A: Policy and Practice, 70. pp. 244-263. ISSN 0965-8564

https://doi.org/10.1016/j.tra.2014.10.022

The Downs-Thomson Paradox with Responsive Transit Service

Fangni Zhang, Hai Yang* and **Wei Liu**

Department of Civil and Environmental Engineering, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, China.

Abstract

Downs (1962) and Thomson (1977) suggested that highway capacity expansion may produce counterproductive effects on the two-mode (auto and transit) transport system (Downs-Thomson Paradox). This paper investigates the occurrence of this paradox when transit authority can have different economic objectives (profit-maximizing or breakeven) and operating schemes (frequency, fare, or both frequency and fare). For various combinations of economic objectives and operating schemes, the interaction between highway expansion and transit service is explored, as well as its impact on travelers' mode choices and travel utilities. Further, for each combination, the conditions for occurrence of the Downs-Thomson Paradox are established. We show that the paradox never occurs when transit authority is profit-maximizing, but it is inevitable when the transit authority is running to maximize travelers' utility while maintaining breakeven. This is because the former transit authority tends to enhance transit service (e.g., raise frequency or reduce fare) when facing an expanded highway; and on the contrary, the latter tends to compromise transit service (e.g., reduce frequency or raise fare). Both analytical and numerical examples are provided to verify the theoretical results.

Keywords: Downs-Thomson Paradox; highway capacity expansion; frequency; fare; profit-maximization; zero-profit.

1. Introduction

Traffic congestion is one of the major concerns in urban planning. Many countries extensively rely on supply-side policies to mitigate urban transportation congestion, such as improving network capacity provision. However, the arguments towards short-sighted capacity expansion are explosive after observing its implementation for about a century. Downs (1962) and Thomson (1977) first claimed that the expansion does not necessarily increase the travelers' travel utility when the public transit serves as a substitute for the highway. This can be explained as follows. The direct consequence of the expansion is attracting more private drivers, and the responsive transit authority would adjust the service frequency and ticket price accordingly. Unfortunately, the adjustment of the transit authority is often to reduce the service

* Corresponding author. Tel.: +852 2358 7178; fax: +852 2358 1534.
E-mail address: cehyang@ust.hk (H. Yang); fzhangad@connect.ust.hk (F. Zhang); wliuaa@connect.ust.hk (W. Liu).

frequency, which unintentionally lead to even fewer passengers (Reinhold, 2008; Bar-Yosef et al., 2013), and even more congested highway. In the end, travelers of both modes would suffer from the highway expansion. Raising fares is another option for a transit authority to maintain the operation when faced with highway expansion, which may also induce a similar vicious cycle as reducing frequency. In fact, the transit authority may adjust the frequency and fare simultaneously, yet it is still possible that its decisions, together with the highway expansion, yield counterproductive results.

The above phenomenon in which the generalized travel costs of both modes rise after highway improvement is named “the Downs-Thomson Paradox”, which is a substantial argument for people who are opposing arbitrary investment in highway systems. Many attempts have been made to investigate the paradox either empirically or theoretically (e.g., Holden, 1989; Arnott and Small, 1994; Abraham and Hunt, 2001; Denant-Boèmont and Hammiche, 2009; Liu et al., 2010; Hsu and Zhang, 2014). Particularly, Mogridge (1997) showed that for highway expansion to reduce travelers’ average travel cost in the two-mode system, the public transport should also be improved. In the same spirit, Arnott and Yan (2000) suggested that if the transit authority fails to take the effects of its decision on highway travelers fully into account, then the travel cost increase due to mode shift would exceed the benefit generated by the highway expansion. Both of their results indicate the substantial role of the transit side in the occurrence of the paradox. Basso and Jara-Díaz (2012) took one step further. They concretely analyze the paradox mechanism by capturing the properties of the average/marginal costs of both travelers and the transit authority, and showed that the paradox survives for transit users in the first-best world. Meanwhile, Bell and Wichiensin (2012) examined the occurrence of the paradox with different transit operators. However, in their study, the transit travel cost was a linear function of patronage with a downward slope, and neither the specified transit operating schemes nor the operation cost was taken into account.

However, when transit authority might have different economic objectives, e.g., profit-maximizing and zero-profit, how it will respond to highway expansion is still not clear. Whether or not, in which cases and under what conditions the Downs-Thomson Paradox would occur is still unknown. This paper tries to fill this gap by comprehensively investigating the occurrence of the Downs-Thomson Paradox when transit authority can have different operating strategies, i.e., the unconstrained case with variable frequency and variable fare, and the constrained cases with variable frequency and fixed fare or fixed frequency and variable fare. The analysis is conducted in the contexts of different economic objectives, i.e., profit-maximizing and zero-profit, which is the key factor for a transit authority in making decisions on its scheduling and pricing schemes. For various combinations of economic objectives and operating schemes, the interaction between highway expansion and transit service is explored, as well as its impact on travelers’ mode choices and travel utilities.

This study also benefits from the literature on the transit service problem (e.g. Mohring, 1972; Jansson, 1980; van Reeven, 2008; Savage and Small, 2010; Basso and Jara-Díaz, 2010; Karamychev and van Reeven, 2010). However, these models were built on the isolated transit system, rarely considering the two-mode interactions. For the two-mode problem, there is indeed a large pool of literature (e.g., Small, 1992; Kraus, 2003, 2012; Small and Verhoef, 2007; Light, 2009; Yang et al., 2009; Ahn, 2009), which provides the theoretical framework for our study here on the Downs-Thomson Paradox. Yet, the existing papers do not focus on the transit responses to the highway expansion and the impacts on the system, which are what we intend to unveil here.

The reminder of this paper is organized as follows. In Section 2, a two-mode transportation system and the concept of Downs-Thomson Paradox are introduced. Section 3 analyzes the impacts of highway expansion on both transit operating schemes and travelers' mode choices with different economic objectives, and then investigates the occurrence of the paradox. Numerical results of each scenario are correspondingly presented in Section 4 to illustrate the essential merits of the proposed models. Conclusions are given in Section 5.

2. Model Formulation

2.1 Problem settings

We consider the corridor with a congested highway running in parallel to an exclusive transit line, linking the residential area and the central business district (CBD), as shown in Figure 1. On a typical day, the total number of travelers commuting from home to the CBD is fixed at d. Travelers choose their travel modes fully based on the generalized travel costs of the two modes: transit and automobile.

Insert Figure 1 here

Transit users have to pay for the transit fare, spend time in the train and on platforms. Thus, the generalized travel cost of travelling by transit is:

$$p_t = \beta\left[w(f) + t_t\right] + \tau_t,$$

where f and τ_t denote the transit frequency and uniform ticket price (transit fare), respectively, and β represents the value of time and t_t stands for the constant in-vehicle travel time. $w(f)$ represents the waiting time at the platform, which is assumed to be a decreasing, convex and twice differentiable function of transit frequency. The operation cost of the single transit authority, $k(f)$, is an increasing, convex and twice differentiable function of frequency.

Throughout the analysis in this paper, it is assumed that the combination of transit fleet size and service frequency is large enough to carry all the travelers waiting on the platform.

For an auto commuter, the generalized travel cost consists of the travel time cost and the monetary cost, i.e.,

$$p_a = \beta t_a(v,c) + \tau_a,$$

where $t_a = t_a(v,c)$ is the highway travel time and τ_a is the monetary cost. It is assumed that $t_a(v,c)$ is homogeneous of degree zero with respect to traffic volume v and highway capacity c, i.e., $t_a(\rho v, \rho c) = t_a(v,c)$ for any $\rho > 0$. It is also assumed that $t'_v = \partial t_a(v,c)/\partial v > 0$, $t'_c = \partial t_a(v,c)/\partial c < 0$, and $t''_{vv} = \partial^2 t_a(v,c)/\partial v^2 > 0$, $t''_{vc} = \partial^2 t_a(v,c)/\partial v \partial c < 0$. The highway capacity c is subject to adjustment and is assumed to be a continuous variable[1], with the current capacity equals c_0. We assume all the automobile commuters drive alone, and thus the traffic volume v is equal to the number of auto commuters.

The monetary cost τ_a is considered as a non-negative constant, which includes, e.g., fuel consumption costs and parking fees. Note that the focus here is on the transit authority's scheduling and pricing responses to the highway capacity changes with different economic objectives, so the highway use is assumed to be free of toll charge (different from Basso and Jara-Díaz, 2012; Bell and Wichiensin, 2012). It is worth mentioning that future research may consider τ_a as a variable to incorporate the impact of highway toll charge on the discussed two-mode system here.

Assumption. Interior Equilibrium.

(i) A too congested highway if all choose to drive: $t_a(d,c) + \tau_a/\beta > w(f) + t_t + \tau_t/\beta$;

(ii) A too congested transit if no one drives: $t_a(0,c) + \tau_a/\beta < w(f) + t_t + \tau_t/\beta$.

Assumption (i) states that the transit mode becomes more attractive if all travelers choose to drive, thus there are at least some traveler choose transit. Similarly, Assumption (ii) means that the auto mode becomes more attractive if no one drives, thus there are at least some travelers choose to drive. The two assumptions, in together, ensure an interior equilibrium where both modes would be used, which is reasonable in reality and common in the literature (e.g. Light, 2009; Nie and Liu, 2010).

2.2 Deterministic mode choice and traffic equilibrium

[1] The assumption of continuity is standard in the theoretical literature which covers highway capacity provision (e.g., Arnott and Yan, 2000; Light, 2009). To focus on the analysis of the impact of transit scheduling and pricing schemes, neither highway investment nor operation cost would be taken into account.

Deterministic user equilibrium is achieved when no one can reduce his or her travel cost by shifting to an alternative travel mode, given the choices of other travelers. Thus, at equilibrium, the generalized travel costs of both modes are identical. Let p_c denote the equilibrium travel cost when the highway capacity is c, and based on the interior equilibrium assumption, p_c is then determined by the following condition:

$$p_c = p_t = p_a \text{, i.e., } p_c = \beta\left[w(f) + t_t\right] + \tau_t = \beta t_a(v,c) + \tau_a . \tag{1}$$

Lemma 1. *Under the interior equilibrium assumption, for any given* f*,* τ_t *and* c*, the equilibrium is unique.*

Proof. According to Eq.(1), for any given f, τ_t and c, at equilibrium, v is pinned down by: $h(v) = 0$, where $h(v) = \beta\left[t_a(v,c) - w(f) - t_t\right] + \tau_a - \tau_t$, which monotonically increases with v. From the interior equilibrium assumption, it is easy to see that $h(0) < 0$ and $h(d) > 0$. Hence v is unique, as well as the equilibrium.

Lemma 2. $E_f^v < 0$, $E_{\tau_t}^v > 0$ and $E_c^v > 0$, where E_f^v, $E_{\tau_t}^v$ and E_c^v are the elasticities[2] of highway travel demand v with respect to transit frequency f, transit fare τ_t and highway capacity c, respectively.

Proof. According to Eq.(1) and Lemma 1, the highway volume v can be regarded as a function of f, τ_t and c, i.e., $v = v(f, \tau_t, c)$, and the partial derivatives of v with respect to f, τ_t and c is given by

$$\frac{\partial v}{\partial f} = \frac{w'}{t'_v} < 0 , \tag{2}$$

$$\frac{\partial v}{\partial \tau_t} = \frac{1}{\beta t'_v} > 0 , \tag{3}$$

$$\frac{\partial v}{\partial c} = -\frac{t'_c}{t'_v} > 0 , \tag{4}$$

where $t'_v > 0$ and $t'_c < 0$ are the partial derivatives of highway travel time with respect to volume and capacity, respectively, and $w' < 0$ is the marginal waiting time with respect to transit frequency. Correspondingly, the sign of the elasticities can thus be determined: $E_f^v < 0$, $E_{\tau_t}^v > 0$ and $E_c^v > 0$.

[2] The elasticity is defined by the notion of point elasticity.

2.3 Downs-Thomson Paradox (D-T Paradox)

The D-T Paradox occurs in the sense that every individual traveler has to suffer from a higher generalized travel cost after the highway capacity expansion due to the direct and indirect effects of the volume shifting. As Downs (1962) and Thomson (1977) claimed, the direct consequence is attracting more private drivers. It is possible that the changes of transit scheduling and pricing schemes, together with the highway capacity expansion, yield a too congested highway and a lower level of transit service. Figure 2 shows the mechanism of the D-T Paradox, which follows Mogridge (1997) and Basso and Jara-Díaz (2012).

Insert Figure 2 here

In Figure 2, c_0 and c_1 are the road capacities before and after expansion ($c_0 < c_1$), v_0 and v_1 are the equilibrium volumes of the highway traffic, respectively. p_t and p_a represent the generalized travel cost of the individual traveler taking transit or driving an automobile, respectively. p_c is the individual travel cost at equilibrium given a certain highway capacity. Undoubtedly, the generalized travel cost of driving a private car increases with the highway traffic volume due to the congestion effect, so the curve of $p_a(c_1)$ is always lower than that of $p_a(c_0)$.

On the other hand, if the generalized travel cost by transit decreases with the total patronage, as shown by the upward dash curve of p_t, then the new equilibrium p_{c_1}, with a greater capacity provision, is higher than p_{c_0}. Figure 2 shows that a sufficient condition for the occurrence of the D-T Paradox is that the generalized travel cost by transit increases when the number of transit users decreases. This highly depends on the transit authority's responses to the passenger loss (e.g., the adjustments on frequency and fare). Before investigating these responses in different settings, we generally defines the occurrence of the D-T Paradox as follows:

Definition. The Downs-Thomson Paradox is said to occur at the equilibrium point of $c = c_0$, if there exists c_1 such that

$$p_{c_0} < p_{c_1}, \text{ with } c_0 < c_1, \tag{5}$$

where p_{c_0} and p_{c_1} are the equilibrium travel costs when the road capacity levels are at c_0 and c_1, respectively.

When the change in c, i.e., $\Delta c = c_1 - c_0$, is sufficiently small, the marginal effect of the change can be mathematically captured. The partial derivative of generalized travel cost with respect to the road capacity is negative, i.e., $\frac{\partial p_c}{\partial c} \cdot \Delta c < 0$, meaning that the direct effect of capacity expansion is to reduce the travel cost. To capture the total effect of the expansion, we evaluate the total derivative of the generalized travel costs of both modes at certain capacity levels (following Abraham and Hunt, 2001). When Δc is sufficiently small, Eq.(5) is equivalent to

$$\left.\frac{dp_c}{dc}\right|_{c=c_0} > 0 . \tag{6}$$

The general notion of the D-T Paradox is represented by Eq.(5), which is a comparison of the travel costs between two disjoint equilibrium states; however, Eq.(6) describes occurrence of the D-T Paradox at a given equilibrium point for a small change of highway capacity. If Eq.(6) holds at any point $c \in [c_0, c_1]$, then the equilibrium travel cost monotonically increases with the road capacity, which implies continuous occurrence of the D-T Paradox in the interval concerned. In this case Eq.(5) holds as well. Evidently, Eq.(5) does not necessarily imply continuous occurrence of the D-T Paradox. As shown in Figure 3, Eq.(5) is valid for all the curves named (a)-(d), but only the solid curve (c) corresponds to the continuous increase of p_c or Eq.(6) holds anywhere in the interval. For ease of examining the occurrence of the D-T Paradox, the following analysis is mainly in line with the definition given by Eq.(6).

Insert Figure 3 here

Lemma 3. The D-T Paradox occurs if and only if there exists $\Delta c > 0$ such that, for any Δf and $\Delta \tau_t$,

$$E_f^v \cdot \frac{v}{f} \cdot \Delta f + E_{\tau_t}^v \cdot \frac{v}{\tau_t} \cdot \Delta \tau_t > 0 . \tag{7}$$

Proof. The highway travel time function, $t_a = t_a(v, c)$, is homogeneous of degree zero with respect to traffic volume, v, and highway capacity, c, i.e.,

$$t_a(\rho v, \rho c) = t_a(v, c) \tag{8}$$

for any $\rho > 0$. Take the derivative of both sides of Eq.(8) with respect to ρ:

$$v \frac{\partial t_a(\rho v, \rho c)}{\partial(\rho v)} + c \frac{\partial t_a(\rho v, \rho c)}{\partial(\rho c)} = 0 . \tag{9}$$

Let $\rho = 1$ in Eq.(9), then

$$\frac{t'_c}{t'_v} = \frac{\partial t_a(v, c)}{\partial c} \Big/ \frac{\partial t_a(v, c)}{\partial v} = -\frac{v}{c} . \tag{10}$$

Therefore, the elasticity of highway travel demand with respect to highway capacity is equal to unit, i.e.,

$$E_c^v = \frac{\partial v}{\partial c} \cdot \frac{c}{v} = 1 . \tag{11}$$

An equivalent condition to the definition of the occurrence of the D-T Paradox at c is

$$p_c < p_{c+\Delta c} \Leftrightarrow t_a(v,c) < t_a(v+\Delta v, c+\Delta c) \Leftrightarrow \frac{v}{c} < \frac{v+\Delta v}{c+\Delta c} \Leftrightarrow \Delta v > \frac{v}{c}\Delta c . \tag{12}$$

And as the outcome of enlarged highway capacity, the change of highway demand is given by

$$\Delta v = E_f^v \cdot \frac{v}{f} \cdot \Delta f + E_{\tau_t}^v \cdot \frac{v}{\tau_t} \cdot \Delta\tau_t + E_c^v \cdot \frac{v}{c} \cdot \Delta c . \tag{13}$$

Therefore, Eqs.(11)-(13) ensure the validity of Eq.(7), which is the sufficient and necessary condition for the occurrence of the D-T Paradox.

Lemma 4. *With a fixed transit fare* τ_t^0*, the D-T Paradox occurs if and only if there exists* $\Delta c > 0$ *such that* $\Delta f < 0$*, i.e.,* $\Delta f \cdot \Delta c < 0$.

Lemma 5. *With a fixed transit frequency* f_0*, the D-T Paradox would occur if and only if there exists* $\Delta c > 0$ *such that* $\Delta\tau_t > 0$*, i.e.,* $\Delta\tau_t \cdot \Delta c > 0$.

Lemma 4 and 5 are the direct results of Lemma 3, where $\Delta\tau_t = 0$, $E_f^v < 0$ and $\Delta f = 0$, $E_{\tau_t}^v < 0$ in Eq.(7) respectively.

3. Paradox Conditions with Responsive Transit Service

In this section, we explore the occurrence of the D-T Paradox with the responsive transit service in different settings. Two typical economic objectives are considered for the transit authority: profit-maximizing and zero-profit (breakeven). Service frequency and fare are the decision variables for the transit authority to achieve certain economic objectives. We first consider the case where the transit authority adjusts the frequency and fare at the same time, which is the unconstrained case. Then we look into the constrained cases where either the service frequency or fare is regulated by the government, and the authority's choice is limited.

3.1 The profit-maximizing transit authority

In this subsection, we look at a transit authority who maximizes its net profit, which is:

$$\pi = \tau_t\left[d - v(f, \tau_t, c)\right] - k(f) \tag{14}$$

In Eq.(14), the first term gives the total fare revenue, which is the product of unit fare τ_t and patronage $(d - v)$, where $v(\cdot)$ is the equilibrium highway volume yielded from Eq.(1). The second term is the operation cost associated with the frequency of transit service.

3.1.1 *Unconstrained scheduling and pricing schemes*

We first consider the unconstrained case where the transit authority can adjust the frequency and fare simultaneously. For any given highway capacity, the optimal frequency and fare solves the following maximization problem with an equilibrium constraint:

$$\max_{f,\tau_t} \ \pi = \tau_t (d - v) - k(f) \tag{15}$$

subject to

$$\beta\left[w(f) + t_t\right] + \tau_t = \beta t_a (v,c) + \tau_a \tag{16}$$

For given c, the optimal $\left(f^*, \tau_t^*\right)$ must satisfy the first-order optimality conditions:

$$f^*: \tau_t \frac{\partial (d - v)}{\partial f} - k' = 0 \tag{17}$$

$$\tau_t^*: \tau_t \frac{\partial (d - v)}{\partial \tau_t} + d - v = 0 \tag{18}$$

Eq.(17) states that at the optimum, marginal revenue equals marginal operation cost if there is a marginal increase in the frequency. Eq.(18) similarly shows that marginal fare revenue gain completely offsets the revenue loss caused by the marginal increase in fare at the optimum.

The existence of the optimal solution is guaranteed by the continuity of the objective function and the compactness of the feasible region. Referring to the mathematical definition of the D-T Paradox given by Eq.(6), one can see that we are interested in the marginal adjustments of frequency and fare of the transit authority facing the marginal expansion of highway, and correspondingly the marginal changes of equilibrium volume and the individual travel cost. Under this consideration, we have the following proposition:

Proposition 1. *If the profit-maximizing transit authority gradually adjusts the service frequency and fare in response to the continuous increase of highway capacity, then with respect to the highway capacity,*
(i) individual travel cost monotonically decreases, implying the D-T Paradox never occurs;
(ii) transit authority's profit monotonically decreases.

Proof. The proof is given in Appendix A.1.

Proposition 1 implies that under the profit-maximizing objective, the transit side is also improved after the highway expansion and thus the travelers' generalized costs of both modes are reduced. This is because, if the transit authority's major concern is its own profit rather than travelers' utility, then the service originally provided by the authority is relatively unattractive to travelers, e.g., with low frequency or high fare, indicating travelers would experience relatively large travel cost (refer to Figure 9 in Section 4, which shows the numerical comparison of individual travel cost between profit-maximizing and zero-profit cases). In this case, when the highway is improved, the profit-maximizing transit authority does not suffer too much by increasing frequency or decreasing fare a little bit, but benefits a lot from maintaining patronage. Therefore, transit service is improved, as well as performance of the overall two-mode system, and D-T Paradox is avoided. However, although the profit is maximized under the capacity after expansion, it is still less than that before the expansion, which means the transit authority would experience a profit loss.

3.1.2 *Constrained scheduling scheme (fixed fare)*

It is often the case that public transport has a relatively fixed charging structure, e.g., some local governments may set a standard fare scheme for the public transit service. This subsection examines such a situation where the fare is fixed, and the transit authority is only able to choose its service frequency in order to maximize its profit. Given highway capacity c and the fixed fare τ_t^0, the optimal frequency solves the following problem:

$$\max_{f} \ \pi = \tau_t^0 (d - v) - k(f) \tag{19}$$

subject to

$$\beta\left[w(f) + t_t\right] + \tau_t^0 = \beta t_a(v, c) + \tau_a \tag{20}$$

Based on Lemma 4, we have the following proposition:

Proposition 2. *If the transit fare is fixed and the profit-maximizing transit authority gradually adjusts the service frequency in response to the continuous increase of highway capacity, then with respect to highway capacity,*
(i) optimal frequency monotonically increases;
(ii) individual travel cost monotonically decreases, implying that the D-T Paradox never occurs;
(iii) transit authority's profit monotonically decreases.

Proof. The proof is given in Appendix A.2.

Proposition 2 implies that the results in the unconstrained case (Proposition 1) sustains in the current case with a fixed fare. This is because, a profit-maximizing transit authority would originally provide relatively low frequent service to save the operation cost (note that fare is fixed). When highway is expanded by a small amount, the transit authority would not suffer too much by increasing frequency by a little bit, since $k'(f)$ is relatively small when f is relatively small, but would benefit a lot from maintaining its patronage. This then leads to improvement of system performance (for the whole two-mode system), and the D-T Paradox is avoided. Moreover, with a higher frequency and lower patronage, the transit authority's profit would evidently shrink when the highway is expanded. However, it is worth mentioning that the constrained case is not a special case of the unconstrained one while the reasoning is similar.

3.1.3 *Constrained pricing scheme (fixed frequency)*

We now turn to the parallel case with a fixed frequency, where the transit authority can only adjust the fare according to the change of highway capacity. Given highway capacity c and fixed frequency f_0, the optimal fare solves:

$$\max_{\tau_t} \ \pi = \tau_t (d - v) - k(f_0) \tag{21}$$

subject to

$$\beta\left[w(f_0) + t_t\right] + \tau_t = \beta t_a(v, c) + \tau_a \tag{22}$$

Based on Lemma 5, we have the following proposition:

Proposition 3. *If the service frequency is fixed and the profit-maximizing transit authority gradually adjusts the fare in response to the continuous increase of highway capacity, then with respect to highway capacity,*
(i) optimal fare monotonically decreases;
(ii) individual travel cost monotonically increases, implying that the D-T Paradox never occurs;
(iii) transit authority's profit monotonically decreases.

Proof. The proof is given in Appendix A.3.

Proposition 3 states that with a predetermined frequency, the transit fare is reduced when the highway is expanded. The reason is similar to that of the constrained case with a fixed fare. Originally, a profit-maximizing transit authority would charge a high fare to obtain a large revenue. When highway is expanded by a small amount, the transit authority would not suffer too much by reducing fare by a little bit, but would benefit a lot from maintaining its patronage. This then leads to improvement of system performance (for the whole two-mode system), and D-T Paradox is avoided. And with the lower fare and patronage, the transit authority inevitably experiences a profit loss.

To summarize, the D-T Paradox never occurs under a profit-maximizing transit authority, no matter its decisions are constrained or not. As mentioned before, it is because the transit authority would always improve its service (e.g., lower fare or higher frequency), in order to compete with the improved highway and maintain patronage. Therefore, travelers will always benefit from highway expansion, while transit authority loses from that even if its profit is maximized after the expansion.

3.2 *The zero-profit transit authority*

If the transit is operated by a revenue-neutral authority (e.g., publicly owned and not profit-seeking), the transit service is then running to maintain breakeven:

$$\pi = \tau_t \left[d - v(f, \tau_t, c) \right] - k(f) = 0 . \tag{23}$$

where $v(f, \tau_t, c)$ is the equilibrium number of auto travelers. Now we examine the responses of such a transit authority to the highway expansion under both unconstrained and constrained cases.

3.2.1 *Unconstrained scheduling and pricing schemes*

In the unconstrained case, the transit authority is able to adjust the frequency and fare simultaneously. Evidently, for given c, f and τ_t cannot be uniquely determined simply by the zero-profit condition Eq.(23). However, when the revenue-neutral transit authority takes the travelers' utility into account, the combination $(f, \tau_t)^*$ that minimizes travelers' travel cost and satisfies the zero-profit condition at the same time would be unique. This is shown in Appendix A.4. We would like to point out that social welfare is likely to be within the concerns of a revenue-neutral authority. Then it is reasonable to consider the situation where the zero-profit transit authority chooses the $(f, \tau_t)^*$ combination to minimize travelers' generalized travel cost. Under this consideration, for any given highway capacity c, the problem for the transit authority is:

$$\min_{f, \tau_t} \ p_c = \tau_t + \beta \left[w(f) + t_t \right] \tag{24}$$

subject to

$$\tau_t (d - v) - k(f) = 0 \tag{25}$$

$$\beta \left[w(f) + t_t \right] + \tau_t = \beta t_a (v, c) + \tau_a \tag{26}$$

The objective (24) is to minimize the travelers' generalized cost. The first constraint (25) ensures the transit authority maintains break-even, and the second constraint (26) corresponds to the equilibrium condition. The existence and uniqueness of the optimal $(f, \tau_t)^*$ are shown in Appendix A.4, and we have the following proposition:

Proposition 4. *If the zero-profit transit authority gradually adjusts the service frequency and fare in response to the continuous increase of highway capacity to minimize individual travel cost of travelers, then the individual travel cost monotonically increases with the highway capacity and the D-T Paradox occurs whenever the highway capacity increases.*

Proof. The proof is given in Appendix A.5.

Proposition 4 implies that, facing the highway expansion, the change in individual travel cost is disagreeable if the zero-profit transit authority minimizes the individual travel cost. When the highway is improved, the transit service loses passengers and the breakeven is no longer maintained under the current frequency and fare. Therefore, the transit authority has to compromise the service quality (reduce frequency or raise fare or both) to keep within budget. Even though the transit authority provides the best service for the travelers given the current highway capacity, the generalized travel cost increases compared with that before highway expansion, i.e., the D-T Paradox occurs.

Recall in the scenario of the unconstrained profit-maximizing transit schemes in Section 3.1.1, the D-T Paradox is avoided. Our results here do not mean that a profit-maximizing transit authority is preferred in terms of social welfare. As shown in the numerical examples (Section 4.2.1), zero-profit transit responses indeed generate more benefit for the travelers, because the absolute value of the individual travel cost produced in this scenario is lower than that with a profit-maximizing transit authority.

However, if revenue-neutral authority is not going to minimize individual travel cost, there would be multiple combinations of f and τ_t that solve Eq.(23). Figure 4 is employed to show a representative case[3] where the generalized costs of all the feasible (f,τ_t) combinations under a given c constitute a circle, while the vertical ranges of different circles are inclusive of each other. Obviously, the point with the lowest generalized cost corresponds to the most socially preferable choice $(f,\tau_t)^*$. If we could figure out the change of $(f,\tau_t)^*$ for different c, then the comparison of other (f,τ_t) combinations is trivial but without too much insight, which is omitted.

Insert Figure 4 here

3.2.2 *Constrained scheduling scheme (fixed fare)*

To maintain the breakeven with a fixed fare, transit frequency must satisfy

$$\pi = \tau_t^0\left[d - v(f,c)\right] - k(f) = 0\,, \tag{27}$$

where $v = v(f,c)$ is derived from Eq.(1) with $\tau_t = \tau_t^0$. Again, the solution of Eq.(27) may not be unique while the number of solutions is finite. For example, if the curves in Figure 4 are projected into the (f,τ_t)-plane, then the projection would have two intersections with the line $\tau_t = \tau_t^0$, and each intersection corresponds to a zero-profit frequency. Firstly, for ease of presentation, we define the following function regarding highway capacity:

[3] The function specifications and parameters used for this example are given in Table 1 and 2 in Section 4.

$$I_1(c) = k' + \tau_t^0 w'/t_v' . \tag{28}$$

According to Eq.(2), w'/t_v' is $\partial v/\partial f$, which can be rearranged as $-\partial(d-v)/\partial f$. Then, we immediately have $I_1(c) = k' - \tau_t^0 \cdot \partial(d-v)/\partial f$, which is the difference between the marginal operation cost and the marginal fare revenue, given the fixed fare τ_t^0 .

Proposition 5. *If the transit fare is fixed at* τ_t^0 *and the authority gradually adjusts the service frequency in response to the continuous increase of highway capacity* c *to maintain zero-profit, then,*

(i) the D-T Paradox occurs if and only if $I_1(c) > 0$ *;*

(ii) given $I_1(c_0) > 0$ *at* $c = c_0$*,* $I_1(c)$ *monotonically decreases with* c*, and the D-T Paradox occurs in and only in the interval of* $[c_0, \hat{c}]$*, where* $\hat{c}$ *is the solution to* $I_1(c) = 0$ *;*

(iii) particularly, if the authority provides the maximal frequency to minimize individual travel cost of travelers, then the individual travel cost monotonically increases with the highway capacity and the D-T Paradox occurs whenever the highway capacity increases.

Proof. The proof is given in Appendix A.6.

Proposition 5(i) establishes the sufficient and necessary condition for the occurrence of the D-T Paradox for the case of a fixed transit fare and a zero-profit transit authority. $I_1(c) > 0$ implies that the marginal operation cost exceeds the marginal fare revenue if the transit authority increases frequency. Therefore, when there is a marginal increase in highway capacity, the transit authority would reduce, rather than increase, its service frequency to maintain zero-profit. Furthermore, Proposition 5(ii) states that the D-T Paradox would continuously occur in the interval of $[c_0, \hat{c}]$ if it occurs at the initial point c_0. Otherwise, it would not occur.

In addition, Proposition 5(iii) indicates the similar result of Proposition 4 can be extended to this case with a fixed fare constraint. Also it reflects the relativity of the D-T Paradox: since the zero-profit transit authority chooses the maximal frequency, the individual travel cost must be the lowest; however, it cannot be further reduced when the highway capacity is enlarged, i.e., the D-T Paradox is inevitable.

3.2.3 *Constrained pricing scheme (fixed frequency)*

In the parallel scenario with a predetermined frequency, the zero-profit transit authority chooses the fare that satisfies:

$$\pi = \tau_t \left[d - v(\tau_t, c) \right] - k(f_0) = 0 , \tag{29}$$

where $v = v(\tau_t, c)$ is derived from Eq.(1) with $f = f_0$. For ease of presentation, we define the following function regarding highway capacity:

$$I_2(c) = (d - v) - \tau_t / \beta t_v' . \tag{30}$$

According to Eq.(3), $1/\beta t_v'$ is $\partial v / \partial \tau_t$, which can be rearranged as $-\partial(d - v)/\partial \tau_t$. Then, we immediately have $I_2(c) = (d - v) + \tau_t \cdot \partial(d - v)/\partial \tau_t$, which is the marginal change in profit if there is a marginal increase in τ_t.

Proposition 6. *If the frequency is fixed at* f_0 *and the transit authority gradually adjusts the fare in response to the continuous increase of highway capacity* c *to maintain zero-profit, then,*

(i) the D-T Paradox occurs if and only if $I_2(c) > 0$*;*

(ii) given $I_2(c_0) > 0$ *at* $c = c_0$*,* $I_2(c)$ *monotonically decreases with* c*, and the D-T Paradox would occur in and only in the interval of* $[c_0, \tilde{c}]$*, where* $\tilde{c}$ *is the solution to* $I_2(c) = 0$*;*

(iii) particularly, if the authority provides the minimal fare to minimize individual travel cost of travelers, then the individual travel cost monotonically increases with the highway capacity and the D-T Paradox occurs whenever the highway capacity increases.

Proof. The proof is given in Appendix A.7.

Proposition 6(i) establishes the sufficient and necessary condition for the occurrence of the D-T Paradox for the case of a fixed transit frequency and a zero-profit transit authority. $I_2(c) > 0$ implies that the profit decreases if the transit authority reduces the fare. Therefore, when there is a marginal increase in highway capacity, the transit authority would increase, rather than reduce, the fare to maintain zero-profit. Furthermore, Proposition 6(ii) states that the D-T Paradox would continuously occur in the interval of $[c_0, \tilde{c}]$ if it occurs at the initial point c_0. Otherwise, it would not occur. Proposition 6(iii) indicates the similar results of Proposition 4 and 5 can be extended to the case with a fixed frequency.

4. An Illustrative Example

To facilitate the presentation of the essential ideas, we employ specific functions to derive analytical solution of the paradox condition and occurrence region under each of the scenarios described in the last section. Consider the two-mode network described in Section 2 with function specifications given in Table 1 and parameter values in Table 2.

Insert Table 1 here

Insert Table 2 here

According to Eq.(1), for any given transit frequency, fare and highway capacity, the equilibrium highway volume is:

$$v = c\left(\frac{1/2\,f + (\tau_t - \tau_a)/\beta + t_t - t_0}{t_0 t_1}\right)^{\frac{1}{\alpha}}. \tag{31}$$

4.1 The profit-maximizing transit authority

4.1.1 *Unconstrained scheduling and pricing schemes*

With function specifications, the profit-maximizing combination of $\left(f^*, \tau_t^*\right)$ solves for:

$$f^*:\quad \frac{\tau_t c}{2\alpha t_0 t_1 f^2}\left(\frac{A_1}{t_0 t_1}\right)^{\frac{1}{\alpha}-1} - (2k_2 f + k_1) = 0,$$

$$\tau_t^*:\quad d - c\left(\frac{A_1}{t_0 t_1}\right)^{\frac{1}{\alpha}} - \frac{\tau_t c}{\alpha\beta t_0 t_1}\left(\frac{A_1}{t_0 t_1}\right)^{\frac{1}{\alpha}-1} = 0.$$

where $A_1 = 1/2\,f + (\tau_t - \tau_a)/\beta + t_t - t_0$. With the parameters in Table 2, Figure 5(a) shows the transit authority's adjustments on frequency and fare for profit maximization when the highway capacity is increased. It is observed that along with the increase of the highway capacity from 2000veh/h to 6000veh/h, the transit fare (the dotted line) declines sharply compared to the decrease of the service frequency (the dash line): the drop of transit fare directly reduces the travel cost from 1200HK$ to nearly zero, while the slight decrement in the service frequency is only equivalent to increasing the waiting time cost from 2.47HK$ to 3.37HK$. Therefore, as the combined impact of highway capacity expansion and transit authority's responses, the individual traveler's generalized travel cost diminishes with the highway capacity as shown in Figure 5(b), which indicates the D-T Paradox is avoided. Meanwhile, the transit profit also decreases as shown by the dash-dotted line.

Insert Figure 5 here

4.1.2 Constrained scheduling scheme (fixed fare)

If the transit fare is fixed at τ_t^0, the first-order condition for the optimal f^* becomes:

$$f^*:\quad \frac{\tau_t^0 c}{2\alpha t_0 t_1 f^2}\left(\frac{A_2}{t_0 t_1}\right)^{\frac{1}{\alpha}-1} - (2k_2 f + k_1) = 0,$$

where $A_2 = 1/2\,f + (\tau_t^0 - \tau_a)/\beta + t_t - t_0$. The relation between f^* and c is captured by:

$$\frac{dc}{df} = \left(\frac{A_2}{t_0 t_1}\right)^{-\frac{1}{\alpha}} \left(\tau_t^0\right)^{-1} \left[12\alpha k_2 A_3 f^2 + \left(4\alpha k_2 + 4\alpha k_1 A_3 + 2k_2\right) f + k_1(\alpha+1)\right] > 0 ,$$

where $A_3 = \left(\tau_t^0 - \tau_a\right)/\beta + t_t - t_0$. According to Lemma 4, this analytically indicates the nonexistence of the D-T Paradox. Numerically, Figure 6 presents the changes of frequency and individual travel cost when the fixed fare is 5HK$ (the dotted line) and the highway capacity increases from 2000veh/h to 6000veh/h. It is found that the frequency (the dash line) is increased, and both the equilibrium travel cost (the solid line) and the transit profit (the dash-dotted line) are correspondingly reduced. It is easy to understand the phenomenon, because the profit-maximizing transit authority has to improve its service to compete with the improving highway.

Insert Figure 6 here

4.1.3 Constrained pricing scheme (fixed frequency)

In parallel, with a fixed frequency f_0, the profit-maximizing τ_t^* solves for:

$$\tau_t^* : \; d - c\left(\frac{A_4}{t_0 t_1}\right)^{\frac{1}{\alpha}} - \frac{\tau_t c}{\alpha\beta t_0 t_1}\left(\frac{A_4}{t_0 t_1}\right)^{\frac{1}{\alpha}-1} = 0$$

where $A_4 = 1/2\, f_0 + \left(\tau_t - \tau_a\right)/\beta + t_t - t_0$. The relation between τ_t^* and c is captured by:

$$\frac{dc}{d\tau_t} = -d\left(\frac{A_4}{t_0 t_1}\right)^{-\frac{2}{\alpha}} \left[1 + \frac{\tau_t c}{\alpha\beta A_4}\right]^{-2}$$

$$\cdot\left[\frac{1}{\alpha\beta t_0 t_1}\left(\frac{A_4}{t_0 t_1}\right)^{\frac{1}{\alpha}-1}\left(1 + \frac{\tau_t c}{\alpha\beta A_4}\right) + \left(\frac{A_4}{t_0 t_1}\right)^{\frac{1}{\alpha}} \frac{c\left(1/2\, f_0 + t_t - t_0 - \tau_a/\beta\right)}{\alpha\beta A_4^2}\right]$$

which is strictly negative, so according to Lemma 5, it analytically indicates the nonexistence of the D-T Paradox. Figure 7 shows the numerical results when the frequency is fixed at 0.5run/min (the dash line) and the highway capacity increases from 2000veh/h to 6000veh/h: the fare (the dotted line) drops, and both the equilibrium travel cost (the solid line) and the transit profit (the dash-dotted line) are correspondingly reduced. Similar to the action of service improvement, the reduction in transit fare is also implemented to attract passengers.

Insert Figure 7 here

4.2 The zero-profit transit authority

4.2.1 *Unconstrained scheduling and pricing schemes*

In the case of zero-profit operation, the transit frequency and fare should satisfy:

$$\left[\beta + 2f\left(\tau_t - \tau_a\right) + 2\beta f\left(t_t - t_0\right)\right] c^{\alpha}\tau_t^{\alpha} - 2\beta t_0 t_1 f\left[\tau_t d - \left(k_2 f^2 + k_1 f^1 + k_0\right)\right]^{\alpha} = 0,$$

and obviously for any given capacity c, the combination of (f, τ) is not unique. Figure 4 in Section 3.2.1 has shown the multiple choices at certain highway capacity. If the zero-profit transit authority chooses the (f, τ_t) combination to minimize the generalized travel cost of travelers that solves problem (24)-(26), then it would reduce the frequency (the dash line) and raise the fare (the dotted line) in response to the highway expansion as shown in Figure 8. As a consequence, the individual travel cost (the solid line) goes upwards with the highway capacity, which indicates that the D-T Paradox continuously occurs. When compared with the results in Section 4.1.1, the absolute value of the individual travel cost in this the current case is much lower than that generated with the profit-maximizing transit authority, and Figure 9 displays the comparison. Note that the dash-dot line is sharply and monotonically decreasing with the highway capacity in the whole range of $c \in [2000, 3000]$ (veh/h), but the absolute value is rather high compared to the normal range, so that part of the dash-dot is cut off from Figure 9 without generality. In the zero-profit scenario, the highway volume reaches the upper-bound (the total demand) when the highway capacity approaches 5900veh/h.

Insert Figure 8 here

Insert Figure 9 here

4.2.2 Constrained scheduling scheme (fixed fare)

With a fixed fare τ_t^0, the relation between f and c derived from Eq.(27) is

$$c(f) = \frac{\tau_t^0 d - \left(k_2 f^2 + k_1 f + k_0\right)}{\tau_t^0 \left(A_2 / t_0 t_1\right)^{1/\alpha}},$$

and the D-T Paradox occurs if and only if

$$\frac{dc}{df} = -\frac{4\alpha k_2 A_5 f^3 + \left[k_2\left(2\alpha+1\right) + 2\alpha k_1 A_5\right] f^2 + k_1\left(\alpha+1\right) f + k_0 - \tau_t^0 d}{2\alpha\tau_t^0 A_2 f^2 \left(A_2/t_0 t_1\right)^{1/\alpha}} < 0$$

$$\Leftrightarrow 4\alpha k_2 A_5 f^3 + \left[k_2\left(2\alpha+1\right) + 2\alpha k_1 A_5\right] f^2 + k_1\left(\alpha+1\right) f + k_0 - \tau_t^0 d > 0 \qquad (32)$$

$$\Leftrightarrow f \in \left(\tilde{f}, \infty\right)$$

where $A_5 = 1/2f - \tau_a/\beta + t_t - t_0$, and $\tilde{f}$ is the unique positive solution to the left-hand side of Eq.(32), which corresponds to the margin of the paradox region. Figure 10 presents the transit frequency response to the highway expansion and the resulting individual travel cost with a fixed fare 5HK$ and a zero-profit transit authority. When the highway capacity is less than 5275veh/h, there are two feasible values of the frequency satisfying the zero-profit condition (the solid and dash lines in Figure 10(a)), but the larger value decreases (the solid line) while the smaller value increases (the dash line) along with the increase of highway capacity. As a result, if the zero-profit authority chooses the higher frequency to minimize the traveler's generalized cost, the individual travel cost would actually increase with the highway capacity (the solid line in Figure 10(b)), which induces the occurrence of the D-T Paradox. On the other hand, the lower but increasing frequency corresponds to the decreasing individual travel cost. However, one should note that the absolute value of the individual travel cost with the lower frequency is much higher than the one with higher frequency, even if it is decreasing and avoids the paradox. Note that when the highway capacity exceeds 5275veh/h, it is large enough to attract all the travelers and there is no feasible frequency that can maintain zero-profit.

Insert Figure 10 here

4.2.3 Constrained pricing scheme (fixed frequency)

In parallel, the relation between τ_t and c with a fixed f_0 is:

$$c(\tau_t) = \frac{\tau_t d - K}{\tau_t}\left(\frac{A_4}{t_0 t_1}\right)^{-\frac{1}{\alpha}},$$

where K denotes the fixed transit operation cost with f_0, and the D-T Paradox occurs if and only if

$$\begin{aligned} &\frac{dc}{d\tau_t} = \frac{K - \tau_t(\tau_t d - K)/(\alpha\beta A_4)}{\tau_t^2 (A_4/t_0 t_1)^{1/\alpha}} > 0 \\ &\Leftrightarrow \tau_t^2 d - (\alpha + K)\tau_t - \alpha\beta(1/2\, f_0 - \tau_a/\beta + t_t - t_0) < 0 \\ &\Leftrightarrow \tau_t \in (0, \tilde{\tau}_t) \end{aligned} \tag{33}$$

where $\tilde{\tau}_t$ corresponds to the margin of the paradox region, and it is the only positive root of the left-hand side of Eq.(33):

$$\tilde{\tau}_t = \frac{(\alpha + K) + \sqrt{(\alpha + K)^2 + 4d\alpha\beta(1/2\, f_0 - \tau_a/\beta + t_t - t_0)}}{2d}.$$

Figure 11 shows the numerical observations when the transit frequency is fixed at 0.5run/min. It is found that when highway capacity is less than 5671veh/h, there are two feasible values of the transit fare satisfying the zero-profit condition (the solid and dash lines in Figure 11(a)), and if the zero-profit authority chooses the lower fare to minimize the traveler's generalized cost, the individual travel cost would actually increase with the highway capacity (the solid line in Figure 11(b)), which induces the occurrence of the D-T Paradox. However, the absolute travel cost with the larger fare is much higher than that with the lower fare. When the highway capacity exceeds 5671veh/h, it is large enough to attract all the demand and there is no feasible fare that maintains zero-profit for the transit.

Insert Figure 11 here

5. Conclusions

This study investigates the occurrence of the well-known "Downs-Thomson Paradox", by taking into account the transit authority with different economic objectives in the two-mode system. In the context of highway expansion, the relation between modes share and transit polices is analysed, and the impact of highway capacity change on transit policies is explored.

In both unconstrained and constrained (either fare or frequency is fixed) cases, it is found that the D-T Paradox never occurs with a profit-maximizing authority. This is because, the transit authority improves its service (e.g., raising frequency, reducing fare) to attract passengers when the highway is expanded. Thus, the overall performance of the two-mode system is improved once highway expands. However, for the transit authority, its profit decreases with the highway capacity, so it has to suffer from profit loss after the highway expansion.

On the contrary, when the transit authority minimizes individual travel cost while maintains break-even, individual travel cost rises whenever highway capacity increases. This is because, transit service is already in relatively good quality (relatively low fare and high efficiency) with such a transit authority. Once the highway expands, it is impossible for the authority to maintain break-even while keeps or even improves service quality. In other words, highway expansion hurts the transit side, as well as the overall system performance. This calls for the comprehensive consideration for highway expansion and development in the multi-modal transportation system. Particularly, the public transit service side should be taken into account.

However, our results do not mean that a profit-maximizing transit authority is preferred in terms of social welfare. As shown in the numerical examples, zero-profit transit responses indeed generate more benefit for the travelers, because the absolute value of the individual travel cost produced in this scenario is lower than that with a profit-maximizing transit authority.

This study conceptually links the Downs-Thomson Paradox with the "Mohring Effect". Mohring (1972) noted that normally the transit authority would provide less frequent service when the demand is lower, which is referred to as the "Mohring Effect". This effect is due to the scale economies of the transit service[4]. In fact, if the transit authority reduces the service frequency when the patronage drops after the highway expansion, it would make all the

[4] The "scale economies" of the transit service refers to the situation where the generalized cost by transit decreases if the number of passenger increases.

travelers suffer. This means that, in the presence of Mohring Effect of transit services, inappropriate actions such as expansion of highway capacity could result in the occurrence of the Downs-Thomson Paradox.

Future research may conduct more elaborate analysis on the relation between the Mohring Effect and the Downs-Thomson Paradox. Besides, it is of our interest to investigate strategies that can effectively prevent the occurrence of the paradox (e.g., congestion pricing, transit subsidy, transit service regulations). Also, travelers' heterogeneity in value of time could be incorporated into the current framework. In that case, the analysis of the paradox occurrence would be more challenging since the changes of the travel costs are no longer identical among travelers.

Acknowledgements. The authors would like to thank Prof. Richard Arnott for providing useful suggestions during his visit to HKUST in December 2013 and Dr. Leonardo J. Basso for his insightful comments at the 2014 ITEA Annual Conference. We are very thankful to the Editor and the anonymous referees for their valuable feedback to improve the quality and clarity of the paper. The work described in this paper was supported by grants from Hong Kong's Research Grants Council under project no. HKUST620712.

Appendix A.

A.1. Proof for Proposition 1

Problem (15)-(16) is equivalent to:

$$\max_{f,\tau_t} \ \pi = \tau_t\left[d - v\left(f,\tau_t,c\right)\right] - k\left(f\right)$$

where $v\left(f,\tau_t,c\right)$ is pinned down by Eq.(1). The necessary conditions for an interior optimal $\left(f^*,\tau_t^*\right)$ are:

FOC: $$f^*: \tau_t w' + t_v' k' = 0\ , \tag{34}$$

$$\tau_t^*: \beta t_v'\left[d - v\left(f,\tau_t,c\right)\right] - \tau_t = 0\ ; \tag{35}$$

SOC: $$\pi'' = \begin{pmatrix} \pi''_{ff} & \pi''_{f\tau_t} \\ \pi''_{f\tau_t} & \pi''_{\tau_t\tau_t} \end{pmatrix} \le 0\ .$$

The Hessian matrix π'' is negative semi-definite if and only if:

$$\pi''_{ff} = \frac{-\tau_t w'' t_v'^2 + \tau_t w'^2 t_v'' - k'' t_v'^3}{t_v'^3} \le 0\ ,$$

$$\pi''_{\tau_t\tau_t} = \frac{\tau_t t_v'' - 2\beta t_v'^2}{\beta^2 t_v'^3} \le 0\ ,$$

$$\pi''_{ff}\pi''_{\tau_t\tau_t} - \pi''^2_{f\tau_t} = \frac{\left(2\beta t_v'^2 - \tau_t t_v''\right)\left(\tau_t w'' + k'' t_v'\right) - \beta^2 t_v'^2 w'^2}{\beta^2 t_v'^4} \geq 0 .$$

Thus necessary condition for an interior optimal $\left(f^*, \tau_t^* \right)$ are:

$$\tau_t w'' t_v'^2 + k'' t_v'^3 - \tau_t w'^2 t_v'' \geq 0$$

$$2\beta t_v'^2 - \tau_t t_v'' \geq 0$$

$$\left(2\beta t_v'^2 - \tau_t t_v''\right)\left(\tau_t w'' + k'' t_v'\right) - \beta^2 t_v'^2 w'^2 \geq 0 .$$

Given the first-order conditions of the optimal $\left(f^*, \tau_t^* \right)$, the local adjustments of $\left(f^*, \tau_t^* \right)$ when there is a marginal change in c can be captured by taking total derivative of the both sides of (34) and (35) with respect to c:

$$\frac{d\tau_t^*}{dc} = \frac{\beta^2 w'^2 t_c' \left(2\beta t_v'^2 - \tau_t t_v''\right)}{\left(2\beta t_v'^2 - \tau_t t_v''\right)\left(\tau_t w'' + k'' t_v'\right) - \beta^2 w'^2 t_v'^2}$$

$$\frac{df^*}{dc} = \frac{-\beta w' t_c' \left(\beta t_v'^2 - \tau_t t_v''\right)}{\left(2\beta t_v'^2 - \tau_t t_v''\right)\left(\tau_t w'' + k'' t_v'\right) - \beta^2 w'^2 t_v'^2} .$$

Thus,

$$\frac{dp_c^*}{dc} = \frac{d\tau_t^*}{dc} + \beta w'\left(f^* \right)\frac{df^*}{dc} = \frac{\beta^3 w'^2 t_v'^2 t_c'}{\left(2\beta t_v'^2 - \tau_t t_v''\right)\left(\tau_t w'' + k'' t_v'\right) - \beta^2 w'^2 t_v'^2} < 0 ,$$

and

$$\frac{d\pi^*}{dc} = \left[(d - v) - \frac{\tau_t}{\beta t_v'} \right]\frac{d\tau_t^*}{dc} - \left(\tau_t \frac{w'}{t_v'} + k' \right)\frac{df^*}{dc} + \frac{t_c'}{t_v'} \cdot \tau_t = \frac{t_c'}{t_v'} \cdot \tau_t < 0 . \tag{36}$$

Therefore, the individual travel cost determined by the optimal frequency and fare monotonically decreases with the highway capacity, and this means that the D-T Paradox would never occur no matter how much the highway capacity is expanded. However, the transit authority would suffer from profit loss.

A.2. Proof for Proposition 2

When the local adjustment of f^* is implemented when there is a marginal change in c, the transit authority is actually optimizing the reaction function $f(c)$ as the following:

$$\max_{f(c)} \pi = \tau_t^0 \left[d - v(f, c) \right] - k(f),$$

where $v(f, c)$ is drawn from the equilibrium constraint. By applying Hamilton method, the optimal $f^*(c)$ can be derived by Euler equation:

$$f^*(c): \ \frac{d}{dc}\pi_{\dot{f}} = \pi_f \Leftrightarrow \tau_t^0 w' + k' t_v' = 0 .$$

According to Lemma 4, to prove the non-existence of the D-T Paradox, it is equivalent to prove

$$\frac{df^*}{dc} = -\frac{k'}{\tau_t^0 w'' + k'' t_v'}\delta > 0 ,$$

where $\delta = t_{vv}''\partial v/\partial c + t_{vc}''$ for all c. Since $k' > 0$ and $\tau_t^0 w'' + k'' t_v' > 0$, it suffices to show that $\delta < 0$. Note that $t_{vv}'' > 0$ and $t_{vc}'' < 0$, and according to Eq.(10),

$$\frac{t_c'}{t_v'} = -\frac{v}{c} .$$

Then we have

$$\delta = t_{vv}''\frac{\partial v}{\partial c} + t_{vc}'' = \frac{t_v' t_{vc}'' - t_c' t_{vv}''}{t_v'} = t_v' \frac{\partial}{\partial v}\left(\frac{t_c'}{t_v'}\right) < 0 .$$

Moreover, according to Eq.(36)

$$\frac{d\pi^*}{dc} = -\left(\tau_t \frac{w'}{t_v'} + k'\right)\frac{df^*}{dc} + \frac{t_c'}{t_v'}\cdot\tau_t = \frac{t_c'}{t_v'}\cdot\tau_t < 0$$

A.3. Proof for Proposition 3

Problem (21)-(22) can be rewritten as

$$\max_{\tau_t(c)} \pi = \tau_t\left[d - v(\tau_t, c)\right] - k(f_0) ,$$

where $v(\tau_t, c)$ is drawn from the equilibrium constraint. By applying Hamilton method, the optimal $\tau_t^*(c)$ can be derived by Euler equation:

$$\tau_t^*(c): \ \frac{d}{dc}\pi_{\dot{\tau}_t} = \pi_{\tau_t} \Leftrightarrow \beta t_v'\left[d - v(\tau_t, c)\right] - \tau_t = 0 .$$

According to Lemma 5, the non-existence of the D-T Paradox can be shown by

$$\frac{d\tau_t^*}{dc} = \frac{\beta\left(t_{vc}'' - t_v'\partial v/\partial c\right)}{\beta t_v' \, \partial v/\partial \tau_t + 1} = \frac{\beta\left(t_{vc}'' + t_c'\right)}{2} < 0 .$$

Moreover, according to Eq.(36)

$$\frac{d\pi^*}{dc} = \left[(d - v) - \tau_t \frac{1}{\beta t_v'}\right]\frac{d\tau_t^*}{dc} + \frac{t_c'}{t_v'}\cdot\tau_t = \frac{t_c'}{t_v'}\cdot\tau_t < 0 .$$

A.4. Proof for the existence and uniqueness of the optimal solution of Problem (24)-(26)

Problem (24)-(26) can be reduced to:

$$\min_{f,\tau_t} \ p_c = \beta t_a\left(d - \tfrac{k(f)}{\tau_t}, c\right) \tag{37}$$

subject to

$$\tau_t + \beta\left[w(f) + t_t\right] - \tau_a - \beta t_a\left(d - \tfrac{k(f)}{\tau_t}, c\right) = 0 \tag{38}$$

The existence of the optimal solution is ensured by the continuity of the objective function and the compactness of the feasible region. The uniqueness of the optimal solution can be shown by contradiction. Suppose there are two distinct optimal solutions $\left(f_1^*, \tau_{t1}^*\right)$ and $\left(f_2^*, \tau_{t2}^*\right)$, such that $f_1^* \neq f_2^*$ or $\tau_{t1}^* \neq \tau_{t2}^*$. Evidently, the objective values of the two optimal solutions should be identical, i.e., $p_{c1}^* = p_{c2}^*$ and $\beta t_a\left(d - \frac{k\left(f_1^*\right)}{\tau_{t1}^*}, c\right) = \beta t_a\left(d - \frac{k\left(f_2^*\right)}{\tau_{t2}^*}, c\right)$. Then it can be obviously concluded from the constraint that $f_1^* = f_2^*$ must hold. The monotonicity of function $t_a\left(d - \frac{k(f)}{\tau_t}, c\right)$ with respect to τ_t implies that the corresponding τ_t^* with identical f^* should also be equal, i.e., $\tau_{t1}^* = \tau_{t2}^*$; otherwise, the two optimal solutions would lead to distinct objective values. This contradicts with the assumption that the two optimal solution s are distinct.

A.5. Proof for Proposition 4

From the zero-profit constraint of problem (24)-(26) it is obvious that transit fare can be expressed as a function of frequency f and highway capacity c, i.e., $\tau_t = \frac{k(f)}{d - v(f,c)}$. Substituting into the equilibrium constraint, we have

$$k(f) + \beta\left[d - v(f,c)\right]\left[w(f) + t_t - t_a\left(v(f,c),c\right)\right] - \tau_a = 0,$$

and the partial derivative of traffic volume $v(f,c)$ with respect to frequency and capacity are:

$$\frac{\partial v}{\partial f} = \frac{k'(f)\left[d - v(f,c)\right] + \beta w'(f)\left[d - v(f,c)\right]^2}{\beta t_v'\left(d - v(f,c)\right)^2 - k(f)},$$

and

$$\frac{\partial v}{\partial c} = \frac{-\beta t_c'\left[d - v(f,c)\right]^2}{\beta t_v'\left[d - v(f,c)\right]^2 - k(f)}.$$

Evidently, $\partial v / \partial c > 0$ is true by the monotonicity of function $t_a(v,c)$, then it is implied that:

$$\varepsilon = \beta t_v'\left[d - v(f,c)\right]^2 - k(f) > 0.$$

Rewrite problem (24)-(26) as

$$\min_{f(c)} H(f,c) = \frac{k(f)}{d - v(f,c)} + \beta\left[w(f) + t_t\right],$$

and apply Hamilton method, then the optimal $f^*(c)$ can be derived by Euler equation:

$$\frac{dH(f,c)}{dc} H_{\dot{f}} = H_f \Leftrightarrow \frac{k'(f)\left[d - v(f,c)\right] + k(f)\dfrac{\partial v}{\partial f}}{\left[d - v(f,c)\right]^2} + \beta w'(f) = 0.$$

It is equivalent to:

$$\gamma = k'(f) + \beta w'(f)\left[d - v(f,c)\right] = 0 . \tag{39}$$

To find the relationship between the optimal $\left(f^*, \tau_t^*\right)$ and c, we take total derivative of both sides of Eq.(39) with respect to c:

$$\frac{df^*}{dc} = -\frac{\beta^2 w'(f) t_c' \left[d - v(f,c)\right]^2}{F_1},$$

$$\frac{d\tau_t^*}{dc} = -\frac{\beta t_c' F_2}{\varepsilon F_1},$$

where

$$F_1 = \left\{k''(f) + \beta w''(f)\left[d - v(f,c)\right]\right\}\varepsilon - \beta w'(f)\left[d - v(f,c)\right]\gamma > 0$$

$$F_2 = \beta^2 w'(f)\left[d - v(f,c)\right]^2 \left\{k'(f) t_v' \left[d - v(f,c)\right] + k(f) w'(f)\right\}$$
$$+ k(f)\left\{k''(f) + \beta w''(f)\left[d - v(f,c)\right]\right\}\varepsilon - \beta k(f) w'(f)\left[d - v(f,c)\right]\gamma$$

According to Lemma 3, the sufficient and necessary condition of the D-T Paradox when both f and τ_t are subject to change is given by

$$\left(E_f^v \cdot \frac{v}{f} \cdot \Delta f + E_\tau^v \cdot \frac{v}{\tau_t} \cdot \Delta\tau_t\right) \cdot \Delta c > 0 . \tag{40}$$

Note that in the current situation, the sufficient and necessary condition is equivalent to:

$$\frac{dp_c^*}{dc} = \beta w'(f)\frac{df^*}{dc} + \frac{d\tau_t^*}{dc} = -\frac{\beta t_c' F_3}{F_1} > 0,$$

where

$$F_3 = \beta w'(f)\left[d - v(f,c)\right]\gamma + \beta k(f) w''(f)\left[d - v(f,c)\right] + k(f) k''(f) > 0 .$$

Therefore, the individual travel cost determined by optimal frequency and fare monotonically increases with highway capacity, and this means that the D-T Paradox would occur whenever there is a capacity expansion.

A.6. Proof for Proposition 5

(i) Since π is strictly decreasing with c, according to Eq.(27), the relation between c and f can be regarded as a function $c = c(f)$, where the derivative is

$$\frac{dc}{df} = -\frac{k'/\tau_t^0 + \partial v/\partial f}{\partial v/\partial c},$$

where $\partial v/\partial f < 0$ and $\partial v/\partial c > 0$ is given by Eq.(2) and (4), and $k' > 0$ is the marginal operation cost. When both Δf and Δc are small, $\Delta f \cdot \Delta c < 0$ is equivalent to $dc/df < 0$. Then according to Lemma 4, the condition for the D-T Paradox is $t_v' > -\tau_t^0 w'/k'$.

(ii) Let $I_1 = t_v' + \tau_t^0 w'/k'$, and assume that the D-T Paradox condition is satisfied at the initial point, such that $I_1\big|_{c=c_0} = t_v' + \tau_t^0 w'/k' > 0$, and that $dc/df\big|_{c=c_0} < 0$. The argument for the (ii) part of Proposition 5 can be shown by the fact that I_1 is strictly decreasing with c:

$$I_1' = t_{vc}'' + \tau_t^0 \frac{df}{dc} \frac{w''k' - w'k''}{k'^2} < 0 .$$

(iii) For given c, the transit authority's problem is:

$$\max_f \quad f$$

subject to

$$g(f) = \beta\left[w(f) + t_t - t_a\left(d - \tfrac{k(f)}{\tau_t^0}, c\right)\right] + \tau_t^0 - \tau_a = 0$$

Similar for problem (24)-(26), the optimal f^* is unique for a given c (refer to Appendix A. 4). Then $g'(f^*) = \frac{\beta}{\tau_t^0}\left[\tau_t^0 w'(f^*) + t_v' k'(f^*)\right] > 0$, and

$$\frac{dp_c^*}{dc} = \frac{\beta \tau_t^0 w'(f^*) t_c'}{\tau_t^0 w'(f^*) + t_v' k'(f^*)} > 0 .$$

A.7. Proof for Proposition 6

(i) Since π is strictly decreasing with c, according to Eq.(29), the relationship between c and τ_t can be regarded as a function $c = c(\tau_t)$, where the derivative is

$$\frac{dc}{d\tau_t} = \frac{(d - v)/\tau_t - \partial v/\partial \tau_t}{\partial v/\partial c},$$

where $\partial v/\partial \tau_t > 0$ and $\partial v/\partial c > 0$ is given by Eq.(3) and (4). When both $\Delta\tau_t$ and Δc are small, $\Delta\tau_t \cdot \Delta c > 0$ is equivalent to $dc/d\tau_t > 0$, and the condition for the D-T Paradox is $(d - v) - \tau_t/\beta t_v' > 0$.

(ii) Let $I_2 = (d - v) - \tau_t/\beta t_v'$, and assume that the D-T Paradox condition is satisfied at the initial point, such that $I_2\big|_{c=c_0} = (d - v) - \tau_t/\beta t_v' > 0$, and that $dc/d\tau_t\big|_{c=c_0} > 0$. The argument is immediate by noting that I_2 is strictly decreasing with c:

$$I_2' = -\beta t_v' \left(\frac{\partial v}{\partial c} + \frac{\partial v}{\partial \tau_t}\frac{d\tau_t}{dc}\right) + \beta t_{vc}''(d - v) - \frac{d\tau_t}{dc} < 0 .$$

(iii) For given c, the transit authority chooses the minimal τ_t subject to the zero-profit and equilibrium constraints according to:

$$\min_{\tau_t} \ \tau_t$$

subject to

$$z(\tau_t) = \beta\left[w(f_0) + t_t - t_a\left(d - \frac{k(f_0)}{\tau_t}, c\right)\right] + \tau_t - \tau_a = 0 .$$

Similar for problem (24)-(26), the optimal τ_t^* is unique for a given c (refer to Appendix A. 4).

Then $z'(\tau_t^*) = \frac{1}{\tau_t^*}\left[\tau_t^* - \beta t'_v\left(d - v^*\right)\right] < 0$, and

$$\frac{dp_c^*}{dc} = \frac{\beta t'_c \tau_t^{*2}}{\tau_t^* - \beta t'_v\left(d - v^*\right)} > 0 .$$

References

Abraham, J., Hunt, J., 2001. Transit system management, equilibrium mode split and the Downs-Thomson paradox. TRB conference preprint. Department of Civil Engineering, University of Calgary.

Ahn, K., 2009. Road pricing and bus service policies. Journal of Transport Economics and Policy 43 (1), 25-53.

Arnott, R., Small, K., 1994. The economics of traffic congestion. American Scientist 82(5), 446–455.

Arnott, R., Yan, A., 2000. The two-mode problem: second-best pricing and capacity. Review of Urban & Regional Development Studies 12, 170-199.

Bar-Yosef, A., Martens, K., Benenson, I., 2013. A model of the vicious cycle of a bus line. Transportation Research Part B: Methodological 54, 37-50.

Basso, L.J., Jara-Díaz, S.R., 2010. The case for subsidisation of urban public transport and the Mohring effect. Journal of Transport Economics and Policy 44(3), 365-372.

Basso, L.J., Jara-Díaz, S.R., 2012. Integrating congestion pricing, transit subsidies and mode choice. Transportation Research Part A: Policy and Practice 46(6), 890-900.

Bell, M.G.H., Wichiensin, M., 2012. Road use charging and inter-modal user equilibrium: The Downs-Thompson Paradox revisited. Energy, Transport, & the Environment. Springer London, 373-383.

Denant-Boèmont, L., Hammiche, S., 2009. Public transit capacity and users' choice: An experiment on Downs-Thomson Paradox, 4th Kuhmo-Nectar Conference "transport and urban economics". Technical University of Denmark, Copenhagen.

Downs, A., 1962. The law of peak-hour expressway congestion. Traffic Quarterly 16(3), 393-409.

Holden, D.J., 1989. Wardrop's third principle: Urban traffic congestion and traffic policy. Journal of Transport Economics and Policy 23(3), 239-262.

Hsu, W.T., Zhang, H., 2014. The Fundamental Law of Highway Congestion Revisited: Evidence from National Expressways in Japan. Journal of Urban Economics 81, 65-76.

Jansson, J.O., 1980. A simple bus line model for optimisation of service frequency and bus size. Journal of Transport Economics and Policy 14(1), 53-80.

Karamychev, V., van Reeven, P., 2010. Oversupply or undersupply in a public transport monopoly? A rejoinder and generalisation. Journal of Transport Economics and Policy 44(3), 381-389.

Kraus, M., 2003. A new look at the two-mode problem. Journal of Urban Economics 54(3), 511-530.

Kraus, M., 2012. Road pricing with optimal mass transit. Journal of Urban Economics 72(2-3), 81-86.

Light, T., 2009. Optimal highway design and user welfare under value pricing. Journal of Urban Economics 66(2), 116-124.

Liu, S., Triantis, K.P., Sarangi, S., 2010. A framework for evaluating the dynamic impacts of a congestion pricing policy for a transportation socioeconomic system. Transportation Research Part A: Policy and Practice 44(8), 596-608.

Mogridge, M.J., 1997. The self-defeating nature of urban road capacity policy: A review of theories, disputes and available evidence. Transport Policy 4(1), 5-23.

Mohring, H., 1972. Optimization and scale economies in urban bus transportation. American Economic Review 62(4), 591–604.

Nie, Y.M., Liu, Y., 2010. Existence of self-financing and Pareto-improving congestion pricing: impact of value of time distribution. Transportation Research Part A: Policy and Practice 44(1), 39-51.

Reinhold, T., 2008. More passengers and reduced costs—the optimization of the Berlin public transport network. Journal of Public Transportation 11(3), 57-76.

Savage, I., Small, K.A., 2010. A comment on 'Subsidisation of Urban Public Transport and the Mohring Effect''. Journal of Transport Economics and Policy 44(3), 373-380.

Small, K.A., 1992. Urban transportation economics. Harwood Academic Publishers, Chur, Switzerland.

Small, K.A., Verhoef, E.T., 2007. The economics of urban transportation. Routledge, London.

Thomson, J.M., 1977. Great cities and their traffic.Victor Gollancz, London.

van Reeven, P., 2008. Subsidisation of urban public transport and the Mohring effect. Journal of Transport Economics and Policy 42(2), 349-359.

Yang, H., Xiao, F., Huang, H.J., 2009. Private road competition and equilibrium with traffic equilibrium constraints. Journal of Advanced Transportation 43(1), 21-45.

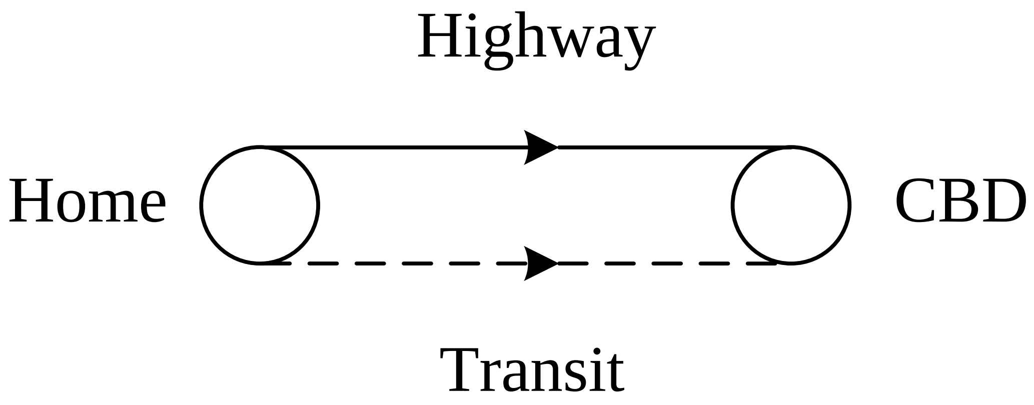


Figure 1. A simple two-mode network

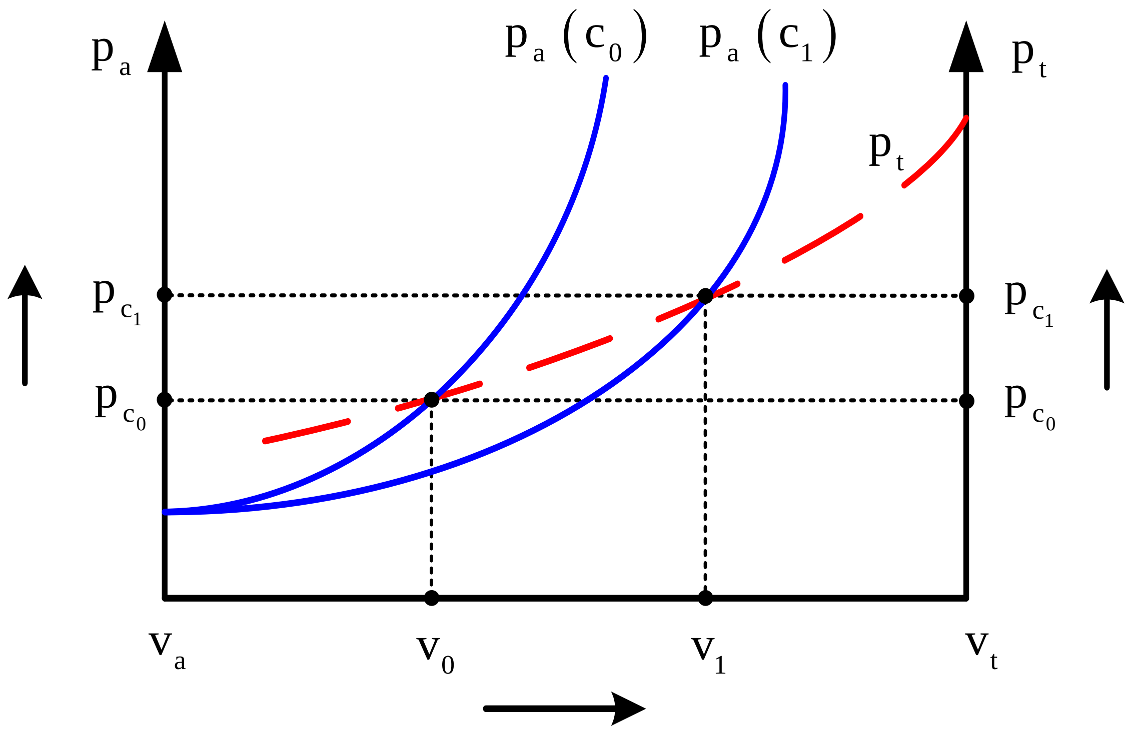


Figure 2. The Downs-Thomson Paradox (equilibrium analysis)

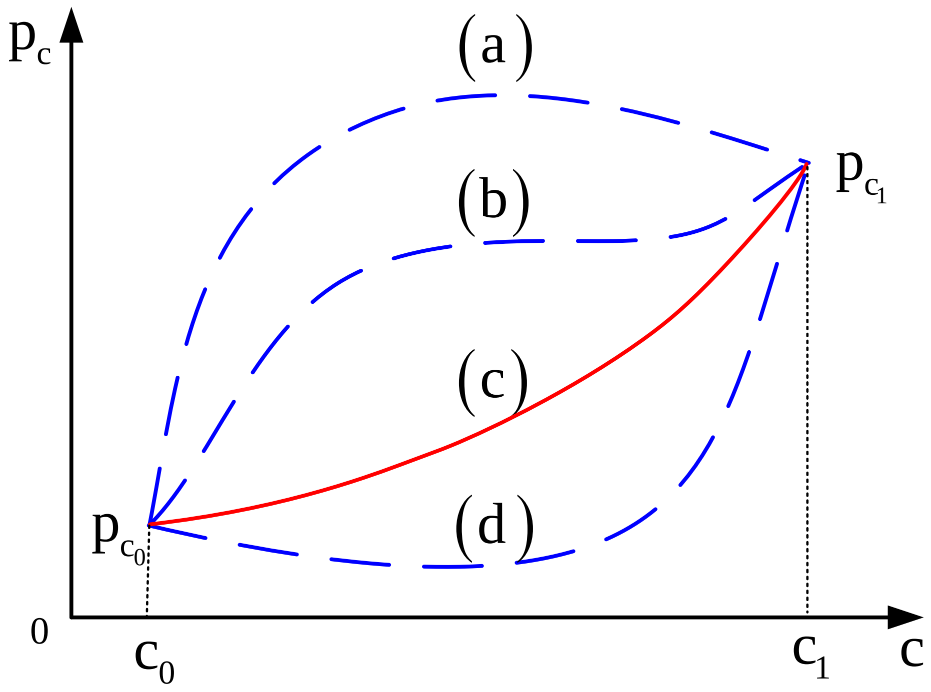


Figure 3. The continuous occurrence of the Downs-Thomson Paradox

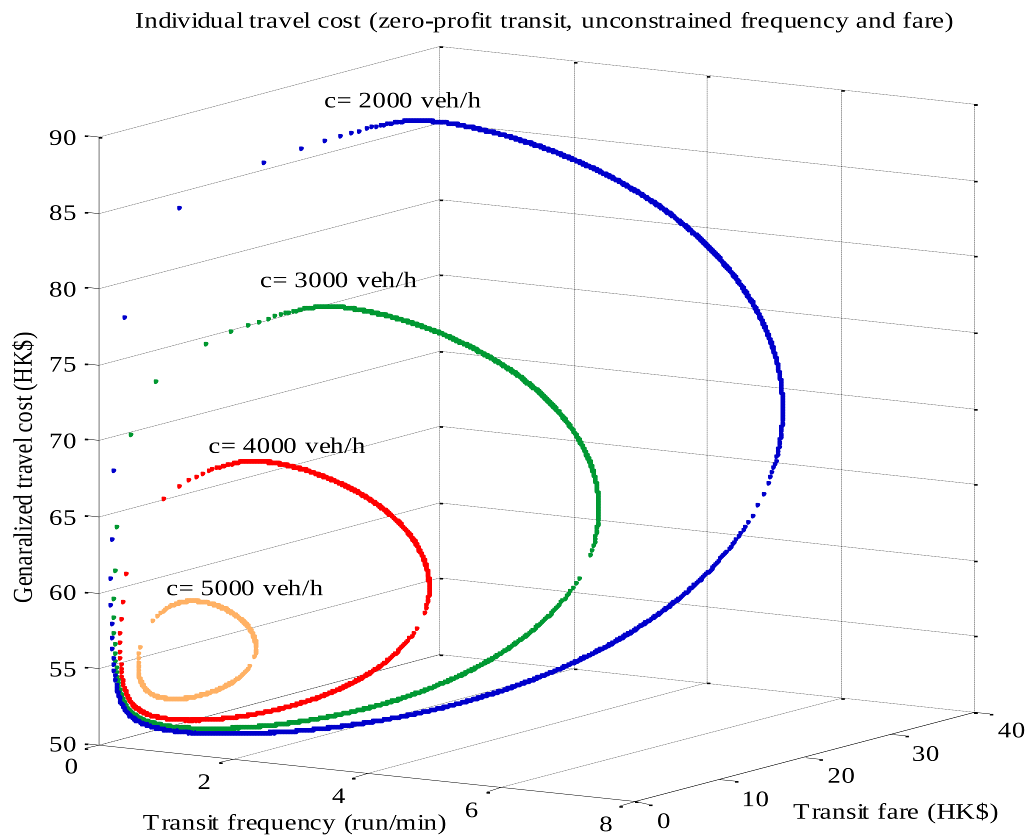
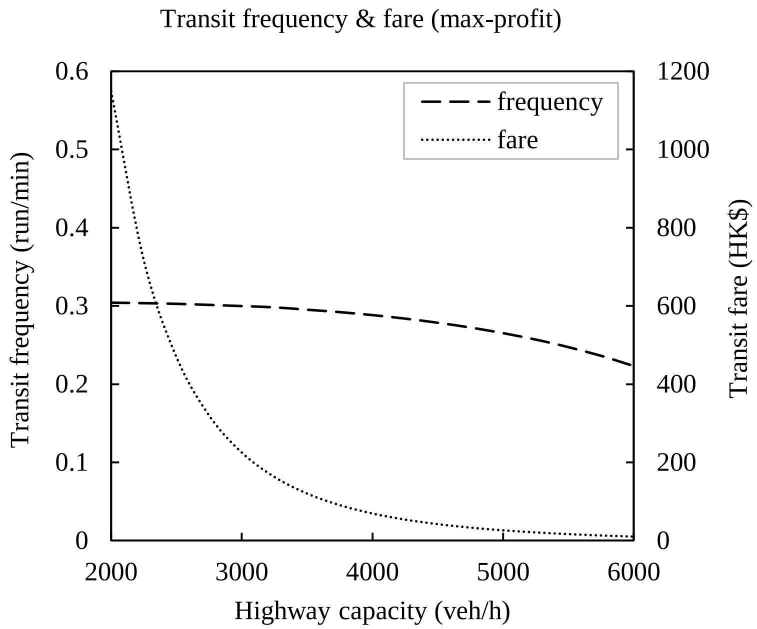
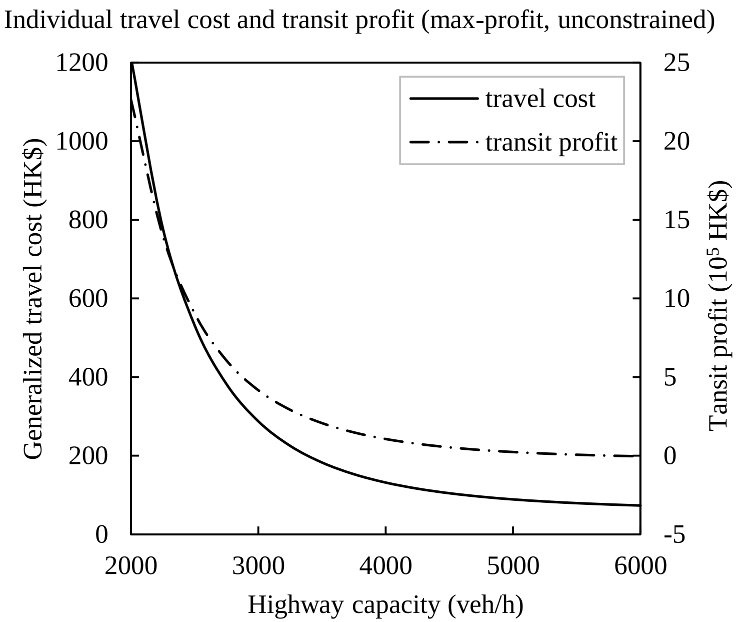


Figure 4. Generalized travel cost with unconstrained and zero-profit transit schemes



(a)



(b)

Figure 5. Unconstrained frequency, fare and individual travel cost with the profit-maximizing transit authority

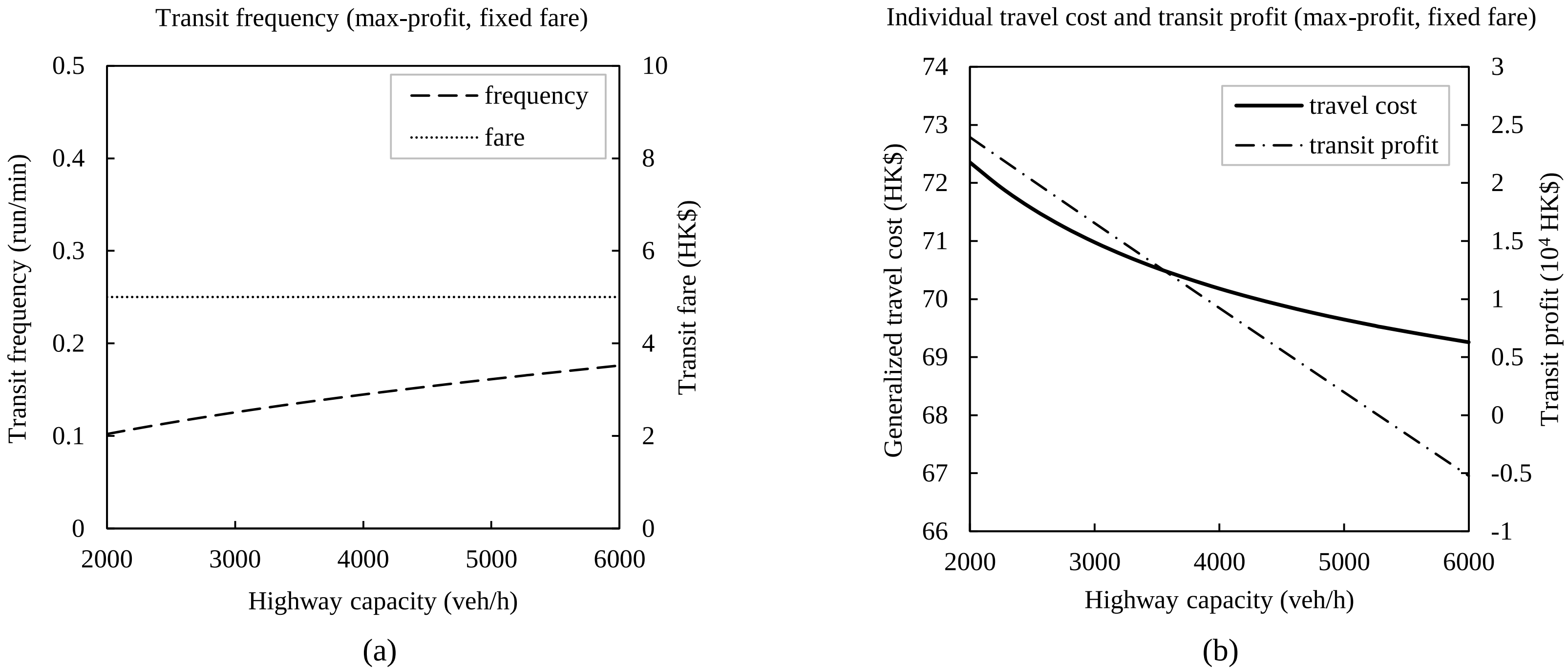


Figure 6. Transit frequency and individual travel cost with a fixed fare and profit-maximizing transit authority

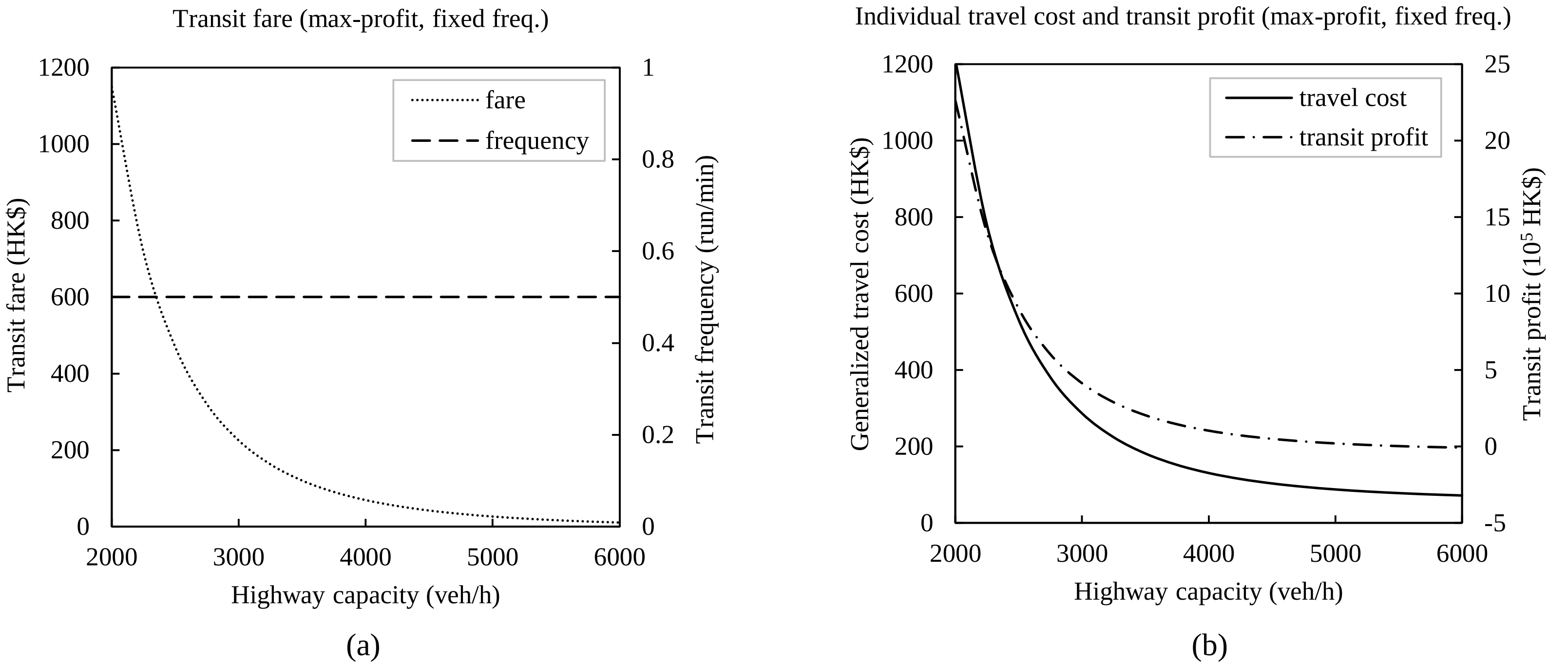


Figure 7. Transit fare and individual travel cost with a fixed frequency and profit-maximizing transit authority

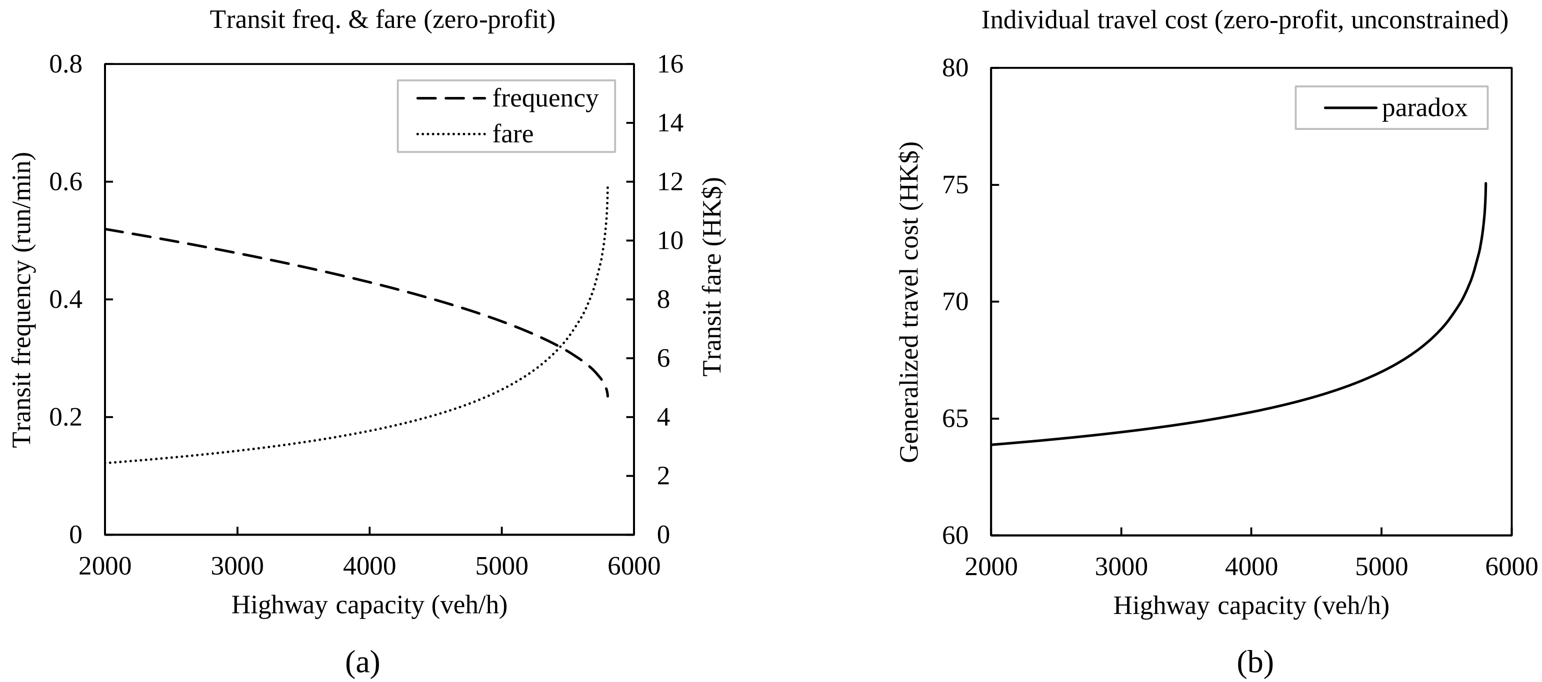


Figure 8. Unconstrained frequency, fare and individual travel cost with a zero-profit transit authority

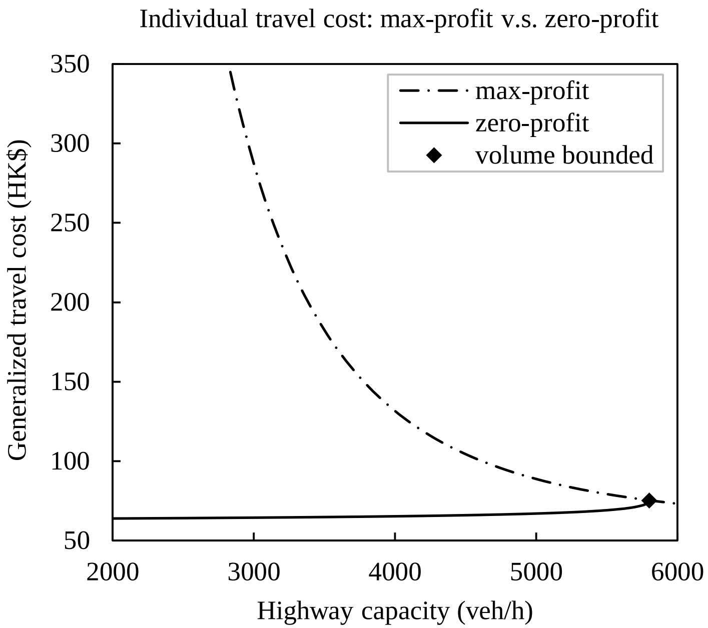
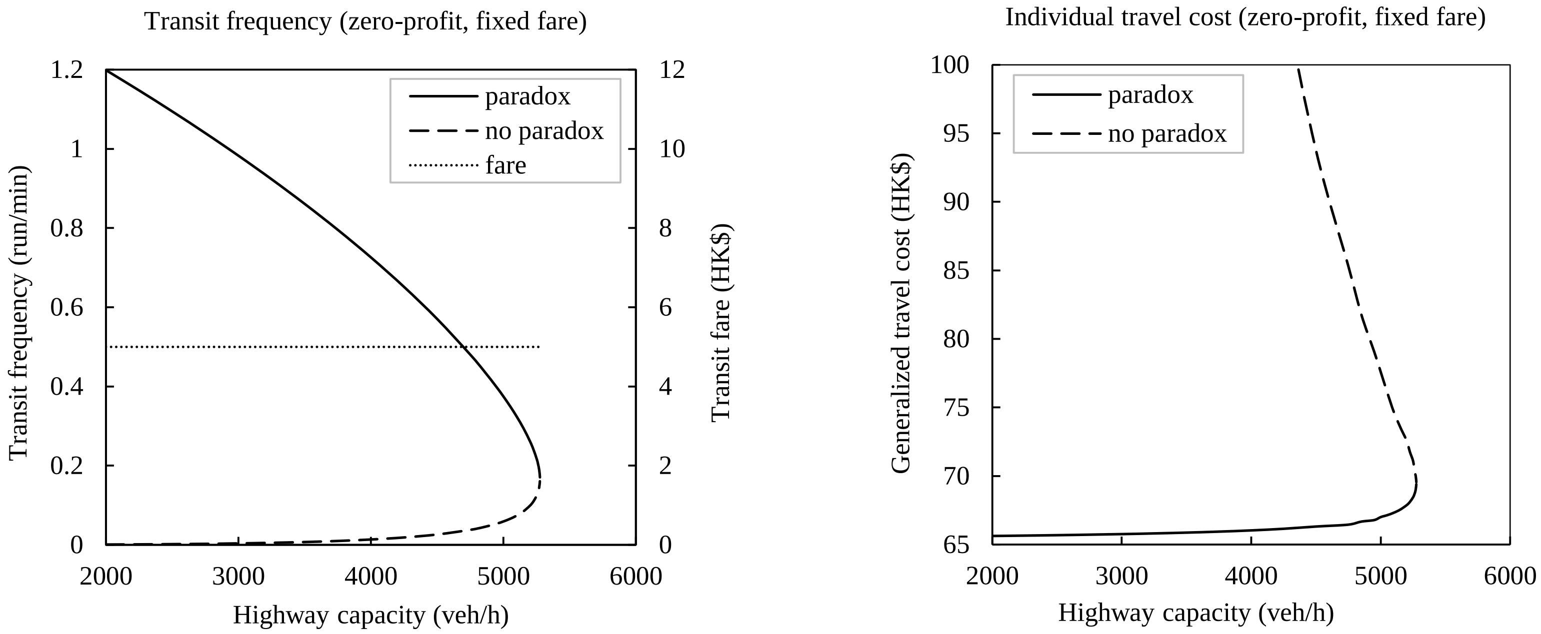


Figure 9. Individual travel cost with unconstrained frequency and fare under different objectives



(a) (b)

Figure 10. Transit frequency and individual travel cost with a fixed fare and zero-profit transit authority

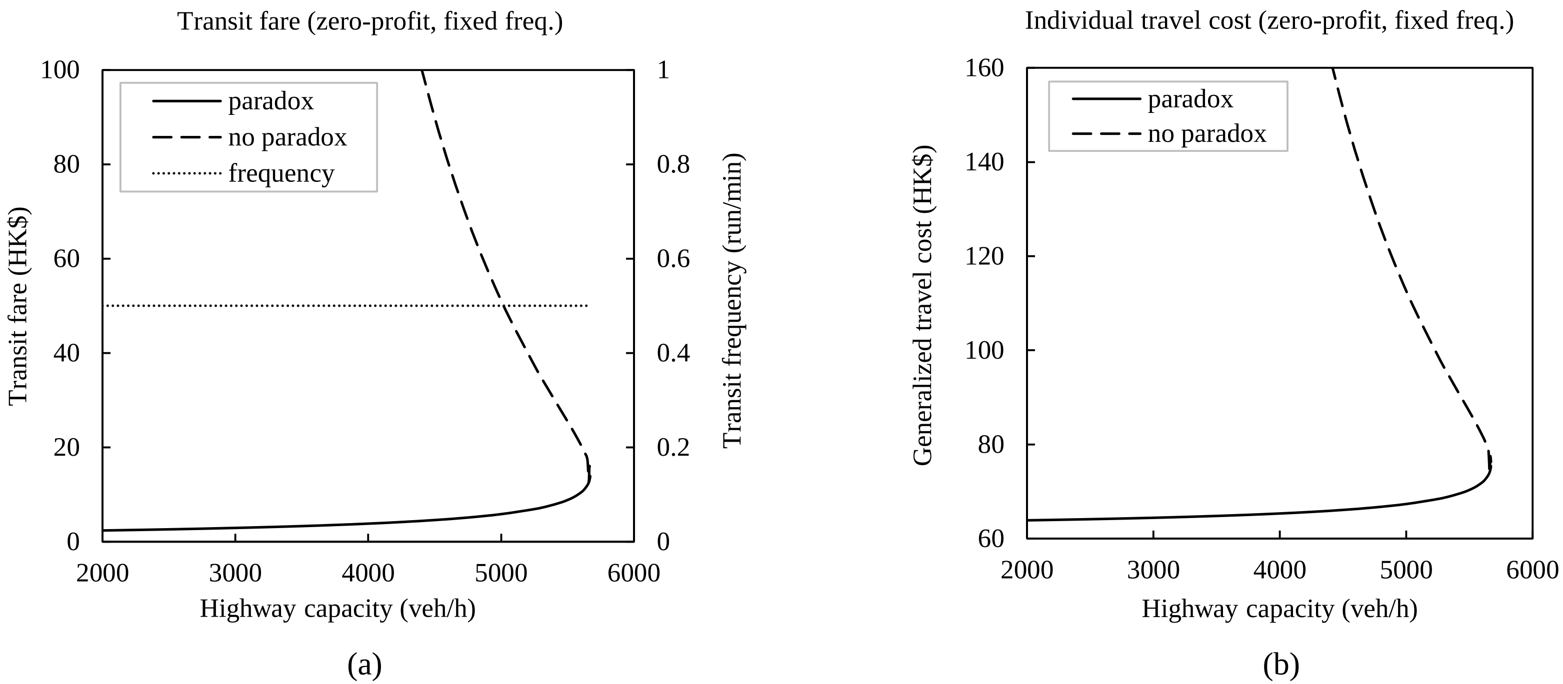


Figure 11. Transit fare and individual travel cost with a fixed frequency and zero-profit transit authority

Table 1. Function specifications

Function	**Specification**
Highway travel time function	$t_a(v/c) = t_0 \cdot \left[1 + t_1 (v/c)^{\alpha}\right]$
Transit operation cost function	$k(f) = k_2 f^2 + k_1 f + k_0$
Transit waiting time function	$w(f) = 1/(2f)$

Table 2. Value of parameter (quantity of one hour)

Parameter	**Value**
Total demand	$d = 10000$ (person)
Value of time	$\beta = 1.5$ (HK$/min)
In-vehicle travel time of transit	$t_t = 40$ (min)
Monetary cost by auto	$\tau_a = 30$ (HK$)
Original highway capacity	$c_0 = 2000$ (vehicle/h)
Fixed fare (in the cases of constrained scheduling scheme)	$\tau_t^0 = 5$ (HK$)
Fixed frequency (in the cases of constrained pricing scheme)	$f_0 = 0.5$ (run/min)
Coefficients in operation cost function	$k_0 = 10000$ (HK$)
	$k_1 = 10000$ (HK$/run)
	$k_2 = 10000$ (HK$/run^2)
Coefficients in highway travel time function	$t_0 = 10$ (min)
	$t_1 = 0.15$
	$\alpha = 4$