Special Issue Article

Advances in Mechanical Engineering

Advances in Mechanical Engineering
2015, Vol. 7(10) 1–12

DOI: 10.1177/1687814015612430
aime.sagepub.com
SAGE

Tradable credit scheme for rush hour travel choice with heterogeneous commuters

Ling-Ling Xiao[1], Hai-Jun Huang[2] and Ronghui Liu[3]

Abstract

This article proposes a tradable credit scheme for managing commuters' travel choices. The scheme considers bottleneck congestion and modal split in a competitive highway–transit network with heterogeneous commuters who are distinguished by their valuation of travel time. The scheme charges all auto travelers who pass the bottleneck during a peak-time window in the form of mobility credits. Those who avoid the peak-time window, by either traveling outside the peak-time window or switching to the transit mode, may be rewarded credits. An artificial market is created so that the travelers may trade these credits with each other. We formulate the credit price and the rewarded and charged credits under tradable credit scheme. Our analyses indicate that the optimal tradable credit scheme can achieve nearly 40% efficiency gains depending on the level of commuters' heterogeneity. In addition, this scheme distributes the benefits among all the commuters directly through the credit trading. Our results suggest that in assessing the efficiency of tradable credit scheme, it is important to take into account the commuters' heterogeneity. Numerical experiments are conducted to examine the sensitivity of tradable credit scheme designs to various system parameters.

Keywords

Bottleneck model, commuter heterogeneity, traffic congestion, tradable credit scheme, welfare efficiency

Date received: 12 July 2015; accepted: 10 September 2015

Academic Editor: Yongjun Shen

Introduction

Congestion in morning commuter traffic has traditionally been modeled as a bottleneck problem. The classic bottleneck model was developed by Vickrey[1] who studied the commuting congestion in a highway between a residential area and a workplace. It shows that there exists an equilibrium departure-time pattern, whereby all commuters incur the same travel cost no matter when they start their trips. Traffic congestion taking the form of queuing behind a bottleneck is a deadweight loss to system efficiency. The bottleneck model offers a flexible framework for investigating the effects of congestion pricing schemes to alleviate the queue behind the bottleneck.[2–4] In this context, congestion pricing is found to be efficient in internalizing the external costs of the traffic by inducing the change in departure pattern.

Since the idea was first put forward by Pigou[5] and Knight,[6] congestion pricing has become a widely known mechanism to regulate traffic congestion. However, it is also known to induce inequity among

[1]School of Economics and Management, Beijing Jiaotong University, Beijing, P.R. China
[2]School of Economics and Management, Beihang University, Beijing, P.R. China
[3]Institute for Transport Studies, University of Leeds, Leeds, UK

Corresponding author:
Ling-Ling Xiao, School of Economics and Management, Beijing Jiaotong University, 3 Shangyuancun, Xizhimenwai, Beijing 100044, P.R. China.
Email: lingsayok@163.com

commuters if the differences in their values of time are not properly taken into account. This issue has been addressed in the literature. The effect of congestion pricing on different commuter groups is demonstrated theoretically with a static network analysis method[7–9] and with the bottleneck modeling method.[10–14] There have been recent interests in the design of more equitable and practical congestion management schemes. An example of such alternative is the "tradable permit system" by which the eligible residents will receive a certain amount of "rights" for the scarce good. Verhoef et al.[15] discussed the possibilities of using tradable permits in the regulation of road transport externalities. Akamatsu et al.[16] proposed a "tradable bottleneck permits" (TBPs) scheme for resolving the morning commute problem. In their model, the road manager issues the "bottleneck permits" to road users, and the permits are tradable in auction markets.

Recently, the cap-and-trade scheme, which involves issuing mobility credits and allowing travelers to trade in a market, seeks to couple quantity restriction with a trading mechanism. It is also known as tradable credit scheme (TCS)[17] and has been extensively studied.[18–20] Xiao et al.[21] analyzed the TCSs in the context of the morning commute problem. Their model replaces Vickrey's toll with a corresponding time-varying credit charge and an initial allocation of credits. Their results indicate that the credit market can achieve the system optimum even when travelers differ in their values of time. Nie[22] adopts tradable credit to replace single step-coarse toll with homogeneous commuters in bottleneck model with respect to three different behavior assumptions.[2,23–26]

Recently, Tian et al.[27] proposed the time-dependent credit charge scheme to manage bottleneck congestion and modal split with heterogeneous users. However, such a complex credit structure is not very well accepted by travelers as they cannot predict the amount of charging they would have to pay in advance. This impels us to develop more practical TCS, which delineates an off-peak credit-rewarding and a peak-time credit-charging systems. In addition, we consider mode choice (between auto and transit modes) in the morning commuting problem as well as heterogeneity in commuters' Value-Of-Time (VOT). We assume that the travelers' VOT is continuously distributed across the population and propose a new tradable credit based on different travelers' behavior assumptions. Compared with the traditional step-coarse tolls, this scheme is revenue neutral and hence less likely to be perceived as another taxation instrument.

Introducing the TCS proposed by Yang and Wang[17] into a mode-choice problem has the potential to lead to revenue-neutral transport pricing and subsidy policy. Inspired by this consideration, this article aims at developing efficient TCS that makes everyone better off and that reduces social cost in a competitive two-mode network. Without loss of generality, we deal with the two-mode morning commuting problem with road bottleneck congestion and incorporate user heterogeneity by assuming that travelers' VOT is continuously distributed across the population. Under the credit scheme, travelers will divide themselves among two modes and choose different departure time for going through the bottleneck according to their own preferences and sell and buy additional credits according to their individual travel needs. The resulting equilibrium price of credits in the trading market and the competitive two-mode traffic equilibrium will thus be ascertained.

The next section presents a bottleneck model of morning commute problem considering the heterogeneity of commuters and discusses the equilibrium without tolling. Section "TCS for a single bottleneck with heterogeneity" proposes a TCS for morning commute problem and looks into the changes in individual travel cost before and after the optimal credit scheme. The two-mode equilibrium with the TCSs is derived in section "Departure time and mode choice in the presence of optimal TCS." In section "Numerical examples," a numerical example is presented to illustrate the effect of commuter heterogeneity on equilibrium solutions with credit scheme. Finally, section "Summary and concluding remarks" draws the conclusions.

No-toll scheme for a single bottleneck with heterogeneity

User heterogeneity can be resulted from income differences between otherwise identical commuters. Following the literature, we represent user heterogeneity in terms of different unit costs of travel time (α), schedule delay early (β), and schedule delay late (γ) across commuters. In most of the existing literature, heterogeneous commuters are grouped into a set of discrete user classes: within each class, commuters have the same unit costs. For example, Arnott et al.[11] define a series of discrete groups of commuters in order of increasing β/α ratios. They provided analytical solutions for no-toll equilibrium for three groups of commuters and compared the welfare effects with optimal time-varying toll for two groups of commuters. In this article, user heterogeneity is incorporated by ranking users according to a continuously increasing VOT function. We provide analytical solutions for no-toll equilibrium and social optimal tolls for these heterogeneous commuters with continuously distributed VOT.

We assume here that everyone has the same work flexibility, and the ratio between the unit penalties for the queuing delay and schedule delay is the same for all commuters, that is, $\beta(x)/\alpha(x) = \eta_1$, $\gamma(x)/\alpha(x) = \eta_2$, and follows a distribution

$$F(\omega) = Pr\{\alpha(x) < \omega\} \tag{1}$$

where η_1 and η_2 are constants and satisfy $0 < \eta_1 < 1 < \eta_2$. Here $\alpha(x)$, $\beta(x)$, and $\gamma(x)$ are the values of queuing delay, early arrival penalty, and late arrival penalty for the xth user, respectively. For convenience, we arrange the commuters in increasing order of α, such that $\alpha(x)$ is monotone and increasing with x, for example, $\alpha(x_1) > \alpha(x_2)$ for $x_1 > x_2$.

Suppose there are two competing modes of a highway and transit line connecting a single pair of origin and destination. It is assumed that a fixed number of N commuters travel from the origin to the destination each morning either with the highway or transit and hope to arrive at the destination at the work starting time t^*, which is normalized to be zero in our analysis. For simplicity and without loss of generality, it is further assumed that the free-flow time on highway is zero. As assumed in the Tabuchi's model, the external demand elasticity is ignored, but internal elasticity is considered by allowing inter-modal competition. Let N_f and N_g denote the number of highway users and transit users, respectively, and $N_f + N_g = N$.

The highway exhibits bottleneck congestion, which is characterized by the standard bottleneck model developed initially by Vickrey.[1] In no-toll scheme, the trip cost of the xth person traversing the bottleneck and taking into account schedule delay early and schedule delay late penalties becomes

$$C(x,t) = \begin{cases} \alpha(x)T(t) + \beta(x)(t^* - t - T(t)), & \text{if } t_e < t \le t_t \\ \alpha(x)T(t) + \gamma(x)(t + T(t) - t^*), & \text{if } t_t < t \le t_l \end{cases} \tag{2}$$

where $T(t)$ denotes the commuting time, and t^* the desired arrive time (which is assumed to be the same work starting time for all commutes). t_e and t_l denote the earliest and the latest departure time and t_t the transition between schedule early and late.

In order to derive the equilibrium based on no-toll scheme, we define the function of generalized queuing time as follows

$$\begin{aligned} \hat{C}_N(t) &= \frac{C(x,t)}{\alpha(x)} \\ &= \begin{cases} T(t) + \eta_1(t^* - t - T(t)), & \text{if } t_e < t \le t_t \\ T(t) + \eta_2(t + T(t) - t^*), & \text{if } t_t < t \le t_l \end{cases} \end{aligned} \tag{3}$$

where t_e and t_l denote the earliest and the latest departure time and t_t the transition between schedule early and late. From the above equations, the schedule delay is translated into an equivalent travel time for the xth commuter, and the generalized travel time is independent to α.

Under user equilibrium without toll, $\hat{C}_N$ should be the same for everyone regardless of the VOT. Thus, at equilibrium, there is no difference in generalized queuing time with respect to departure time, that is, $\mathrm{d}\hat{C}_N(t)/\mathrm{d}t = 0$. Then we have

$$T'(t) = \begin{cases} \dfrac{\eta_1}{1-\eta_1}, & \text{if } t_e < t \le t_t \\ \dfrac{-\eta_2}{1+\eta_2}, & \text{if } t_t < t \le t_l \end{cases} \tag{4}$$

Since the first and last departures confront no queue at bottleneck, the cumulative departure flow at the last time t_l equals to the traffic demand N, that is, $t_l = t_e + N/s$, where s is the capacity of the bottleneck. Using $\hat{C}_N(t_e) = \hat{C}_N(t_l) = \hat{C}_N(t_e + N/s) = \eta_1(t^* - t_e)$, we have

$$t_e = t^* - \frac{N\delta}{s\eta_1}, \quad t_l = t^* + \frac{N\delta}{s\eta_2}, \quad \text{and} \quad t_t = t^* - \frac{N\delta}{s}$$

where $\delta = \eta_1\eta_2/(\eta_1 + \eta_2)$. Rearranging the above results, we obtain the equilibrium trip cost of every commuter as follows

$$C(x) = \alpha(x)\delta\frac{N}{s} \tag{5}$$

The system trip cost calculated in monetary unit is obtained by integrating the individual costs throughout the whole population

$$STC_N = \int_0^N \alpha(x)\delta\frac{N}{s}dx = \delta\frac{N}{s}\int_0^N \alpha(x)dx \tag{6}$$

For homogeneous travelers, the system total cost becomes $STC_N = \alpha\delta N^2/s$

TCS for a single bottleneck with heterogeneity

Here, we focus on step-coarse toll under proportional heterogeneity. And let us first explain how a step-coarse toll works in a single bottleneck. A key question in analyzing a step-coarse toll is how to deal with discontinuously taking place at the boundary of the peak-time window. Such discontinuity forces users arriving at the boundary to have different travel delays, depending on whether or not they pay the toll. Arnott et al.[2] argued that because the first person who pays the toll must have a lower travel delay compared to his or her immediate predecessor who escape the toll, he or she must arrive at the bottleneck later by ρ/α, which in turn implies that there must be a period of time during which the arrival rate at the bottleneck is zero (see Figure 1). The discontinuity between the last person who pays the toll and his or her immediate successor leads to the following behavioral assumptions:

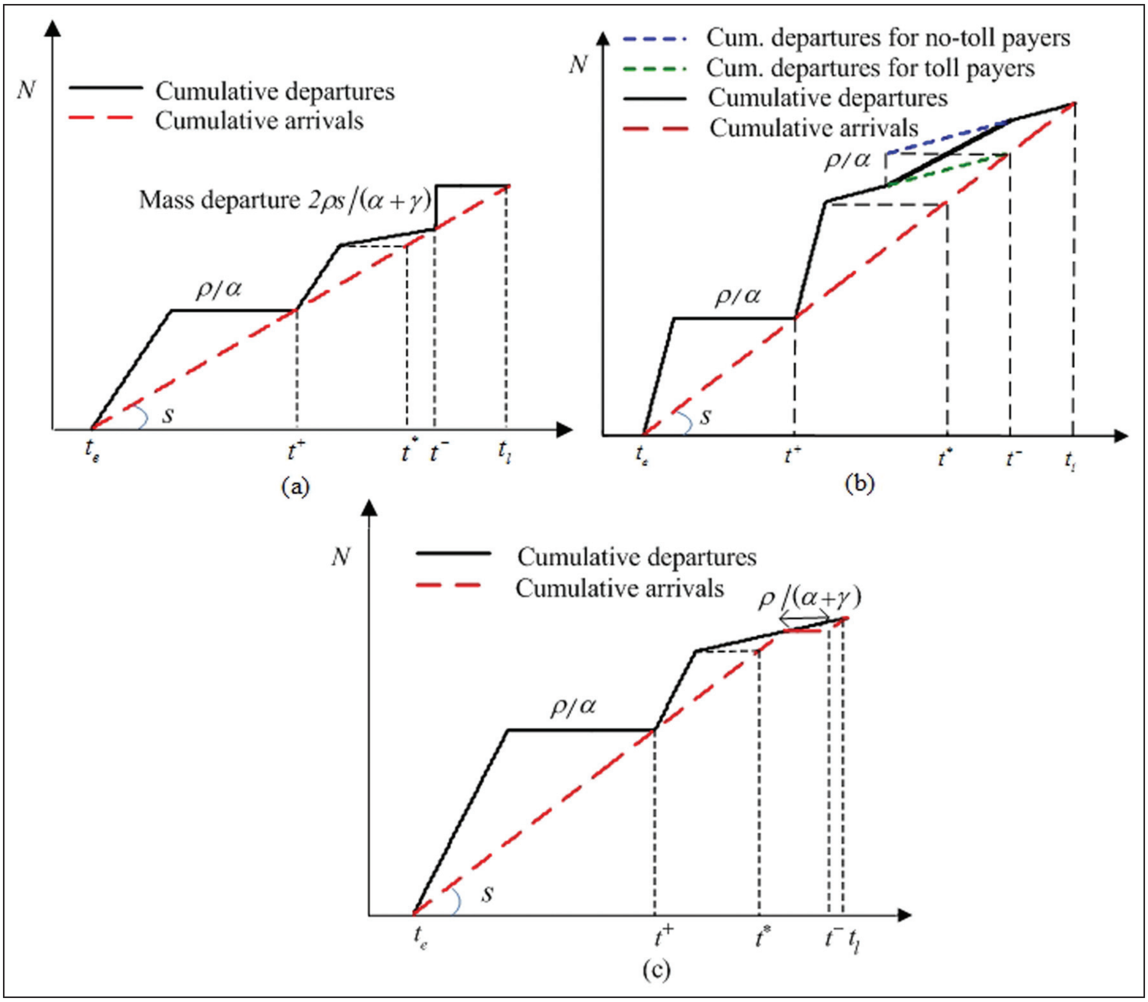

Figure 1. Equilibrium with a coarse toll based on three different behavioral assumptions: (a) MD, (b) SW, and (c) BI.

Mass departure (MD),[2] which assumes that a MD at the bottleneck occurs immediately after the toll is lifted (see Figure 1(a), at time t^-), and that the commuters in the mass experiences an identical expected general cost which offsets the toll.

Separated waiting (SW),[23,24] in which commuters who choose to pass the bottleneck after a tolling period can wait on a set of secondary lanes (the dotted blue curve in Figure 1(b)) without impeding other drivers who do pass the bottleneck in that tolling period.

Braking-induced idling (BI),[25] which assumes that commuters would slow down or stop just before reaching a tolling point and wait until the toll is lowered from one step to the next step before proceeding. The equilibrium departure-time patterns as derived by the above authors, under homogeneous conditions, are reproduced in Figure 1(c). Some of the key characteristics of these assumptions are further described in this section.

Under heterogeneous preferences, the optimal coarse tolling of congestion corresponding to the above assumptions is analyzed (see Van den Berg,[28] for details). Based on the proportional heterogeneity, total cost will be minimized with respect to the number of un-tolled users (V) and then there are $N-V$ tolled users. The level of the coarse toll (ρ) is such that at the start (t^+) and end (t^-) of the tolled period, the queue is zero. For a given number of un-tolled commuters V, the rushing hour period and the tolling window are given as follows

$$t_e = t^* - \frac{\delta}{\eta_1}\frac{N}{s} + \omega\frac{V}{s}; \quad t_l = t^* + \frac{\delta}{\eta_2}\frac{N}{s} + \mu\frac{V}{s} \tag{7}$$

$$t^+ = t_e + \frac{\sigma}{\eta_1}\frac{V}{s}; \quad t^- = t_l - \frac{\sigma}{\lambda}\frac{V}{s}; \quad \rho = \sigma\frac{V}{s}\alpha(V) \tag{8}$$

where σ, ω, and μ are parameters dependent on the behavior assumptions, as specified in Table 1.

The number of commuters who pay the toll is

$$N_t = (t^- - t^+)s = N + \left(\mu - \omega - \frac{\sigma}{\lambda} - \frac{\sigma}{\eta_1}\right)V \quad (9)$$

Also, the system travel cost excluding toll is

$$\begin{aligned} STC &= \int_0^V \beta(x)(t^* - t_e)dx + \int_V^N \beta(x)(t^* - t^+)dx \\ &= A\left(\delta\frac{N}{s} - \omega\eta_1\frac{V}{s}\right) - A_2\frac{\sigma V}{s} \end{aligned} \quad (10)$$

and

$$\begin{aligned} &A_1 = \int_0^V \alpha(x)dx, \quad A_2 = \int_V^N \alpha(x)dx, \\ &A = A_1 + A_2 = \int_0^N \alpha(x)dx \end{aligned} \quad (11)$$

By taking $\partial STC(V)/\partial V = 0$, we obtain the first-order optimality condition in which only V is variable

$$\alpha(V)V - \int_V^N \alpha(x)dx = \frac{\omega\eta_1}{\sigma}\int_0^N \alpha(x)dx \quad (12)$$

Equation (12) implicitly determines the amount of commuters V who are traveling outside the tolling period after toll is imposed, as long as the VOT distribution of the commuter population is known. Generally, we cannot obtain the explicit expression of V except for simple forms of VOT distribution. For instance, if VOT follows a uniform distribution, that is, $\alpha(x) = ax + b$, $a>0$, and $b>0$, we can obtain that

$$V = \frac{\sqrt{4b^2 + 6a\left(\frac{\omega\eta_1}{\sigma} + 1\right)\left(\frac{a}{2}N^2 + bN\right)} - 2b}{3a} \quad (13)$$

Substituting the parameters of Table 1 into equation (13), it is clear that $V_M > V_S > V_B$, here subscripts M, S, B denote the behavior assumptions.

In the TCS, authority delineates a peak-time window $[t^+, t^-]$ and requires all travelers who pass the bottleneck within that window to pay κ units of mobility credits. The authority rewards r units of credits to those who travel during the off-peak time, that is, $[t_e, t^+] \cup [t^-, t_l]$. A market is created such that those who need to pay credits can purchase them from those who acquired them. Assume that this market is cleared in the sense that the total number of credits earned by the off-peak users equal the number of credits used by the peak-time users, which leads to

$$(N - V)\kappa = Vr \quad (14)$$

or

$$\frac{r}{\kappa} = \frac{N - V}{V} \quad (15)$$

which represents the ratio between the unit of reward for travel in the off-peak to the unit of pay in peak travel. It is also easy to see that $(\partial r/\kappa)/\partial V = -N/V^2 < 0$, implying that the ratio of reward to pay decreases with the increase in the number of uncharged commuters.

The trading between the off-peak and peak users can be translated into a peak-time step-coarse toll

$$\rho = P(r + \kappa) \quad (16)$$

where P represents a market clearing price of the credits.

Adopting the same uniform VOT distribution $\alpha(x) = ax + b$, we examine the ratio between reward and pay r/κ with the optimal coarse toll of this section and under the different behavioral assumptions.

For heterogeneous users with the special case $a \neq 0$ and $b = 0$, substituting equation (13) into equation (15), we note that the SW assumption leads to $r/\kappa = 0.73$ and $r/\kappa = 0.66$ for the MD assumption and $r/\kappa = 0.87$ for the BI assumption. For homogeneous users, that is, $a = 0$ and $b \neq 0$, taking the optimal V from equation (13), we obtain $r/\kappa = 1$ for SW, $r/\kappa = 0.85$ for MD, and $r/\kappa = 1.34$ for BI assumptions.

Departure time and mode choice in the presence of optimal TCS

A TCS was first proposed by Yang and Wang,[17] and may be used to replace Vickrey's time-varying toll, as

Table 1. Parameters used in the analytical solutions for step-coarse toll models.

Scenarios	σ	ω	μ	λ
MD	$\frac{(1+\eta_2)\eta_1}{1+2\eta_1+\eta_2}$	$\frac{(\eta_2-1)\sigma}{(1+\eta_2)(\eta_1+\eta_2)}$	$\frac{(\eta_2-1)\sigma}{(1+\eta_2)(\eta_1+\eta_2)}$	$\frac{1+\eta_2}{2}$
SW	δ	0	0	η_2
BI	$\frac{(1+\eta_2)\eta_2\eta_1}{\eta_1+\eta_2(1+\eta_2)}$	$-\frac{\eta_2\sigma}{(1+\eta_2)(\eta_1+\eta_2)}$	$\frac{\eta_1\sigma}{(1+\eta_2)(\eta_1+\eta_2)}$	$\frac{\eta_2(1+\eta_2)}{1+2\eta_2}$

MD: mass departure; SW: separated waiting; BI: braking-induced idling.

shown in Xiao et al.[29] (for homogeneous commuters) and Tian et al.[27] (with heterogeneous commuters). While social optimum (SO) TCS is capable of eliminating queuing delays completely, it requires a continuously adjusting credit charges which can be difficult to implement in practice. Recognizing the importance of simplicity, Nie[22] proposed a TCS that aims at replicating the effect of a step-coarse toll with homogeneous commuters. Therefore, it is straightforward to start the analysis with user heterogeneity in the context of TCS and in particular its welfare and distributional effects.

The toll scheme formulated in section "TCS for a single bottleneck with heterogeneity" keeps the toll rate constant in a time window to reallocate the peak demand on the congested highway. Such a pricing structure is not very well accepted by travelers as it may lead to issues of fairness and the lack of alternative mode of travel. Redistributing the toll revenue to the auto travelers can deal with some of these issues, so can provide attractive policies and services for rewarding travel by public transport. Here, we consider a policy based on the TCS for reducing congestion costs in a two-mode network. We examine the two subsystems of TCS: a credit charging and a credit distribution schemes.

Analysis of mode choice with TCS

Unlike the auto mode, we assume that (1) the capacity of transit mode is sufficiently large, (2) the transit pattern ignore that schedule late or early cost, and (3) the cost of commuting by transit depends only on the in-vehicle travel time cost. Since we assume the free-flow travel time for highway is zero, then T can be regard as additional constant travel time in relation to the auto mode. Commuters can travel on the highway either by car, experiencing queuing and schedule delays, or by transit, experiencing a longer travel time but no delays. Note that a necessary condition for both modes to be used is

$$T = \frac{\theta N \delta}{s}, \quad \theta \in (0, 1) \tag{17}$$

Under no-toll scheme, $\bar{N}_f = \theta N$, that is, travel cost for the $\bar{N}_f$th commuter is indifferent between choosing either auto or transit mode. Under TCS, because only the ratio between the rewarded and charged credits matters, the number of charged credits κ is set to 1. Furthermore, the transit users will be rewarded r_b credits. Let N_f be the total number of users who use the highway route, and N_u be the portion of N_f that do not use the bottleneck within the charging window. We define

$$p = \frac{N_u}{N_f} \tag{18}$$

and naturally, $p \in [0, 1]$, with the new notation, and noting from equation (8) that

$$\rho = \sigma \frac{N_u}{s} \alpha(N_u) \tag{19}$$

The total system cost can be rewritten as follows

$$STC = A_1 \eta_1 (t^* - t_e) + A_2 \eta_1 (t^* - t^+) + BT \tag{20}$$

and

$$A_1 = \int_0^{N_u} \alpha(x)dx, \quad A_2 = \int_{N_u}^{N_f} \alpha(x)dx, \quad B = \int_{N_f}^{N} \alpha(x)dx \tag{21}$$

Substituting equations (7) and (8) into equation (20), and using equation (18), the optimal flow distribution (N_f) and p can be solved from the following optimization problem

$$\min STC = A(\delta - \omega \eta_1 p) \frac{N_f}{s} - A_2 \sigma p \frac{N_f}{s} + B \frac{\delta \theta N}{s} \tag{22}$$
$$\text{Subject to} : 0 \le p \le 1; \ 0 \le N_f \le N$$

where ω and σ are defined in Table 1. Equation (22) implicitly determines the optimal N_f and p, as long as the VOT distribution of the commuter population is known. Then, P, r, and r_b can be obtained by solving the following equation system

$$\alpha(N_f) T - P r_b = \alpha(N_f) \frac{(1-p)\delta N_f}{s} + P \tag{23}$$

$$(N - N_f) r_b + N_u r = N_f - N_u \tag{24}$$

$$P = \frac{\rho}{1 + r} \tag{25}$$

In equation (23), N_f represents the watershed dividing users between the two modes. Note that the travel cost for the users N_f is indifferent between choosing either auto or transit mode. Equation (24) states the credit conservation condition, while equation (25) as derived from equation (16) states the relationship between the credit price and the coarse toll.

The number of peak-time highway users $N_t = (1-p)N_f$, and accordingly the cost of peak and non-peak-time highway users are

$$c_t(x) = \alpha(x) \frac{\delta N_t}{s} + P, \quad c_u(x) = \alpha(x)(\delta - \omega \eta_1 p) \frac{N_f}{s} - Pr \tag{26}$$

Note that if $N_f = N$, that is, the transit is not used, then r_b should equal to zero, and r and P are obtained by solving equations (24) and (25).

The above analysis also provides a framework to evaluate the efficiency of a given TCS defined by r and

r_b. Specifically, one can solve N_f, p, and P from equations (23) to (25) and hence obtain the system travel cost, once r and r_b are given.

Changes in individual travel cost

Once the TCS is implemented, the whole traveling population can then be divided into four groups, as follows:

1. *The group with high VOT who remains on transit service*. Their travel costs will never increase because the transit mode is free of congestion, and these commuters benefit most from selling their earned credits to other commuters. The cost change is

$$\Delta c(x) = -Pr_b < 0 \tag{27}$$

2. *The group of commuters who are forced to shift from auto to transit*. Those commuters have relatively low VOTs, and by shifting modes, they experience a longer travel time, although the credit charge is avoided. In addition, they receive a subsidy to cover the additional travel time cost. The cost change is

$$\Delta c(x) = \alpha(x)T_t - Pr_b - \alpha(x)\frac{\bar{N}_f}{s}\delta = -Pr_b < 0 \tag{28}$$

3. *The group of commuters who remain on highway during peak-time window*. These travelers pay the credits and enjoy a travel time reduction because the highway is less congested with less demand. The travel costs can be higher or lower depending on whether the reduction of delay cost covers the credit charge. Comparing the two costs before and after the scheme, the cost change is

$$\Delta c(x) = \alpha(x)\frac{(1-p)\delta\hat{N}_f}{s} + P - \alpha(x)\frac{\delta\bar{N}_f}{s} \tag{29}$$

where $\bar{N}_f$ and $\hat{N}_f$ are give by equations (17) and (22). Taking the first-order derivative, we have $\Delta c'(x) < 0$, which implies that the commuters with higher VOT will experience more travel cost decrease.

4. *The group of commuters who remain on highway during non-peak-time window*. As will be seen later, in some cases, this group of commuters can benefit from the scheme, whereas in other cases, some of them will not

$$\Delta c(x) = \alpha(x)\frac{\delta}{s}\left(\hat{N}_f - \bar{N}_f\right) - Pr \tag{30}$$

The changes in the travel costs of commuters after implementing the TCS are summarized by the following piecewise function

$$\Delta c(x) = \begin{cases} \alpha(x)\frac{\delta}{s}\left(\hat{N}_f - \bar{N}_f\right) - Pr, & 0 < x < N_u \\ \alpha(x)\frac{\delta}{s}\left((1-p)\hat{N}_f - \bar{N}_f\right) + P, & N_u < x < \hat{N}_f \\ -Pr_b, & \hat{N}_f < x < N \end{cases} \tag{31}$$

where $\bar{N}_f$ is derived from equation (17), $\hat{N}_f$ and N_u are give by equations (22) and (18). The first equation of equation (31) is for the non-peak-time highway commuters, the second for the peak-time highway commuters, and the last for the transit users.

In the following section, we only consider the SW assumption and a uniform VOT distribution for better analytical tractability. Finding the analytical solutions for the other assumptions is tedious but relatively straightforward following the procedure used in SW. Hence, we shall examine the impact of different behavior assumptions in numerical experiment.

Special case: heterogeneous commuters

Finding an analytical solution to the above problem is tedious, and we shall examine the impacts of different behavior assumptions and congestion effects in numerical experiment. Below, we show how such an analysis can be conducted when VOT follows a uniform distribution, $\alpha(x) = ax$, and the SW assumption is employed.

Substituting $\alpha(x) = ax$ into equation (22) and using the first-order derivation equals to zero, that is, $\partial STC/\partial N_f = 0$, we have

$$N_f = \frac{2Ts}{3\delta(1-p+p^3)} \tag{32}$$

Using equation (17), then the above equation can be rewritten as follows

$$N_f = \frac{2\theta}{3(1-p+p^3)}N \tag{33}$$

Now, from equation (8), we can obtain the equilibrium credit price as follows

$$P = \frac{\rho}{1+r} = \frac{2N\theta\delta p}{3s(1-p+p^3)(1+r)}\alpha(N_u) \tag{34}$$

The credits reward to the transit riders should be set such that the income from selling them would be able to offset the cost difference between the two modes, that is

$$\alpha(N_f)T - Pr_b = \alpha(N_f)\frac{(1-p)N_f\delta}{s} + P \tag{35}$$

Finally, the credit conservation condition dictates that

$$(N - N_f)r_b + N_u r = N_f - N_u \tag{36}$$

Combining equations (33), (35), and (36), we can find

$$r = \frac{p^2\theta^*}{(\theta^* + p - 1)(\theta^* - \theta) + \theta p^3} - 1, \quad r_b = \frac{(\theta^* + p - 1)\theta^*}{(\theta^* + p - 1)(\theta^* - \theta) + \theta p^3} - 1 \tag{37}$$

where

$$\theta^* = \frac{3(1 - p + p^3)}{2} \tag{38}$$

Proposition 1. Under the system optimal tradable credit scheme (SO-TCS), (1) if $\theta \geq \theta^*$, only the highway should be used, and $p = 1/\sqrt{3}$ for auto users; (2) if $0 < \theta < \theta^*$, both transit and the highway should be used and $p = 1/\sqrt{3}$ for auto users; and (3) if $\theta = 0$, only transit should be used.

Proof. Condition 3 is trivial: if $T = 0$, the total system cost will be zero and everyone will use transit. To prove condition 1, first note that when both the primary route and transit are used, the minimum system cost can be written as

$$STC^* = \frac{a}{2}\frac{\theta\delta}{s}N^3\left(1 - \frac{4\theta^2}{27(1 - p + p^3)^2}\right)$$

If only the highway is used, then the minimum system cost is obtained when $p = 1/\sqrt{3}$ as

$$STC^{**} = \frac{a}{2}\frac{\delta N^3}{s}\left(1 - p + p^3\right)$$

We prove condition 1 by contradiction now. Suppose when $\theta \geq \theta^*$, both the highway and transit have to be used in order to minimize the total cost. This implies $STC^* < STC^{**}$, which in turn implies

$$\left(1 + p - p^3\right) < \frac{2\theta}{3}\sqrt{\frac{\theta}{3(\theta - 1) + 2/\sqrt{3}}}$$

Combining this inequality with equation (33) yields

$$N_f = \frac{2\theta}{3(1 - p + p^3)}N > \sqrt{3 - \frac{2}{\theta}\theta^*}N \geq N$$

The last inequality holds because $\theta \geq \theta^*$. This is a contradiction. Hence, when $\theta \geq \theta^*$, traveler would only use the highway, and the system cost is minimized when $p = 1/\sqrt{3}$.

The proof of Proposition 1 shows clearly that both the highway and the transit service should be used when $\theta < \theta^*$. It is easy to verify that to minimize STC^* with respect to p, p must equal to $1/\sqrt{3}$. This completes the proof.

Proposition 1 assures that at the system optimum, $p = 1/\sqrt{3}$. With p being determined, the key design variables for the TCS can now be obtained as follows

$$r = \frac{\theta - 1/2}{3/2 - \theta}, \quad r_b = \frac{\theta}{3/2 - \theta}, \quad P = \frac{a\delta N^2}{3s}\left(\frac{\theta}{\theta^*}\right)^2\left(\frac{3}{2} - \theta\right) \tag{39}$$

when $\theta \in (0, (3/2) - (1/\sqrt{3}))$, we have

$$-\frac{1}{3} \leq r \leq \sqrt{3} - 1, \quad 0 < r_b < \frac{3}{2}\sqrt{3} - 1, \quad 0 < P < \frac{a\delta N^2}{3\sqrt{3}s} \tag{40}$$

Note that the nonpeak users on highway may receive negative credits, that is, r becomes negative when the travel time of using transit is smaller than half of the no-toll generalized travel time on the highway ($\theta < 0.5$). This is intuitive since when transit service becomes so attractive such that its travel time is less than half of the $N\delta/s$, travelers should not be given credits for staying on the highway, even during the nonpeak time.

Hereafter, we examine the welfare effects of the system travel cost with TCS and compare it to the system trip cost under no-toll equilibrium. Equation (17) indicates that the number of highway users $\bar{N}_f = \theta N$. Therefore, the system trip cost under no-toll scheme can be formulated as $STC_N = a\delta N^3\theta/2s$, if $\theta \leq 1$. When $\theta > 1$, all commuters will stay on the transit line, and the system trip cost is $STC_N = a\delta N^3/2s$.

For SO-TCS, the analysis in Proposition 1 shows that when $\theta < \theta^*$, both the highway route and the transit line will be used, with the number of highway uses given as equation (33). When $\theta > \theta^*$, all users will stay on the transit line. Thereafter, the system trip cost under TCS can be formulated as follows

$$STC^* = \frac{a}{2}\frac{\delta N^3}{s}\begin{cases}\theta - \frac{4\theta^3}{27(1-p+p^3)^2}, & \text{if } \theta \leq \theta^* \\ 1 - p + p^3, & \text{if } \theta^* < \theta\end{cases} \tag{41}$$

where $\theta^* = (3/2)(1 - p + p^3)$ and $p = 1/\sqrt{3}$. Note that the efficiency gains of the SO-TCS with heterogeneous travelers is

$$e(\theta) = \frac{STC_N - STC^*}{STC_N} = 1 - \max\left\{1 - \frac{4\theta^2}{27(1 - p + p^3)}, \frac{1 - p + p^3}{\theta}\right\} \tag{42}$$

It is clear that the efficiency gains first increases from 0% to 20.5% as θ is increased from 0 to θ^* and then reached the top value 38.5% when θ reached 1.

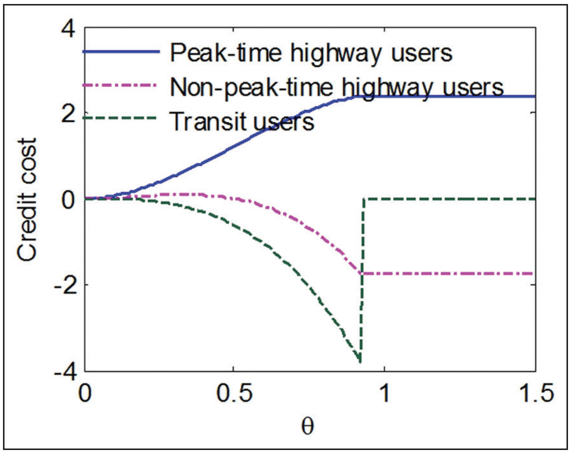

Figure 2. SO-TCS solutions with SW assumption.

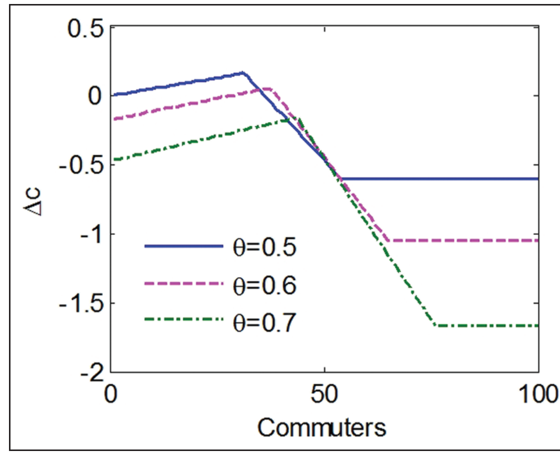

Figure 3. Change in travel cost with TCS.

Numerical examples

In this section, we provide numerical examples to illustrate the differences under the TCS discussed in the previous sections between homogeneous and heterogeneous traveling populations. We assume the demand $N = 100$, and the capacity of the bottleneck $s = 50$ (veh/h). The value attached to queuing delay α follows a uniform distribution $\alpha(x) = ax + b$ and is distributed in the interval $[\bar{\alpha} - \varepsilon, \bar{\alpha} + \varepsilon]$, while $\bar{\alpha}$ is the mean value and ε defined as the standard error. Thus, $a = 2\varepsilon/N$ and $b = \bar{\alpha} - \varepsilon$. Following Arnott et al.,[2] we set the mean values $\bar{\alpha} = 6.4\,\$/\text{h}$, $\bar{\beta} = 3.9\,\$/\text{h}$, and $\bar{\gamma} = 15.21\,\$/\text{h}$, and from which, we let $\eta_1 = 0.609$ and $\eta_2 = 2.377$.

Figure 2 shows how the SO credit costs under SW assumption changes with parameter θ for the three groups. The peak-time highway users incur the largest cost to acquire credits, whereas the transit users benefit most from selling their credits to other commuters. However, the credit cost of peak-time highway users invariably peaks and that of transit users reaches zero when θ is around 0.92 (i.e. all users will be on the highway).

We further investigate the difference in equilibrium trip cost between the TCS and the no-toll scheme with SW assumption, as denoted by Δc. Figure 3 shows the variation of Δc with different values of θ. It can be seen that all three curves are composed of three piecewise linear functions (i.e. three commuter groups), and the value of Δc can be either positive or negative when $\theta = 0.5$, implying that the TCS can either increase or decrease the commuter trip cost. One can observe from this figure that the slop of the peak-time highway users is decreasing with VOT. This means that the commuters with higher VOT are better off more quickly than lower value commuters, while the curve of the transit users is negative and benefit most from selling their earned credits to other commuters.

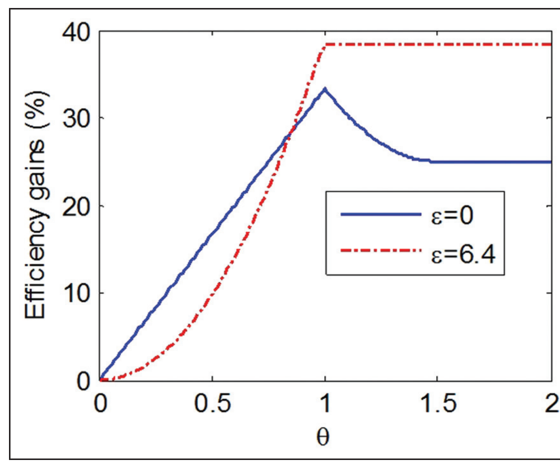

Figure 4. Percentage reduction in system travel cost due to TCS as a function of θ.

Figure 4 shows the percentage reduction in system travel cost (excluding credit) generated by TCS based on homogeneous commuters and a special case of heterogeneity, as a function of θ. It can be seen that the TCS can achieve up to 33.3% efficiency gains in the homogeneous case, while for the heterogeneous commuters, the efficiency gains increase with increasing θ value and reach the top 38.5% at $\theta = 1$. For small θ value, the efficiency gains have been overestimated with homogeneity, while a higher θ would actually increase the efficiency gains in the context of heterogeneity.

Figure 5 shows how the design variable of SO-TCS as well as the corresponding equilibrium flow pattern changed with T and ε. Note that ε implies the distribution effect. We vary the ε value between 0 and 6.4, while ε equals to zero represents the homogeneous case. Figure 5(a) shows that the system travel cost (i.e. STC) changes monotonically with T value. For homogeneous case (i.e. $\varepsilon = 0$), when T is around 1.46 (i.e. $\theta = 1.5$),

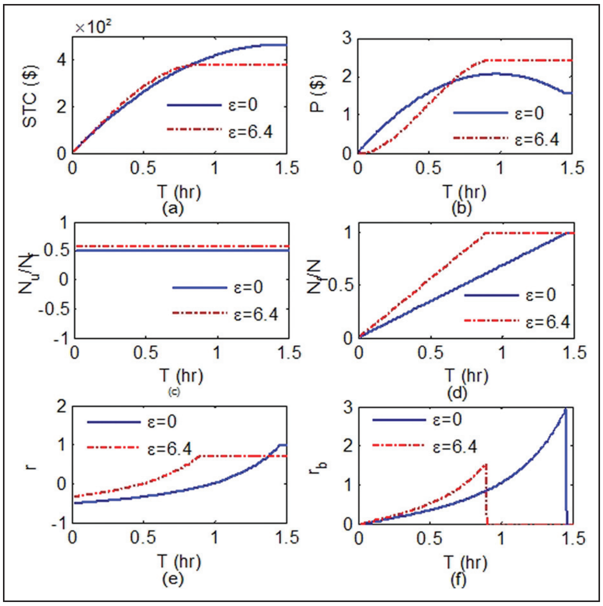

Figure 5. SO-TSC solutions changed with T under SW assumption.

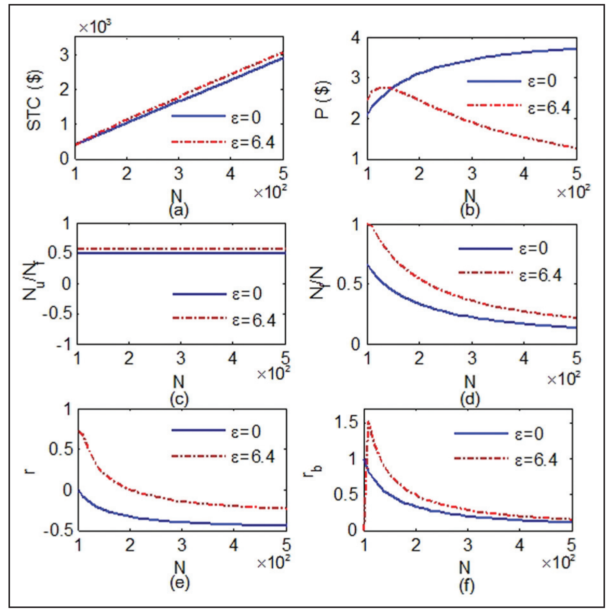

Figure 6. SO-TCS solutions changed with N under SW assumption.

the system travel cost invariably peaks. This is because no traveler would use the transit line even under SO-TCS. In addition, for $\varepsilon = 6.4$, only the highway line should be used when T is not less than 0.90.

Figure 5(b) shows the credit price (i.e. P) is a concave function of T for homogeneous case. For small T value, however, a higher ε would actually lower credit price. Moreover, under heterogeneous case, the credit price is underestimated with homogeneity for high T value.

Figure 5(c) shows that the percentage of the non-peak-time users (i.e. p) is increasing with ε value. This is as expected since increasing the ε value is equivalent to increasing the difference in VOT among commuters and pushing more commuters to having higher or lower VOTs; therefore, more commuters will choose to depart from home outside the peak time to avoid the increased overall trip cost (including credit charge). Figure 5(d) suggests that ε plays an important role in determining flow distribution between the two alternatives. However, a higher ε always pushes more users to the highway.

Figure 5(e) reveals that the off-peak-time users ought to be rewarded more credits (i.e. r) when the travel time T on transit is larger. In addition, a higher ε would actually increase the number of reward credits. Interestingly, the same trend is observed for the transit users, as shown in Figure 5(f).

Figure 6 repeats the above sensitivity analysis for when the demand N is increased gradually from 100 to 500. Figure 6(a) shows that the system travel cost (i.e. STC) increase with N. As expected, larger ε values correspond to larger system travel cost, and the relative influence of ε is stronger for larger T. Figure 6(b) shows that the credit price (i.e. P) increases with N when $\varepsilon = 0$, while for $\varepsilon = 6.4$, the result turns reverse. Figure 6(d) shows that larger N consistently reduces the percentage of highway users, although the trend is slowed down by larger ε on commuters distribution. Similar to Figure 5(e) and (f), Figure 6(e) and (f) suggests that more credits should be rewarded in respond to larger ε.

Figure 7 shows the sensitivity analysis results when capacity s is changed from 20 ($\theta = 0.2$) to 100 ($\theta = 2$). Note the patterns described in Figure 7(c)–(f) are remarkably similar to the corresponding plots in Figure 5. Figure 7(a) shows that the system travel cost (i.e. STC) decrease monotonically as s increases, and larger ε values correspond to larger system travel cost for small s. However, when s is increasing, a higher ε would actually lower the system travel cost. Figure 7(b) shows that variance ε seems to lower the credit price for small s.

Figure 8 compares the results for the three behavior assumptions when $\varepsilon = 6.4$ (heterogeneous case). Figure 8(a) shows that the MD and BI assumptions lead to slightly lower and higher system travel costs, respectively. Figure 8(b) shows that the MD assumption gives higher credit price, while the BI assumption gives lower credit price. Figure 8(c) shows that the percentage of non-peak-time highway user (i.e. Nu/Nf) is slightly affected by the behavior assumption. The percentage is about 60.1% for the MD assumption, 57.7% for SW, and about 53.4% for BI. The percentages, however, are not affected by T. Figure 8(e) and (f) shows that the differences in the credit given by these assumptions are not noticeably larger, especially for small T.

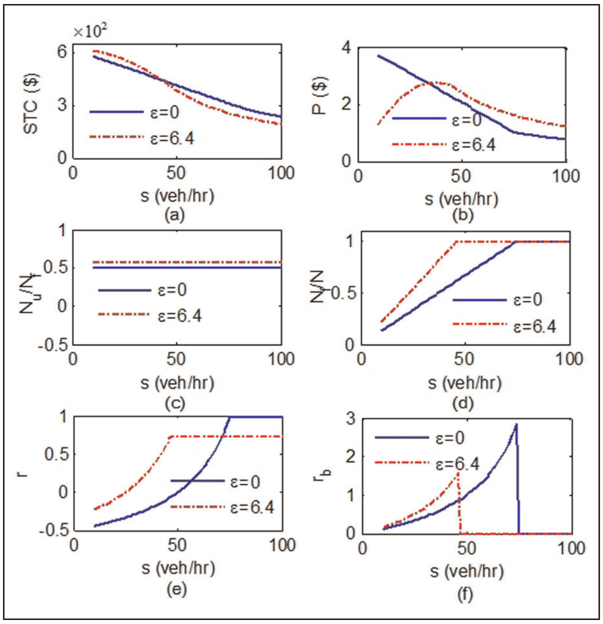

Figure 7. SO-TCS solutions changed with *s* under SW assumption.

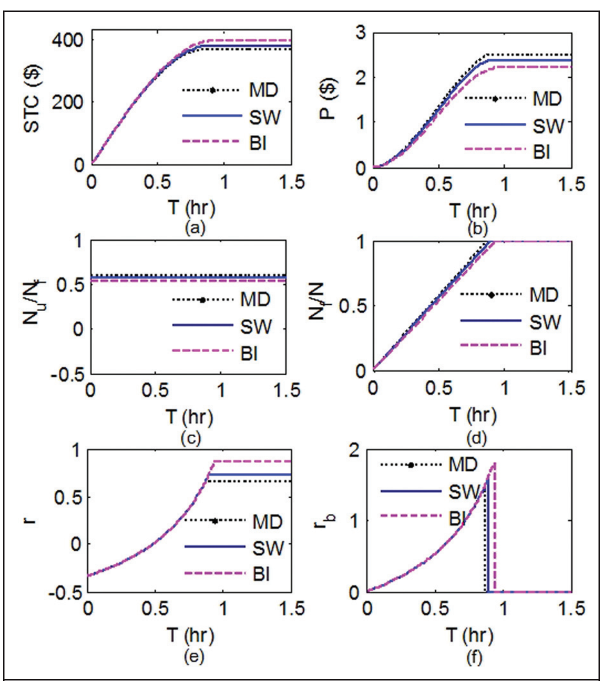

Figure 8. SO-TCS solutions under different behavior assumptions ($\varepsilon = 6.4$).

Summary and concluding remarks

In this article, the economics of a morning commute problem on a single bottleneck is investigated for a heterogeneous commuter population and with a choice of auto and transit mode of travel. First, we derived the equilibrium departure-time profiles for the heterogeneous commuters, under no-toll condition. The corresponding individual trip cost and the total system trip cost are presented. For heterogeneous commuters, queuing delay is a pure loss, and it remains that congestion toll is effective for reducing the queues behind the bottleneck. We proposed a TCS as an alternative demand management strategy to replace the step-coarse toll. TCS works like a single-coarse toll, except that the price of the credit is determined by the competitive market. It is a combination of two subschemes: a credit charging and a credit distribution schemes. We analyzed TCS under three different behavior assumptions, namely, MD, SW, and BI. With each behavior assumption, we presented the optimum ratio of the unit of reward and charge credits. For TCS with a coarse toll alternative, the departure pattern of commuters with heterogeneous commuters is similar to the homogeneous case. Still, a difference that arises for the ratio of the unit of reward and charge credits, which is due to the pattern of distribution effect, is different with respect to VOT.

The main focus of this article is the two-mode problem under a tradable travel credit scheme, with a variety of assumptions about commuter' behavior in response to the discontinuous credit charge introduced at the boundary of the peak-time window. In the simple case (when VOT follows a linear function for the SW assumption), the design of SO-TCS, including the choice of the peak-time windows as well as the number of credit rewarded to or charged on different groups of commuters, can be described using simple formulas. These analytical results indicate that the equilibrium solutions are only affected by the relative generalized travel time of the two routes (θ). To better deal with welfare estimation under user heterogeneity, different θ value is introduced for auto and transit modes to capture the realistic aspect associated with the two-mode equilibrium. We show that SO-TCS can either increase or decrease the commuter trip cost: the peak-time highway commuters with higher VOT are better off more quickly than lower VOT users, while the commuters at the highest end of the VOT distribution who take transit are made better off by SO-TCS.

With a general VOT distribution of heterogeneous users, the characteristics of the SO-TCS under the above three different behavioral assumptions are compared with those with homogeneous commuters. It showed that the characteristics of traffic patterns for homogeneous and heterogeneous commuters are markedly different. Without considering commuter heterogeneity, the number of peak-time highway commuters, the credit reward-to-charge ratio and the transit users are all overestimated. The ε value on the VOT distribution increases the difference in VOT among commuters and pushing more commuters to having higher or lower VOTs; therefore, more commuters will choose to depart from home outside the peak-time interval, and more commuters will switch into highway from transit line to avoid the increased overall trip cost (including credit charge). In addition, the ε value can lead to higher system cost and higher credit price and can cause more credits to be rewarded to both the off-peak-time on highway and transit users. Different assumptions about commuters' behavior do affect most SO-TCS design variables, and the influences tend to be slightly magnified by the ε value on VOT.

Overall, our findings suggest that it is important to take into account the heterogeneity of commuters in assessing the impacts of tradable credit under two-mode system. The analyses presented in this article still leave out a few important real-world features, of which the most critical is perhaps corridor network with multiple bottlenecks. In our future work, we plan to further extend the application of the credit scheme to solve a multi-route and multi-mode problem and to manage traffic demands, parking spots, and emissions.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was financially supported by the National Basic Research Program of China (2012CB 725401) and the Talent Fund of Beijing Jiaotong University (2014RC021).

References

1. Vickrey WS. Congestion theory and transport investment. *Am Econ Rev* 1969; 59: 251–260.
2. Arnott R, de Palma AD and Lindsey R. Economics of a bottleneck. *J Urban Econ* 1990; 27: 111–130.
3. Qian Z and Zhang HM. Modeling multi-modal morning commute in a one-to-one corridor network. *Transport Res C: Emer* 2011; 19: 254–269.
4. Xiao LL, Huang HJ and Liu R. Congestion behavior and tolls in a bottleneck model with stochastic capacity. *Transport Sci* 2015; 49: 46–65.
5. Pigou AC. *The economics of welfare*. 1st ed.London: Macmillan and Company, 1920.
6. Knight FH. Some fallacies in the interpretation of social cost. *Quart J Econ* 1924; 38: 582–606.
7. Evans AW. Road congestion pricing: when is it a good policy? *J Transp Econ Policy* 1992; 26: 213–244.
8. Hau TD. Congestion pricing and road investment. In: Button KJ and Verhoef E (eds) *Road pricing, traffic congestion and the environment*. Cheltenham: Edward Elgar, 1998, pp.39–78.
9. Yang H and Zhang X. Multiclass network toll design problem with social and spatial inequity constraints. *J Transp Eng: ASCE* 2002; 128: 420–428.
10. Cohen Y. Commuter welfare under peak-period congestion tolls: who gains and who loses? *Int J Transp Econ* 1987; 14: 239–266.
11. Arnott R, de Palma AD and Lindsey R. The welfare effects of congestion tolls with heterogeneous commuters. *J Transp Econ Policy* 1994; 28: 139–161.
12. Huang HJ. Fares and tolls in a competitive system with transit and highway: the case with two groups of commuters. *Transport Res E: Log* 2000; 36: 267–284.
13. Huang HJ. Pricing and logit-based mode choice models of a transit and highway system with elastic demand. *Eur J Oper Res* 2002; 140: 562–570.
14. Qian Z and Zhang HM. The morning commute problem with heterogeneous travelers: the case of continuously distributed parameters. *Transport A: Transport Sci* 2013; 9: 178–203.
15. Verhoef ET, Nijkamp P and Rietveld P. Tradeable permits: their potential in the regulation of road transport externalities. *Environ Plann B* 1997; 24: 527–548.
16. Akamatsu T, Sato S and Nguyen LX. Tradable time-of-day bottleneck permits for morning commutes. *Doboku Gakkai Ronbunshu D* 2007; 62: 605–620.
17. Yang H and Wang X. Managing network mobility with tradable credits. *Transport Res B: Meth* 2011; 45: 580–594.
18. Zhang XN, Yang H and Huang HJ. Improving travel efficiency by parking permits distribution and trading. *Transport Res B: Meth* 2011; 45: 1018–1034.
19. Liu W, Yang H and Yin Y. Expirable parking reservations for managing morning commute with parking space constraints. *Transport Res C: Emer* 2014; 44: 185–201.
20. Liu W, Yang H, Yin Y, et al. A novel permit scheme for managing parking competition and bottleneck congestion. *Transport Res C: Emer* 2014; 44: 265–281.
21. Xiao F, Qian Z and Zhang HM. Managing bottleneck congestion with tradable credit. *Transport Res B: Meth* 2013; 56: 1–14.
22. Nie Y. A new tradable credit scheme for the morning commute problem. *Netw Spat Econ*. Epub ahead of print 19 June 2013. DOI: 10.1007/s11067-013-9192-8.
23. Laih CH. Queueing at a bottleneck with single- and multi-step tolls. *Transport Res A: Pol* 1994; 28: 197–208.
24. Laih CH. Effects of the optimal step toll scheme on equilibrium commuter behavior. *Appl Econ* 2004; 36: 59–81.
25. Lindsey R, van den Berg V and Verhoef ET. Step tolling with bottleneck queuing congestion. *J Urban Econ* 2012; 72: 46–59.
26. Ren HL, Xue Y, Long JC, et al. A single-step-toll equilibrium for the bottleneck model with dropped capacity. *Transport B: Transport Dyn*. Epub ahead of print 20 October 2014. DOI: 10.1080/21680566.2014.967327.
27. Tian LJ, Yang H and Huang HJ. Tradable credit schemes for managing bottleneck congestion and modal split with heterogeneous users. *Transport Res E: Log* 2013; 54: 1–13.
28. Van de Berg V. Coarse tolling with heterogeneous preferences. *Transport Res B: Meth* 2014; 64: 1–23.
29. Xiao F, Qian Z and Zhang HM. The morning commute problem with coarse toll and nonidentical commuters. *Netw Spat Econ* 2011; 11: 343–369.