Transportation Research Part C 151 (2023) 104121

Contents lists available at ScienceDirect

Transportation Research Part C

journal homepage: www.elsevier.com/locate/trc

Market design for tradable mobility credits

Siyu Chen [a], Ravi Seshadri [b,*], Carlos Lima Azevedo [b], Arun P. Akkinepally [c], Renming Liu [b], Andrea Araldo [d], Yu Jiang [b], Moshe E. Ben-Akiva [a]

[a] Civil and Environmental Engineering Department, Massachusetts Institute of Technology, Cambridge, MA, USA
[b] Department of Technology, Management and Economics, Technical University of Denmark, Denmark
[c] Caliper Corporation, MA, USA
[d] Samovar, Télécom SudParis, Institut Polytechnique de Paris, France

ARTICLE INFO

Keywords:
Tradable mobility credits
Demand management
Traffic management
Simulation

ABSTRACT

Tradable mobility credit (TMC) schemes are an approach to travel demand management that have received significant attention in recent years as a promising means to mitigate the adverse environmental, economic, and societal effects of urban traffic congestion. This paper proposes and analyzes alternative market models for a TMC system – focusing on market design aspects such as allocation/expiration of credits, rules governing trading, transaction fees, and regulator intervention – and develops a methodology to explicitly model the dis-aggregate behavior of individuals within the market. Extensive simulation experiments are conducted within a combined mode and departure-time context for the morning commute problem to compare the performance of the alternative designs relative to congestion pricing and a no-control scenario.

The results indicate that small, fixed transaction fees can effectively mitigate undesirable speculation in the market without a significant loss in efficiency (total welfare) whereas proportional transaction fees are less effective, both in terms of efficiency and in avoiding undesirable speculation. Further, an allocation of credits in continuous time can be beneficial in dealing with non-recurrent events and avoiding concentrated trading activity. In the presence of income effects, despite small, fixed transaction fees, the TMC system yields a marginally higher social welfare than congestion pricing while attaining revenue neutrality. Moreover, it is more robust in the presence of forecasting errors and non-recurrent events due to the adaptiveness of the market. Finally, as expected, the TMC scheme is more equitable (when revenues from congestion pricing are not redistributed) although it is not guaranteed to be Pareto-improving when credits are distributed equally.

1. Introduction

Historically, transportation network inefficiencies and externalities such as congestion and vehicular emissions have been addressed through road pricing, which although used in several cities worldwide, is plagued by issues of inequity and public acceptability (Tsekeris and Voß, 2009; de Palma and Lindsey, 2011). An alternative approach to travel demand management that has received increasing attention in the transportation domain in recent years is quantity control — in particular, tradable mobility credit (TMC) schemes (Fan and Jiang, 2013; Grant-Muller and Xu, 2014; Dogterom et al., 2017). Within a TMC scheme, the regulator provides an initial endowment of mobility credits or tokens to potential travelers. In order to use the road network or transportation system, users need to spend a certain number of tokens (i.e., pay a tariff) that could vary with the attributes or performance of the

* Corresponding author.
E-mail address: ravse@dtu.dk (R. Seshadri).

https://doi.org/10.1016/j.trc.2023.104121
Received 7 July 2022; Received in revised form 19 January 2023; Accepted 30 March 2023
Available online 19 April 2023

specific mobility alternative used. The tokens can be bought and sold in a market at a price determined endogenously by token demand and supply.

In principle, TMC schemes are appealing since they offer a means of directly controlling quantity, they are revenue neutral in that there is no transfer of money to the regulator, and they are viewed as being less vertically inequitable than congestion pricing (de Palma and Lindsey, 2020). Despite these promises, several important questions remain with regard to the design and functioning of the market within TMC schemes, an aspect critical to the effective operationalization of these schemes. For instance, how should the allocation and expiration of tokens be designed? What rules should govern trading behavior in the market so as to avoid undesirable speculation and trading (see Brands et al. (2020) for more on this), and yet ensure efficiency and revenue neutrality? How should the regulator intervene in the market in the presence of special or non-recurrent events? What is the role and impact of transaction fees? Despite the large body of literature on TMCs, issues of market design, market dynamics and market behavior have received relatively little attention although being critical to the successful real-world deployment of a TMC scheme.

This paper aims to address these issues and contributes to the existing literature in several respects. First, we propose alternative market models (focusing on all aspects of market design including allocation/expiration of credits, rules governing trading, transaction fees, regulator intervention, and price dynamics) for a TMC system and develop a methodology that explicitly models the dis-aggregate behavior of individuals within the market. Second, we conduct extensive simulation experiments within a departure-time and mode choice context for the morning commute problem to compare the performance of the alternative designs relative to congestion pricing and a no-control scenario. The simulation experiments employ a day-to-day assignment framework wherein transportation demand is modeled using a logit-mixture model (with income effects) and supply is modeled using a standard bottleneck model. The experiments yield insights into market design and the comparative performance of the TMC system relative to congestion pricing.

The results indicate that small, fixed transaction fees can effectively mitigate undesirable speculation in the market without a significant loss in efficiency (total welfare) whereas proportional transaction fees are less effective both in terms of efficiency and in avoiding undesirable speculation. Further, an allocation of tokens in continuous time provides the TMC system with additional flexibility (compared to a *lump-sum* allocation) and can be beneficial in dealing with non-recurrent events. With regard to the relative performance vis-a-vis congestion pricing, the results indicate that the TMC scheme attains a marginally higher social welfare (under income effects and a small fixed transaction fee). Further, the TMC scheme is more robust in the presence of forecasting errors (during the optimization of the toll profiles) and can achieve a higher welfare than congestion pricing through the adjustment of token allocation. Finally, as expected, the TMC scheme is more equitable (when revenues from congestion pricing are not redistributed) although it is not guaranteed to be Pareto-improving when tokens are distributed equally.

The paper addresses a growing and imminent need to develop methodologies to realistically model TMCs that are suited for real-world deployments and can help us better understand the performance of these systems — and the impact in particular, of market dynamics.

The rest of the paper is organized as follows. Section 2 presents a review of the literature and identifies contributions of our paper. Sections 3 and 4 propose a market design for TMCs and a framework for modeling market behavior, respectively. Section 5 introduces the simulation model, including demand, supply and day-to-day learning. Section 6 describes findings from extensive simulation experiments and Section 7 provides concluding remarks and directions for further research.

2. Review of literature

Although early work on the use of tradable mobility credits (TMCs; also termed TCS or Tradable Credit Schemes in the literature) in transportation dates back several years (Verhoef et al., 1997; Raux, 2007; Goddard, 1997), formulations of the market and network equilibrium for TMCs are more recent, pioneered by the work of Yang and Wang (2011) who proposed a user equilibrium variant to model a TMC. Their work, along with advancements in technology and the widely recognized limitations of congestion pricing, has spurred interest in TMCs for transportation network management. Extensive reviews may be found in Grant-Muller and Xu (2014), Fan and Jiang (2013) and Dogterom et al. (2017). We provide a brief summary of existing literature, limiting our attention to that of mobility management (in the context of both entire networks and single bottlenecks) although applications may also be found in parking.

In the model of Yang and Wang (2011), the regulator distributes a pre-specified number of credits to travelers, charges a link-specific credit tariff and allows trading of credits within a market. They identify conditions under which the network and market equilibrium are unique. Extensions to their model have been proposed to incorporate heterogeneity in the value of time (Wang et al., 2012) and multiple user classes (Zhu et al., 2015) using variational inequality formulations to establish existence and uniqueness of the equilibrium. He et al. (2013) employed a similar equilibrium approach considering allocations of credits to not just individual travelers, but to transportation firms such as logistics companies and transit agencies. The effect of transaction costs in a TMC scheme with two types of markets (auction-based and negotiated) is considered by Nie (2012). In contrast with the aforementioned TMC schemes, Kockelman and Kalmanje (2005) and Gulipalli and Kockelman (2008) proposed a system of credit-based congestion pricing (termed CBCP) where credits are allowances used to pay tolls.

TMC schemes have also been studied in the context of managing congestion at a single bottleneck (or simple two route networks) by achieving peak spreading. Nie and Yin (2013) modeled a tradable credit scheme that manages commuters' travel choices and attempts to persuade commuters to spread their departure times evenly within the rush hour and between alternative routes (see also Nie (2015)) whereas Tian et al. (2013) investigated the efficiency of a tradable travel credit scheme for managing bottleneck congestion and modal split in a competitive highway/transit network with heterogeneity. Along related lines, Xiao et al. (2013)

studied a tradable credit system (consisting of a time-varying credit charge at the bottleneck wherein the credits can be traded and the price is determined by a competitive market) to manage morning commute congestion with both homogeneous and heterogeneous users. More recently, Bao et al. (2019) examined the existence of equilibria under tradable credit schemes using different models of dynamic congestion and Akamatsu and Wada (2017) proposed a tradable bottleneck credit scheme where the regulator issues link- and time-specific credits required for passing through a certain link or bottleneck in a pre-specified time period. Liu et al. (2023) considered distance-based token tariffs in a TMC scheme within a departure-time setting and examined its performance using a trip-based MFD supply model.

In contrast with the previously described literature that largely focuses on variants of the standard user equilibrium under TMC schemes, a related stream of research examines the design of the TMC schemes using bi-level optimization formulations in different contexts (Wu et al., 2012; Bao et al., 2017; Wang et al., 2014). On the other hand, the comparison of efficiency properties of tradable credits and congestion pricing has received relatively lesser attention. de Palma et al. (2018) performed a comparative analysis of the two instruments in a simple static transportation network (see also Seshadri et al. (2022) for a within-day dynamic setting) and showed that as long as there is no uncertainty, price and quantity regulation are equivalent as in the regular market case studied by Weitzman (1974). In the presence of uncertainty and strongly convex congestion costs, the TMC instrument outperforms the pricing instrument in efficiency terms (see also de Palma and Lindsey (2020) for comparisons in the case with one route and time period under elastic demand). Akamatsu and Wada (2017) reached similar conclusions (see also Shirmohammadi et al. (2013)), demonstrating the equivalence of the tradable permit system and a congestion pricing system when the road manager has perfect information of transportation demands. On the behavior side, several stated preference studies have highlighted the importance of key factors from the perspective of behavioral economics and cognitive psychology towards tradable credits (Dogterom et al., 2017).

To the best of our knowledge, Brands et al. (2020) is the only study thus far to examine issues of market design for tradable credits. They conducted a lab-in-the-field experiment in a parking context and examined performance of the credit system empirically in terms of several criteria including undesirable speculation, price stability and transaction costs.

In summary, despite the large body of research on TMCs, several important issues remain to be studied. First, the modeling of the market has received little attention and almost all the studies employ an equilibrium approach to model the credit market (with the notable exception of Ye and Yang (2013) who model the price and flow dynamics of a tradable credit scheme). The literature has – to the best of our knowledge – thus far not attempted to realistically model the disaggregate behavior of individuals within the market. This would enable the consideration of empirically observed phenomena such as loss aversion, endowment effects, mental accounting, and day-to-day learning (Dogterom et al., 2017). Second, despite being a critical step towards real-world deployment, design aspects of the credit market have received little attention. In particular, features such as token allocation/expiration, trading, intervention, and transaction fees, and their impact on efficiency and market behavior remain to be studied. Finally, income effects, which impact both efficiency and equity, have received relatively little attention (with the exception of Wu et al. (2012) who consider it in a route choice setting). This paper aims to address these gaps by proposing and analyzing alternative market designs of the TMC system and investigating their performance relative to congestion pricing using realistic models of traveler behavior (with heterogeneity and income effects) and congestion.

3. Market design

In this section, we focus on market design for a tradable credit scheme. Within the TMC scheme, the regulator provides a token endowment to all potential travelers. The application we explore involves a daily commute context where, in order to use the road network at a particular time-of-day (e.g., for a given departure time interval), travelers have to pay a pre-specified toll in tokens that does not vary from day to day. In other words, the toll in tokens is dynamic and varies by time-of-day, but is fixed across days. The rationale for this assumption is that modifying the toll in tokens from day to day would involve communicating the tariff or toll structure on a daily basis, which is complicated, particularly in large general networks (for instance, the electronic road pricing or ERP scheme in Singapore includes dynamic tolls, which are revised only every three months or longer).

In general, the design choices we employ are motivated by the need to ensure that the TMC system is practical, acceptable, and transparent without significant loses in efficiency. This involves keeping transaction costs to a minimum and price volatility low, and preventing undesirable speculation (for example, profiting from selling tokens and buying them back later). In this regard, our design draws on the experiences of the lab-in-the-field experiment of Brands et al. (2020).

3.1. Market setup and role of the regulator

The regulator operates a market where tokens can be bought and sold at a prevailing market price and may also levy pre-specified transaction fees for buying and selling. All transactions take place between an individual and the regulator directly, who guarantees all buying and selling requests. This central market with a regulator who acts as a price setting intermediary is similar to the *virtual bank* in Brands et al. (2020), who observe that such a market can significantly reduce transaction costs (associated with information acquisition, negotiation, finding a potential buyer or seller etc.) compared to designs that involve consumer to consumer trading (and over existing designs such as Dutch and English auctions, sealed-bid auctions and Vickrey auction markets). It is clearly desirable to reduce transactions costs, which can cause the system to deviate from desired equilibria (Nie, 2012).

With regard to modeling consumer to consumer trading, there exists a rich literature on mechanism design. This includes simple bilateral trades with a single buyer and seller (Myerson and Satterthwaite, 1983), double auctions in which every seller can trade a single item with every buyer (McAfee, 1992), and in the transportation context — car sharing auctions (Hara and Hato, 2017), double sided-auctions in ride-sourcing platforms (Zhang et al., 2015) and peer-to-peer ride-sharing systems (Tafreshian and Masoud, 2022). We omit further discussions of this given our assumption of trading with the regulator.

3.2. Token allocation and expiration

Moving to the token allocation or endowment, we adopt a 'continuous time' approach wherein tokens are acquired (provided by the regulator) at a certain rate over the entire day and each token has a lifetime (i.e., it expires after a certain period specified by the regulator). The expiration of tokens avoids undesirable consequences of the TMC system that can compromise public acceptability such as speculative behavior and hedging in the market. The 'continuous' allocation on the other hand avoids concentrated trading activities and excessive trading near a boundary (a time period when a large amount of tokens expire at the same time, such as for instance, in a *lump-sum* allocation). It also provides more degrees of freedom for the regulator to intervene than that of a 'lump sum' allocation which distributes tokens at the beginning of each day. A comparison between the two allocation approaches will be performed through numerical experiments presented in Section 6.6.

As a result, each individual acquires tokens at a constant rate r over the entire day (credited into a *wallet*) and each token has a lifetime L to avoid speculation and hoarding. Let $x_n^d(t)$ denote traveler n's token account (or wallet) balance at time t on day d. A full wallet state indicates that the number of tokens in the wallet has reached a maximum (Lr), and in the absence of traveling or selling, does not change since the acquisition of new tokens is balanced by an expiry of old tokens. Thus, a full wallet implies that the oldest token in an individual's account has an age of L. In contrast, when the account is not in a full wallet state, it increases by an amount $r\Delta_t$ in a unit time interval Δ_t.

3.3. Rules governing buying and selling

Several additional assumptions regarding market design are noteworthy — these serve to avoid quantity buildup and market manipulation. First, travelers can only buy tokens from the regulator at the time of traveling for immediate use, i.e., only if they wish to travel and are short of tokens. This prevents one type of undesired speculation wherein users buy tokens (without intending to use them for travel) only to sell them later in order to make a profit. Second, when they sell tokens to the regulator, they have to sell all tokens in their wallet. Third, buying and selling cannot happen at the same time, i.e., travelers can sell all tokens anytime except at the time of buying. The second assumption differs from the design of Brands et al. (2020), who assume that tokens can be traded per piece, and implications of this assumption warrant more investigation, particularly when the market prices vary within-day. Since a large part of our experiments do not involve within-day dynamic prices and given that it considerably simplifies the modeling of selling behavior, we defer the relaxation of this assumption to future research.

3.4. Account evolution

Let $T(t)$ denote the toll in tokens to travel at time t, $\tilde{t}_n^d$ represent the departure time of traveler n on day d and D represent the duration of one day. Note that in the simulation framework (Section 5), time will be discretized into intervals of a specified size; for now, we treat it as continuous. Let r denote the allocation rate, L denote token lifetime, and $x_n^d(t)$ denote traveler n's token account balance at time t on day d. At time t on day d, traveler n can perform one and only one of the following actions:

1. Perform a trip if $t = \tilde{t}_n^d$.
   - If $x_n^d(t) \geq T(t)$, she consumes $T(t)$ tokens. Her account balance at $t + \Delta_t$, $x_n^d(t + \Delta_t)$, can be written as:
   $$x_n^d(t + \Delta_t) = \min\left(x_n^d(t) - T(t) + r\Delta_t, Lr\right), \tag{1}$$
   where the cap Lr ensures that tokens with life greater than L expire.
   - If $x_n^d(t) < T(t)$, she needs to buy $T(t) - x_n^d(t)$ tokens. Her account balance $x_n^d(t + \Delta_t)$ becomes:
   $$x_n^d(t + \Delta_t) = r\Delta_t, \tag{2}$$
   since all of $x_n^d(t)$ and the newly bought tokens are used to travel.
2. Does nothing. Her account balance $x_n^d(t + \Delta_t)$ becomes:
   $$x_n^d(t + \Delta_t) = \min\left(x_n^d(t) + r\Delta_t, Lr\right). \tag{3}$$
3. Sells all tokens $x_n^d(t)$. Her account balance becomes:
   $$x_n^d(t + \Delta_t) = r\Delta_t. \tag{4}$$

3.5. Token price dynamics

The allocation of tokens and the resulting evolution of account balances described in 3.2 and 3.4 occur 'within-day'. In contrast, we assume that the adjustment of the token price happens only at end of each day (day-to-day adjustment) and hence, prices are constant within-day. Note that the frequency of price adjustment is another design consideration within the TMC system. In the most general case, the price adjustment may occur after each transaction, which is the design adopted in Brands et al. (2020). In the application that we consider in Sections 5 and 6, since we have time-dependent tolls that vary within-day, we deem the

Table 1
Market elements for the tradable mobility credits system.

Elements	Design	Motivation
Allocation	Lump-sum Continuous	Simple; automated trading Avoid concentrated trading; additional control
Expiration	Lifetime	Avoid quantity buildup
Transaction fee	Proportional Fixed	Avoid undesirable speculation (e.g., frequent selling)
Price adjustment	Day to day constant adjustment	Balance demand and supply
Market rules governing trading		

assumption of constant within-day prices reasonable. Thus, throughout the model description in Section 5, we will assume that token prices are constant within-day and only change from day to day.

The marketplace dictates the token price p^d on day d, which is adjusted according to an apriori rule established by the regulator to achieve revenue neutrality. The price p^d is modified daily with a deterministic rule considering the regulator revenue K^{d-1} (net revenue from all buying and selling transactions of users) from the previous day as follows

$$p^d = \begin{cases} p^{d-1} & K^{d-1} \in [-\bar{K}, \bar{K}] \\ p^{d-1} + \Delta p & K^{d-1} < -\bar{K} \\ p^{d-1} - \Delta p & K^{d-1} > \bar{K}, \end{cases} \tag{5}$$

where Δp is a constant parameter representing the price change. The regulator revenue K^{d-1} is an outcome of the disaggregate token buying and selling decisions of users over the course of day $d-1$, which in turn depend on the mobility choices of the users. These dependencies are explained in more detail in Sections 4 and 5.

$\bar{K}$ is a constant parameter representing a regulator revenue threshold to adjust the price and ensures that price will not fluctuate for small regulator revenues close to zero. Price is ensured to be positive and below a certain cap p_m as follows:

$$p^d = \max\left(0, \min\left(p^d, p_m\right)\right). \tag{6}$$

3.6. Transaction fees for buying and selling

We assume that the regulator levies a two-part (fixed and proportional) transaction fee for both buying and selling transactions. Let F_S^P, F_B^P ($F_S^P, F_B^P \geq 0$) denote the proportional part of selling and buying transaction fees (this component of the transaction fee is proportional to the amount of the trade), and F_S^F, F_B^F ($F_S^F, F_B^F \geq 0$) denote the fixed part of selling and buying transaction fees. The effect of transaction fees on market behavior and efficiency will be examined in Section 6.4.

The revenue obtained from selling y tokens ($y \leq Lr$) with transaction fees on day d at time t can be written as,

$$S(y) = y p_s^d - F_S^F, \tag{7}$$

where $p_s^d = p^d(1 - F_S^P)$ is the token market price adjusted for the proportional selling transaction fee. Transaction fees and price are not expressed in function inputs for conciseness.

The cost of buying y tokens ($y \leq Lr$) with transaction fees at time t on day d can be written as,

$$B(y) = y p_b^d + F_B^F, \tag{8}$$

where $p_b^d = p^d(1 + F_B^P)$ is the token market price adjusted for the proportional buying transaction fee.

3.7. Regulator intervention

In principle, the regulator may choose to intervene in the token market within the day by controlling token market price, token allocation rate, and transaction fees to manage non-recurrent events. Thus, although these parameters are assumed to be constant within the day throughout most of the paper, in Section 6.6, we will briefly consider the case where the regulator may intervene in the market to adjust some or all of these parameters within-day in the presence of unusual events. For example, if road capacity drops because of an accident, or if demand increases due to a concert, the regulator can intervene, increasing token price in a certain period to discourage travel and reduce congestion. Numerical experiments are conducted to study this in Section 6.6.

In summary, the various market elements discussed in this section are described in Table 1.

4. Market behavior

As buying behavior is governed by the previously specified buying rule, this section primarily discusses individual selling behavior. It is assumed that the individual selling decision and mobility decision (departure time and mode) are inter-dependent. In other words, selling decisions are made conditional on a departure time/mode chosen at the beginning of the day, which in turn is based on a forecast of the account balance over the entire day. This forecast is based on historical travel and selling decisions of the user and his/her past experience (described in more detail in Section 5.1.3). One could think of the selling behavior as a strategy or an automated operation performed through (or programmed into) for example, a smartphone application, since in practice, it may be onerous to expect users to constantly make these selling decisions 'manually'. In this respect, one may also view it as an element in the overall design of the tradable credit scheme. Nevertheless, our modeling framework does not preclude the use of an actual behavioral model of selling in the market (in place of the selling strategy we formulate next), which would require the collection of empirical data; we defer this to future research.

From the perspective of simply maximizing profit (which is a reasonable selling strategy), the decision to sell can be formulated as a dynamic programming or optimal control problem, where the optimal selling strategy is characterized by Bellman's equation (Kirk, 2004). However, this is complicated, both from the standpoint of computational complexity and system design, and instead, we derive a simpler heuristic approach to characterize an individual's selling strategy.

At time t on day d, assume traveler n has an upcoming planned trip at a time denoted by $\tilde{t}_n$, where $\tilde{t}_n = \tilde{t}_n^d$ if $t \le \tilde{t}_n^d$, and $\tilde{t}_n = \tilde{t}_n^{(d+1)}$, if $t > \tilde{t}_n^d$. Given the next trip, a conditional profit function $\Pi_n^d(t)$, which represents the profit obtained by selling all tokens at time t (with no further selling until the next departure $\tilde{t}_n$) can be written as follows,

$$\begin{aligned}\Pi_n^d(t) &= S\left(x_n^d(t)\right) - \mathbb{I}\left(T(\tilde{t}_n) > \hat{x}_n(\tilde{t}_n)\right) \cdot B\left(T(\tilde{t}_n) - \hat{x}_n(\tilde{t}_n)\right) \\ &= x_n^d(t) p_s^d - F_S^F - \mathbb{I}\left(T(\tilde{t}_n) > \hat{x}_n(\tilde{t}_n)\right) \cdot \left(\left(T(\tilde{t}_n) - \hat{x}_n(\tilde{t}_n)\right) p_b^d + F_B^F\right),\end{aligned} \tag{9}$$

where $\hat{x}_n(\tilde{t}_n)$ represents the expected account balance at the time of the next trip $\tilde{t}_n$. Since it is assumed there will be no further selling until the next departure $\tilde{t}_n$, it can be written as,

$$\hat{x}_n(\tilde{t}_n) = \min\left[(\tilde{t}_n - t)r, Lr\right]. \tag{10}$$

Other notation used in the conditional profit function $\Pi_n^d(t)$ includes $T(\tilde{t}_n)$, which represents the toll in tokens of traveling at departure time $\tilde{t}_n$. A buying cost is incurred only if the toll at $\tilde{t}_n$ is greater than traveler n's expected account balance (i.e., $T(\tilde{t}_n) > \hat{x}_n(\tilde{t}_n)$), which is represented by the indicator function. Note that in defining the profit function above, we have made the critical assumption that if a decision to sell at the current time is made, no further selling will occur until the next trip. This simplification allows us to derive an optimal selling strategy analytically and is partly justifiable given that we also assume that during selling, an individual needs to sell all tokens in her wallet, and that prices do not vary within-day. However, observe that the selling strategy we derive, when applied, involves a decision made at every time point t, implying that it does not preclude the possibility of an individual making multiple selling decisions in the time period until the next trip if this is beneficial.

Under our assumptions, at time t on day d, traveler n will consider selling tokens only if the profit value is positive, i.e., $\Pi_n^d(t) > 0$. If the profit value is positive, she may still decide to wait if the derivative of the profit function is positive (meaning that the profit is expected to increase if she defers the decision to sell). Therefore, the selling strategy depends on both the profit function and its derivative, which can be analyzed in the following three cases:

1. $T(\tilde{t}_n) < \hat{x}_n(\tilde{t}_n)$ (no tokens need to be bought for the next trip)
   The profit function $\Pi_n^d(t)$ can be written as

$$\Pi_n^d(t) = x_n^d(t) p_s^d - F_S^F, \tag{11}$$

   and the derivative can be written as

$$\frac{d\Pi_n^d(t)}{dt} = \begin{cases} 0 & x_n^d(t) = Lr \\ rp_s^d & \text{otherwise,} \end{cases} \tag{12}$$

   which implies that profit will continue to increase until a full wallet is reached. It does not make sense to wait longer at a full wallet because newly acquired tokens simply replace expired tokens. Hence, selling should be at a full wallet.
   However, it is worth noting that, without fixed transaction fees, the selling revenue at full wallet is the same as that obtained from selling every time one receives new tokens. In fact, as long as one avoids token expiration, any selling strategy is equivalent in the absence of fixed transaction fees. It is fixed transaction fees that prevent frequent selling.
2. $T(\tilde{t}_n) > \hat{x}_n(\tilde{t}_n)$ (tokens need to be bought for the next trip)
   The profit function $\Pi_n^d(t)$ can be written as

$$\Pi_n^d(t) = x_n^d(t) p_s^d - F_S^F - \left(\left(T(\tilde{t}_n) - \hat{x}_n(\tilde{t}_n)\right) p_b^d + F_B^F\right), \tag{13}$$

   and its derivative can be written as

$$\frac{d\Pi_n^d(t)}{dt} = \begin{cases} -rp_b^d & x_n^d(t) = Lr \\ rp_s^d - rp_b^d & \text{otherwise,} \end{cases} \tag{14}$$

which is always negative since $p_s^d < p_b^d$ given that F_B^P or F_S^P is greater than 0. This implies that profit obtained from waiting and selling at any time in the future (until the next trip) is guaranteed to be less than the profit from selling now. Hence, she should sell now if the profit is positive.

Without transaction fees, the profit function $\Pi_n^d(t)$ can be written as

$$\Pi_n^d(t) = x_n^d(t)p^d - \left(T(\tilde{t}_n) - \hat{x}_n(\tilde{t}_n)\right)p^d, \tag{15}$$

and its derivative can be written as

$$\frac{d\Pi_n^d(t)}{dt} = \begin{cases} -rp^d & x_n^d(t) = Lr \\ 0 & \text{otherwise,} \end{cases} \tag{16}$$

which means that as long as account balance is not full, it does not matter whether one sells now or later. However, once we introduce fixed transaction fees, it is better to sell at a full wallet to minimize the number of transactions. With additional proportional transaction fees, it is better to sell immediately and not worth waiting anymore as the derivative is always negative.

3. $T(\tilde{t}_n) = \hat{x}_n(\tilde{t}_n)$ (the expected account balance is just enough to cover the toll of the next trip)
The profit function $\Pi_n^d(t)$ can be written as

$$\Pi_n^d(t) = x_n^d(t)p_s^d - F_S^F, \tag{17}$$

but its derivative does not exist because the conditional profit function is discontinuous at t due to the transaction fees of buying. To avoid any buying transaction fees (either fixed or proportional), it is optimal to sell immediately if profit $\Pi_n^d(t)$ is positive. Without transaction fees, similarly, it does not matter whether one sells now or later as long as token expiration is avoided.

Based on the analysis in this section, the effect of fixed transaction fees is to prevent multiple transactions while the effect of proportional transaction fees is to make one sell as soon as possible when the conditional profit is positive (if tokens need to be bought for the next trip). The proportional transaction fee is not preferable because it does not prevent frequent selling but instead prevents selling at a full wallet. Numerical experiments in Section 6 will provide further justification for the use of only a fixed transaction fee from an efficiency perspective.

The selling strategy for an individual n at any time t on day d considering positive transaction fees is summarized in Algorithm 1.

Algorithm 1: Selling Rule

input: $d, t, n, p^d, \tilde{t}_n, x_n^d(t), L, r, F_S^P, F_B^P, F_S^F, F_B^F$
At time t on day d, calculate $\Pi_n^d(t)$;
and expected account balance $\hat{x}_n(\tilde{t}_n) = \min\left[(\tilde{t}_n - t)r, Lr\right]$;
if $\Pi_n^d(t) > 0$ **then**
  if $T(\tilde{t}_n) \geq \hat{x}_n(\tilde{t}_n)$ **then**
    Sell now;
  else
    if $x(t) = Lr$ **then**
      Sell now;
    else
      Do nothing;
    end
  end
else
  Do nothing;
end

5. Simulation framework

This section describes the modeling and simulation framework for evaluating the performance of the designed instruments including a no-toll or no-control benchmark, referred to as NT, congestion pricing, referred to as CP, and the tradable mobility credit scheme, termed TMC. The overall simulation framework is shown in Fig. 2.

N travelers perform a daily commute between a single origin–destination pair. For the sake of simplicity, each traveler performs a single morning trip and a single evening trip. The morning commute trip is explicitly simulated whereas the evening trip is assumed to be a mirror of the morning trip. We consider a standard bi-modal transportation network (similar to Liu and Szeto (2020) and Liu et al. (2017); see Fig. 1) where travelers choose between driving and a public transit alternative. If they drive, they use a path containing a bottleneck of fixed capacity and choose their time of departure. Along the lines of Liu et al. (2017), we do not consider the departure time dimension for the transit alternative.

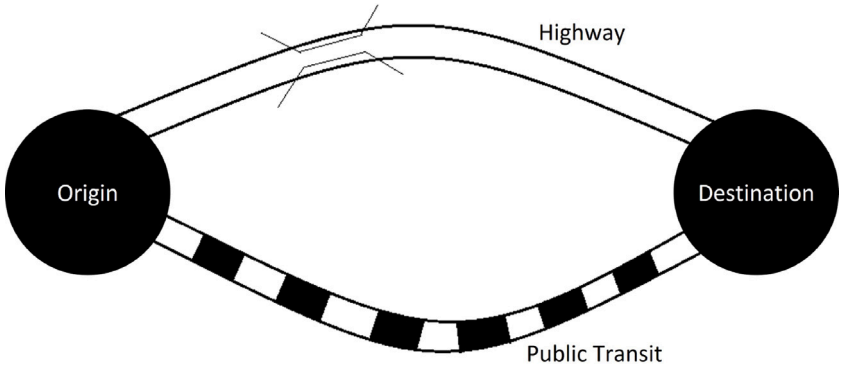


Fig. 1. Bi-modal transportation network.

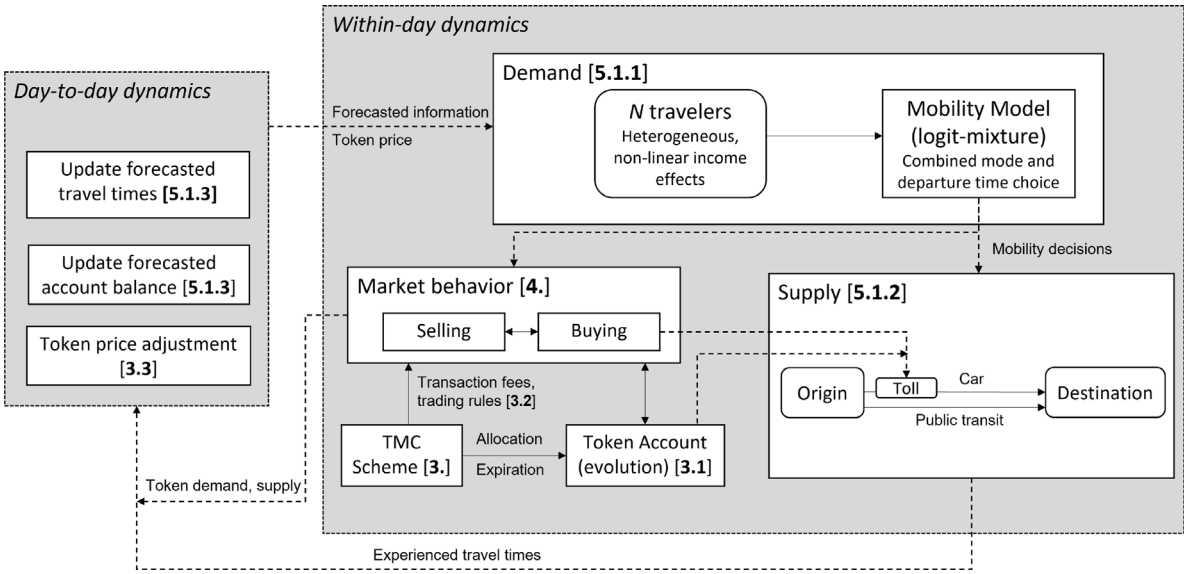


Fig. 2. Modeling and simulation framework.

At the beginning of each day, every traveler uses forecasted information of travel times, schedule delays and their account balance over the entire day to make a *pre-day mobility decision*, which is the combination of a choice of mode (between car and public transit, hereafter PT) and departure-time (over an individual specific choice set) for their morning commute trip. Travelers who choose to drive are subject to a time-of-day toll. For the TMC scheme, the time-of-day toll levied is in units of tokens. Note that mobility credits can only be used for the toll payment. The individual mobility decision is modeled using a logit-mixture model allowing for heterogeneity and non-linear income effects.

Next, traffic dynamics along with trading decisions – which occur over the entire day (i.e., are within-day) – are simulated on the bi-modal network connected by a single driving path and an alternative public transit (PT) line. Congestion (for driving) is modeled by a point queue model (bottleneck of finite capacity), in which a queue develops once flow exceeds capacity. Travel time of PT is assumed to be constant.

Day-to-day dynamics and travelers' learning are modeled through an exponential smoothing filter (Cantarella and Cascetta, 1995) that updates forecasts of travel time and account balance over the day. The day-to-day framework in Fig. 2 is used to simulate the evolution of the system state (mode and departure flows, travel times) until a measure of convergence has been reached. The performance measures (welfare, distribution of user benefits, congestion, and mode shares) at convergence are used to evaluate the different instruments. The model is a doubly dynamical system (Guo et al., 2018a), which considers the day to day evolution of a within-day dynamic system involving departure-time and mode choices. In our case, for the TMC scheme, the within-day dynamics has the added dimension of token selling behavior and the day-to-day dynamics has the added dimension of the token price adjustment. Doubly dynamical systems of the type we consider have been widely studied in the recent literature (Guo et al., 2018b,a; Liu and Geroliminis, 2017; Liu et al., 2017; Liu and Szeto, 2020) although they date back several decades (Ben-Akiva et al., 1984, 1986).

In the following sections, we first describe the models of demand, supply and day-to-day dynamics, termed the *system model*, in more detail. Next, we discuss social welfare computation and the simulation-based toll optimization problem (to determine optimal tolls) for the different instruments. Relevant notation is summarized in Table C.6 in Appendix C.

5.1. System model

As noted previously, the setting we consider involves N users traveling between a single origin–destination pair connected by a path containing a bottleneck of finite capacity and a PT line. Users wish to arrive at the destination within a certain "preferred arrival time window" in the morning, and can choose between PT and car. If they decide to drive, they can adjust their departure times to avoid congestion (similar to the model in Ben-Akiva et al. (1984), which is a dynamic extension of De Palma et al. (1983)).

The mobility demand model, network model, and day-to-day dynamics are discussed in detail next.

5.1.1. Demand model

The demand model (preday mobility decision) is a combined model of departure time and mode choice. Unless otherwise specified, the discussion in this subsection pertains to a specific day d (we omit d in all quantities for notational convenience). We reintroduce the superscript d in Section 5.1.3 when discussing day-to-day dynamics. The day is discretized into $h = 1 \dots H$ time intervals of size Δ_h (let the set of all time intervals in the day be denoted by $\mathcal{H} = \{1, \dots, h, \dots, H\}$), and it is assumed that each individual n has a preferred/desired arrival time $\hat{t}_n$ (more specifically, users are assumed to wish to arrive within a time window of size $2\Delta_a$ centered around $\hat{t}_n$; this is discussed in more detail later). The day is also discretized into smaller time intervals of size $\bar{h} = 1 \dots \bar{H}$ of size $\Delta_{\bar{h}}$, which is the resolution of the supply model and trading (selling) decisions.

The choice set of mode for individual n is defined as $M_n = \{C, PT\}$, where C represents car and PT represents transit. The choice set of feasible departure time intervals by car $\mathcal{H}_n \subset \mathcal{H}$ is individual-specific and defined as $\mathcal{H}_n = \{\tilde{t}_{0n} - \eta\Delta_h, \tilde{t}_{0n} - (\eta-1)\Delta_h, \dots, \tilde{t}_{0n} + \eta\Delta_h\}$, where η is a parameter, and $\tilde{t}_{0n}$ represents the initial departure time interval on day 0, which is computed based on the preferred arrival time $\hat{t}_n$ and the free flow travel time. Thus, the departure time choice set $\mathcal{H}_n$ consists of 2η time intervals of size Δ_h centered around the preferred departure time interval on day 0, $\tilde{t}_{0n}$. Note that because we model income effects, the individual departure time choice set is also subject to a budget constraint (i.e., an individual cannot choose a departure time that is not affordable). Thus, we define the set of feasible departure time intervals under instrument j ($j = NT, P, M$ for the No Toll scenario, congestion pricing, and the tradable mobility credit scheme respectively), as $\mathcal{H}_n^j \subseteq \mathcal{H}_n$. Under the No Toll scenario, $\mathcal{H}_n^{NT} = \mathcal{H}_n$. We will revisit the budget constraint later when discussing income effects. For the transit alternative, we consider only one departure time interval (since travel times and headways are constant) that will result in an arrival time closest to the preferred arrival time. This is denoted by h_n^{PT} and formally defined later. Let $i = \{m, h\} \in \mathcal{I}_n$ represent an individual's mobility decision as a combination of mode and departure time choice (where $\mathcal{I}_n = \{C, h | h \in H_n^j\} \bigcup \{PT, h_n^{PT}\}$).

Each individual is assumed to be rational and wishes to maximize her money-metric utility from the choice situation. The utility of the mobility decision i for individual n is denoted by U_{in}, which consists of two parts: a systematic utility V_{in} which is a function of observable variables and a random utility component ϵ_{in} that represents the analyst's imperfect knowledge. ϵ_{in} is assumed to follow an i.i.d. extreme value distribution with zero mean and individual specific scale parameter μ_n. When considering day-to-day dynamics, it is assumed that the individual random error component is perfectly correlated across days and across instruments (i.e., remains the same before the 'change' and after the 'change') assuming before and after periods are not too far apart (e.g., McFadden (2001) and de Palma and Kilani (2005)). This assumption can be relaxed in future work (see for example, Delle Site and Salucci (2013) and Zhao et al. (2008)).

The systematic money-metric utility for individual n departing in time interval h by car under instrument j is denoted by $V_{in}(\tilde{\boldsymbol{\phi}}_i^{j,n})$, where $i \in \{m = C, h | h \in H_n^j\}$. $\tilde{\boldsymbol{\phi}}_i^{j,n}$ is a vector of forecasted information in the systematic utility that affects the choice of departure time interval for driving and consists of five components. The first is forecasted/expected travel time $\tilde{\tau}_i^j$, which determines the expected schedule delay early (second component) and schedule delay late (third component). The fourth component is expected cost $\tilde{c}_{in}^j$, which is explained in more detail next. The last component is remaining income, which is equal to the disposable income for transportation I_n minus expected cost $\tilde{c}_{in}^j$.

The marginal utility of an additional unit of travel time for individual n is denoted by α_n. For simplicity, we assume travelers have common knowledge of forecasted travel times (more on this in Section 5.1.3). The desired arrival time window for individual n is defined as $[\hat{t}_n - \Delta_a, \hat{t}_n + \Delta_a]$, where $\hat{t}_n$ represents the center of the period and Δ_a represents arrival flexibility. If she arrives outside of the desired time period, she incurs a schedule delay. The marginal utility of an additional unit of schedule delay early is β_{En} and an additional unit of schedule delay late is β_{Ln}, where $\beta_{En} \leq \alpha_n \leq \beta_{Ln}$ from empirical evidence (e.g., Small (1982)).

The expected cost $\tilde{c}_{in}^j$ warrants additional discussion. Under the No Toll (NT) scenario, it is equal to the operational cost c_f (fuel cost). Under pricing ($j = P$), it is equal to the toll in dollars charged for departing in time interval h, $T^P(h)$, plus the operational cost c_f, which can be written as

$$\tilde{c}_{in}^P = T^P(h) + c_f. \tag{18}$$

Under the TMC ($j = M$) scheme, it depends on an individual's expected opportunity cost of tokens $\tilde{R}_{in}$ (which can be negative if one has a net revenue from selling tokens) plus the operation cost c_f as follows:

$$\tilde{c}_{in}^M = \tilde{R}_{in} + c_f. \tag{19}$$

Recall that the selling revenue of y tokens with transaction fees F_S^F, F_S^P and token price p, can be written as (selling revenue function),

$$S(y) = yp\left(1 - F_S^P\right) - F_S^F, \tag{20}$$

and similarly, the buying cost of y tokens can be written as (buying cost function),

$$B(y) = yp\left(1 + F_B^P\right) + F_B^F. \tag{21}$$

Let t_h be the start time of interval h, $\tilde{x}_n(t_h)$ be the expected account balance at time t_h, the beginning time of the time interval h. If a traveler does not need to pay any toll, she can sell the entire day's token allocation completely. Hence, the opportunity cost (or negative opportunity benefit) is equal to the negative of selling revenue of the entire day's allocation, $-S(Lr)$.

If a traveler needs to pay a toll $T^M(h)$ in h, and the expected account balance $\tilde{x}_n(t_h)$ is greater or equal to $T^M(h)$ (no buying), her opportunity cost is equal to the negative of selling revenue of the one-day allocation Lr minus the toll in tokens $T^M(h)$, which can be written as

$$\tilde{R}_{in} = -S\left(Lr - T^M(h)\right). \tag{22}$$

However, if she does not have enough account balance to cover the toll $T^M(h)$, she has to buy additional tokens equal to $T^M(h) - \tilde{x}_n(t_h)$ in order to travel in h. The amount of tokens she can sell for profit is equal to the one-day allocation Lr minus her expected account balance $\tilde{x}_n(t_h)$ since all of her tokens will be used for toll payment if she departs in h. The opportunity cost can be written as

$$\tilde{R}_{in} = -S\left(Lr - \tilde{x}_n(t_h)\right) + B\left(T^M(h) - \tilde{x}_n(t_h)\right). \tag{23}$$

In summary, the expected opportunity cost $\tilde{R}_{in}$ of departing by car in interval h depends on an individual's forecasted account balance $\tilde{x}_n(t_h)$, market price p, the toll in tokens $T^M(h)$ and transaction fees as follows:

$$\tilde{R}_{in} = \begin{cases} -S\left(Lr - T^M(h)\right) & \tilde{x}_n(t_h) \geq \mathrm{T}^M(h) \\ -S\left(Lr - \tilde{x}_n(t_h)\right) + B\left(T^M(h) - \tilde{x}_n(t_h)\right) & \text{otherwise.} \end{cases} \tag{24}$$

Note that if transaction fees are zero, the opportunity cost in Eq. (24) reduces to the one-day allocation minus the toll in tokens times token price, i.e., $\tilde{R}_{in} = -\left(Lr - \mathrm{T}^M(h)\right)p$. In the absence of non-linear income effects, Lrp can be ignored because it is a constant (appearing in all alternatives) that does not affect the choice, and the expression reduces to $\mathrm{T}^M(h)p$, which is intuitive.

Regarding the income effect, the diminishing marginal utility of income suggests that as an individual's income increases, the extra benefit to that individual decreases. It is thus natural to model this nonlinear effect of remaining income by a quasiconcave function (as per McFadden, 2017). Hence, we add the remaining income plus a natural log of the remaining income to the systematic money-metric utility.

The utility of an individual n driving and departing in time interval h (choosing a mobility decision $i \in \{C, h | h \in H_n^j\}$) under instrument j can thus be written as,

$$\begin{aligned} U_{in}\left(\tilde{\boldsymbol{\phi}}_i^{j,n}\right) =& V_{in}\left(\tilde{\boldsymbol{\phi}}_i^{j,n}\right) + \epsilon_{in} \\ =& -2\alpha_n \tilde{\tau}_i^j - \beta_{En} SDE\left(h, \hat{t}_n, \tilde{\tau}_i^j\right) - \beta_{Ln} SDL\left(h, \hat{t}_n, \tilde{\tau}_i^j\right) \\ &+ I_n - 2\tilde{c}_{in}^j + \lambda ln\left(\gamma + I_n - 2\tilde{c}_{in}^j\right) + \epsilon_{in}, \end{aligned} \tag{25}$$

where

$$SDE\left(h, \hat{t}_n, \tilde{\tau}_i^j\right) = \max\left(0, \hat{t}_n - \Delta_a - (t_h + \tilde{\tau}_i^j)\right), \tag{26}$$

$$SDL\left(h, \hat{t}_n, \tilde{\tau}_i^j\right) = \max\left(0, (t_h + \tilde{\tau}_i^j) - \hat{t}_n - \Delta_a\right). \tag{27}$$

Schedule delay of the evening trip is ignored because it is assumed to be more flexible.

The systematic money-metric utility function of user n who departs in time interval h_n^{PT} by PT is denoted as $V_{in}(\tilde{\boldsymbol{\phi}}_i^{j,n})$, where $i = \{PT, h_n^{PT}\}$. Since the travel time and headway of PT are constant, we only need to consider one departure time interval $h_n^{PT} = \lfloor \hat{t}_n - \tau_{PT} \rfloor$, which has a corresponding arrival time closest to the desired arrival time $\hat{t}_n$. For PT, the vector $\tilde{\boldsymbol{\phi}}_i^{j,n}$ for the systematic utility consists of four components: PT travel time τ_{PT}, expected waiting time W_{PT}, expected PT cost $\tilde{c}_{in}^j$ and remaining income $I_n - \tilde{c}_{in}^j$.

The marginal utility of an additional unit of PT travel time of individual n is assumed to be the same as that of car travel time, α_n. The marginal utility of an additional unit of waiting time is β_{Wn}.

The expected PT cost $\tilde{c}_{in}^j$ is equal to the PT fare c_{PT} under the No Toll (NT) scenario and pricing. Under the TMC scheme, it depends on an individual's expected opportunity cost of tokens $\tilde{R}_{in}$ and the PT fare c_{PT}, where $\tilde{R}_{in}$ is equal to the negative of selling revenue of a full wallet Lr since travelers who choose PT can sell all of their tokens acquired in one day for maximum return. It can be written as $\tilde{R}_{in} = -S(Lr)$.

Hence, the expected PT cost $\tilde{c}_{in}^j$ under the TMC scheme can be written as

$$\tilde{c}_{in}^M = \tilde{R}_{in} + c_{PT}. \tag{28}$$

The utility of an individual n using PT who departs in interval h^{PT} (choosing a mobility decision $i = \{PT, h^{PT}\}$) can be thus written as,

$$\begin{aligned} U_{in}\left(\tilde{\boldsymbol{\phi}}_i^{\boldsymbol{j,n}}\right) &= V_{in}\left(\tilde{\boldsymbol{\phi}}_i^{\boldsymbol{j,n}}\right) + \epsilon_{in} \\ &= -2\alpha_n \tau_{PT} - 2\beta_{Wn} W_{PT} + I_n - 2\tilde{c}_{in}^{j} + \lambda ln\left(\gamma + I_n - 2\tilde{c}_{in}^{j}\right) + \epsilon_{in}. \end{aligned} \tag{29}$$

Thus, the probability of individual n choosing alternative i under tolling instrument j is given by:

$$\omega_i^{j,n}\left(\tilde{\boldsymbol{\phi}}^{\boldsymbol{j,n}}\right) = \frac{exp\left(\mu_n V_{in}\left(\tilde{\boldsymbol{\phi}}_i^{j,n}\right)\right)}{\sum_{k\in\mathcal{I}_n} exp\left(\mu_n V_{kn}\left(\tilde{\boldsymbol{\phi}}_k^{j,n}\right)\right)}, \tag{30}$$

where $i \in \mathcal{I}_n$ and $\tilde{\boldsymbol{\phi}}^{\boldsymbol{j,n}} = (\tilde{\boldsymbol{\phi}}_i^{\boldsymbol{j,n}}, \forall i \in \mathcal{I}_n)$.

The vector of departure flows by car is denoted by $\boldsymbol{f}^{\boldsymbol{j}} = (f_h^j, h \in \mathcal{H})$ where,

$$\mathbb{E}[f_h^j] = \sum_{n:\, h\in\mathcal{H}_n} \omega_i^{j,n}\left(\tilde{\boldsymbol{\phi}}^{\boldsymbol{j,n}}\right), \tag{31}$$

and $i = \{C, h\}$.

5.1.2. Supply model

The network is assumed to be a single origin–destination pair connected by a single path containing a bottleneck of fixed capacity s (Arnott et al., 1990). A first-in-first-out (FIFO) queue develops once the flow of travelers exceeds s. The free flow travel time is t_f and the extra delay for a traveler departing from home at time t is $t_v(t)$. Thus, the total travel time for a traveler departing from home at time t is:

$$\tau(t) = t_v(t) + t_f. \tag{32}$$

Let $Q(t)$ be the number of travelers in the queue at time t. The delay at time t is derived from the deterministic queuing model as follows:

$$t_v(t) = \frac{Q(t)}{s}, \tag{33}$$

where $Q(t) = 0$ and $t_v(t) = 0$ when there is no congestion.

Note that within the simulation framework for the supply model, time t is discretized into time intervals $\bar{h} = 1...\bar{H}$ of size $\Delta_{\bar{h}} (< \Delta_h)$. The exact time of departure of a traveler within the supply model is randomly (uniformly) drawn within the chosen departure time interval h. The travel time for a given departure time interval h is obtained by averaging the travel times of all travelers departing in h.

Thus, if $\bar{\boldsymbol{f}} = (\bar{f}_{\bar{h}}, \bar{h} = 1...\bar{H})$ denotes the departure flows by car in the time intervals $\bar{h} = 1...\bar{H}$ (directly obtained from $\boldsymbol{f}$ defined in Eq. (31)), the queue length in time interval $\bar{h}$ can be written explicitly as a function of $\bar{\boldsymbol{f}}$ as follows (see Ramadurai et al. (2010)):

$$Q_{\bar{h}}(\bar{\boldsymbol{f}}) = \begin{cases} \left[\bar{f}_{\bar{h}} - s\Delta_{\bar{h}}\right]_+ & \bar{h} = 1 \\ \left[Q_{\bar{h}-1}(\bar{\boldsymbol{f}}) + \bar{f}_{\bar{h}} - s\Delta_{\bar{h}}\right]_+ & \bar{h} = 2....\bar{H}, \end{cases} \tag{34}$$

where $[.]_+ = \max(., 0)$. Thus, the experienced travel time by car in interval $\bar{h}$ is given by,

$$\tau_{\bar{h}}(\bar{\boldsymbol{f}}) = \frac{Q_{\bar{h}}(\bar{\boldsymbol{f}})}{s} + t_f. \tag{35}$$

With a slight abuse of notation, we will later use τ to also denote travel times aggregated at the level of the departure time interval.

The alternative PT line has a constant travel time τ_{PT}. Its headway is also constant, which is equal to twice the expected waiting time W_{PT}.

5.1.3. Day-to-day dynamics

Let $\tau_i^{d,j}$ denote the actual or experienced car travel time on day d for choice i under instrument j, where $i \in \{C, h | h \in \mathcal{H}\}$. As per the demand model, travelers are assumed to make their choices of departure time according to forecasted car travel times $\tilde{\tau}_i^{d,j}$, $i \in \{C, h | h \in \mathcal{H}\}$ from their memory and learning. We use an exponential smoothing filter, a type of homogeneous filter (Cantarella and Cascetta, 1995), to model the learning and forecasting process by weighting actual and forecasted costs of the previous day as follows:

$$\tilde{\tau}_i^{d,j} = (1-\theta_\tau)\tilde{\tau}_i^{d-1,j} + \theta_\tau \tau_i^{d-1,j}(\boldsymbol{f}^{\boldsymbol{d-1,j}}), \forall i \in \{C, h | h \in \mathcal{H}\}, \tag{36}$$

where $\theta_\tau \in [0, 1]$ is a learning weight given to the previous day's realized travel time, $\boldsymbol{f}^{\boldsymbol{d-1,j}} = (f_h^{d-1,j}, \forall h \in \mathcal{H})$ denotes the vector of departure flows by car on day $d-1$ under instrument j, and the dependence of $\tau_i^{d-1,j}$ on $\boldsymbol{f}^{\boldsymbol{d-1,j}}$ is described in Section 5.1.2.

An additional component of day-to-day dynamics in the case of the TMC scheme is the daily adjustment of token prices as per Eq. (5), where the regulator revenue on day $d-1$, K^{d-1} is a function of the dis-aggregate mobility decisions and buying/selling decisions, and is given by:

$$K^{d-1} = \sum_{n=1}^{N} \left(\sum_{\bar{h} \in \{1 \ldots \bar{H}\}} \left(B\left(T^M\left(\bar{h}\right) - x_n^{d-1}(\bar{h})\right) \mathbb{I}_n^{B,d-1}\left(\bar{h} | \boldsymbol{T}^M\right) - S\left(x_n^{d-1}(\bar{h})\right) \mathbb{I}_n^{S,d-1}(\bar{h} | \boldsymbol{T}^M)\right) \right), \tag{37}$$

where $B(\cdot)$ is cost of buying function and $S(\cdot)$ is revenue of selling function; $\mathbb{I}_n^{B,d-1}\left(\bar{h} | \boldsymbol{T}^M\right)$ and $\mathbb{I}_n^{S,d-1}\left(\bar{h} | \boldsymbol{T}^M\right)$ are indicators of buying or selling at $\bar{h}$ on day $d-1$ (for individual n) given the toll in tokens $\boldsymbol{T}^M$.

Further, under the TMC scheme, recall that the individual forecasted account balance on day d for individual n, denoted by $\tilde{x}_n^d(\bar{h})$, is required to compute the expected cost in Eq. (24). To compute $\tilde{x}_n^d(\bar{h}), \forall \bar{h}$, an individual forecasted departure time $\tilde{t}_n^d$ is first computed by once again, applying an exponential smoothing filter as follows:

$$\tilde{t}_n^d = (1-\theta_t)\tilde{t}_n^{d-1} + \theta_t t_n^{d-1}(\tilde{\boldsymbol{\phi}}^{\boldsymbol{d-1,n,j}}(\tilde{\boldsymbol{\tau}}^{\boldsymbol{d-1,j}})), \tag{38}$$

where $j = M$ (TMC system), $\theta_{\tilde{t}} \in [0,1]$, t_n^{d-1} is the chosen departure time of individual n on day $d-1$, and $\tilde{\boldsymbol{\tau}}^{\boldsymbol{d-1,j}} = (\tilde{\tau}_h^{d-1,j}, \forall h \in \mathcal{H})$. Next, the selling model presented in Section 3.6 is applied using the individual forecasted departure time to determine their forecasted account balance over the entire day, which is used to compute the expected toll costs under the TMC scheme through Eq. (24).

For all instruments j, from Eq. (31), we have,

$$\mathbb{E}[f_h^{d,j} | \boldsymbol{f}^{\boldsymbol{d-1,j}}] = \sum_{n: h \in \mathcal{H}_n} \omega_i^{n,j} \left(\tilde{\boldsymbol{\phi}}^{\boldsymbol{d,j,n}}(\tilde{\boldsymbol{\tau}}^{\boldsymbol{d,j}}) \right), \forall h \in \mathcal{H}, \tag{39}$$

where $i = \{C, h\}$. Eqs. (36) and (39) (along with Eqs. (38) and (5) for the TMC scheme) completely describe the day-to-day dynamical system.

If the day-to-day stochasticity in flows is ignored, i.e., $f_h^{d,j} \equiv \mathbb{E}[f_h^{d,j} | \boldsymbol{f}^{\boldsymbol{d-1,j}}]$, the Eqs. (36) and (39) (Eqs. (36), (38), (39), and (5) under the TMC scheme) constitute a deterministic process (Cantarella and Cascetta, 1995). For an analysis of the properties of the fixed-point of this deterministic day-to-day dynamical system, we refer the reader to Cantarella and Cascetta (1995). In our model, under the TMC scheme, the lack of a closed-form for the regulator revenue (due to the dis-aggregate selling model) complicates this analysis. Nevertheless, one approach for demonstrating existence of a fixed-point is to show that the travel time and schedule delay (and hence, utility) functions in our model are Lipschitz continuous with respect to the departure flows along the lines of Guo et al. (2018b). Although an interesting avenue for future research, this is beyond the scope of our paper. It should also be pointed out that Guo et al. (2018a) identify conditions under which the evolution of a day-to-day dynamic system with departure time choice (for the standard bottleneck model) will not converge to the fixed-point (equilibrium). Nevertheless, they observe that this result does not necessarily preclude the possibility of convergence occurring under preference heterogeneity in complex networks.

In contrast, if the day-to-day stochasticity in departure flows is explicitly modeled, as we do in our simulation framework, the Eqs. (36) and (39) (Eqs. (36), (38), (39), and (5) under the TMC scheme) constitute a stochastic process. For details on the conditions that ensure regularity of the stochastic process, refer Cantarella and Cascetta (1995). Given the complexity of the model, we numerically examine stationarity of the stochastic process model under different starting conditions in Section 6.2 and Appendix A.

5.2. Simulation-based optimization

The problem of determining the optimal toll in dollars for congestion pricing, $T^P(h), \forall h \in \mathcal{H}$ and the optimal toll in tokens for the TMC scheme, $T^M(h), \forall h \in \mathcal{H}$ can be formulated as a simulation-based optimization problem with the objective of maximizing total social welfare (SW). The social welfare of the CP and TMC instruments is calculated relative to the NT scenario and is equal to the sum of user benefits (UB) and regulator revenue (K). All quantities are calculated upon convergence of the day-to-day system, and hence, we drop the superscript d in this Section. We continue to use the notation K to denote total revenue, adding a superscript to denote the specific instrument. Under congestion pricing (P), the regulator revenue is given by,

$$K^P = \sum_{n=1}^{N} \sum_{i \in M_n \times H_n^P} \hat{c}_{in}^P \mathbb{I}_n\left(i | \boldsymbol{T}^P\right), \tag{40}$$

where i is the mobility decision, which is a combination of mode and departure time choice; $\mathbb{I}_n(i | \boldsymbol{T}^P)$ is an indicator if traveler n chooses mobility choice i given toll vector $\boldsymbol{T}^P = \{T^P(h) | h \in \mathcal{H}\}$; $\hat{c}_{in}^P$ is equal to the toll payment for driving ($T^P(h)$) or the PT fare payment for PT (c_{PT}); and H_n^P is the set of feasible departure time intervals taking into account budget constraints.

Under the TMC scheme, the regulator revenue K^M consists of two parts. The first part is the sum of PT fare payments and the second part is the revenue associated with the buying and selling transactions over one day, given by Eq. (37). Thus, K^M can be written as,

$$K^M = \sum_{n=1}^{N} \left(c_{PT} \mathbb{I}_n\left(PT | \boldsymbol{T}^M\right) + \sum_{\bar{h} \in \{1 \ldots \bar{H}\}} \left(B\left(T^M\left(\bar{h}\right) - x_n(\bar{h})\right) \mathbb{I}_n^B\left(\bar{h} | \boldsymbol{T}^M\right) - S\left(x_n(\bar{h})\right) \mathbb{I}_n^S(\bar{h} | \boldsymbol{T}^M)\right) \right). \tag{41}$$

Note that the price adjustment mechanism described in Section 3.5 is designed to ensure that the regulator revenue of the TMC scheme (second part in Eq. (41)) at equilibrium is close to zero, thereby achieving revenue neutrality.

The user benefits, which are a measure of consumer well being, can be computed in three ways in the presence of non-linear income effects (we refer the reader to McFadden (2017) for a detailed discussion): Market Compensating Equivalent (MCE), Hicksian Compensating Variation (HCV), and Hicksian Equivalent Variation (HEV). MCE is equal to the difference in indirect utilities between a "but for" scenario and "as is" scenario, scaled to money metric units by dividing by the marginal utility of income (MUI) at the "as is" scenario. It differs from the commonly used Marshallian consumer surplus (MCS) only in the MUI scaling factor. It can be easily computed when the indirect utility function and its derivatives are known. HCV is equal to the net decrease in the "but for" scenario income that equates utility in the two scenarios while HEV is equal to the net increase in the "as is" scenario income that equates utility in the two scenarios.

A potential drawback of these three measures is that their ethical implications are not defensible as pointed out by Blackorby and Donaldson (1990). Well-being measured in units of income treat increases in income as equally socially valuable no matter who receives them. This is not the case with net utility improvement, since the nonlinear effect of income improvement is captured by the income effect term in the utility specification (lower income users have a higher marginal utility of income). Hence, we measure user benefits (Z^j) under instrument j as the sum of all users' net experienced utilities relative to NT denoted as z_n^j (along the lines of De Palma and Lindsey (2004)). Since the utilities adopted in this study are money-metric, the net utility amount serves as a meaningful measurement of improvement directly. An individual n's net experienced utility is the difference between maximum utility under instrument j and under NT, which can be written as,

$$z_n^j = \max_{i \in M_n \times H_n^j} \left(U_{in}\left(\boldsymbol{\phi}_i^{j,n}\right)\right) - \max_{i \in M_n \times H_n^{NT}} \left(U_{in}\left(\boldsymbol{\phi}_i^{NT,n}\right)\right), \tag{42}$$

where $\boldsymbol{\phi}_i^{j,n}$ is a vector of experienced variables under instrument j and $\boldsymbol{\phi}_i^{NT,n}$ is a vector of experienced variables under NT. Hence, the user benefits Z^j can be written as

$$Z^j = \sum_{n=1}^{N} z_n^j. \tag{43}$$

As noted before, in the case of the CP scheme, we determine the toll in dollars which maximizes social welfare SW, computed at the equilibrium (or after convergence of the day-to-day model). For the TMC system, in addition to the toll in tokens, which is optimized, other parameters like the allocation rate r and transaction fees are set exogenously (explained in more detail in Section 6.3). This is formulated as,

$$\begin{aligned} \max_{\boldsymbol{T}^j} \quad & Z^j + K^j \\ \text{s.t.} \quad & Z^j, K^j = SM\left(\boldsymbol{T}^j, \boldsymbol{\xi}, \boldsymbol{\psi}\right) \\ & \boldsymbol{T}^j = \{T^j(h) | h \in \mathcal{H}\} \\ & \boldsymbol{T}^j \geq 0, \end{aligned} \tag{44}$$

where j can be either P or M; the toll profile $\boldsymbol{T}^j$ is a set of toll values over the entire day. $\boldsymbol{\xi}$ represents all input data for the simulation, such as individual income, preferred arrival time, and choice attributes. $\boldsymbol{\psi}$ represents all model parameters, such as the demand model coefficients, bottleneck capacity, user learning weights, and market parameters for the TMC scheme. The $SM(\cdot)$ function is the system model discussed in Section 5.1. The toll function that we consider is a step toll profile (of the kind implemented in Singapore and Stockholm), which consists of five step toll values and six break points.

Clearly, the optimization problem has no closed-form since the objective function for a given toll profile is the outcome of a simulation of the stochastic process (a simulation-based optimization problem), or more specifically, the system model presented in 5.1, which includes traveler behavior, regulator states and actions, and the resulting network and market conditions. In order to solve this simulation-based optimization problem, a differential evolution (DE) algorithm is adopted as it is derivative-free and performs well for global optimization problems of this kind (Storn and Price, 1997).

5.2.1. Differential evolution algorithm

In this section, we illustrate the application of the DE algorithm. Let $\mathbf{X}$ represent the decision variables of the simulation-based optimization problem (i.e., the parameters of the step toll profile). The DE algorithm essentially has three iterative operators, mutation, crossover, and selection, to iteratively improve candidate solutions.

The mutation operator uses individuals from the current solution population to generate variant vectors. The gth variable of vector i at generation k, $\mathbf{Y}_{g,i}^k$, is given by

$$\mathbf{Y}_{g,i}^k = \mathbf{X}_{g,r1}^k + F \cdot (\mathbf{X}_{g,r2}^k - \mathbf{X}_{g,r3}^k), \tag{45}$$

where $r1,\ r2,\ r3 \in [1, NP]$, $i \neq r1 \neq r2 \neq r3$, F is a scale factor and NP is the solution population size.

Next, the crossover operator creates a trial vector $\mathbf{U}_i^k$ by combining the variant vector and original vector as follows,

$$\mathbf{U}_{g,i}^k = \begin{cases} \mathbf{Y}_{g,i}^k & \text{if } rand(0,1) < CR \text{ or } g = rg \\ \mathbf{X}_{g,i}^k & \text{otherwise,} \end{cases} \tag{46}$$

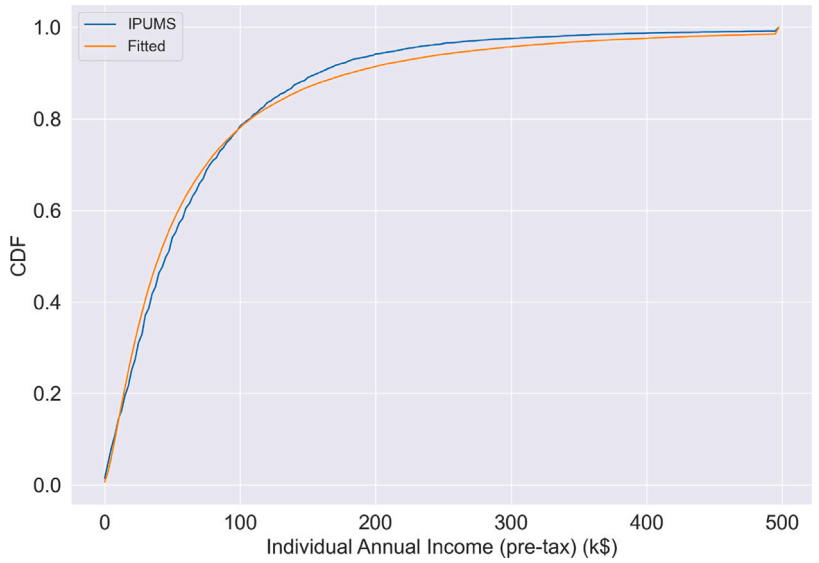


Fig. 3. Individual pre-tax annual income distribution.

where $CR \in [0, 1]$ is the crossover rate, $rand(0, 1)$ represents a random uniformly distributed variable within $(0, 1)$, and rg is a random integer in $[1, \|\mathbf{X}\|]$ ensuring at least one variable of the trial vector $\mathbf{U}_i^k$ is from the variant vector $\mathbf{Y}_i^k$.

Finally, the selection operator produces the next generation of vectors by comparing the original vector $\mathbf{X}_i^k$ and the trial vector $\mathbf{U}_i^k$ in terms of social welfare as follows,

$$\mathbf{X}_i^{k+1} = \begin{cases} \mathbf{U}_i^k & \text{if } SW(\mathbf{U}_i^k) > SW(\mathbf{X}_i^k) \\ \mathbf{X}_i^k & \text{otherwise.} \end{cases} \tag{47}$$

6. Numerical experiments

This section describes numerical experiments that examine the performance of the TMC scheme using the simulation framework described in Section 5. Data and parameters of the demand model, supply model, and the day-to-day learning model are introduced in Section 6.1. The parameters are calibrated based on empirical evidence to represent a realistic base case. Selected parameters will be varied in later experiments. The objectives of the experiments are to: (1) examine the overall performance of the TMC system under alternative market designs — more specifically, the effect of transactions fees in avoiding undesirable market behavior whilst retaining efficiency, (2) compare the efficiency and equity of the TMC scheme (market design based on experiments under (1) and CP under varying levels of congestion, heterogeneity, income effects, and (3) examine the robustness (adaptiveness) of the TMC scheme with a lump sum allocation versus a continuous allocation in the presence of unusual events.

6.1. Data and model parameters

The demand model requires both choice attributes and individual characteristics as input data. The individual characteristics include disposable income I_n and preferred arrival time $\hat{t}_n$. Recall that disposable income I_n in our context is defined as personal net income after taxes and after subtracting necessary living expenses (e.g., housing, health, food). The individual pre-tax annual income is assumed to follow a lognormal distribution and is fitted using the Integrated Public Use Microdata Series (IPUMS) 2019 census data (Ruggles et al., 2021). The cumulative distribution functions (CDF) of IPUMS data and the fitted data are shown in Fig. 3. Note that all annual incomes greater than 500 thousand dollars are grouped together. As we can see, the lognormal distribution fits the income distribution reasonably well.

Individual daily income is computed as the annual income divided by 260 working days per year and the individual hourly wage rate is computed by dividing daily income by 8 working hours per day. The minimum wage rate is set to $7.25 per hour as per the U.S. Department of Labor. It is less straightforward to obtain disposable income after taxes since necessary living expenses could vary significantly based on income, and disaggregate data on this is sparse. Based on average data from the Bureau of Labor Statistics (U.S. Bureau of Labor Statistics, 2019), we assume that each traveler's daily disposable income after taxes and necessary living expenses is equal to 60% of their pre-tax daily income.

The preferred departure time distribution (preferred arrival time $\hat{t}_n$ minus the free flow travel time) is based on a recent empirical study of road users in Stockholm (Kristoffersson and Engelson, 2018), and is shown in Fig. 4. For simplicity, the size of the preferred arrival window Δ_a is set to 0, which implies that individuals have a single preferred arrival time as in the standard Vickrey model.

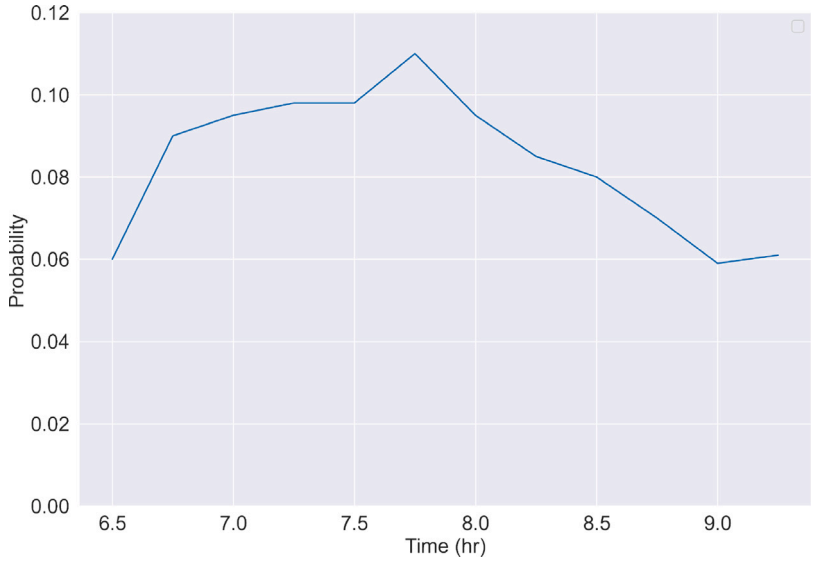


Fig. 4. Individual preferred departure times distribution.

The departure time window size parameter η is set to 30, which means the individual departure time choice set H_n ranges over a 60-minute interval.

Empirical evidence (e.g., Small et al. (2005)) indicates that there is significant heterogeneity in the value of time across drivers. We assume that the individual value of time α_n is perfectly correlated with income and is one third of the wage rate (White, 2016). Note that this assumption will be relaxed in some experimental scenarios to consider different levels of heterogeneity.

Values of schedule delay early β_{En} and late β_{Ln} are also likely to be distributed across individuals. Due to the lack of empirical data on this, the literature on bottleneck models incorporate heterogeneity by making assumptions on the ratios between values of schedule early/late to values of time. Proportional heterogeneity (first considered by Vickrey (1973)) assumes that values of time and schedule delays vary proportionally or in other words, the ratio of the parameters is identical for all individuals (Van den Berg and Verhoef, 2011). Ratio heterogeneity assumes that the values of schedule delays are fixed and only values of time are distributed (De Palma and Lindsey, 2002). As a result, ratios of parameters are distributions. Van Den Berg and Verhoef (2011) considers a more general heterogeneity, assuming the ratio of values of time to values of schedule delay early follows a symmetric triangular distribution from 1 to 3 based on intuition and the ratio of values of schedule delay late to values of schedule delay early is a constant 3.9 based on Arnott et al. (1990). Along similar lines, we assume that the ratio of values of schedule delay early β_{En} to values of time α_n follows a triangular distribution from 0.1 to 1 with a mode at 0.5. The ratio of values of schedule delay late β_{Ln} to α_n is assumed to follow a triangular distribution from 1 to 3 with a mode at 2 (the selection of the modes as 0.5 and 2 are based on Small (2012)). The bounds are set based on the empirical relationship $\beta_{En} \leq \alpha_n \leq \beta_{Ln}$ (Small, 2012).

Further, as pointed out by Small (2012), waiting times are onerous compared to in-vehicle times by multiples of two to three by most assessments. For simplicity, the ratio of values of time α_n to values of waiting time β_{Wn} is assumed to be a constant equal to 3. With regard to the income effect, γ is set to 2 and λ is calibrated to be 3 to have the highest marginal utility of income to be less than 1.34 (Layard et al., 2008).

The scale parameter μ_n is known to be confounded with the systematic utility and inversely related to error variance within the choice data (Ben-Akiva et al., 1985). As pointed out in the literature (e.g., Louviere and Eagle (2006)), the modeled heterogeneity can come from heterogeneity in individual coefficients and scale heterogeneity that is shared across coefficients. We assume the scale parameter follows a lognormal distribution. The mean value is calibrated based on price elasticity (discussed subsequently) and the coefficient of variation is set to 0.5 based on judgement.

Next, we calibrate the mean value of the scale parameter to ensure that the aggregate price elasticities of the mobility model (for departure time choice) are reasonable and accord with empirical evidence. The price elasticity of peak hour demand (7 AM–8 AM) is computed assuming there exists a flat toll profile during the period from 6:30–9:30 AM.

From the literature, the aggregate elasticities of peak hour travel vary greatly from case to case as they are dependent on the model structure, physical environment, activity type, initial toll levels and many other factors. Ding et al. (2015) estimated the elasticity of departing during the peak in Washington D.C. to be −0.0906 for driving alone. Sasic and Habib (2013) estimated that the elasticity of departing by car in the AM peak for work trips in Toronto is between −0.067 to −0.12. Holguin-Veras et al. (2005) found price elasticity of using crossings (tunnels and bridges) in NYC ranges from −0.31 to −1.97 for weekdays depending on the time-of-day. When there is no initial toll, the price elasticity simply represents fuel price elasticity. Lipow (2008) estimated fuel price elasticity as −0.17 and Gillingham (2014) estimated fuel price elasticity in California as −0.15.

Table 2
Price elasticities across income groups by toll levels.

Toll	≤25%	25% to 50%	50% to 75%	75% to 90%	90% ≤	Total
0	−0.34	−0.29	−0.12	0.00	0.00	−0.19
2.5	−1.14	−0.59	−0.10	−0.04	−0.03	−0.38
5	−1.57	−1.07	−0.20	−0.09	−0.06	−0.53

Table 3
Model and simulation parameters.

Variables	Description	Values
N	Population	7500
Δ_t	Duration of a simulation time step	1 min
Δ_h	Duration of a departure time interval	5 min
Δ_a	Size of desired arrival window	0 min
η	Departure time window size parameter	30
λ	Coefficient of nonlinear income effect	3
γ	Nonlinear income effect adjustment parameter	2
s	Bottleneck capacity (per min)	39
t_0	Free flow travel time	24 min
c_f	Operation cost of car	$3.13
τ_{PT}	PT travel time	43 min
W_{PT}	Expected waiting time	5 min
c_{PT}	Operation cost of PT	$2
$\theta_\tau/\theta_{\bar{t}}$	Learning weights	0.1

From calibration, the mean of scale parameter is determined to be 0.5. The corresponding price elasticities across different income groups and initial toll levels are presented in Table 2. As we can see, low income users are more sensitive to price than high income users, and when there is no toll, the aggregate price elasticity is similar to the empirical fuel elasticity. As the toll level increases, the aggregate price elasticity also increases and is similar to empirical values found in Holguin-Veras et al. (2005).

Regarding the supply model, the free flow speed of car is set to be 45 mph (Ali et al., 2007) and the one way driving distance is assumed to be 18 miles (free flow travel time of 24 min). The operational cost of driving is assumed to be composed of only fuel cost, which is equal to $3.13 (driving distance times 1/23 gallon per mile times 4 dollars per gallon). For public transit, based on the New York City MTA data, the fare is set to $ 2, average speed is 25 mph, and headway is 10 min. The PT distance is also assumed to be 18 miles, and the resulting PT travel time is 43 min since both headway and travel time of PT are constant. The expected waiting time is assumed to be 5 min.

The bottleneck capacity s is determined based on a calibration of the travel time index (TTI), which represents the ratio between actual travel time and free flow travel time. The base case capacity is determined to be 2340 vehicles per hour to have a reasonable level of congestion with a TTI of 1.68 (Chen, 2010) under the no toll (NT) scenario.

As described in Section 5.1.3, an exponential smoothing filter is adopted to update travel time information and individual departure time. The greater the learning weights are, the more unstable the system becomes (Cantarella and Cascetta, 1995). The learning weights θ_τ and $\theta_{\bar{t}}$ are assumed to be 0.1.

Recall that we focus on the morning commute and hence, we simulate half a day (12 h) with a simulation time interval Δ_t of 1-minute, yielding 720 time intervals, $\bar{t} = 0 \ldots 719$. The market elements (allocation, expiration, and price adjustment) and trading behavior are also simulated for the first half. The second half of a day is assumed to be a mirror of the first half. The departure time interval (Δ_h) is assumed to be 5 min and the population size N is 7500. Descriptions and values of key parameters are summarized in Table 3.

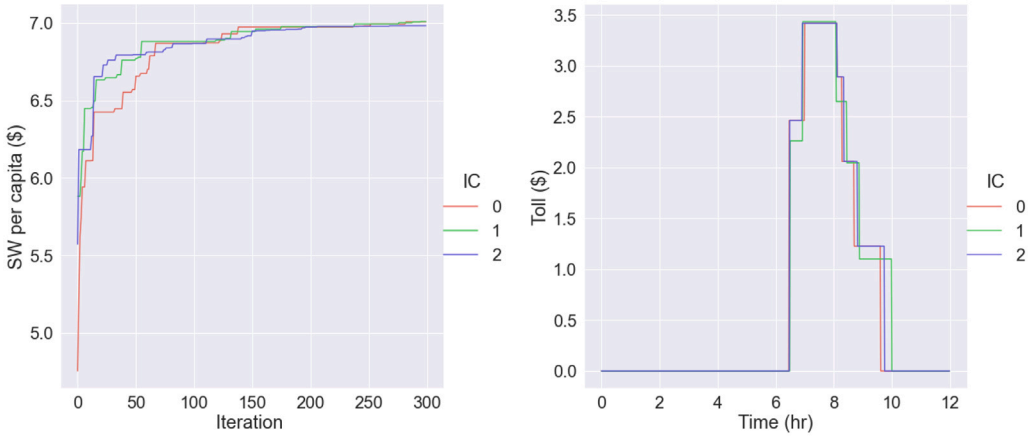
6.2. Convergence of the day-to-day dynamic model

As noted in Section 5.1.3, the day-to-day dynamic model described in Section 5.1.3 is a stochastic process. We numerically examine properties of the model via simulations with different initial conditions, which suggest that the day-to-day model converges to the same stationary distribution of departure flows. The tests on stationarity are summarized in Appendix A.

Note that the travel time of driving is used as a representative measure of the system state partly because it is central to the day-to-day learning process of travelers (alternatively, departure flows could also be used). The infinity norm (supremum norm) of day $d-1$'s travel time vector of driving and day d's travel time vector of driving is calculated and used as a measure of convergence across days. This is given by,

$$\|\boldsymbol{\tau}^{d-1} - \boldsymbol{\tau}^{d}\|_\infty = \sup\{|\tau_i^{d-1} - \tau_i^{d}| : i \in \{m = C, h | h \in \mathcal{H}\}\} \tag{48}$$

where $\boldsymbol{\tau}^{d-1}$ is the vector of travel times of day $d-1$ defined over the set of time intervals $\mathcal{H}$.



(a) Convergence of the social welfare (b) Optimal toll profile

Fig. 5. Convergence of social welfare and optimal time-of-day step toll profile of congestion pricing for different initial populations.

Table 4
Market parameters for tradable mobility credits (base case).

Variables	Description	Values
r	Token allocation rate	0.00285 per min
L	Token lifetime	720 min
F_B^F / F_S^F	Proportional transaction fee of buying/selling	$0
F_B^P / F_S^P	Fixed transaction fee of buying/selling	$0
p^0	Initial token price	$1
Δp	Price change	$0.05
K_t	Regulator revenue threshold	$300

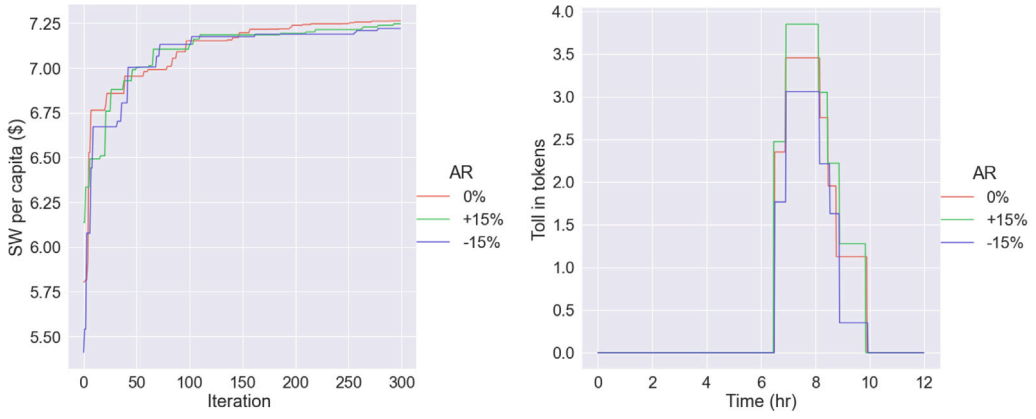
6.3. Optimization of tolls in TMC and congestion pricing

As discussed in Section 5.2.1, the time-of-day step toll profile is optimized using a type of metaheuristic algorithm known as Differential Evolution (DE). Metaheuristic algorithms have been shown to work well for nonconvex and nonlinear toll design problems (e.g., Shepherd and Sumalee (2004) and Zhang and Yang (2004)). The population size of the DE algorithm NP is set to 15. Using data and parameters discussed in Section 6.1, we first examine the performance of the optimization algorithm for the congestion pricing (CP) instrument using three different initial populations (with 15 candidates). Next, the performance of the optimized TMC instrument and convergence properties are examined.

Under congestion pricing, the convergence of social welfare along with the optimal time-of-day step toll profile with three different initial populations are shown in Fig. 5. The algorithm converges to the same optimal welfare (within a tolerance of $0.01) for three different initial populations. The optimal toll profiles are near identical although there are minor differences at the locations when the toll changes ('steps').

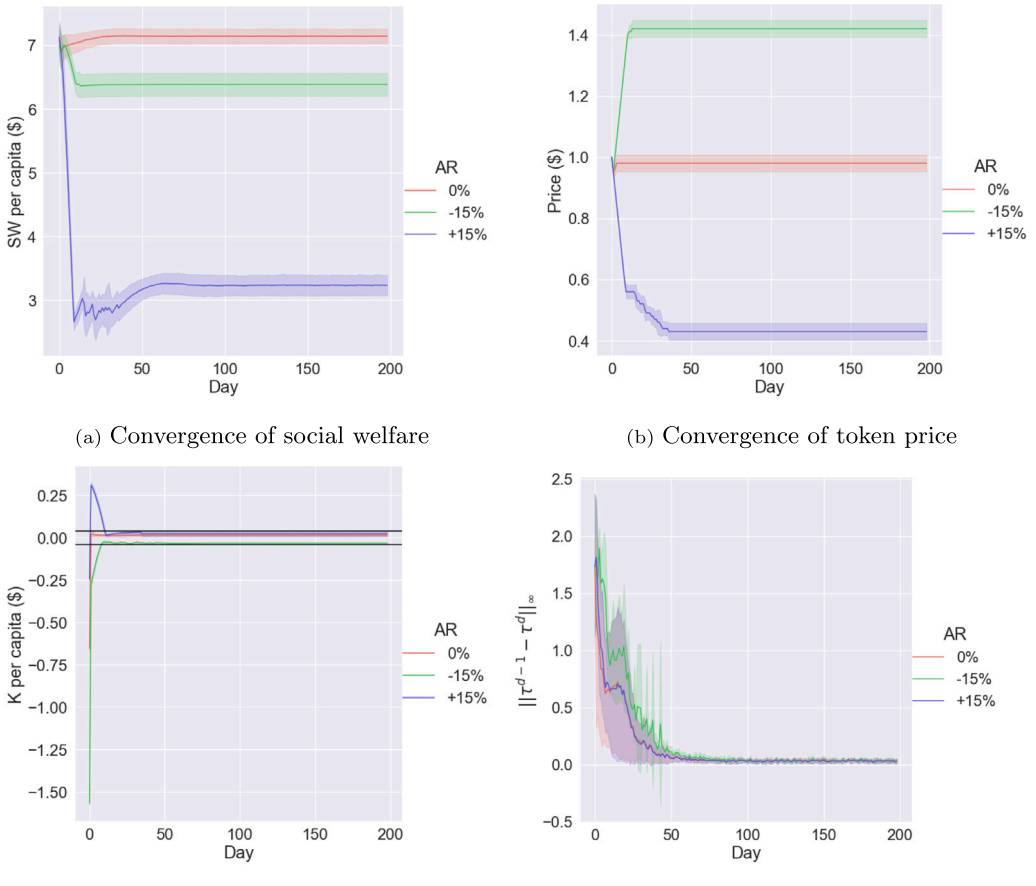
For the TMC instrument, we set the total daily allocation of tokens for each individual as the per capita regulator revenue from congestion pricing. This should in theory yield a 1 dollar equilibrium market price and a toll in tokens that is identical to the optimal toll in dollars from CP, assuming there are no transaction costs and income effects. Note that the daily allocation can be set arbitrarily. Next, we optimize the toll profile in tokens given the specified token allocation. We assume that tokens are distributed uniformly to everyone and are allocated in continuous time. Every traveler is assumed to have a random account balance at the beginning of the simulation (to reflect different times of entry into the system; more on this in Section 6.4). Recall also that we assume the evening period is a mirror of the morning period, and hence, we only simulate half a day, and assume the lifetime L of tokens is 720 min. For simplicity, in these experiments, transaction fees are set to zero. Other parameters of the TMC are summarized in Table 4.

The convergence of the optimized objective values (social welfare) along with the optimal time-of-day step toll profile in tokens with three different allocation rates ranging from 15% less than the baseline to 15% more than the baseline are shown in Fig. 6. As we can see, the different allocation rates converge to the same social welfare (within a tolerance of $0.02) at the end of 300 iterations. The token price of the baseline allocation rate is close to $1 (as expected) while the lower allocation rate has a higher token price of $1.1 and the higher allocation rate has a lower token price equal to $0.9. This is consistent with our expectation that the lower allocation rate leads to a higher token price due to less supply and vice versa. The optimal toll profiles in tokens in Fig. 6(b) show that the higher allocation rate leads to the overall higher tolls in tokens and vice versa, again, as expected. Interestingly, due to the non-linear income effects, the social welfare of the TMC scheme is slightly greater than that of congestion pricing (we return to this later).



(a) Convergence of the social welfare (b) Optimal toll profile in tokens

Fig. 6. Convergence of social welfare, and optimal time-of-day step toll profile in tokens and in dollars of TMC for three allocation rates.

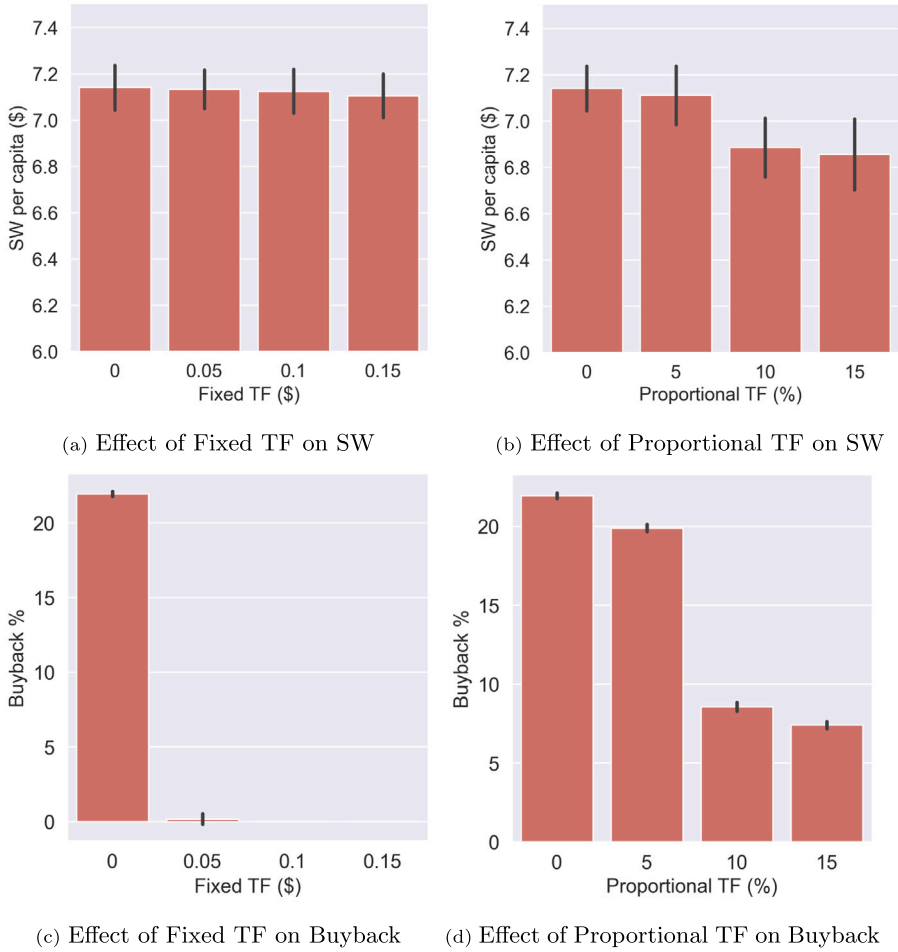


(a) Convergence of social welfare (b) Convergence of token price

(c) Convergence of regulator revenue (d) Convergence of travel time of driving

Fig. 7. Convergence of social welfare, token price, regulator revenue, and travel time of driving for different allocation rates r.

Given that disaggregate market behavior is modeled, it is also worthwhile to examine stability and convergence of the market prices. We examine this under different allocation rates r where the toll in tokens is not varied and is set to the optimal profile under the baseline allocation rate. The convergence of social welfare, market price, regulator revenue, and travel time of driving are shown in Fig. 7. As we can see, since tolls are not re-optimized, different allocation rates lead to different social welfare and price values. The results also indicate stability of the credit market, and that the variation of equilibrium price and welfare with allocation



(a) Effect of Fixed TF on SW (b) Effect of Proportional TF on SW

(c) Effect of Fixed TF on Buyback (d) Effect of Proportional TF on Buyback

Fig. 8. Effect of fixed and proportional transaction fees on social welfare and buyback behavior.

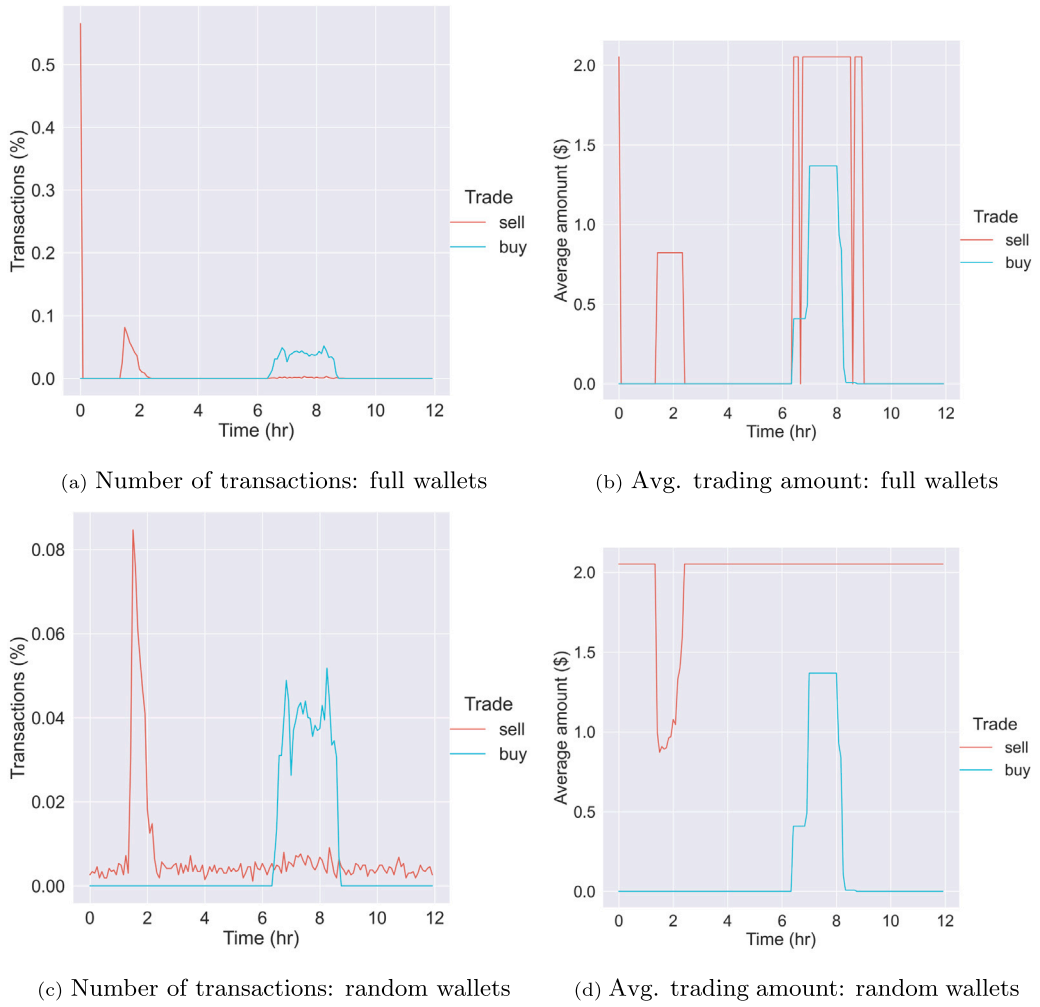
rate is intuitive. The effect of various initial market prices on convergence of token price and social welfare is also examined in Appendix B.

6.4. TMC design and market behavior

In the literature, a transaction fee has been used to prevent undesirable market behavior like frequent selling. For example, Brands et al. (2020) apply a small transaction fee of 0.01 euro to prevent frequent selling in their experiment. However, it has also been shown that transaction fees could reduce system efficiency (Nie, 2012). Our analysis in Section 4 shows that a fixed transaction fee has the effect of preventing multiple transactions while the proportional transaction fees has the effect of making one sell as soon as possible when the conditional profit is positive (if buying is required at the time of the next trip). As an alternative to transaction fees, the regulator may also impose a minimum threshold (in dollar amounts) below which transactions are not permitted.

Numerical experiments in this section examine the effect of proportional and fixed transaction fees on social welfare and undesirable behavior. Specifically, undesirable behavior is defined as buying back tokens sold previously. We deem this undesirable because we would like to have users strictly being either sellers or buyers (not both). Sellers are the ones who travel in the off peak (or by transit) and sell their tokens while buyers are the those with a high willingness to pay to travel by car during the peak period.

For simplicity, the fixed transaction fees of buying and selling are varied together with the proportional transaction fees set to zero and vice versa. The effects of fixed and proportional transaction fees on social welfare and buyback behavior are shown in Fig. 8. From the simulation experiments, a small fixed transaction fee (5 cents in this study) is seen to be able to eliminate buyback behavior in Fig. 8(c) and reduce welfare only slightly in Fig. 8(a). On the other hand, there are higher social welfare losses in the case of the proportional transaction fees in Fig. 8(b), which are also less effective in reducing buyback behavior in Fig. 8(d). This is consistent with the findings from the literature on efficiency losses from proportional transaction costs (Nie, 2012).



(a) Number of transactions: full wallets

(b) Avg. trading amount: full wallets

(c) Number of transactions: random wallets

(d) Avg. trading amount: random wallets

Fig. 9. The effects of full and random initial account balances on the transaction numbers and amount by time-of-day at equilibrium.

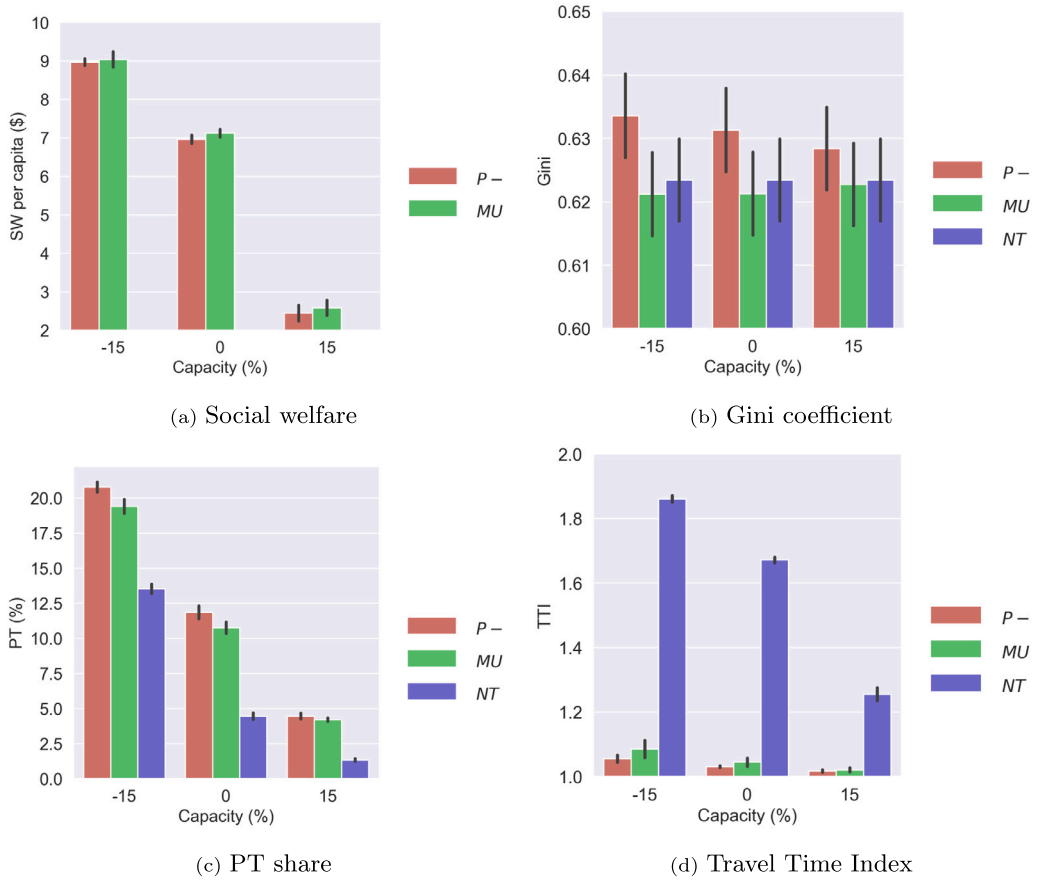
Next, we examine the effect of full initial account balances and random initial account balances on the transaction numbers and amount by time-of-day at equilibrium. The plots are based on simulations with a particular random seed because stochasticity can make the visualizations hard to interpret.

In Fig. 9(a), the numbers of buying and selling transactions as a percentage of the corresponding total number by time-of-day at equilibrium for full initial account balances are plotted. As we can see, buying transactions only happen in the peak hour because travelers can only buy tokens at time of traveling if they are short of tokens. In contrast, selling transactions happen at the beginning of the day, in the early morning, and peak hour, which can be explained by the plot of the average transaction amount by time-of-day for full initial account balances in Fig. 9(b). For travelers selling at the beginning of the day, all of them sell at full wallets as shown in Fig. 9(b) (2.052 tokens as equilibrium token price is $1) because they travel in the off peak and do not need to use tokens. For travelers that sell in the early morning (around 2 AM), their account balances at time of selling are not full because their future token allocations until their departure times can cover their toll and it is optimal for them to sell now. Finally, for travelers that sell in the peak period, they sell at full wallets because their account balances reach the full wallet after paying small toll charges. The selling behavior is consistent with the derived selling strategy but the excessive trading at the beginning of the day may be undesirable and avoiding this was in fact one of the motivations of the continuous allocation.

In practice, it is plausible that travelers will register for the program at different times in the day (one may think of the system as being implemented via a smartphone app). As a result, their account balances at the beginning of the day will be different, and we now assume the initial account balances are distributed uniformly between 0 and the maximum account balance (2.052 tokens). As shown in Fig. 9(c), the selling transactions are now spread across the day with a relatively mild peak in the early morning (around 2 AM). Apart from these travelers who sell in the early morning not at full wallets, other travelers sell only at full wallets as shown in Fig. 9(d). Under this assumption, we see a much more desirable pattern of transactions over the day, and the usefulness of the continuous allocation. However, the fact that different initial account states lead to different patterns of selling behavior at

Table 5
Factor levels for experiments.

Factor	Level 1	Level 2	Level 3
Capacity (s)	−15%	**0%**	15%
Income Effect (λ)	0	**3**	6
Heterogeneity (c.o.v)	0.2	0.9	**1.6**



(a) Social welfare (b) Gini coefficient

(c) PT share (d) Travel Time Index

Fig. 10. Variation of performance measures with capacity.

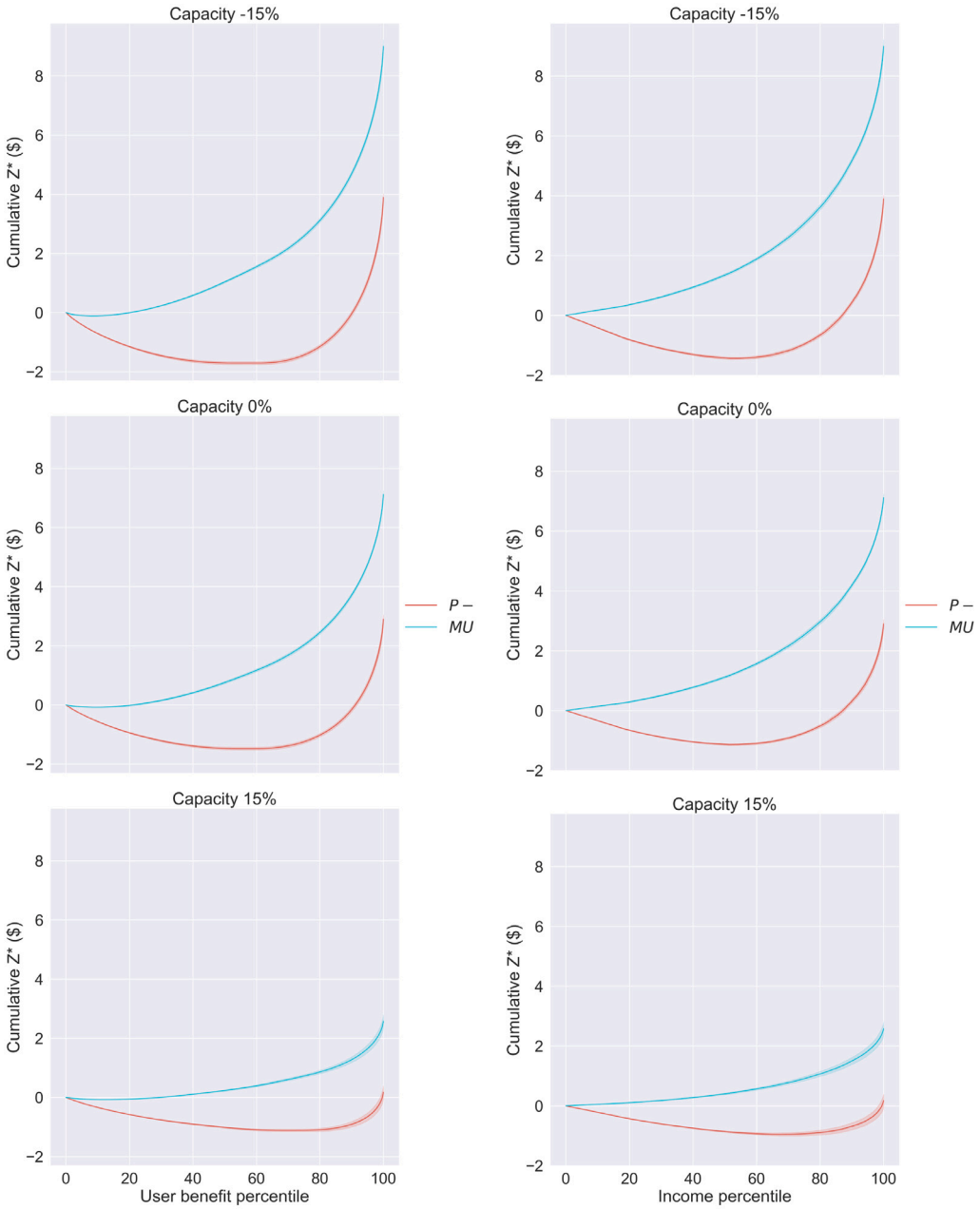
equilibrium is a problematic property of the system (despite the fact that the optimal welfare and associated market prices and flows are unique), and one that deserves further investigation.

6.5. *Performance of the TMC scheme under varying levels of congestion, heterogeneity and income effects*

We next examine the performance of the TMC scheme relative to congestion pricing at varying levels of three important experimental factors: capacity, income effect and heterogeneity. For the TMC scheme, fixed transaction fees are set to \$0.05 and proportional transaction fees are set to 0 based on the experiments in the previous section. The factors are varied one at a time across three levels as presented in Table 5. Values used in the base case are highlighted in bold (when varying a given factor, other factors are fixed at the base level). With regard to capacity, bottleneck capacity s is varied from 15% less capacity than the baseline to 15% more capacity than the baseline; for the income effect, the nonlinear income effect coefficient in the utility specification λ is varied from 0 to 6; for heterogeneity, the coefficient of variation of value of time α_n is varied from 0.2 to 1.6. In the following discussion, the TMC system is denoted by MU (U denotes the uniform allocation of tokens) and congestion pricing is denoted by P- to indicate that we do not assume a redistribution of toll revenues in any form. The No-toll scenario is denoted by NT. For each scenario in the experimental design, the two instruments and NT are simulated with five different random seeds until convergence.

6.5.1. *Capacity*

The comparative performance of the various instruments under varying levels of capacity in terms of social welfare (relative to the NT scenario), Gini coefficient, PT share and travel time index (TTI) are shown in Fig. 10. First, we observe that the overall

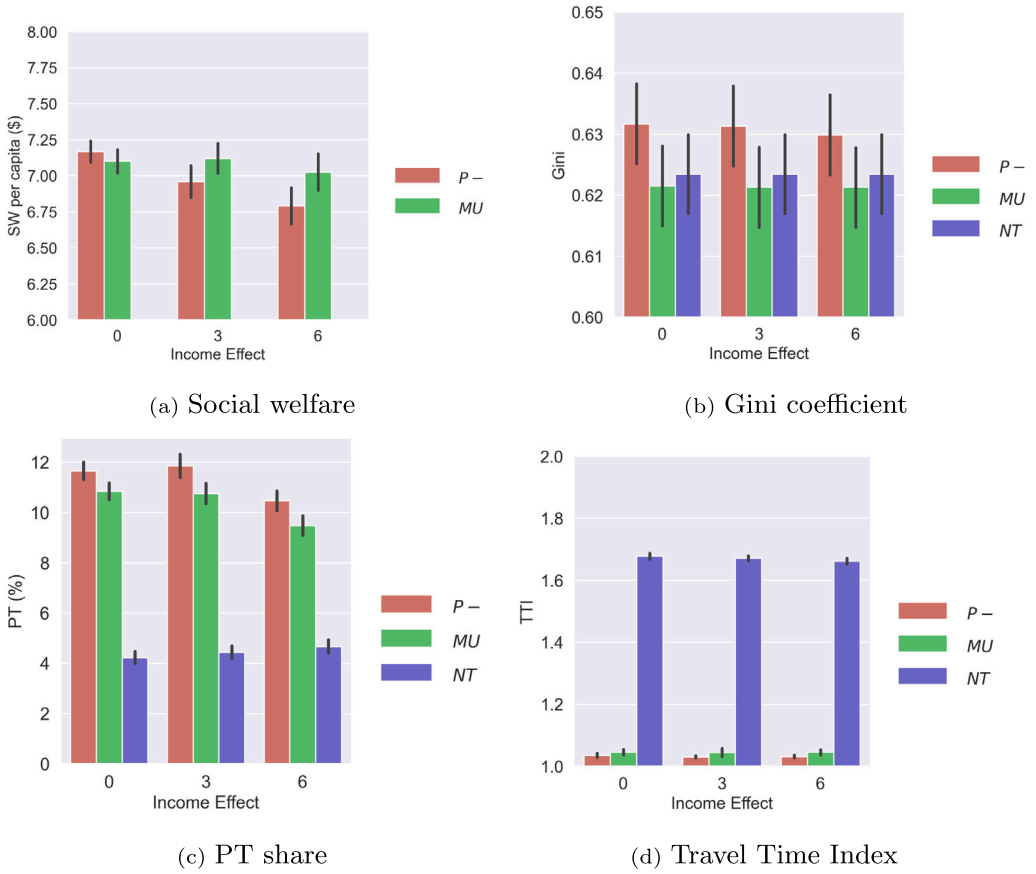


(a) Ordered by benefit (Lorenz curve) (b) Ordered by income

Fig. 11. Variation of cumulative user benefits with capacity.

welfare of pricing and the TMC scheme are similar at all levels of capacity, and in fact, marginally higher for the TMC scheme despite the small fixed transaction fees (due to the income effect). This is in line with the general finding that both pricing and tradable credits are equivalent in terms of efficiency under deterministic demand/supply and in the absence of transaction costs and income effects (Yang and Wang, 2011; de Palma et al., 2018). As expected, overall welfare gains decrease as the capacity increases and congestion effects are less severe.

Next, Fig. 10(b) shows that when toll revenues are not redistributed, the congestion pricing scheme is regressive, as seen by the increase in the Gini coefficient or GC (computed based on individual disposable income I_n plus user benefit z_n, see Eq. (43)) relative to the NT case. Observe also that the GC ($P-$) increases as capacity level decreases, which implies that $P-$ becomes less equitable



(a) Social welfare (b) Gini coefficient

(c) PT share (d) Travel Time Index

Fig. 12. Variation of performance measures with income effect level.

because as capacity decreases, the tolls increase to deal with the increasing congestion leading to the greater losses of low income users.

In contrast, the TMC scheme improves the GC relative to the NT case, due to the free uniform allocation of tokens to all travelers (a uniform allocation increases the proportion of cumulative benefits obtained by the travelers with lower values of time). The regressive nature of the congestion pricing scheme can also be observed in Fig. 11(a) which plots cumulative user benefits (normalized by population size) as a function of the user benefit percentile and Fig. 11(b) which plots cumulative user benefits (normalized by population size) as a function of income percentile. Clearly, at all capacity levels, one can observe that a large proportion of users are worse off from pricing (negative user benefits) whereas in the case of the TMC scheme the proportion of 'losers' is significantly smaller. Although not clearly visible in the plots, there are still some 'losers' with small negative benefits in the TMC scheme. Thus, under a uniform allocation of tokens, a tradable credit scheme does not necessarily guarantee Pareto improvement. This conclusion accords with the finding in Arnott et al. (1994) that under pricing, even with a uniform revenue rebate, some users may still be worse off. Fan et al. (2022) discuss the conditions under which Pareto Improvement is guaranteed for a tradable credit scheme using the standard bottleneck model with homogeneous users.

It should be pointed out that the above discussion on the regressiveness of pricing is premised on the assumption that value of time is correlated with income and that there is a one-to-one relationship between VOT and income. Empirical studies have identified other correlates of income and hence, Verhoef and Small (2004) have cautioned against viewing the VOT distribution as simply representing the income distribution (see also Lehe (2020) on this). Moreover, as noted in Eliasson and Mattsson (2006), ultimately, the distributional outcomes of congestion pricing depend largely on how the toll revenues – which can be significantly larger than net user benefits – are used. The ratio between tolls revenues and user benefits under pricing in our experiments are in the range of the empirical values reported in Eliasson and Mattsson (2006).

In terms of network performance, significant improvements in the TTI and larger transit shares are observed under both instruments (relative to NT) at all capacity levels (Figs. 10(c) and 10(d)).

6.5.2. *Income effect*

Next we examine the impact of income effects, measured by the parameter λ. As noted previously, non-linear income effects do impact behavior and the relative efficiency of pricing and TMCs. This impact on behavior along with the fact that the valuation of

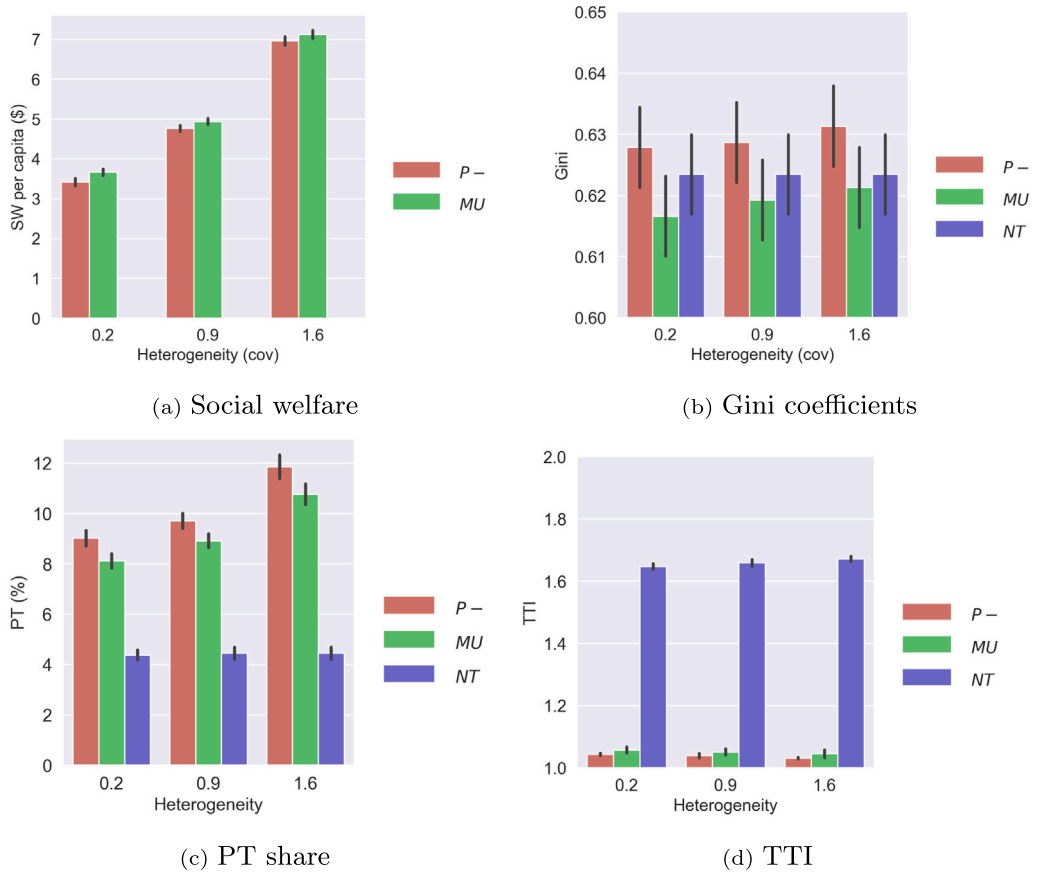


Fig. 13. Variation of performance measures with heterogeneity levels.

the token allocation is higher for lower income users (who have a higher marginal utility of income — recall that our measure of user benefit is directly the money metric utility) results in increasing welfare differences as λ increases (this can be seen in Fig. 12(a)). However, we caution that in terms of magnitudes, these differences are still small. As expected, when income effects are absent, the welfare of the TMC is marginally lower than P- due to transaction fees. Observe also that the social welfare of $P-$ decreases as the income effect increases since users are more sensitive to the tolls.

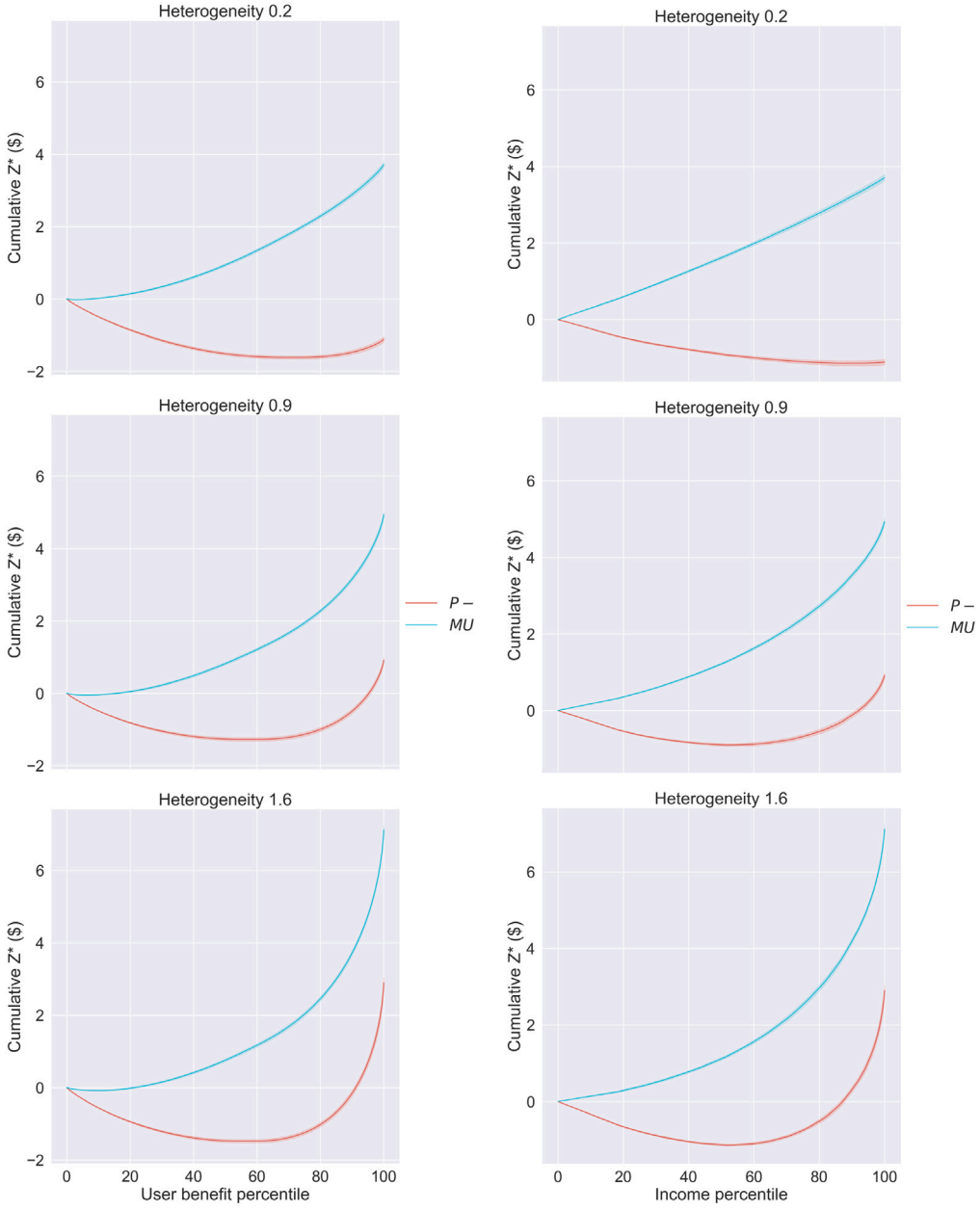
The impact on behavior can also be seen in the lower PT shares with increasing λ in Fig. 12(c). In terms of distributional impacts, we observe similar trends in the GC and patterns in user benefits as described in Section 6.5.1 (which show the regressiveness and larger proportion of 'losers' under pricing; see the base capacity case in Figs. 11(a) and 11(b)). There are small variations across λ and hence, for brevity, we omit plots of user benefits.

6.5.3. Heterogeneity

As can be seen in Fig. 13(a), as the extent of heterogeneity increases, the social welfare of $P-$ increases, which is consistent with findings in the literature (e.g., Van Den Berg and Verhoef (2011)). Neglecting heterogeneity underestimates the benefits of pricing (Verhoef and Small, 2004). The relative performance of the two instruments does not change appreciably with the extent of heterogeneity, with the TMC scheme having marginally higher welfare and significantly better GC at all levels of heterogeneity. In terms of distributional impacts, at low levels of heterogeneity (COV of 0.2), it can be seen in Fig. 14(a) that all users are worse off under P- whereas most of them are better off under the credit system.

6.6. Robustness

In practice, toll profiles may often be sub-optimal because of changing conditions, forecast errors and uncertainty. Practically, it is difficult to update these toll profiles (especially at the network level) regularly in practice. For example, Singapore updates the ERP scheme once every three months. In contrast, some market elements of the TMC scheme (e.g., allocation rate) are easier to adapt and have the potential to influence travelers' behavior (through the market price) to recover efficiency losses. In this section, two scenarios of a sub-optimal toll profile are investigated, including forecast error and non-recurrent events.



(a) Ordered by benefit (Lorenz curve) (b) Ordered by income

Fig. 14. Variation of cumulative user benefits with heterogeneity level.

The first scenario is forecast error wherein actual road capacity is assumed to be 15% less than the anticipated road capacity used to optimize the toll profile. The social welfare of pricing and TMC with this sub-optimal toll profile (based on anticipated road capacity) are plotted in Fig. 15 and denoted by $P{-}_S$ and MU_S. The social welfare of pricing and TMC with the associated optimal toll profiles are also plotted and denoted as $P{-}_O$ and MU_O. As we can see, MU_S (TMC with sub-optimal tolls) is able to recover efficiency losses through a reduction in the allocation rate, which reduces token supply and increases token price. The optimal allocation rate is determined using a grid search and is found to be 15% lower than the original allocation rate.

The second scenario is a non-recurrent event. Specifically, it is assumed that there is a sudden within-day capacity drop by 15% (e.g., due to an accident or incident) from 7 AM to 8:30 AM on the 10th day after the system has reached an equilibrium. The social welfare across days for the three instruments are plotted in Fig. 16. The first instrument is pricing without distribution and denoted

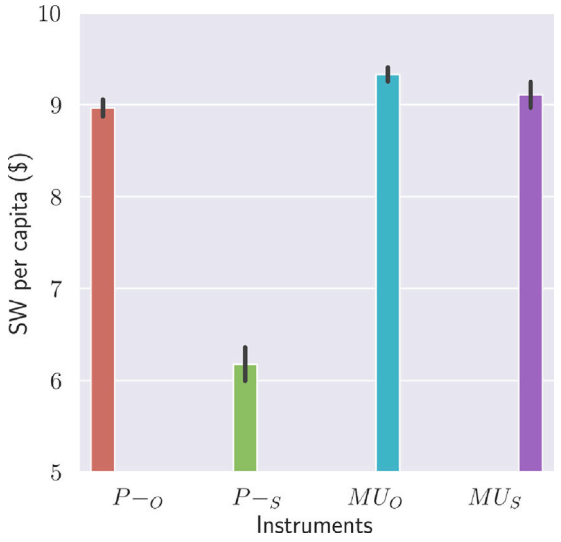


Fig. 15. Social welfare of pricing and TMC with sub-optimal and optimal toll profiles.

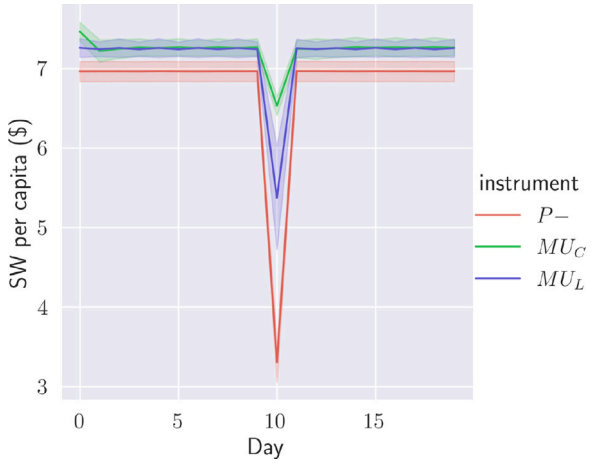


Fig. 16. Social welfare of pricing and TMC with sub-optimal and optimal toll profiles.

as $P-$. The second one is TMC with a lump-sum allocation and denoted by MU_L. The third instrument is TMC with a continuous allocation and denoted as MU_C.

Under MU_L, travelers receive the entire day's token allocation at the beginning of the day in the form of a 'lump-sum' allocation. This form of token allocation is the standard design of TMC schemes in the literature (e.g., Yang and Wang (2011) and Brands et al. (2020)). Regarding trading, they can buy additional tokens at the time of traveling for immediate use and redeem unused tokens at the end of the day. Since trading is automated, there is no transaction fee considered under the lump-sum allocation. Once the incident occurs, the regulator has three market parameters to control including token price, regulation starting time and ending time. It cannot control allocation rate as all the tokens have already been allocated at the beginning of the day. Note that thus far, we have treated the token price as fixed within-day. Incorporating a truly within-day dynamic token price in the day-to-day departure time choice model involves additional assumptions about users' forecasts of the dynamic prices, which would not be known when the departure time choice is made. Instead, we use a simpler ad-hoc approach, assuming that users who have not yet departed at the time the new token price takes effect, will reevaluate their decisions based on the information of the new prices. We assume the same behavioral model used in the pre-day decision applies (with updated token prices and transaction fees). Using the DE algorithm, we determine that it is optimal for the regulator to increase the token price to $1.8 between 6:55 AM and 9:15 AM. As shown in Fig. 16, it performs better than $P-$.

Under MU_C, the regulator can control not only token price, regulation starting and ending time, but also the allocation rate (which affects the forecasted account balance of users over the rest of the day) and transaction fees. We optimize the fixed transaction fees of buying and selling together. Through optimization, between 7:05 AM and 8:50 AM, the regulator should set token price equal

to \$1.25, allocation rate r equal to 0 and fixed transaction fee equal to \$0.5. It performs better than the lump-sum allocation MU_L as shown in Fig. 16. This is intuitive because the travel behavior is impacted by both the allocation rate and transaction fees, which provide the regulator more degrees of freedom to intervene. This demonstrates the advantages of a continuous allocation of tokens over a lump-sum allocation of tokens.

7. Conclusions

This paper proposes and analyzes alternative market models for a tradable mobility credit system focusing on allocation/expiration of tokens, rules governing trading, transaction fees, and regulator intervention. We develop a methodology to explicitly model the dis-aggregate behavior of individuals within the market. Extensive simulation experiments are conducted within a combined mode and departure time context for the morning commute problem using a day-to-day assignment framework wherein transportation demand is modeled using a logit-mixture model with income effects and supply is modeled using a standard bottleneck model.

The results show that small, fixed transaction fees can effectively mitigate undesirable speculation in the market without a significant loss in efficiency whereas proportional transaction fees are less effective both in terms of efficiency and in avoiding undesirable market speculation. The market design we adopt is shown to yield stable and convergent market prices and is revenue neutral. We also show that an allocation of tokens in continuous time can be beneficial in dealing with non-recurrent events. One undesirable property of the system we do observe is that the selling behavior of individuals at equilibrium is dependent on the initial account states (although optimal welfare, flows and market prices are unique), and this deserves further investigation. With regard to the comparative performance relative to congestion pricing, in the presence of income effects, the TMC system yields a marginally higher social welfare. Finally, the TMC scheme is more equitable (when revenues from congestion pricing are not redistributed) although it is not guaranteed to be Pareto-improving when tokens are distributed equally.

There are several promising directions for future research. First, given that a uniform allocation of tokens does not guarantee Pareto-improvement, it is clear that personalization is necessary to ensure that no individual loses under the TMC scheme. The design of personalized token allocation schemes is an important direction of future research. Second, more systematic experiments on day-to-day variability and its impact on the robustness of the TMC scheme and market are warranted. Finally, large scale simulations on real-world networks with disaggregate agent-based models is a natural next step in studying market design and other aspects of tradable credit schemes towards real-world deployments.

CRediT authorship contribution statement

Siyu Chen: Conceptualization, Methodology, Visualization, Investigation, Formal analysis, Software, Writing – original draft, Writing – review & editing. **Ravi Seshadri:** Conceptualization, Methodology, Investigation, Formal analysis, Writing – original draft, Writing – review & editing, Supervision. **Carlos Lima Azevedo:** Conceptualization, Methodology, Investigation, Supervision, Funding, Writing – review & editing. **Arun P. Akkinepally:** Conceptualization, Methodology, Investigation, Supervision. **Renming Liu:** Methodology, Investigation, Formal analysis, Software, Writing – review & editing. **Andrea Araldo:** Conceptualization, Methodology, Investigation, Supervision. **Yu Jiang:** Methodology, Investigation, Supervision, Writing – review & editing, Funding. **Moshe E. Ben-Akiva:** Conceptualization, Methodology, Supervision, Funding.

Acknowledgments

This research was funded by the U.S. National Science Foundation (Award CMMI-1917891) and the NEMESYS project funded by the DTU (Technical University of Denmark)-NTU (Nanyang Technical University) Alliance.

Appendix A. Convergence of the day-to-day dynamic model

In this section, we examine convergence of the day-to-day model (Section 5). Using data and parameters discussed in Section 6.1, a No Toll (NT) scenario with different sets of initial travel time information for the driving alternatives is simulated. Five independent replications or draws of the error terms in the utility (Eq. (25)) are performed for each scenario. The four different sets of initial travel time information of driving are plotted in Fig. A.17(a). Note that the initial travel time information serves as the basis for the departure time decisions in day 0 of the day-to-day simulation. The initial travel time set 0 represents free flow travel times across the entire day; the initial travel time set 1 consists of equilibrium travel times of driving obtained from the simulation using the initial travel time set 0 for a particular random seed; the initial travel time set 2 is 0.6 times the travel time set 1; finally, the initial travel time set 3 represents a constant 30 min travel time across the entire day.

The corresponding convergence of travel times of driving is plotted in Fig. A.17(b). The lines and bands in the plot represent averages and standard deviations, respectively, across the five replications. As we can see from the plot, the infinity norm (Eq. (48)) converges to within a threshold of 0.01 by about 50 days, indicating acceptable convergence of the day-to-day model. We perform a commonly used test for stationarity, the Augmented Dickey–Fuller (ADF) test (Cheung and Lai, 1995), on the infinity norm. The p-value is significantly smaller than 0.05 (i.e., the tested data does not have a unit root) indicating that stationarity has been achieved.

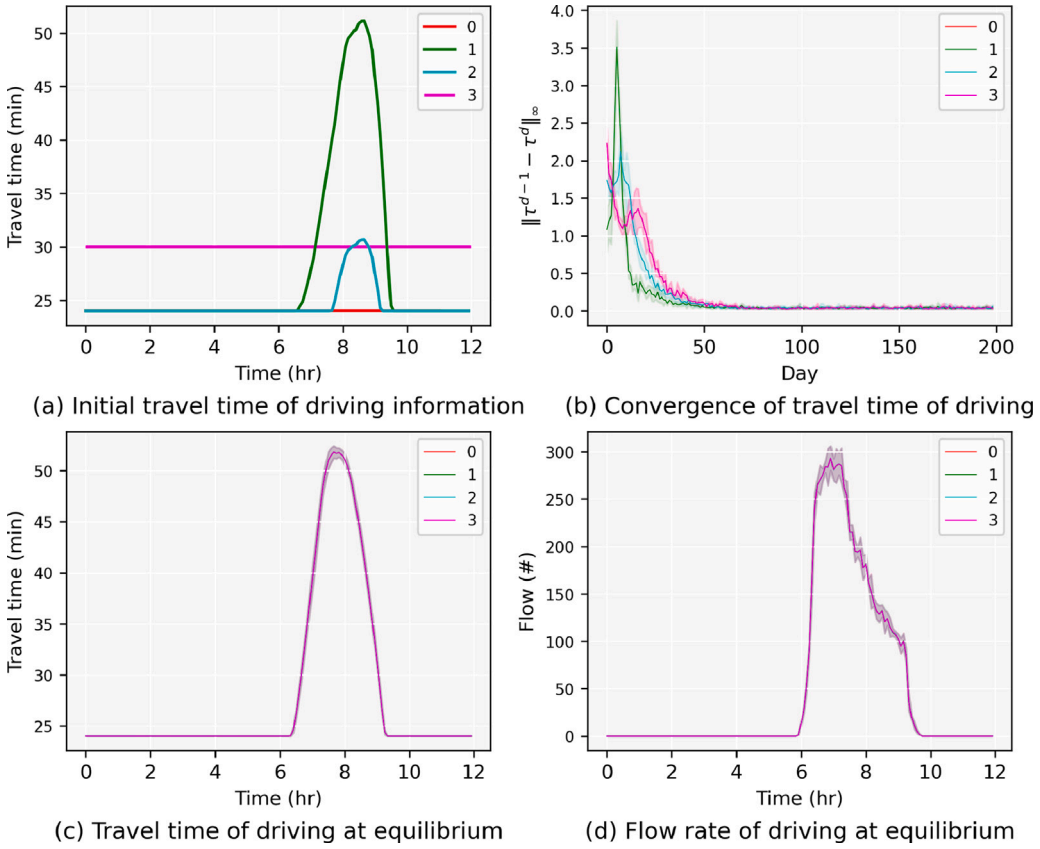


Fig. A.17. Convergence of the day-to-day model (base capacity).

In order to examine regularity of the stochastic process, the travel time vectors and flow vectors of driving under the four different starting conditions are plotted in Fig. A.17(c)–(d). The results indicate that simulations with different initial travel times of driving converge to the same distribution of travel times and departure flows. We conduct two-sample t-tests for the mean departure flow (and travel time) in each interval (using the Bonferroni correction for multiple comparisons) across every pair of the four scenarios. The tests indicate that the mean flows and travel times (at convergence) under the different starting conditions are statistically not different ($\alpha = 0.01$). As only two mode choices are considered and the travel time and headway of PT are constant, once the travel time of driving converges, the resulting departure flow of PT also converges.

We also examine convergence of the NT scenario under two additional capacity levels as presented in Table 5 (i.e., increasing and decreasing the bottleneck capacity by 15%). All experiment settings are the same as before, except that for the -15% capacity case, the initial travel time set 2 is 0.8 times the travel time set 1. Results are shown in Figs. A.18 and A.19, respectively. In both cases, it is observed that the infinity norms converge and the travel time vectors and flow vectors of driving under different starting conditions converge to the same distribution of travel times and departure flows. The ADF tests show the stationarity of the infinity norms (p-values are smaller than 0.01), and the t-tests confirm that the mean flows and the travel times across time intervals are statistically the same.

Appendix B. Convergence of market prices

The effect of various initial market prices p^0 on convergence of token market price and social welfare are examined in Fig. A.20. The price and social welfare converge to values that are not statistically significantly different at a significance level of 0.05 regardless of the initial price. The regulator revenues under the three initial prices converge to be within the regulator revenue threshold band (the black lines) as shown in Fig. A.20(c). Travel times under the three initial prices converge too as shown in Fig. A.20(d). Similar experiments under a range of scenarios showed that the price adjustment scheme leads to convergence (in terms of travel times, flows and token prices) of the day-to-day model to the same stationary distribution for different initial market prices.

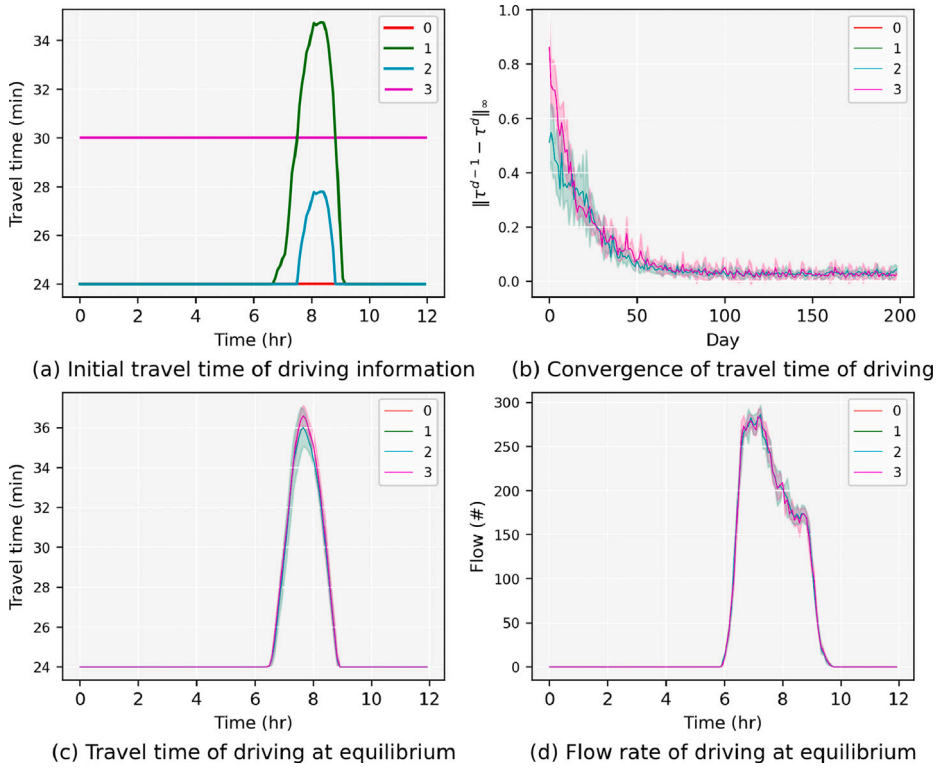


Fig. A.18. Convergence of the day-to-day model (15% higher capacity).

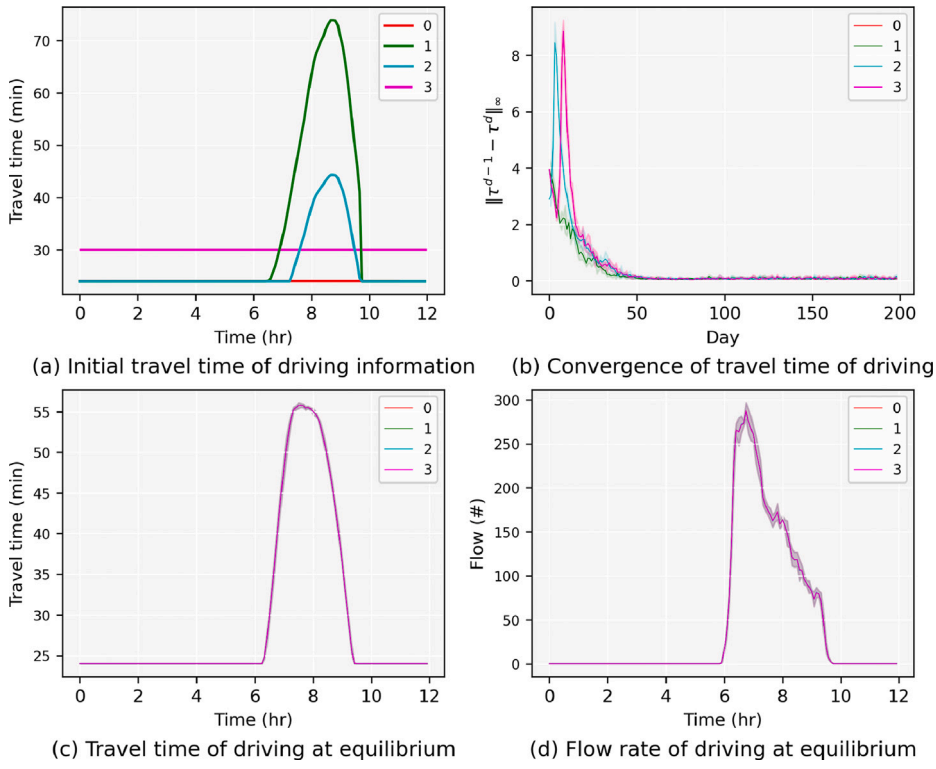
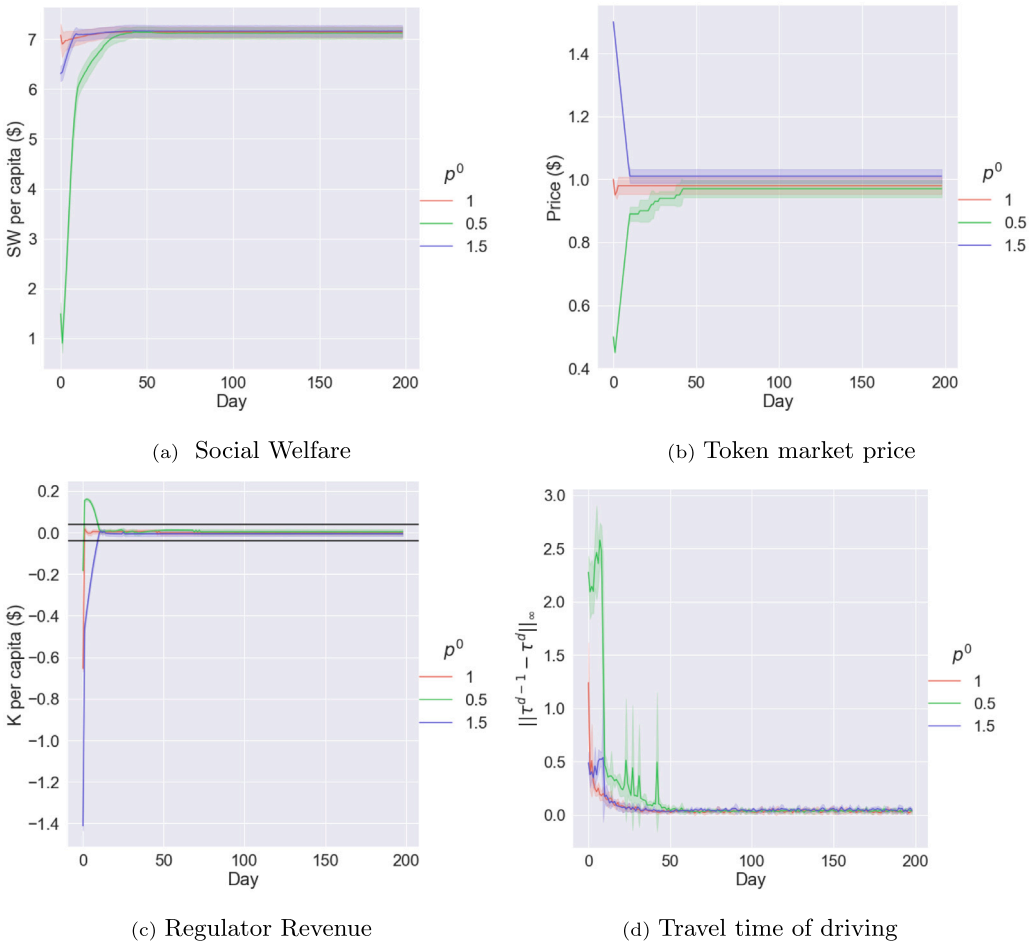


Fig. A.19. Convergence of the day-to-day model (15% lower capacity).



(a) Social Welfare

(b) Token market price

(c) Regulator Revenue

(d) Travel time of driving

Fig. A.20. Convergence of various metrics with different initial prices p^0.

Table C.6
Notation.

Variables	Description
h	Departure time interval
$\bar{h}$	Time interval for supply and selling model
t	Continuous time
d	Day d
t_h	Start time of interval h
Δ_h	Duration of departure time interval h
$\Delta_{\bar{h}}$	Duration of interval $\bar{h}$
Δ_a	Size of desired arrival window
n	Individual n
α_n	Value of time of individual n
β_{En}/β_{Ln}	Value of schedule delay early/late of individual n
λ	Coefficient of nonlinear income effect
γ	Nonlinear income effect adjustment parameter
μ_n	Random component scale parameter of individual n
ϵ_{in}	Random utility component for mobility decision i of individual n
I_n	Disposable income of individual n
H_n	Departure time choice set of individual n
M_n	Mode choice set of individual n

(*continued on next page*)

Appendix C. Notation

See Table C.6.

Table C.6 (*continued*).

Variables	Description
$\hat{t}_n$	Desired arrival time of individual n
η	Departure time window size parameter
p	Market price
$\mathrm{T}^j(h)$	Toll of instrument j in h
$\tilde{\tau}_i$	Forecasted travel time of choice i
$\tilde{c}_{in}$	Expected cost for mobility decision i
$x_n^d(t)$	Account balance of individual n at time t
L	Token lifetime
r	Token allocation rate
$F_S^P,\ F_B^P$	Proportional transaction fees for selling and buying
$F_S^F,\ F_B^F$	Fixed transaction fees for selling and buying
t_f	Free flow travel time
$t_v(t)$	Delay in queue at t
$Q(t)$	Number of drivers in queue at t
θ_τ/θ_t	Weights on previous day's forecasts
U_{in}	Utility of individual n for mobility alternative i
$(\cdot)^j$	Variable associated with instrument j

References

Akamatsu, T., Wada, K., 2017. Tradable network permits: A new scheme for the most efficient use of network capacity. Transp. Res. C 79, 178–195.

Ali, A.T., Flannery, A., Venigalla, M.M., 2007. Prediction Models for Free Flow Speed on Urban Streets. Technical Report.

Arnott, R., De Palma, A., Lindsey, R., 1990. Economics of a bottleneck. J. Urban Econ. 27 (1), 111–130.

Arnott, R., De Palma, A., Lindsey, R., 1994. The welfare effects of congestion tolls with heterogeneous commuters. J. Transp. Econ. Policy 139–161.

Bao, Y., Gao, Z., Yang, H., Xu, M., Wang, G., 2017. Private financing and mobility management of road network with tradable credits. Transp. Res. A 97, 158–176.

Bao, Y., Verhoef, E.T., Koster, P., 2019. Regulating dynamic congestion externalities with tradable credit schemes: Does a unique equilibrium exist? Transp. Res. B 127, 225–236.

Ben-Akiva, M., Cyna, M., De Palma, A., 1984. Dynamic model of peak period congestion. Transp. Res. B 18 (4–5), 339–355.

Ben-Akiva, M., De Palma, A., Kanaroglou, P., 1986. Dynamic model of peak period traffic congestion with elastic arrival rates. Transp. Sci. 20 (3), 164–181.

Ben-Akiva, M.E., Lerman, S.R., Lerman, S.R., et al., 1985. Discrete Choice Analysis: Theory and Application to Travel Demand, Vol. 9. MIT Press.

Van den Berg, V., Verhoef, E.T., 2011. Congestion tolling in the bottleneck model with heterogeneous values of time. Transp. Res. B 45 (1), 60–78.

Blackorby, C., Donaldson, D., 1990. A review article: The case against the use of the sum of compensating variations in cost-benefit analysis. Can. J. Econ. 471–494.

Brands, D.K., Verhoef, E.T., Knockaert, J., Koster, P.R., 2020. Tradable permits to manage urban mobility: market design and experimental implementation. Transp. Res. A 137, 34–46.

Cantarella, G.E., Cascetta, E., 1995. Dynamic processes and equilibrium in transportation networks: towards a unifying theory. Transp. Sci. 29 (4), 305–329.

Chen, M., 2010. Travel time based congestion measures for freeway corridors.

Cheung, Y.-W., Lai, K.S., 1995. Lag order and critical values of the augmented Dickey–Fuller test. J. Bus. Econom. Statist. 13 (3), 277–280.

De Palma, A., Ben-Akiva, M., Lefevre, C., Litinas, N., 1983. Stochastic equilibrium model of peak period traffic congestion. Transp. Sci. 17 (4), 430–453.

De Palma, A., Lindsey, R., 2002. Congestion pricing in the morning and evening peaks: A comparison using the Bottleneck Model. In: Transportation Visioning-2002 and beyond (Vision D'Avenir Des Transports-2002 et au-Dela), Canadian Transportation Research Forum, Proceedings of the 37th Annual Conference.

De Palma, A., Lindsey, R., 2004. Congestion pricing with heterogeneous travelers: A general-equilibrium welfare analysis. Netw. Spat. Econ. 4 (2), 135–160.

Delle Site, P., Salucci, M.V., 2013. Transition choice probabilities and welfare analysis in random utility models with imperfect before–after correlation. Transp. Res. B 58, 215–242.

Ding, C., Mishra, S., Lin, Y., Xie, B., 2015. Cross-nested joint model of travel mode and departure time choice for urban commuting trips: Case study in maryland–washington, dc region. J. Urban Plann. Dev. 141 (4), 04014036.

Dogterom, N., Ettema, D., Dijst, M., 2017. Tradable credits for managing car travel: a review of empirical research and relevant behavioural approaches. Transp. Rev. 37 (3), 322–343.

Eliasson, J., Mattsson, L.-G., 2006. Equity effects of congestion pricing: quantitative methodology and a case study for Stockholm. Transp. Res. A 40 (7), 602–620.

Fan, W., Jiang, X., 2013. Tradable mobility permits in roadway capacity allocation: Review and appraisal. Transp. Policy 30, 132–142.

Fan, W., Xiao, F., et al., 2022. Managing bottleneck congestion with tradable credits under asymmetric transaction cost. Transp. Res. E 158, 102600.

Gillingham, K., 2014. Identifying the elasticity of driving: evidence from a gasoline price shock in California. Reg. Sci. Urban Econ. 47, 13–24.

Goddard, H.C., 1997. Using tradeable permits to achieve sustainability in the world's large cities: policy design issues and efficiency conditions for controlling vehicle emissions, congestion and urban decentralization with an application to Mexico City. Environ. Resour. Econ. 10 (1), 63–99.

Grant-Muller, S., Xu, M., 2014. The role of tradable credit schemes in road traffic congestion management. Transp. Rev. 34 (2), 128–149.

Gulipalli, P.K., Kockelman, K.M., 2008. Credit-based congestion pricing: A Dallas-Fort Worth application. Transp. Policy 15 (1), 23–32.

Guo, R.-Y., Yang, H., Huang, H.-J., 2018a. Are we really solving the dynamic traffic equilibrium problem with a departure time choice? Transp. Sci. 52 (3), 603–620.

Guo, R.-Y., Yang, H., Huang, H.-J., Li, X., 2018b. Day-to-day departure time choice under bounded rationality in the bottleneck model. Transp. Res. B 117, 832–849.

Hara, Y., Hato, E., 2017. A car sharing auction with temporal-spatial OD connection conditions. Transp. Res. Procedia 23, 22–40.

He, F., Yin, Y., Shirmohammadi, N., Nie, Y.M., 2013. Tradable credit schemes on networks with mixed equilibrium behaviors. Transp. Res. B 57, 47–65.

Holguin-Veras, J., Ozbay, K., de Cerreno, A.L., et al., 2005. Evaluation Study of Port Authority of New York and New Jersey's Time of Day Pricing Initiative. Technical Report, University Transportation Research Center.

Kirk, D.E., 2004. Optimal Control Theory: An Introduction. Courier Corporation.

Kockelman, K.M., Kalmanje, S., 2005. Credit-based congestion pricing: a policy proposal and the public's response. Transp. Res. A 39 (7–9), 671–690.

Kristoffersson, I., Engelson, L., 2018. Estimating preferred departure times of road users in a large urban network. Transportation 45 (3), 767–787.
Layard, R., Mayraz, G., Nickell, S., 2008. The marginal utility of income. J. Public Econ. 92 (8–9), 1846–1857.
Lehe, L.J., 2020. Winners and losers from road pricing with heterogeneous travelers and a mixed-traffic bus alternative. Transp. Res. B 139, 432–446.
Lipow, G.W., 2008. Price–Elasticity of Energy Demand: A Bibliography. Carbon Tax Center, (www.carbontax.org).
Liu, R., Chen, S., Jiang, Y., Seshadri, R., Ben-Akiva, M., Azevedo, C.L., 2023. Managing network congestion with a trip- and area-based tradable credit scheme. Transportmetr. B 11 (1), 434–462. http://dx.doi.org/10.1080/21680566.2022.2083034, arXiv:https://doi.org/10.1080/21680566.2022.2083034.
Liu, W., Geroliminis, N., 2017. Doubly dynamics for multi-modal networks with park-and-ride and adaptive pricing. Transp. Res. B 102, 162–179.
Liu, W., Li, X., Zhang, F., Yang, H., 2017. Interactive travel choices and traffic forecast in a doubly dynamical system with user inertia and information provision. Transp. Res. C 85, 711–731.
Liu, W., Szeto, W.Y., 2020. Learning and managing stochastic network traffic dynamics with an aggregate traffic representation. Transp. Res. B 137, 19–46.
Louviere, J.J., Eagle, T., 2006. Confound it! That pesky little scale constant messes up our convenient assumptions. In: Sawtooth Software Conference. Sawtooth Software Inc.
McAfee, R.P., 1992. A dominant strategy double auction. J. Econom. Theory 56 (2), 434–450.
McFadden, D., 2001. Economic choices. Amer. Econ. Rev. 91 (3), 351–378.
McFadden, D., 2017. Foundations of Welfare Economics and Product Market Applications. Technical Report, National Bureau of Economic Research.
Myerson, R.B., Satterthwaite, M.A., 1983. Efficient mechanisms for bilateral trading. J. Econom. Theory 29 (2), 265–281.
Nie, Y.M., 2012. Transaction costs and tradable mobility credits. Transp. Res. B 46 (1), 189–203.
Nie, Y.M., 2015. A new tradable credit scheme for the morning commute problem. Netw. Spat. Econ. 15 (3), 719–741.
Nie, Y.M., Yin, Y., 2013. Managing rush hour travel choices with tradable credit scheme. Transp. Res. B 50, 1–19.
de Palma, A., Kilani, K., 2005. Switching in the logit. Econom. Lett. 88 (2), 196–202.
de Palma, A., Lindsey, R., 2011. Traffic congestion pricing methodologies and technologies. Transp. Res. C 19 (6), 1377–1399.
de Palma, A., Lindsey, R., 2020. Tradable permit schemes for congestible facilities with uncertain supply and demand. Econ. Transp. 21, 100149.
de Palma, A., Proost, S., Seshadri, R., Ben-Akiva, M., 2018. Congestion tolling-dollars versus tokens: A comparative analysis. Transp. Res. B 108, 261–280.
Ramadurai, G., Ukkusuri, S.V., Zhao, J., Pang, J.-S., 2010. Linear complementarity formulation for single bottleneck model with heterogeneous commuters. Transp. Res. B 44 (2), 193–214.
Raux, C., 2007. Tradable driving rights in urban areas: their potential for tackling congestion and traffic-related pollution. In: The Implementation and Effectiveness of Transport Demand Management Measures. p. 95.
Ruggles, S., Flood, S., Foster, S., Goeken, R., Pacas, J., Schouweiler, M., Sobek, M., 2021. IPUMS USA: Version 11.0. IPUMS, Minneapolis, MN, http://dx.doi.org/10.18128/D010.V11.0.
Sasic, A., Habib, K.N., 2013. Modelling departure time choices by a Heteroskedastic Generalized Logit (Het-GenL) model: An investigation on home-based commuting trips in the Greater Toronto and Hamilton Area (GTHA). Transp. Res. A 50, 15–32.
Seshadri, R., de Palma, A., Ben-Akiva, M., 2022. Congestion tolling — Dollars versus tokens: Within-day dynamics. Transp. Res. C 143, 103836. http://dx.doi.org/10.1016/j.trc.2022.103836, URL: https://www.sciencedirect.com/science/article/pii/S0968090X2200256X.
Shepherd, S., Sumalee, A., 2004. A genetic algorithm based approach to optimal toll level and location problems. Netw. Spat. Econ. 4 (2), 161–179.
Shirmohammadi, N., Zangui, M., Yin, Y., Nie, Y., 2013. Analysis and design of tradable credit schemes under uncertainty. Transp. Res. Rec. 2333 (1), 27–36.
Small, K.A., 1982. The scheduling of consumer activities: work trips. Am. Econ. Rev. 72 (3), 467–479.
Small, K.A., 2012. Valuation of travel time. Econ. Transp. 1 (1–2), 2–14.
Small, K.A., Winston, C., Yan, J., 2005. Uncovering the distribution of motorists' preferences for travel time and reliability. Econometrica 73 (4), 1367–1382.
Storn, R., Price, K., 1997. Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces. J. Global Optim. 11 (4), 341–359.
Tafreshian, A., Masoud, N., 2022. A truthful subsidy scheme for a peer-to-peer ridesharing market with incomplete information. Transp. Res. B 162, 130–161.
Tian, L.-J., Yang, H., Huang, H.-J., 2013. Tradable credit schemes for managing bottleneck congestion and modal split with heterogeneous users. Transp. Res. E 54, 1–13.
Tsekeris, T., Voß, S., 2009. Design and evaluation of road pricing: state-of-the-art and methodological advances. NETNOMICS: Econ. Res. Electron. Netw. 10 (1), 5–52.
U.S. Bureau of Labor Statistics, 2019. Consumer expenditures - 2019. p. 5, https://www.bls.gov/news.release/pdf/cesan.pdf. Accessed Dec. 29, 2021.
Van Den Berg, V., Verhoef, E.T., 2011. Winning or losing from dynamic bottleneck congestion pricing?: The distributional effects of road pricing with heterogeneity in values of time and schedule delay. J. Public Econ. 95 (7–8), 983–992.
Verhoef, E., Nijkamp, P., Rietveld, P., 1997. Tradeable permits: their potential in the regulation of road transport externalities. Environ. Plan. B: Plann. Des. 24 (4), 527–548.
Verhoef, E.T., Small, K.A., 2004. Product differentiation on roads. J. Transp. Econ. Policy (JTEP) 38 (1), 127–156.
Vickrey, W., 1973. Pricing, Metering, and Efficiently using Urban Transportation Facilities. 476.
Wang, G., Gao, Z., Xu, M., 2014. An MPEC formulation and its cutting constraint algorithm for continuous network design problem with multi-user classes. Appl. Math. Model. 38 (5–6), 1846–1858.
Wang, X., Yang, H., Zhu, D., Li, C., 2012. Tradable travel credits for congestion management with heterogeneous users. Transp. Res. E 48 (2), 426–437.
Weitzman, M.L., 1974. Prices vs. quantities. Rev. Econom. Stud. 41 (4), 477–491.
White, V., 2016. Revised Departmental Guidance on Valuation of Travel Time in Economic Analysis. Office of the Secretary of Transportation, US Department of Transportation, Available at: transportation.gov/sites/dot.gov/files/docs/2016%20Revised%20Value%20of%20Travel%20Time%20Guidance.pdf.
Wu, D., Yin, Y., Lawphongpanich, S., Yang, H., 2012. Design of more equitable congestion pricing and tradable credit schemes for multimodal transportation networks. Transp. Res. B 46 (9), 1273–1287.
Xiao, F., Qian, Z.S., Zhang, H.M., 2013. Managing bottleneck congestion with tradable credits. Transp. Res. B 56, 1–14.
Yang, H., Wang, X., 2011. Managing network mobility with tradable credits. Transp. Res. B 45 (3), 580–594. http://dx.doi.org/10.1016/j.trb.2010.10.002.
Ye, H., Yang, H., 2013. Continuous price and flow dynamics of tradable mobility credits. Transp. Res. B 57, 436–450.
Zhang, J., Wen, D., Zeng, S., 2015. A discounted trade reduction mechanism for dynamic ridesharing pricing. IEEE Trans. Intell. Transp. Syst. 17 (6), 1586–1595.
Zhang, X., Yang, H., 2004. The optimal cordon-based network congestion pricing problem. Transp. Res. B 38 (6), 517–537.
Zhao, Y., Kockelman, K.M., Karlstrom, A., 2008. Welfare calculations in discrete choice settings: Role of error term correlation. In: Proceedings of the Transportation Research Board 87th Annual Meeting. Citeseer.
Zhu, D.-L., Yang, H., Li, C.-M., Wang, X.-L., 2015. Properties of the multiclass traffic network equilibria under a tradable credit scheme. Transp. Sci. 49 (3), 519–534.