THEMA

théorie économique, modélisation et applications

THEMA Working Paper n°2021-12
CY Cergy Paris Université, France

Congestion Tolling−Dollars versus Tokens: Within-day Dynamics

Ravi Seshadri, André de Palma, Moshe Ben-Akiva

June 2021

CERGY PARIS UNIVERSITÉ

cnrs

Congestion Tolling – Dollars versus Tokens: Within-day Dynamics

Ravi Seshadri

Department of Technology, Management and Economics, Technical University of Denmark, DTU, 2800 Kgs. Lyngby, Denmark

André de Palma

CY Cergy Paris Université THEMA 33 Boulevard du Port, 95000 Cergy-Pontoise. France

Moshe Ben-Akiva

Massachusetts Institute of Technology (MIT), Room 1-181, 77 Massachusetts Avenue, Cambridge, MA, 02139

Abstract

Tradable permit schemes (or tolling in tokens) are a form of quantity control, which promise to be an appealing alternative to congestion pricing (or tolling in dollars) owing to considerations of revenue neutrality, equity, reduced infrastructure costs, and political acceptability. The comparative performance of the two instruments under uncertainty in demand and supply has only recently received attention in the transportation setting, despite being widely studied for emission markets. In this paper, we add to this literature by considering a tradable permit scheme in a departure time context wherein users are provided an initial endowment of tokens by the regulator and incur a token charge (determined prior to all departures) to travel in a specific time period. Tokens can be bought and sold within a marketplace at a price determined by a market clearing mechanism in each time period. A key feature of the market model is that the selling decisions of users are explicitly considered, which enables us to study the impact of selling behavior on performance of the permit system. Travel demand is modeled using a logit mixture model and supply consists of static congestion.

In the case of uncertain demand/supply wherein the tolls (in dollars and tokens) can be adapted from day to day (or alternatively demand/supply are deterministic), the two instruments can be shown analytically to be equivalent. In contrast, when the tolls are not day to day adaptive, the comparison of the two instruments is performed numerically. Our experiments over a wide range of demand and supply scenarios show that although neither instrument is consistently superior in terms of efficiency (overall social welfare), tolling in tokens outperforms tolling in dollars when congestion effects are more severe (e.g. realistic BPR models and steep congestion functions, high demand levels and high day-to-day variability). Importantly, we find that the token system is robust in efficiency terms (social

Email addresses: `ravse@dtu.dk` (Ravi Seshadri), `andre.de-palma@cyu.fr` (André de Palma), `mba@mit.edu` (Moshe Ben-Akiva)

welfare) with respect to selling behavior in the market, although there can be welfare losses in the quantity control system when selling behavior in the market is too irrational (relative to a quantity control system implementing rational selling behavior). Moreover, when the supply of tokens can be adapted from day to day, the permit system was found to be superior in all tested scenarios in which the selling behavior of individuals is rational. Finally, even in the case when toll revenues in the price instrument are equally redistributed (often difficult in practice), tolling in tokens (when tokens are equally distributed) is marginally more equitable in scenarios where congestion effects are more severe. These findings make a case for tolling in tokens.

Keywords: Tolls, tradable mobility permits, congestion, dynamic models, efficiency, equity

JEL codes: R, R48

---

1. Introduction

Congestion is a pervasive problem in most transportation networks worldwide, and the standard approach to address this issue has been to internalize congestion externalities through a toll in dollars (congestion pricing). Extensive reviews may be found in de Palma and Lindsey (2011) and Santos and Verhoef (2011). Pricing, however, has long been beset by issues of inequity, complexity, high infrastructure costs and public and political acceptability, notwithstanding the redistribution of toll revenues, whose benefits may take years to be realized (see also Jaensirisak et al. (2005) and de Palma and Lindsey (2020) for more on the acceptability of road pricing). In contrast, tradable permit schemes (or tolling in tokens; also called tradable mobility credits) are a form of quantity control typically characterized by the following features: (Fan and Jiang, 2013): 1) a fixed total number of tokens (mobility credits) or 'quota' is pre-specified by the regulator, 2) an initial endowment allocates or distributes the tokens to a selected population (all individuals may not receive tokens), 3) individuals are allowed to buy and sell tokens in a market, 4) use of the road network requires tokens and can be differentiated by time of day, geography, vehicle type etc., and 5) enforcement is necessary to ensure valid trading/consumption of tokens.

Tradable permit schemes have several potential advantages over pricing. First, they are revenue neutral and hence, may not be faced with similar public opposition, more so if the tokens are handed out for free. Second, they are viewed as being less vertically inequitable than pricing (de Palma and Lindsey (2020)). Since the number of tokens each user receives may differ, any regressive effect (very well documented for congestion pricing), which may trigger political opposition will not occur with tokens. In other words, lower income users who tend to travel less by car can obtain monetary gains by selling their excess permits. Third, implementation costs may be low given developments in information and communications technology. Finally, they provide the ability to directly control quantity, which may be beneficial in some situations, for example, when the price elasticity of demand in the short or medium term is low. Note that the distribution of tokens allows the regulator to cap the maximum number of travellers. This is a pure version of quantity control, which appears to be more flexible than conventional policies (such as the odd and even licence plate policy, or reduced speed) to reduce environmental costs.

Despite the large body of literature on tradable credit schemes in the transportation literature, relatively little attention has been paid to the comparison of the price and quantity control instruments under uncertainty in a transportation context. In this paper, we consider a tradable permit scheme in a departure time context wherein users are provided an initial endowment of tokens by the regulator and incur a token charge (endogenous and determined prior to all departures) to travel in a specific time period. Tokens can be bought and sold within a marketplace at a price determined by a market clearing mechanism in each time period. A key feature of the market model is that the selling decisions of users are explicitly considered, which enables us to study the impact of selling behavior on performance of the permit system. Travel demand is modeled using a logit mixture model and supply consists of static congestion specific to each time period.

In the case of uncertain demand/supply wherein the tolls (in dollars and tokens) can be adapted from day to day (or alternatively demand/supply are deterministic), the two instruments are equivalent. In contrast, when the tolls are not day to day adaptive, we find that the quantity control instrument is superior in welfare terms when congestion effects are more severe, i.e. steep congestion functions (realistic BPR models), high demand levels and high day-to-day variability. The optimal network usage is relatively similar across states for quantity control whereas the optimal toll in dollar amounts varies significantly across states. Further, non-rational selling behavior, which has the effect of equalizing token supply across time intervals leads to a deterioration in the performance of the quantity instrument, although in general the token system is robust (in efficiency terms) with respect to selling behavior in the market. Moreover, when the token supply can be adapted from day-to-day, the quantity instrument is superior in all scenarios where selling behavior is rational. Finally, even in the case when toll revenues in the price instrument are equally redistributed (often difficult in practice), tolling in tokens (when tokens are equally distributed) is marginally more equitable in scenarios where congestion effects are more severe. These findings make a potential case for tolling in tokens.

The remainder of the paper is organized as follows. Section 2 reviews the existing literature on tradable credits and describes our contributions. Section 3 describes the basic model of supply, demand and equilibrium in the case of deterministic demand (and supply) for both instruments (dollars and tokens). The comparison of the two instruments in this case can be performed analytically. Following this, Section 4 describes the model for stochastic demand for which, the comparisons must be performed numerically. The numerical experiments and findings are described in Section 5. Finally, Section 6 provides concluding remarks and directions for future research.

2. Review of Literature

Although discussions on the use of tradable permits in the transportation sector date back to Raux (2004), Verhoef et al. (1997) and Goddard (1997), they have received significant attention only in the recent past (detailed reviews may be found in Fan and Jiang (2013) and Grant-Muller and Xu (2014)). In particular, a variety of tradable permit schemes have been proposed (utilizing largely network and market equilibrium approaches) in the context of mobility management (at the network level) and bottleneck management, and these are discussed in turn.

In the context of mobility management, Yang and Wang (2011) propose a scheme wherein the social planner initially distributes a certain number of credits to all potential travelers, charges a link-specific number of credits for a given link, and allows trading of the credits among travelers. Supply is modeled using static congestion (separable link performance functions) and travelers are assumed to be homogenous. They demonstrate that for a given set of tolls in tokens (or credit rates) in a general network, the user equilibrium (UE) link flow pattern is unique under standard assumptions, and the credit price at the market equilibrium is unique under some relatively mild additional assumptions (i.e if all equilibrium path flow patterns contain at least two paths with different credit charges connecting the same O–D pair). The proposed network equilibrium formulation is a variant of the standard UE model with the additional network-wide credit feasibility constraint, which simply states that the total consumption of tokens at equilibrium is less than or equal to the total credit endowment. Extensions that consider heterogeneity in the value of time and multiple user classes are proposed in Zhu et al. (2015) and Wang et al. (2012), whereas He et al. (2013) consider allocations of credits to not just individual travelers, but to transportation firms such as logistics companies and transit agencies. In a similar vein, Pareto-improving credit based congestion management schemes on a general two mode network are investigated in Liu and Nie (2017).

Nie (2012) examines the effect of transaction costs in a tradable permit scheme for two types of markets: an auction market in which users purchase all of the needed mobility credits through a competitive bidding process, and a negotiated market in which users initially receive certain amount of mobility credits from the government and trade with each other through negotiation to meet their needs. A brokerage service is built into both markets to facilitate transactions and accordingly, the users have to pay a commission fee proportional to the value of trade. The modified UE formulation of Yang and Wang (2011) is extended to incorporate transaction costs for both the auction and negotiated markets. Based on numerical experiments on a toy network, it is shown that an auction market can achieve an equilibrium allocation of mobility credits if the government sets the price suitably and the unit transaction cost is lower than the price that the market would reach in the absence of transaction costs. The work also highlights the fact that in case of the negotiated market, the initial allocation of mobility credits may affect the final equilibrium even when marginal transaction costs are constant.

A related stream of research on the use of tradable credits schemes for mobility management examines the design of these schemes using bi-level optimization formulations (mathematical programming problems with equilibrium constraints). For instance, Wang et al. (2014b,a) formulate the continuous network design problem with a tradeable credit scheme as a bi-level programming problem, where the decision variables for the upper level problem are capacity enhancements for selected links whereas the lower level problem determines equilibrium link flows and the credit price. Along similar lines Wu et al. (2012) proposed a framework that considers decisions of mode/route choice and trip generation on a multimodal transportation network to design efficient and equitable congestion pricing and tradable credit schemes (considering a measure including both net social benefit and equity). They find that the Pareto frontier (with respect to the two aforementioned objectives) of the credit scheme strictly dominates that of congestion pricing although the two schemes achieve the same level of maximum net benefits. Further, their results suggest that tradable credit schemes can be progressive whereas congestion pricing schemes are largely regressive for the tested network. Finally, the literature on mobility management also includes a series of studies on credit-based congestion pricing (CBCP), where credits in CBCPs are allowances used to pay tolls (Kalmanje and Kockelman, 2004; Kockelman and Kalmanje, 2005). The studies involved the use of destination, mode, and departure time choice models to examine the potential impacts of using a CBCP scheme.

Researchers have also examined the use of tradeable mobility credits to manage bottleneck congestion and achieve peak spreading in an efficient manner. Nie and Yin (2013) developed an analytical framework to model a tradable credit scheme that manages commuters' travel choices in a simple transportation system consisting of two parallel routes. The scheme attempts to persuade commuters to spread their departure times evenly within the rush hour and between primary and alternative routes to mitigate traffic congestion. It defines a 'peak-time' window within which users are charged mobility credits to travel on the primary route and those that avoid either the peak-time window or the primary route may be rewarded with credits (see also Nie (2015)). Tian et al. (2013) investigate the efficiency of a tradable travel credit scheme for managing bottleneck congestion and modal split in a competitive highway/transit network with a continuously distributed value of time. They propose a tradable credit scheme which emulates the bottleneck congestion pricing and transit subsidy in a revenue-neutral manner and demonstrate that both the modal split and credit charge at equilibrium are unique.

Xiao et al. (2013) examined the efficiency of a tradable credit system in managing morning commute congestion with both homogenous and heterogeneous users. Similar to other studies in this stream, the TMC system consists of credits that are universal with regards to time, a time-varying credit is charged at the bottleneck; the credits can be traded and the price is determined by a competitive market. They show that even in the presence of heterogeneity, an optimal credit scheme that eliminates the bottleneck queue always exists under the assumption that late arrival is not allowed. More recently, Akamatsu and Wada (2017) proposed a tradable bottleneck credit scheme where the regulator issues link- and time-specific credits permitting passage through a certain link or bottleneck in a pre-specified time period. They develop a model to describe time-dependent flow patterns at equilibrium under a system of tradable bottleneck permits for general networks and show that the equilibrium obtained under this system is efficient in that it minimizes the social transportation cost. Bao et al. (2019) show that the equilibrium with a tradable credit scheme may not be unique for particular models of traffic congestion, including the first-best solution for the conventional Vickrey's bottleneck model. Finally, Brands et al. (2020) conduct an interesting lab-in-the-field experiment of tradable credit schemes with virtual mobility behavior and real financial incentives. They adopt a market design, which lets users trade with a price setting intermediary, termed a virtual bank. An incremental price adjustment scheme is adopted and their experiments suggest that it ensures that the price stays largely within the equilibrium range. Overall, their results are promising and indicate that tradable permits can be a viable alternative to pricing in a parking setting.

In contrast, comparisons of price and quantity control under uncertainty in a transportation context are relatively sparse (for other contexts see Weitzman (1974); Laffont (1977)). Note that in the emissions context, agents are not directly impacted by the externality they generate, while in the transportation context they are (as far as congestion is concerned). For this reason, their results on stochastic demand and stochastic supply are not directly applicable here, even if they are useful as a general guideline. Shirmohammadi et al. (2013) examine the performance of tradable permit systems under demand and supply uncertainty using a toy network. Specifically, they compare the performance of a link differentiated toll system (in dollars) and a mobility credit system that is differentiated by link. However, their analysis does not focus on measures of efficiency but rather examines performance relative to a given target volume of cars. They find strong variations in the permit prices are required to ensure demand matches the specified volume targets.

de Palma et al. (2018) compare the performance (in efficiency terms) of congestion pricing and tradable mobility credit schemes under uncertainty using a simple road network in a stochastic route choice setting (including a public transit alternative). They find that when the tolls (in either dollars or tokens) cannot be adapted from day to day, the credit scheme performs better typically when the slope of the congestion function is steep. Further, when the token supply can be adapted from day to day, the token system always outperforms congestion pricing.

More recently, de Palma and Lindsey (2020) ranked the efficiency of permits and tolls for one route, one time period and elastic demand. They consider additive and multiplicative demand and cost (capacity) shocks and show they may lead to qualitatively different results. Their approach is analytical. They consider linear and non-linear demand (and show the role of the convexity of demand). They also study the impact of the correlation between demand and supply on the ranking. Rezaeinia et al. (2021) consider the comparison of tolls and permits in a radial network, within the Vickery frameswork. User select mode, departure time and route based on a Nested Logit continuous/discrete framework, and congestion is computed using the METROPOLIS software. Tolls and permits are independent of the time of the day. They found that tolls outperform permits if capacities shocks are perfectly correlated. However, the ranking is reverse if capacities shocks are independent.

This paper is an extension of de Palma et al. (2018), who consider static congestion (with parallel routes) and one time period (i.e. it is a pure static model). Here instead, users decide when to travel within a multi-period setting, and a market clearing mechanism exists in every time period. Thus, we consider within-day dynamics and multiple time intervals. Further, a key feature of the market model is that selling decisions of users are now explicitly considered (this to the best of our knowledge has not yet been addressed in the literature), which enables us to study the impact of selling behavior on performance of the permit system. Compared to de Palma and Lindsey (2020) (they consider a single route with elastic demand), our model is more complex and hence, comparisons of the two instruments cannot be performed analytically.

3. Multiperiod Model: Deterministic Demand

The transportation network of interest consists of a single origin-destination pair connected by a single route (this can be extended to multiple OD pairs and routes with no methodological difficulty). We consider commuting trips performed within a time period T, which is partitioned into three sub-periods T_1, T_2, T_3 (early morning, peak, off-peak) so that $T_1 \bigcup T_2 \bigcup T_3 = T$ (this may be easily extended to model a larger number of time periods). Note that T_1, T_2, T_3 may represent any three time periods within a day, and are not necessarily contiguous. There are a total of N users who wish to travel, and each user performs a single trip or activity during the day in any one of the three intervals or chooses to stay at home (denoted by T_0). A glossary of notation can be found in Appendix A.

We assume that the network is subject to time dependent congestion justifying the need for congestion control in the form of either a price instrument (tolling in dollars) or a quantity instrument (tolling in tokens). Under the price control instrument, users have to pay a toll in dollars τ_i to travel in time period T_i ($i = 1 \ldots 3$). In case of the quantity control system, the regulator distributes a certain number of permits (tokens) M (known in advance) to each potential user at the beginning of the time period T. The tokens expire at the end of time period T, or in other words their market value is zero at the end of time period T. Tokens cannot be banked or traded across days (i.e. across periods T). Users are required to spend a certain number of tokens to travel in time interval $T_i, i = 1 \ldots 3$, given by δ_i (toll in tokens). Further, tokens can be bought and sold within a marketplace. The price of the token in time interval T_i is denoted p_i, and is determined endogenously by the demand and supply of tokens in the market in time interval T_i. To improve acceptability, we assume that all transactions take place at the beginning of the period T. Further, note that the regulator has the flexibility to institute any desired token allocation scheme including ones wherein users receive an unequal number of tokens. Given that user choices are unaffected by the token allocation (ignoring income effects), this implies that in principle any desired distribution of equity can be achieved through the initial token allocation.

In this section, we consider the case of deterministic demand (i.e. the number of users N is deterministic and known). The transportation model is first described followed by the two instruments (price and quantity) in turn and a comparison with respect to individual benefits, social benefits, and equity.

3.1. Transportation Model: Demand, Supply and Equilibrium

The money-metric utility of an individual n to travel in a time period $T_i, i = 0 \ldots 3$ ($i = 0$ denotes the stay at home option) is given by,

$$U^n(T_0) = B_0^n + \mu_n \epsilon_0, \tag{1}$$

$$U^n(T_i) = B_i^n - \alpha^n t_i(X_i) - p_i \delta_i + \mu_n \epsilon_i, \;\; i = 1 \ldots 3, \text{ (quantity control)}$$

$$U^n(T_i) = B_i^n - \alpha^n t_i(X_i) - \tau_i + \mu_n \epsilon_i, \;\; i = 1 \ldots 3, \text{ (price control)}$$

where X_i is the flow in time period T_i, ϵ_i is an i.i.d. Gumbel disturbance term and μ_n, B_i^n, α^n are individual specific parameters with $\log(\mu_n), B_i^n, \log(\alpha^n)$ normally distributed. The alternative specific benefit B_i^n incorporates time-period specific scheduling preferences, or alternatively, time-period specific schedule delay costs. α^n and μ_n are the value of time and scale parameter, respectively, for individual n.

A standard BPR type function is assumed to model the travel time $t_i(X_i)$ in time period i,

$$t_i(X_i) = t_i^{FF} \left(1 + \alpha_i \left(X_i/C_i\right)^{\beta_i}\right), \tag{2}$$

where t_i^{FF} is the free flow time in period i, C_i is a capacity associated with time period i and α_i, β_i are function parameters. Note that we assume that the three time periods are independent in the sense that congestion does not spillover from one period to the next (in other words we have three static models).

Further, let $\boldsymbol{\theta_n} = (\mu_n, B_i^n, \alpha^n)$ denote the vector of parameters for individual n. The probability of individual n choosing to travel in time period T_i is given by,

$$p(T_i|\boldsymbol{\theta_n}) = \frac{\exp[V^n(T_i, \boldsymbol{\theta_n})]}{\sum_{j=0\ldots3}\exp[V^n(T_j, \boldsymbol{\theta_n})]}, \tag{3}$$

where $V^n(T_i, \boldsymbol{\theta_n})$ is given by,

$$V^n(T_0, \boldsymbol{\theta_n}) = (1/\mu_n)\,(B_0^n)\,, \tag{4}$$

$$V^n(T_i, \boldsymbol{\theta_n}) = (1/\mu_n)\,(B_i^n - \alpha^n t_i(X_i) - p_i\delta_i)\,,\quad i = 1\ldots3,\ \text{(quantity)}$$

$$V^n(T_i, \boldsymbol{\theta_n}) = (1/\mu_n)\,(B_i^n - \alpha^n t_i(X_i) - \tau_i)\,,\quad i = 1\ldots3,\ \text{(price)}$$

The number of travellers travelling in the three periods X_i $(i = 1\ldots3)$, given τ_i $(i = 1\ldots3)$ in the price system and $p_i\delta_i$ $(i = 1\ldots3)$ in the quantity system, are obtained by solving the fixed point problem (note that $X_0 = N - \sum_{i=1}^{i=3} X_i$),

$$X_i = \sum_{n=1}^{n=N} \frac{\exp[V^n(T_i, \boldsymbol{\theta_n})]}{\sum_j \exp[V^n(T_j, \boldsymbol{\theta_n})]}, i = 1\ldots3. \tag{5}$$

Since the set of demand feasible flows $\mathcal{D} = (\boldsymbol{X} : \sum_{i=0\ldots3} X_i = N)$ forms a closed and convex set, and the right hand side of Equation 5 is a continuous function of flows, Brouwer's fixed point theorem implies that a solution exists to the fixed point problem in Equation 5.

3.2. Tolls in Dollars: Price Control

In the price control system, the regulator is assumed to have knowledge of the demand N and sets the tolls in dollars τ_i $(i = 1\ldots3)$ to maximize total welfare (defined as the sum of consumer surplus and regulator revenue), formulated as the following optimization problem,

$$\underset{\tau_1,\tau_2,\tau_3}{\text{Max}}\quad \Omega_p = \sum_{n=1}^{N} \mu_n \log\left(\sum_{j=0\ldots3}\exp[V^n(T_j, \boldsymbol{\theta_n})]\right) + \sum_{j=1\ldots3}\tau_j X_j \tag{6}$$

s.t

$$X_i = \sum_{n=1}^{N} \frac{\exp[V^n(T_i, \boldsymbol{\theta_n})]}{\sum_j \exp[V^n(T_j, \boldsymbol{\theta_n})]}, i = 1\ldots3,$$

$$\sum_{i=0\ldots3} X_i = N,\quad X_i \geq 0, i = 0\ldots3.$$

The optimum welfare obtained by solving 6 above and the corresponding optimum welfare and tolls in dollars are denoted by Ω_p^* and $\boldsymbol{\tau^*} = (\tau_i^*, i = 1\ldots3)$ respectively.

3.3. Tolls in Tokens: Quantity Control

Recall that in the quantity control system, the regulator distributes a certain number of permits (tokens) M to each potential user at the beginning of a time period time period T, and the number of tokens required to travel in time interval $T_i, i = 1 \dots 3$ is given by δ_i.

3.3.1. Demand for Tokens

For a given set of tolls in tokens and market prices δ_i, p_i $(i = 1 \dots 3)$, the total demand for tokens in interval T_i is $X_i \delta_i$ and the total amount of tokens possessed by people travelling in period T_i is $X_i M$, where M is the initial token endowment. Thus, the demand for tokens in interval T_i $(i = 1 \dots 3)$, is given by,

$$D_i = X_i \text{Max}(0, (\delta_i - M)). \tag{7}$$

where X_i $(i = 1 \dots 3)$ are obtained from the solution to Equation 5. Note that in the above we assume that the token endowment to each user is equal (M). This assumption can be relaxed.

3.3.2. Supply of Tokens

We assume that the decision to sell tokens is made after the mobility decision and hence, users sell all unused tokens. Note also that it is assumed that there is some scarcity in the system, namely that the number of tokens available is less than what would be consumed if the tokens were free. Formally, if $\bar{X}_i$ $(i = 0 \dots 3)$ denotes the equilibrium flows in the absence of tolls, we assume that $\sum_{i=1 \dots 3} \bar{X}_i \delta_i > MN$.

Lemma: In the case that there is no congestion (i.e t_i is constant), we have $\delta_i = 0$ $(i = 1 \dots 3)$, or equivalently, $\delta_i < M$ $(i = 1 \dots 3)$, and hence, the price of tokens p_i $(i = 1 \dots 3)$ is zero and the tokens have no effect/value.

We focus on the case where congestion effects are present, and hence, it should be the case that atleast one of $\delta_1, \delta_3, \delta_3$ is larger than M (note that $\delta_0 = 0$), which implies that $p_i > 0$, for at least one interval T_i. Let the subset of time periods where $\delta_i > M$ be denoted by $\tilde{I}$, then the subset of time periods where $\delta_i < M$ is $I \setminus \tilde{I}$.

In order to avoid speculation, we assume that tokens can only bought if they are needed for travel, and can only be used for travelling in the chosen time interval. Further, we assume that the regulator requires the user to sell all her unused tokens in one time interval. This assumption is made from a practical standpoint to ensure simplicity of the system. Since the tokens are worthless at the end of the last time interval, no user will keep tokens unused.

Note that sellers are individuals who have chosen to travel in an interval $T_j, j \in I \setminus \tilde{I}$ whereas buyers are individuals who have chosen to travel in $T_i, i \in \tilde{I}$. We assume that the token price in interval $T_i, i \in \tilde{I}$ perceived by a seller is given by,

$$\tilde{p}_i = p_i + \epsilon_i, \tag{8}$$

where p_i is the true market price of the token and ϵ_i is an i.i.d. Gumbel error term with scale parameter $\bar{\mu}$. The rationale for the additive error term is that since all transactions happen at the beginning of the time period T (i.e. at the beginning of the day), travelers may have imperfect forecasts or perceptions of the prevailing market price.

For a seller, we model the choice of a selling interval using a simple logit model based on Equation 8, where the systematic utility of selling in interval $T_i, i \in \tilde{I}$ is the true market price in the interval. Consider a user n who has decided to travel in period $T_j, j \in I \setminus \tilde{I}$. We assume that the amount of tokens to sell does not influence the time period of selling. The probability of the user n selling his/her token in interval $T_i, i \in \tilde{I}$, Q_i is assumed to be,

$$Q_i = \text{Prob}[\tilde{p}_i > \max_{j \neq i; j \in \tilde{I}} \tilde{p}_j] = \frac{\exp(p_i/\bar{\mu})}{\sum_{j \in \tilde{I}} \exp(p_j/\bar{\mu})}. \tag{9}$$

The supply of tokens from users travelling in $T_j, j \in I \setminus \tilde{I}$ who have decided to sell in period $T_i, i \in \tilde{I}$ is given by,

$$S_i = \sum_{j \in I \setminus \tilde{I}} Q_i X_j (M - \delta_j) + Q_i X_0 M.$$

Or,

$$S_i = Q_i \left(\sum_{j \in I \setminus \tilde{I}} X_j (M - \delta_j) + X_0 M \right) \quad i \in \tilde{I}. \tag{10}$$

The total supply of tokens is given by,

$$\begin{aligned} S = \sum_{i \in \tilde{I}} S_i &= \sum_{i \in \tilde{I}} Q_i \left(\sum_{j \in I \setminus \tilde{I}} X_j (M - \delta_j) + X_0 M \right) \\ &= \sum_{j \in I \setminus \tilde{I}} X_j (M - \delta_j) + X_0 M. \end{aligned} \tag{11}$$

We next study the market clearing condition.

3.3.3. Market Clearing

The market clearing conditions in each interval $T_i, i \in \tilde{I}$ imply ,

$$S_i = D_i, \forall i \in \tilde{I}$$

$$\Rightarrow Q_i \left(\sum_{j \in I \setminus \tilde{I}} X_j \left(M - \delta_j \right) + X_0 M \right) = X_i \left(\delta_i - M \right), \forall i \in \tilde{I}. \tag{12}$$

Further, the total demand for tokens is given by,

$$D = \sum_{i \in \tilde{I}} D_i = \sum_{i \in \tilde{I}} X_i \left(\delta_i - M \right). \tag{13}$$

Thus, market clearing requires $D = S$, or,

$$\sum_{i \in \tilde{I}} X_i \left(\delta_i - M \right) = \sum_{j \in I \setminus \tilde{I}} X_j \left(M - \delta_j \right) + X_0 M$$

$$\implies \sum_i X_i \delta_i = M \left(\sum_i X_i + X_0 \right) = MN. \tag{14}$$

The market clearing conditions in all periods is satisfied, but the demand for tokens and supply of tokens are not necessarily the same in each time period. The price adjustment of tokens will guarantee that each market will clear.

3.3.4. Price Adjustment

For a given vector of tolls in tokens ($\delta_i, i = 1 \ldots 3$), the market clearing price can be computed through the following iterative process, where the price in iteration $w + 1$ is given by,

$$p_i^{w+1} = p_i^w + h(D_i - S_i), i = 1 \ldots 3, \tag{15}$$

where $h'(.) > 0$. This is a standard cobweb adjustment process. It should be noted that the price adjustment process above is merely a numerical method to compute the market clearing prices and does not imply that the market actually operates in this manner. Otherwise, users could discover this and make use of the knowledge of this process strategically, which will become a complex game theoretic problem outside the scope of this paper (and possibly behaviorally unrealistic in any case).

Conjecture: The equilibrium prices for the three time intervals satisfy $lim_{\bar{\mu} \to 0}\, p_i = p^*, i = 1 \ldots 3$ for any $\delta_1, \delta_2, \delta_3$.

Given this result, when we optimize $\delta_i, i = 1 \ldots 3$, the adjustment process has reached a stationary state so that $p_1^* = p_2^* = p_3^*$, when $\bar{\mu} \to 0$.

3.3.5. Optimization

As before, let $\tilde{I}$ denote the subset of time periods where $\delta_i > M$. We assume that the regulator has knowledge of the demand N and sets the tolls in tokens for time periods 1 and 3, and the supply of tokens M (δ_2 is normalized to 1 token without loss of generality) that maximizes total welfare, formulated as the following optimization problem,

$$\underset{\delta_1,\delta_3,M}{\text{Max}} \;\; \Omega_q = \sum_{n=1}^{N} \mu_n \log\left(\sum_{j=0\ldots3} \exp[V^n(T_j, \boldsymbol{\theta_n})] \right) + \sum_{j=1\ldots3} p_j \delta_j X_j \tag{16}$$

s.t

$$X_i = \sum_{n=1}^{N} \frac{\exp[V^n(T_i, \boldsymbol{\theta_n})]}{\sum_j \exp[V^n(T_j, \boldsymbol{\theta_n})]}, i = 1\ldots3,$$

$$\sum_{i=0\ldots3} X_i = N; \quad X_i \geq 0, i = 0\ldots3,$$

where the equilibrium prices $\mathbf{p} = (p_1, p_2, p_3)$ satisfy the market equilibrium conditions (and Q_i is given by Equation 9),

$$Q_i \left(\sum_{j\in I\setminus\tilde{I}} X_j \left(M - \delta_j\right) + X_0 M \right) = X_i \left(\delta_i - M\right), \forall i \in \tilde{I}.$$

The optimum welfare obtained by solving 16 above and the corresponding optimal tolls in tokens are denoted by Ω_q^* and $\boldsymbol{\delta^*} = (\delta_i^*, i = 1\ldots3)$ respectively. The associated market clearing prices are denoted by $p_i^*, i = 1\ldots3$.

3.4. Comparison

In the deterministic case, the comparison of the two instruments is trivial and can be performed analytically. The two instruments, when optimally chosen, yield identical social welfare. This is shown in proposition 1 below.

Proposition 1. *Under deterministic demand and supply, the two instruments, price and quantity, when optimally chosen, are equivalent.*

Proof. Let $\Omega_p\left(\boldsymbol{\tau^*}\right)$ and $\Omega_q\left(\boldsymbol{\delta^*, p^*}\right)$ denote the optimum welfare attained by the price and quantity instruments respectively, where $\boldsymbol{\tau}^*$ is the optimum vector of tolls (assume $\boldsymbol{\tau}^* > \mathbf{0}$ without loss of generality), $\boldsymbol{\delta}^*$ is a vector of optimum number of tokens required for each time interval, and $\boldsymbol{p}^*$ is the vector of market clearing prices. Note that for simplicity (w.l.o.g), we do not adopt the normalization of $\delta_2 = 1$ and instead assume that M is fixed arbitrarily, and the regulator optimizes $\boldsymbol{\delta}$. Further, let the optimum flows obtained under the price instrument be denoted by $\boldsymbol{X^p} = (X_i^p, i = 0\ldots3)$. We wish to show that $\Omega_p\left(\boldsymbol{\tau^*}\right) = \Omega_q\left(\boldsymbol{\delta^*, p^*}\right)$.

Since $\boldsymbol{\tau}^*$ is the optimum toll vector, we have

$$\Omega\left(\boldsymbol{\tau}^*\right) > \Omega\left(\boldsymbol{\tau}\right) \;\; \forall \boldsymbol{\tau} \neq \boldsymbol{\tau}^*. \tag{17}$$

First, assume that $\boldsymbol{p^*}$ is given exogenously. Clearly, if we set $\delta_i^* = \tau_i^*/p_i^*, i = 1\dots3$, the flows under the quantity instrument satisfy $\boldsymbol{X^q(\delta^*, p^*)} = \boldsymbol{X^p}$, and hence, $\Omega_q\left(\boldsymbol{\delta^*, p^*}\right) = \Omega_p\left(\boldsymbol{\tau^*}\right)$. The prices p^* can be determined from the market clearing conditions in Equation 12, which imply that (note that since $\boldsymbol{\tau}^* > \mathbf{0}$, it must be the case that $\delta_i > M, i = 1\dots3$ and hence, $\boldsymbol{p^*} > \mathbf{0}$),

$$Q_i X_0^p M = X_i^p \left(\delta_i - M\right), \forall i \in I. \tag{18}$$

Substituting $\delta_i^* = \tau_i^*/p_i^*, i = 1\dots3$, we obtain the market clearing prices,

$$p_i^* = \frac{\tau_i^*}{M\left(1 + \dfrac{Q_i X_0^p}{X_i^p}\right)}, \forall i \in I. \tag{19}$$

Note that M can be set arbitrarily, the prices will adjust accordingly. Thus, to summarize, by setting $\delta_i^* = \tau_i^*/p_i^*, i = 1\dots3$, where p_i^* is given by Equation 19, we have $\Omega_q\left(\boldsymbol{\delta^*, p^*}\right) = \Omega_p\left(\boldsymbol{\tau^*}\right)$.

Now, assume that there exists $\boldsymbol{\delta} = \boldsymbol{\delta}'$ and $\boldsymbol{p} = \tilde{\boldsymbol{p}}$ such that $\Omega_q\left(\boldsymbol{\delta}', \tilde{\boldsymbol{p}}\right) > \Omega_p\left(\boldsymbol{\tau^*}\right)$. Then, setting $\tau_i' = \tilde{p}_i \delta_i'$ $(i = 1\dots3)$, we have $\Omega_p\left(\boldsymbol{\tau}'\right) > \Omega_p\left(\boldsymbol{\tau^*}\right)$. This contradicts Equation 17 and hence, the result follows. ■

4. Multiperiod Model: Stochastic Demand

In the discussion thus far, transportation demand and supply were assumed to be deterministic. We now turn our attention to the case where the demand is stochastic.

4.1. Transportation Model: Demand, Supply and Equilibrium

Assume that there exist two days or states of nature $s1$ and $s2$ (denoted by $sk, k = 1, 2$), where the alternative specific benefit to travel for individual n, varies across the days, taking values $B_i^{n,s1}$ for period i $(i = 1, 2, 3)$ with probability q and values $B_i^{n,s2}, i = 1, 2, 3$ with probability $1 - q$. In other words, the source of day-to-day variability or stochasticity is on the demand side and arises due to fluctuations in the scheduling preferences of travelers (note that the total number of users is fixed). The variability in scheduling preferences may be due to special events, weather etc. leading to a higher number of users who wish to travel during the peak period. Stochasticity may also arise from external factors affecting supply such as incidents and accidents, or factors affecting both demand and supply. The methodological framework can be extended in a straightforward manner to model these cases as well.

The systematic utilities to travel in time period T_i on day $sk, k = 1, 2$ (denoted by T_i^{sk}) are given by,

$$V_{sk}^n(T_0^{sk}, \boldsymbol{\theta_n}) = \left(1/\mu_n\right)\left(B_0^n\right), k = 1, 2 \tag{20}$$

$$V^n_{sk}(T^{sk}_i, \boldsymbol{\theta_n}) = (1/\mu_n)\left(B^{n,sk}_i - \alpha^n t_i(X^{sk}_i) - p^{sk}_i\delta_i\right), \;\; i = 1\ldots3, k = 1,2 \text{ (quantity)}$$

$$V^n_{sk}(T^{sk}_i, \boldsymbol{\theta_n}) = (1/\mu_n)\left(B^{n,sk}_i - \alpha^n t_i(X^{sk}_i) - \tau_i\right), \;\; i = 1\ldots3, k = 1,2 \text{ (price)},$$

where p^{sk}_i and X^{sk}_i are the token price and number of individuals travelling in time period T^{sk}_i respectively. Note that in case of the price system, the terms $p^{sk}_i\delta_i$ are replaced by τ_i. As before, for a given set of tolls in tokens ($\delta_i, i = 1...3$) and token prices ($p^{sk}_i, i = 1...3; k = 1,2$) — or tolls in dollars ($\tau_i, i = 1...3$) in the case of price control—, X^{sk}_i for $k = 1,2$ can be determined by solving the following fixed point problem (note that $X^{sk}_0 = N - \sum_{i=1}^{i=3} X^{sk}_i$),

$$X^{sk}_i = \sum_{n=1}^{n=N} \frac{\exp[V^n_{sk}(T^{sk}_i, \boldsymbol{\theta_n})]}{\sum_j \exp[V^n_{sk}(T^{sk}_j, \boldsymbol{\theta_n})]}, i = 1\ldots3. \tag{21}$$

4.2. Tolls in Dollars: Price Control

In case of the price instrument, we assume that the regulator may not wish to change the tolls from day to day for reasons of acceptability and ease of implementation (or may not have knowledge of the specific realization of the state of nature). For instance, in the ERP system of Singapore, tolls are revised only once every few months and do not vary from day to day. Thus, in the case of stochastic demand, the regulator sets the tolls in tokens for the three time periods that maximizes *expected* total welfare, formulated as the following optimization problem,

$$\begin{aligned}\underset{\tau_1,\tau_2,\tau_3}{\text{Max}} \;\; & q\left\{\sum_{n=1}^{N}\mu_n \log\left(\sum_{j=0\ldots3}\exp[V^n_{s1}(T^{s1}_j, \boldsymbol{\theta_n})]\right) + \sum_{j=1\ldots3}\tau_j X^{s1}_j\right\}\\ & +(1-q)\left\{\sum_{n=1}^{N}\mu_n \log\left(\sum_{j=0\ldots3}\exp[V^n_{s2}(T^{s2}_j, \boldsymbol{\theta_n})]\right) + \sum_{j=1\ldots3}\tau_j X^{s2}_j\right\}\end{aligned} \tag{22}$$

s.t

$$X^{sk}_i = \sum_{n=1}^{n=N} \frac{\exp[V^n_{sk}(T^{sk}_i, \boldsymbol{\theta_n})]}{\sum_j \exp[V^n_{sk}(T^{sk}_j, \boldsymbol{\theta_n})]}, i = 1\ldots3; k = 1,2$$

$$\sum_{i=0\ldots3} X^{sk}_i = N, k = 1,2; \quad X^{sk}_i \geq 0, i = 0\ldots3, k = 1,2.$$

4.3. Tolls in Tokens: Quantity Control

We distinguish two configurations of the quantity control system. First, in the case of adaptive token supply, the supply of tokens can vary by day and is denoted M^{s1} and M^{s2}, whereas in the case of fixed token supply, it is assumed that the total supply of tokens is fixed across days i.e. $M^{s1} = M^{s2} = M$. From the standpoint of implementation, adapting the token supply is likely to be far easier than adapting the tolls in tokens (or dollars), which may involve communicating a complex tariff structure (in a general network) to commuters.

The market clearing conditions in Equation 12 now apply in each time interval for both days and are given by (the same notation as before is used with the added superscript or subscript sk to denote the day),

$$Q_i^{sk}\left(\sum_{j\in I\setminus\tilde{I}^{sk}} X_j^{sk}\left(M^{sk}-\delta_j\right)+X_0^{sk}M^{sk}\right)=X_i^{sk}\left(\delta_i-M^{sk}\right),\forall i\in\tilde{I}^{sk},k=1,2. \tag{23}$$

In the case of fixed token supply, we assume that the regulator does not have knowledge of the specific realization of the state of nature (or day) and hence, sets the tolls in tokens for time periods 1 and 3, and the supply of tokens M (as before δ_2 is normalized to 1 token without loss of generality) that maximizes *expected* total welfare, formulated as the following optimization problem,

$$\begin{aligned}\underset{\delta_1,\delta_3,M}{\text{Max}}\ & q\left\{\sum_{n=1}^{N}\mu_n\log\left(\sum_j\exp[V_{s1}^n(T_j^{s1},\boldsymbol{\theta_n})]\right)+\sum_j p_j^{s1}\delta_j X_j^{s1}\right\}\\ &+(1-q)\left\{\sum_{n=1}^{N}\mu_n\log\left(\sum_j\exp[V_{s2}^n(T_j^{s2},\boldsymbol{\theta_n})]\right)+\sum_j p_j^{s2}\delta_j X_j^{s2}\right\}\end{aligned} \tag{24}$$

s.t

$$X_i^{sk}=\sum_{n=1}^{n=N}\frac{\exp[V_{sk}^n(T_i^{sk},\boldsymbol{\theta_n})]}{\sum_j\exp[V_{sk}^n(T_j^{sk},\boldsymbol{\theta_n})]},i=1\ldots3;k=1,2$$

$$\sum_{i=0\ldots3}X_i^{sk}=N,k=1,2;\quad X_i^{sk}\geq 0,i=0\ldots3,k=1,2,$$

where the equilibrium prices $\mathbf{p}^{sk}=(p_1^{sk},p_2^{sk},p_3^{sk}),k=1,2$ satisfy the market equilibrium conditions:

$$Q_i^{sk}\left(\sum_{j\in I\setminus\tilde{I}^{sk}} X_j^{sk}\left(M-\delta_j\right)+X_0^{sk}M\right)=X_i^{sk}\left(\delta_i-M\right),\forall i\in\tilde{I}^{sk},k=1,2.$$

In the case of adaptive token supply, we assume that the regulator has knowledge of the specific realization of the state of nature (or day) and sets the tolls in tokens for time periods 1 and 3, and the supply of tokens M^{s1},M^{s2} (as before δ_2 is normalized to 1 token without loss of generality) that maximizes *expected* total welfare, formulated as the following optimization problem,

$$\begin{aligned}\underset{\delta_1,\delta_3,M^{s1},M^{s2}}{\text{Max}}\ & q\left\{\sum_{n=1}^{N}\mu_n\log\left(\sum_j\exp[V_{s1}^n(T_j^{s1},\boldsymbol{\theta_n})]\right)+\sum_j p_j^{s1}\delta_j X_j^{s1}\right\}\\ &+(1-q)\left\{\sum_{n=1}^{N}\mu_n\log\left(\sum_j\exp[V_{s2}^n(T_j^{s2},\boldsymbol{\theta_n})]\right)+\sum_j p_j^{s2}\delta_j X_j^{s2}\right\}\end{aligned} \tag{25}$$

s.t

$$X_i^{sk}=\sum_{n=1}^{n=N}\frac{\exp[V_{sk}^n(T_i^{sk},\boldsymbol{\theta_n})]}{\sum_j\exp[V_{sk}^n(T_j^{sk},\boldsymbol{\theta_n})]},i=1\ldots3;k=1,2$$

$$\sum_{i=0\ldots 3} X_i^{sk} = N, k = 1, 2; \quad X_i^{sk} \geq 0, i = 0 \ldots 3, k = 1, 2.$$

where the equilibrium prices $\mathbf{p}^{sk} = (p_1^{sk}, p_2^{sk}, p_3^{sk}), k = 1, 2$ satisfy the market equilibrium conditions,

$$\Rightarrow Q_i^{sk} \left(\sum_{j \in I \setminus \tilde{I}^{sk}} X_j^{sk} \left(M^{sk} - \delta_j \right) + X_0^{sk} M^{sk} \right) = X_i^{sk} \left(\delta_i - M^{sk} \right), \forall i \in \tilde{I}^{sk}, k = 1, 2.$$

4.4. Comparison

In contrast with the deterministic case, when demand (or supply) is stochastic, the comparison of the price and quantity control instruments cannot be performed analytically. Hence, we perform the comparison numerically.

5. Numerical Experiments: Stochastic Demand

5.1. Experimental Design

The two instruments are compared using a synthetic example across a wide range of demand and supply inputs. The setting considered is identical to the formulation in Section 4, wherein there are N potential travelers who may choose to either travel in one of three time periods ($T_i, i = 1 \ldots 3$) or cancel their trip (option T_0). The stochasticity or variability in demand –as noted in Section 4– is modeled by varying the levels of the alternative specific benefit to travel (in time periods) on the two days, $B_i^{n,sk}, i = 1 \ldots 3, k = 1, 2$ (refer Equation 20). Thus, the source of day-to-day variability or stochasticity is on the demand side and arises due to fluctuations in the scheduling preferences of travelers, which may arise due to special events, weather etc. leading to a higher number of users who wish to travel. The mean and standard deviation of the alternative specific benefit to cancel trip ($B_0^{n,sk} k = 1, 2$) are normalized to zero, and the probability q is assumed to be 0.5.

Table 1: Fixed Factors

Parameter	Time Period		
	T_1	T_2	T_3
$B_i^{n,s1}$ [Mean] (\$)	7.5	10.0	7.5
$B_i^{n,s1}$ [SD] (\$)	0.25	0.25	0.25
Free flow time (min)	13	13	13
Capacity (vehicles/time period)	350	350	350
α (BPR parameter)	0.175	0.15	0.2

The values of the fixed factors are shown in Table 1. The capacities are set based on the range of demand values (varies with scenario, see Table 2) to yield a ratio of congested to free flow travel time (in the absence of tolls) in the range $1.25 - 2.5$. The free flow travel time is set to be 13 minutes (assuming a free flow speed of 60 km/hr, this corresponds to a trip length of 13 km, which is in the range of average trip lengths in typical urban transportation networks). The mean of the alternative specific benefit is assumed to be higher in period two to represent peaking effects and commute behavior (note that the table describes the distribution of the alternative specific benefits in the different time periods for day $s1$; the values on day $s2$ vary with the scenario, and are part of the experimental design, which is described later in the section). The coefficient of variation of the alternative specific benefit is assumed to be lower in period T_2 reflecting a morning commute context where work start times are largely in this time interval. Further, we introduce some asymmetry in periods T_1 and T_3 through the BPR congestion function, which could potentially reflect choices of different routes in these periods. The mean and standard deviation of the value of time are assumed to be 0.33\$ per min (around 20\$ per hr) and 0.067\$ per min (around 4\$ per hr) respectively (refer Prato et al. (2014); Hess et al. (2005); Cirillo and Axhausen (2006) for empirical evidence; note that the literature reports a wide range of values for the coefficient of variation, we adopt a conservative value of 0.2).

Table 2: Variable Factors and Levels

Factors	Levels			
COV of mobility model scale μ_n	0.0	0.2	0.33	0.5
Scale of selling model ($\bar{\mu}$)	1	1.5	2	100000
Number of travelers (N)	1400	1550	1700	
BPR Congestion coefficient (β)	3	4	6	
Benefit Difference in \$ (Δ)	3	4	5	

In the experimental design, five factors are varied, which include the coefficient of variation (COV) of the scale parameter μ_n in the mobility model (the mean of μ_n is fixed at 1.5), the scale parameter of the selling model $\bar{\mu}$, total number of users N, the congestion coefficient β, and the benefit difference between the two days Δ. The factor levels are shown in Table 2. A total of 432 test instances or scenarios ($4^2 \times 3^3$) were simulated.

Several additional points are noteworthy. First, in all the scenarios wherein μ_n is deterministic (in other words, COV of μ_n is zero), the standard deviations of all other randomly distributed parameters (i.e. $\alpha^n; B_i^{n,sk}, i = 1 \dots 3, k = 1, 2$) are also set to zero. Thus, this subset of scenarios represents the setting with no heterogeneity in the mobility model. Second, in order to set the values of the alternative specific benefits for a given scenario with a benefit difference Δ, the values of $B_i^{n,s1}, i = 1 \dots 3$, are first sampled (based on the mean and standard deviation in Table 1), and $B_i^{n,s2}, i = 1 \dots 3$ is given by, $B_i^{n,s2} = B_i^{n,s1} + \Delta, i = 1 \dots 3$. Third, the scenarios with $\bar{\mu} = 100000$ represent non-rational market behavior or a purely random selling model (i.e. $\bar{\mu} \to \infty$).

The two instruments (tolling in dollars and tolling in tokens) are compared across the 432 test instances based on the optimum social welfare obtained by solving the optimization problems in Equations 22, 24 and 25. In case of the price control system, the optimization problem in Equation 22 is solved as a bi-level problem using MATLAB. The *fmincon* routine is used for the upper level (the sequential quadratic programming algorithm is applied which is known to work well for non-convex problems, Meng et al. (2004)) and the *lsqnonlin* routine is used for the lower level equilibrium problem. Given the non-convexity of the problem, 25 randomly generated starting points are used for the optimization algorithm (the value of 25 was arrived at empirically based on preliminary experiments wherein it was found that increasing the number of starting points beyond 25 did not yield improvements in the objective value). A similar approach is used for the quantity control system. At the upper level, the *fmincon* routine is used, and for each candidate solution of the tolls in tokens $\delta_i, i = 1...3$, a simple bisection method is applied to compute the prices that ensure market clearing in all intervals (for a given vector of tolls in tokens and market prices, the *lsqnonlin* routine is used to solve the lower level equilibrium problem).

5.2. Results and Discussion

The results from the numerical experiments and their implications are discussed in this section. After a description of the overall results in terms of optimum social welfare, the effects of congestion, extent of day to day variability, selling model, and heterogeneity are discussed in turn.

5.2.1. Welfare

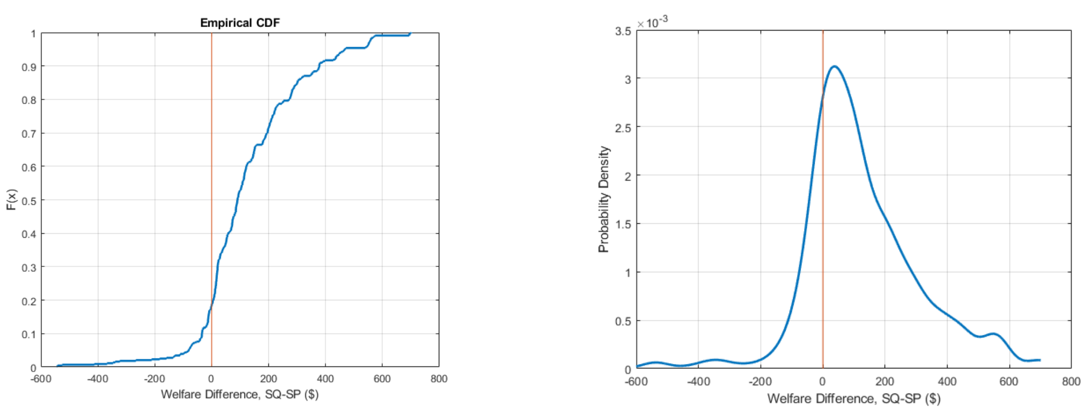
Summary statistics (across the 432 scenarios) of the welfare differences between various instruments under stochastic demand are presented in Table 3. The following abbreviations are used: **NT** for the no-toll equilibrium, **SP** for the price system or tolling in dollars, **SQ** for the quantity system or tolling in tokens. We also include a benchmark (abbreviated **ADP**) in which the tolls (in either dollars or tokens) are adaptive across the two days and set by the regulator based on the realization of demand. Clearly, in this case, the price and quantity instruments are equivalent (as shown in Section 3.4), and this benchmark represents the maximum welfare that can be attained in case of stochastic demand. N_S denotes the number of scenarios.

The results show that in the case of the fixed token supply, neither instrument is consistently superior across all scenarios (column SQ-SP in Table 3). The quantity system is superior in around 81% of the tested scenarios, with the absolute welfare difference (SQ-SP) ranging between -545\$ and 700.7\$, and mean and median values of 126.6 \$ and 90.4 \$ respectively. To put these differences in

Table 3: Summary Statistics: Welfare Differences

Statistic	Welfare Difference ($)					Percentage Diff.		Welfare ($)	
	ADP-NT	SP-NT	SQ-NT	SQ-SP	ADP-SP	(SQ-SP)/NT	(SQ-SP)/ADP	NT	ADP
Mean	1776.1	1554.7	1681.3	126.6	221.5	2.0	1.4	7347.3	9123.4
Median	1569.1	1416.1	1501.1	90.4	169.1	1.2	1.0	7442.9	9131.7
Min	575.4	516.6	104.4	-545.0	38.5	-6.8	-6.2	4645.1	8023.9
Max	4667.1	3889.7	4590.4	700.7	777.5	13.9	7.2	9021.2	10336.3
25th per.	1003.0	867.2	874.4	15.1	103.4	0.2	0.2	6613.1	8668.3
75 per.	2180.9	1895.6	2075.6	213.3	285.5	3.3	2.4	8131.2	9540.3
N_S(>0)	432	432	432	351	432	351	351	-	-
N_S	432	432	432	432	432	432	432	432	432
%>0	100	100	100	81.3	100	81.3	81.3	-	-

context, the total welfare of the no toll equilibrium ranges between 4645.1$ and 9021.2$ while that of the benchmark ranges between 8023.9$ and 10336.3$; the total toll revenue in the benchmark system ranges between 4239.6$ and 7026.8$. The percentage difference in welfare (SQ-SP) relative to the welfare of the no-toll equilibrium ranges between -6.8% and 13.9%, with mean and median percentage differences of 2% and 1.2% respectively. Note that all welfare values (ADP, SQ, SP, NT) can be considered as being relative to a situation where all travelers stay at home (i.e for example, due to very large travel times), which will yield zero welfare due to the normalization of the utility of the cancel trip option. The percentage differences need to be interpreted in this context.



(a) Empirical CDF

(b) Kernel density

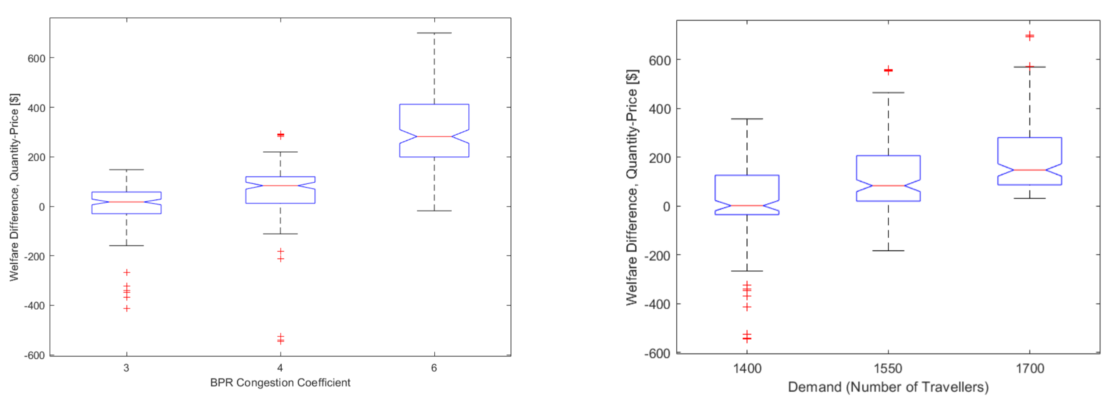
Figure 1: Distribution of Welfare Differences

Moreover, an examination of the average welfare difference between the price and quantity instrument (column SQ-SP in Table 3) relative to the average welfare difference between the price instrument and the adaptive benchmark (column ADP-SP in Table 3) suggests that the added flexibility of the permit market allows us to recover a little over 57% of the welfare lost due to the tolls in dollars and tokens being fixed across days. This is also evident when looking at the welfare differences between the price and quantity instruments relative to the adaptive benchmark (column (SQ-SP)/ADP in Table 3), which ranges between -6.2% and 7.2% with a mean value of 1.4%.

The overall distribution of welfare differences (SQ-SP) is shown in Figures 1a and 1b (the kernel density is plotted assuming a normal kernel function), which as noted before indicate that when the supply of tokens is fixed across days, neither instrument is consistently superior in terms of efficiency. In order to gain more insights into the conditions under which the quantity instrument is superior, we next examine the impacts of the shape of the congestion function, selling behavior, the benefit difference across days and the extent of heterogeneity.

5.2.2. Effect of Congestion Function

In order to gain insights into the scenarios where the quantity control instrument is superior, we first examine the nature of the congestion function, and draw on recent theoretical insights from de Palma and Lindsey (2020), who study tradable permit schemes in a setting with homogeneous agents and a single congestible facility. They conjecture (and explore through simple numerical examples) that in the case of variable demand and a fixed, but nonlinear cost function, the performance of a quantity control system dominates that of a price control system when the cost function is more steeply curved. This relates to their general finding that a quantity control system is relatively efficient if optimal usage levels are similar across states whereas a congestion fee achieves high efficiency if the first-best fee varies little over states.



(a) BPR Congestion Coefficient

(b) Demand

Figure 2: Welfare Difference: Effect of BPR Congestion Coefficient and Demand

The results from our experiments support these findings. First, the BPR congestion coefficient has a statistically significant effect (level of significance $\alpha = 0.01$) on the difference in total welfare between the quantity and price control systems. The average difference increases from \$2.1 at $\beta = 3$

to \$63.8 at $\beta = 4$, and \$313.9 at $\beta = 6$. Moreover, the percentage of scenarios where the quantity control system is superior increases from 61.8% at $\beta = 3$ to 82.6% at $\beta = 4$ and 99.3% at $\beta = 6$. The effect of β on the welfare difference is shown in the box plot in Figure 2a where it can be seen that the quantity control mechanism is superior in terms of total welfare typically when congestion curve is steeper or more convex (high value of $\beta = 4, 6$). In the box plot, the lower and upper edges of the blue box represent the 25th and 75th percentile respectively, the red line represents the median, and the notch represents a 95% confidence interval for the median.

A similar trend of increasing and statistically significant ($\alpha = 0.01$) welfare differences (SQ-SP) is observed as the demand level increases (under fixed capacity, i.e. congestion levels increase). The average difference increases from \$26.9 at $N = 1400$ to \$138.6 at $N = 1550$, and 214.4\$ at $N = 1700$. Moreover, the percentage of scenarios where the quantity control system is superior increases from 50.7% at $N = 1400$ to 93.1% at $N = 1550$ and 100% at $N = 1700$. The effect of total demand (number of travelers) on the welfare differences between the quantity and price instruments is shown in the box plot in Figure 2b.

Table 4: Illustrative Scenarios: Effect of BPR Congestion Coefficient β

Instrument		Flows				Optimal/Equivalent Tolls		
		T_1	T_2	T_3	T_0	T_1	T_2	T_3
Scenario 1								
No Toll (NT)	s1	377	542	367	115	0	0	0
	s2	419	567	407	8			
Stochastic Price (SP)	s1	326	404	315	356	2.59	4.31	2.63
	s2	436	498	421	45			
Stochastic Quantity (SQ)	s1	374	459	351	216	1.16	2.67	1.33
	s2	395	428	365	212	5.85	8.07	6.11
Benchmark (ADP)	s1	338	416	327	319	2.26	3.96	2.30
	s2	411	474	397	118	4.60	6.36	4.64
Scenario 2								
No Toll (NT)	s1	389	469	382	160	0	0	0
	s2	441	498	432	28			
Stochastic Price (SP)	s1	287	337	281	495	3.17	4.90	3.21
	s2	419	443	410	128			
Stochastic Quantity (SQ)	s1	343	401	340	316	1.52	2.84	1.63
	s2	360	372	355	313	6.22	8.27	6.42
Benchmark (ADP)	s1	326	366	319	389	2.59	4.30	2.63
	s2	372	401	364	263	5.29	7.06	5.33

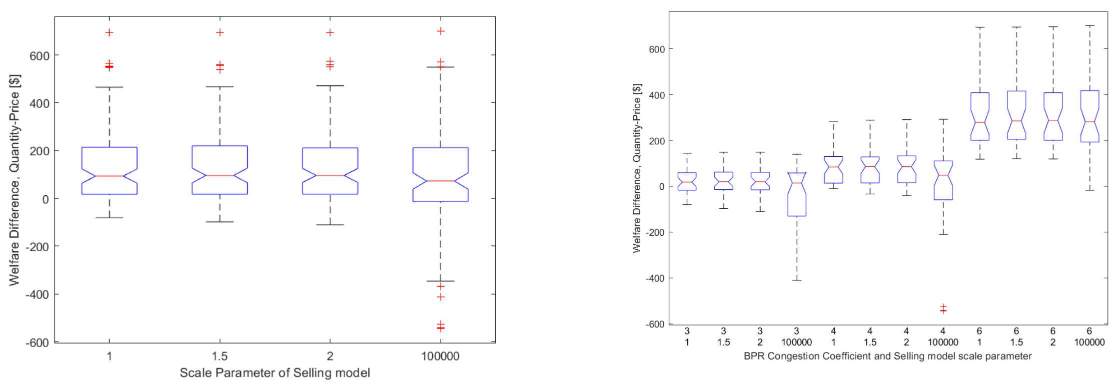
In order to gain more insight into the effect of the BPR congestion coefficient and the total demand (number of users), we examine several illustrative scenarios. First, we compare two scenarios (referred to as 1 and 2) with $\beta = 3, N = 1400, \Delta = 5$ and $\beta = 6, N = 1400, \Delta = 5$, respectively. All other factors including the scale of the mobility model and selling model are the same. The welfare difference between the quantity and price instruments (SQ-SP) are -72.9$ and 271.0$ in scenario 1 and 2 respectively.

Table 5: Illustrative Scenarios: Effect of Demand

Instrument		Flows T_1	Flows T_2	Flows T_3	Flows T_0	Optimal Tolls T_1	Optimal Tolls T_2	Optimal Tolls T_3
Scenario 1								
No Toll (NT)	s1	377	542	367	115	0	0	0
	s2	419	567	407	8			
Stochastic Price (SP)	s1	326	404	315	356	2.59	4.31	2.63
	s2	436	498	421	45			
Stochastic Quantity (SQ)	s1	374	459	351	216	1.16	2.67	1.33
	s2	395	428	365	212	5.85	8.07	6.11
Benchmark (ADP)	s1	338	416	327	319	2.26	3.96	2.30
	s2	411	474	397	118	4.60	6.36	4.64
Scenario 2								
No Toll (NT)	s1	446	586	432	235	0	0	0
	s2	522	638	505	35			
Stochastic Price (SP)	s1	321	399	311	668	3.65	5.39	3.69
	s2	493	549	476	182			
Stochastic Quantity (SQ)	s1	399	498	387	416	1.85	3.04	1.86
	s2	416	469	401	414	6.58	8.50	6.62
Benchmark (ADP)	s1	362	435	350	553	2.84	4.57	2.88
	s2	441	500	426	333	5.77	7.55	5.80

Table 4 summarizes the flows in different time periods and the tolls in dollar amounts (note that for the price instrument this is directly the toll in dollars whereas for the quantity instrument it is the product of the toll in tokens and the token market price). First, observe that the for both scenarios, as expected, under the price instrument (SP), the number of individuals traveling (total flow in periods T_1, T_2, T_3) varies significantly across the days s1 and s2 (also evident from the number of travelers cancelling trip, i.e. flow in T_0) whereas the toll in dollar amounts is fixed. In contrast, under the quantity instrument, the number of travelers traveling is roughly the same across the two days whereas the toll in dollar amounts varies significantly. Next, we see that the optimal usage of the network (or number of people traveling) under the adaptive benchmark ADP varies more across the states S1 and S2 in scenario 1 than in scenario 2 (difference in the optimal flows for T_0 across days is $319 - 118 = 201$ for scenario 1 versus $389 - 263 = 126$ in scenario 2). Conversely, looking at optimal tolls under the benchmark ADP, one can see that the toll difference across s1 and s2 is higher in scenario 2 compared to scenario 1 (in interval T_2, $7.06 - 4.30 = 2.76$ for scenario 2 versus $6.36 - 3.96 = 2.40$ for scenario 1). Thus, the results suggest that at higher BPR coefficients or steeper congestion functions, the optimal usage levels of the network are relatively more similar across states leading to superiority of the quantity instrument. More intuition for this is provided analytically by de Palma and Lindsey (2020) who look at the welfare losses of the two instruments using a single congested alternative under linear and non-linear demand.

Along similar lines, to examine the effect of the total number of travelers N, we compare two different scenarios (referred to as 1 and 2) with $\beta = 3, N = 1400, \Delta = 5$ and $\beta = 3, N = 1700, \Delta = 5$ respectively. All other factors including the scale of the mobility model and selling model are the same. The welfare difference between the quantity and price instruments (SQ-SP) are -72.9\$ and 88.73\$ in scenario 1 and 2 respectively. Table 5 presents the flows in different time periods and the tolls in dollar amounts for these two scenarios. We observe – as before in the case of higher β – that when the overall demand level is higher (scenario 2), the difference in optimal toll rates across s1 and s2 is higher compared to scenario 1 (in interval T_2, $7.55 - 4.57 = 2.98$ for scenario 2 versus $6.36 - 3.96 = 2.40$ for scenario 1). Thus, at higher demand levels (and hence, more severe congestion effects) the optimal toll rates vary more across states, once again leading to superior performance of the quantity control instrument relative to scenarios with lower demand levels.



(a) Selling model scale parameter $\bar{\mu}$

(b) Interaction between $\bar{mu}$ and BPR β

Figure 3: Welfare Difference: Effect of Selling Model

5.2.3. Selling Model

The explicit treatment of selling behavior is an important characteristic of the proposed model, and allows us to examine the impact of the selling decisions on the performance of the quantity control system. Figure 3a presents a box plot of the effect of the scale parameter of the selling model on the difference between the quantity and price instruments. The value of $\bar{\mu} = 100000$, which corresponds to $\bar{\mu} \to \infty$ represents a purely random selling model (or a non-rational market) and has the impact of equalizing the supply of tokens across the three time periods (consequently, the demand of tokens as well). As the results show, this has the effect of a deterioration in the performance of the quantity control system, reflected in the mean difference in welfare between the two instruments (Quantity

- price), which takes a mean value of 85.7 S$ at $\bar{\mu} = 100000$ versus 140.2 S$ for $\bar{\mu} \in [1, 2]$. Note that within the range of $\bar{\mu} \in [1, 2]$, performance of the quantity control does not vary substantially. However, interestingly, even with in the case of a non-rational market, the quantity instrument remains superior to the price instrument in cases where congestion effects are severe (e.g. $\beta = 6$ in Figure 3b). In other words, the advantages of the quantity control system noted in Section 5.2.2 remain even if users are not perfectly rational in the selling market, although the extent of welfare difference is marginally lower.

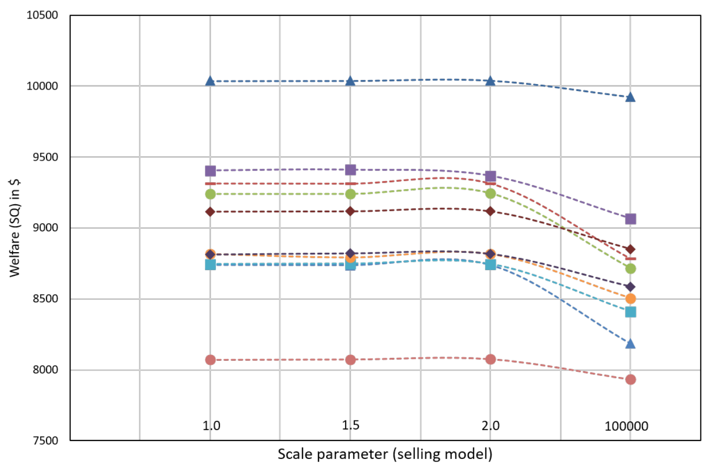


Figure 4: Effect of $\bar{\mu}$ on welfare of quantity control (SQ))

The impact of market behavior is illustrated in Figure 4 for a few selected scenarios which yield the highest deterioration in welfare for $\bar{\mu} = 100000$ compared to the corresponding scenarios with $\bar{\mu} = 1, 1.5, 2$. Each marker type or series represents scenarios where all parameters are identical except $\bar{\mu}$. Note that the interpolating lines between each point are not meant to be indicative of the actual trend but are used to simply make the figure more legible. As noted above, we observe that within the range $\bar{\mu} \in [1, 2]$, the differences in welfare of the quantity instrument are negligible whereas the welfare deteriorates if users are more irrational in the selling market. Thus, the findings suggest that market design aspects of the quantity control instrument are important and can have effects on efficiency. This is explained in more detail next.

In order to gain more intuition into the effect of selling behavior in the market, we examine two illustrative scenarios, one with $\bar{\mu} = 1$ and the second with $\bar{\mu} = 100000$ ($\beta = 4, N = 1550, \Delta = 3$ in both scenarios). Thus, the two scenarios are identical in all respects except the scale parameter of the selling model. The welfare of the scenario with $\bar{\mu} = 1$ is higher than that with $\bar{\mu} = 100000$ by 104.2\$. Table 6 summarizes the flows in different time periods, the tolls in dollar amounts (note that is the product of the toll in tokens and the token market price) and the demand and supply of tokens for all

Table 6: Illustrative Scenarios: Effect of selling behavior

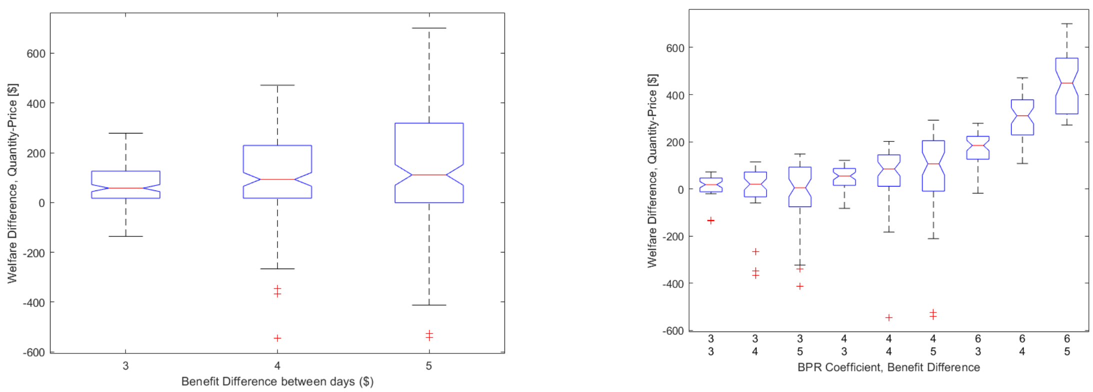
Instrument		Flows				Optimal Tolls		
		T1	T2	T3	T0	T1	T2	T3
Stochastic Quantity (SQ)	s1	366	440	353	391	2.22	3.77	2.29
$\bar{\mu} = 1.0$	s2	381	418	365	386	5.00	7.10	5.09
Stochastic Quantity (SQ)	s1	355	378	341	476	2.72	5.06	2.82
$\bar{\mu} = 100000$	s2	355	378	341	476	5.72	8.06	5.82
Benchmark (ADP)	s1	346	409	336	459	2.26	3.96	2.30
	s2	391	444	380	335	4.60	6.36	4.64
		Demand for Tokens				Supply of Tokens		
		T1	T2	T3	T0	T1	T2	T3
Stochastic Quantity (SQ)	s1	53.89	148.93	56.04	-	53.89	148.93	56.04
$\bar{\mu} = 1.0$	s2	56.07	141.53	57.96	-	56.05	141.53	57.95
Stochastic Quantity (SQ)	s1	111.81	111.81	111.81	-	111.81	111.81	111.81
$\bar{\mu} = 100000$	s2	111.81	111.81	111.81	-	111.81	111.81	111.81

time intervals and both days s1 and s2. First, note that non-rational selling behavior or $\bar{\mu} = 100000$ has the effect of equalizing the probability of selling in all three time intervals and hence, equalizes the supply of tokens for all three intervals (last two rows in Table 6). Interestingly, this also causes the token supply to be equal on both days s1 and s2 (for all three intervals). Thus, we see that the optimal tolls (SQ with $\bar{\mu} = 100000$) on days s1 and s2 differ by an additive constant of 3$ (which is exactly equal to the benefit difference between the two days, $\Delta = 3\$$) resulting in identical token supply and also, identical flows on both days s1 and s2. Moreover, the equal token supply results in lower flows in interval T_2 (for example, 440 on s1 for $\bar{\mu} = 1$ versus 378 on s1 for $\bar{\mu} = 100000$). This results in a significantly higher number of travelers cancelling trip (choosing T_0) in the scenario with $\bar{\mu} = 100000$ (476) versus the scenario with $\bar{\mu} = 1$ (391 and 386) leading to a loss in welfare. In summary, we see that the quantity control instrument is robust with respect to selling behavior in the market and is still superior even with irrational sellers when congestion effects are more severe. However, there is a deterioration in welfare when the behavior of sellers is more irrational, which causes the equilibrium price of tokens to increase leading to less travel.

5.2.4. Difference between states of nature

The benefit difference between the two days or states of nature (Δ) is a measure of the extent of day to day variability, and the results indicate —similar to the congestion coefficient and demand— that it significantly affects the relative performance of the two instruments ($\alpha = 0.05$). The average difference increases from \$81 at $\Delta = 3$ to \$125 at $\Delta = 4$, and 175.4\$ at $\Delta = 5$. Further, the variance in difference between the two instruments also increases as is evident from the boxplot in Figure 5a. This can be better understood by examining the interaction effects of the benefit difference with the

BPR congestion coefficient. Thus, when the congestion curve is steeper, a larger degree of day of day variability results in a greater advantage for the quantity control system ($\beta = 6, \Delta = 5$ versus $\beta = 6, \Delta = 3$ in Figure 5b). On the hand, congestion effects are less severe, an increase in day to day variation ($\beta = 3, \Delta = 5$ versus $\beta = 3, \Delta = 3$ in Figure 5b) results in a poorer performance of the quantity control system.



(a) Benefit Difference (Δ)

(b) Interaction between Δ and BPR Coefficient β

Figure 5: Welfare Difference: Effect of Benefit Difference and Interaction between Benefit Difference/Beta

This can be explained once again by the fact that at higher levels of the benefit difference (or day to day variability) and BPR coefficient, the optimal toll rates (in dollar amounts) tend to vary significantly across the states s1 and s2, leading to superior performance of the quantity instrument.

5.2.5. Heterogeneity

The impact of the coefficient of variation (COV) of the scale parameter μ_n in the mobility model (note that μ_n is assumed to be lognormally distributed across the population of travelers) is shown in the boxplot in Figure 6. First, it can be observed that as the COV increases from $\mu_n = 0.2$ to $\mu_n = 0.5$, the mean difference in welfare between the quantity and price instruments increases only marginally from 109.8 \$ to 123.1 \$ (statistically insignificant at $\alpha = 0.05$).

In contrast, when we examine the effect of overall heterogeneity in the mobility model (recall that the scenarios with $\mu_n = 0$ represent the homogeneous case where all other parameters in the mobility model are also assumed to be deterministic), we see a significant effect ($\alpha = 0.05$). The mean difference between the quantity and price control systems for the homogeneous scenarios is in fact higher at 158\$ (scenarios with COV of $\mu_n = 0$ in Figure 6) compared to 115\$ when heterogeneity is considered (scenarios with COV of $\mu_n = 0.2, 0.33, 0.5$ in Figure 6) . This has important implications and suggests that ignoring heterogeneity can potentially overestimate the benefits of the quantity

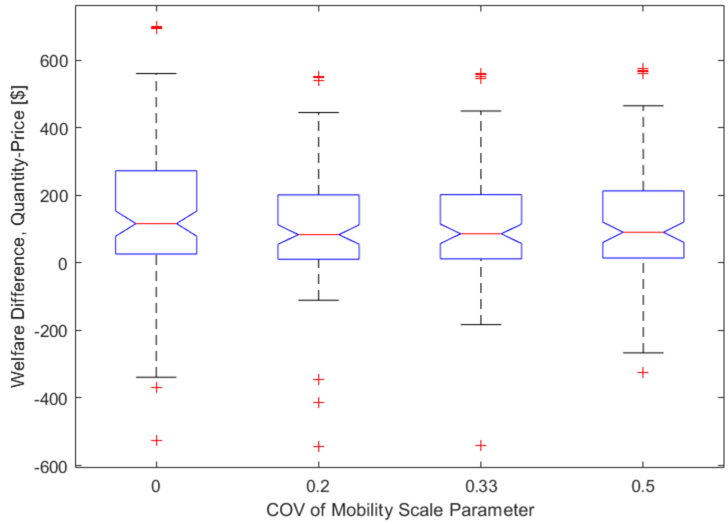


Figure 6: Welfare Difference: Effect of Heterogeneity and COV of mobility scale

control system. This contrasts with the findings in de Palma et al. (2018) where heterogeneity was found to slightly increase the average welfare difference between the two instruments. The intuition for these differences is hard to arrive at, one potential cause may be differences in the nature of variability, which arises from scheduling preferences in our case as opposed to the total number of users in their case.

5.2.6. Equity

Equity is a key consideration in the comparison between the price and quantity control instruments. In this section, we compare the two instruments using the Gini coefficient computed based on the logsum (a measure of user benefits). For a detailed discussion of measures of inequality and welfare in the transportation context, we refer the readers to Trannoy (2011) and Delle Site et al. (2021). We first discuss the computation of the Gini coefficient followed by a discussion of findings.

Consider the population of N travelers ($k = 1 \dots N$), and let $UB(k)$ denote the user benefit of individual k in \$ amounts. In the case of the no toll equilibrium (denoted NT) and tolling in dollars with no redistribution of toll revenues (denoted SPN), $UB(k)$ is simply the logsum of individual k. In the case of tolling in dollars with an equal redistribution of toll revenues (denoted SP), $UB(k)$ is the logsum of individual k plus the average toll revenue per individual. Finally, in the case of tolling in tokens (denoted SQ), $UB(k)$ is the logsum of individual k plus the market value of the initial token endowment (since tokens are distributed for free and no tokens are unused at the end of the day).

Assume that drivers arranged in increasing order of their user benefit and let $x = k/N$. Define,

$$g(x) = \frac{\sum_{j=1}^{xN} UB(j)}{\sum_{j=1}^{N} UB(j)}. \tag{26}$$

where $g(x)$ represents the Lorenz curve, which is the cumulative share of total user benefits (based on the logsum measure) obtained by the bottom xN individuals in the population (note that xN is an integer). The Gini coefficient of user benefits (denoted GC) is computed as,

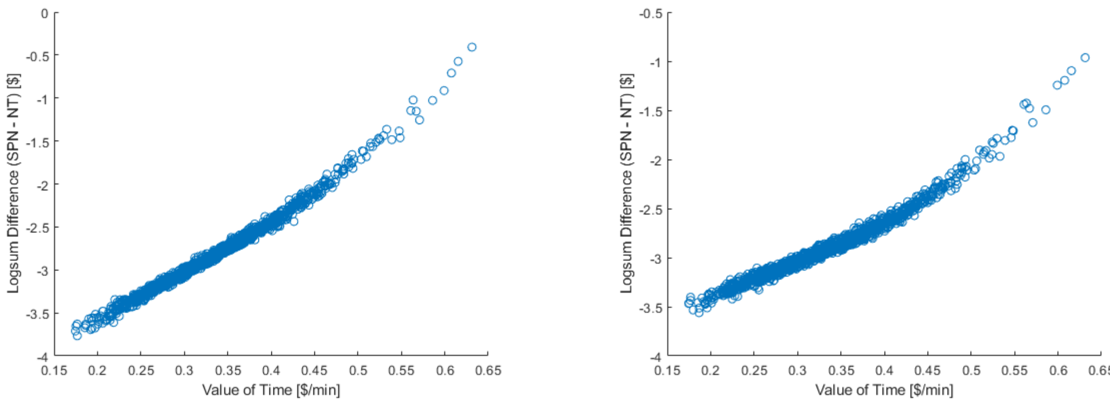
$$GC = \frac{\left|0.5 - \int_0^1 g(x)dx\right|}{0.5}. \tag{27}$$

The Gini coefficient is a measure of equity and takes a value between 0 and 1; a value equal to zero implies total equity and a value of 1 indicates total inequity. The larger it is, the more inequitable is the policy.

Table 7: Summary Statistics: Gini Coefficient

Statistic	GC_{NT}	Percentage Difference in Gini Coefficient			
		(SPN-NT)/NT	(SP-NT)/NT	(SQ-NT)/NT	(SQ-SP)/SP
Mean	0.183	28.7	−43.7	−43.7	−0.6
Median	0.175	32.3	−43.0	−43.9	0.0
Minimum	0.123	−2.4	−65.9	−68.9	−11.2
Maximum	0.275	46.2	−22.9	−1.9	40.1
25th Percentile	0.155	21.7	−51.8	−52.2	−3.3
75th Percentile	0.208	38.1	−34.2	−34.3	1.4
Scenarios >0	324	312	0	0	156
% Scenarios >0	100	96.3	0	0	48.1
N_S	324				

Table 7 summarizes the distribution of the Gini coefficient (across the 324 scenarios with heterogeneity) for the NT equilibrium (denoted by GC_{NT}) and percentage differences between the Gini coefficient for the different instruments. First, observe that with tolling in dollars wherein toll revenues are not redistributed, in a majority of the scenarios (96.3%), the Gini coefficent increases (i.e is more inequitable) relative to the No Toll equilibrium, and is on average 28.7% higher (column three of Table 7). This is in line with the general observation that pricing is vertically inequitable and benefits the rich (here the individuals with high value of time) more than the poor. The scatter plots (Figure 7) of logsum difference (between SP and NT) versus value of time (a proxy for income) corroborate this observation, where we see that the benefits clearly increase with an increase in value of time. The plots represent two illustrative scenarios for the s2 day and each point in the plot represents an individual. Interestingly, there are a small number of scenarios (3.7%), where the Gini coefficient reduces even when toll revenues are not redistributed. This occurs in scenarios where the congestion effects are the most severe (BPR coefficient of 6 and highest demand level).



(a) Scenario 1 ($\beta = 6, N = 1400, \Delta = 5$)

(b) Scenario 2 ($\beta = 3, N = 1550, \Delta = 4$)

Figure 7: User Benefits versus Value of Time

Our second observation is that both in case of the quantity instrument and the price instrument (when toll revenues are equally redistributed), there is a significant improvement of equity, by an average of 43.7%, relative to the No Toll equilibrium (column four and five of Table 7). This is a key finding and implies that both instruments favour the poor (low value of time individuals) since there is a large reduction in the Gini coefficient across all scenarios relative to the No Toll equilibrium. The primary reason is that an equal redistribution of toll revenues (price instrument) and the equal allocation of tokens (quantity instrument) results in an increase in the cumulative share of benefits obtained by lower income travelers, leading to an improvement in equity. Further, note that in case of the quantity instrument, a further improvement in equity can be achieved through any progressive allocation of the tokens.

Finally, comparing the Gini coefficient for the quantity instrument and the price instrument with equal redistribution of toll revenues (column six of Table 7), we see that neither instrument is consistently superior in terms of equity, although the quantity instrument is on average marginally better (average difference of 0.6%). Moreover, similar to the comparative performance with respect to welfare, we find that the quantity instrument is superior in terms of equity in scenarios with more severe congestions effects (high BPR congestion coefficient of 4 and 6) and when the selling behavior of individuals is rational. These findings make an additional case for tolling in tokens.

5.3. Adaptive Token Supply

The experiments in Section 5.2 consider a quantity control system wherein the token supply is fixed across days and the results indicate the tolling in tokens is not consistently superior to tolling in dollars. The results also suggest that the price system is typically superior when congestion effects are less severe (slope of the congestion function is less steep, demand is lower). In these cases, as seen in the illustrative scenarios in Section 5.2.2, the quantity targets may be too lax on the s1 day and thus, the performance of the quantity control system can be improved by allowing the token supply to be adapted across days in response to the realization of demand (Equation 25).

In this section, we examine the comparative performance of the two instruments when the token supply is adaptive. Note that, in this case certain parameters of the quantity control system (M^{s1}, M^{s2} in Equation 25) are dependent on the state of nature. This is in contrast with the adaptive tradable permit system (TPS) considered in de Palma and Lindsey (2020) where, the regulator issues a certain number of permits, but in addition, offers to sell *further* permits at a price s, and buy permits at a price r, where $r < s$ This limits the price of permits to the range $[r, s]$, where r and s are fixed and state independent.

Table 8: Summary Statistics: Welfare Differences (Adaptive token supply)

Statistic	Welfare Difference ($)					Percentage Diff.	NT Welfare ($)
	SP-NT	SQ_A-NT	SQ-SP	SQ_A-SP	ADP-SP	(SQ_A-SP)/NT	
Mean	1554.7	1744.6	126.7	190.0	221.5	2.9	7347.3
Median	1416.1	1560.8	90.4	157.6	169.1	2.1	7442.9
Min	516.6	348.0	-545.0	-177.4	38.5	-2.0	4645.1
Max	3889.7	4656.2	700.7	766.6	777.5	15.2	9021.2
25th per.	867.2	952.4	15.1	81.6	103.4	1.0	6613.1
75th per.	1895.6	2143.2	213.3	262.5	285.5	4.0	8131.2
$N_S > 0$	432	432	351	410	432	410	-
%>0	100	100	81.3	94.9	100	94.9	-

The results are summarized in Table 8 and as expected, indicate that the quantity control system with adaptive supply is superior to that with fixed supply in all scenarios. The mean welfare difference between the quantity system with adaptive supply (denoted SQ_ A) and the price system is 190.0 $ compared to 126.6 $ with fixed supply (refer columns SQ-SP and SQ_A - SP). Moreover, a comparison of these numbers against the mean difference of 221.5$ between the adaptive benchmark and the price system (ADP-SP) reveals the extent of welfare improvements that can be attained by adapting the token supply across days. Thus, while the quantity instrument with fixed token supply recovers a little over 57% of the welfare loss due to fixing the tolls (in dollars and tokens) across days, the quantity instrument with adaptive token supply recovers almost 86% of this welfare loss.

However, contrary to intuition, even with adaptive token supply, the quantity control system is still not consistently superior to the price control system although it yields a higher welfare in 94.9% of tested the tested scenarios. This is in contrast with the findings in de Palma et al. (2018) for a single period setting where the quantity control with adaptive token supply is consistently superior. A more detailed examination shows that the scenarios where the price control is superior are in fact all scenarios where the selling behavior is non-rational or completely random (i.e $\bar{\mu} \to \infty$) as shown

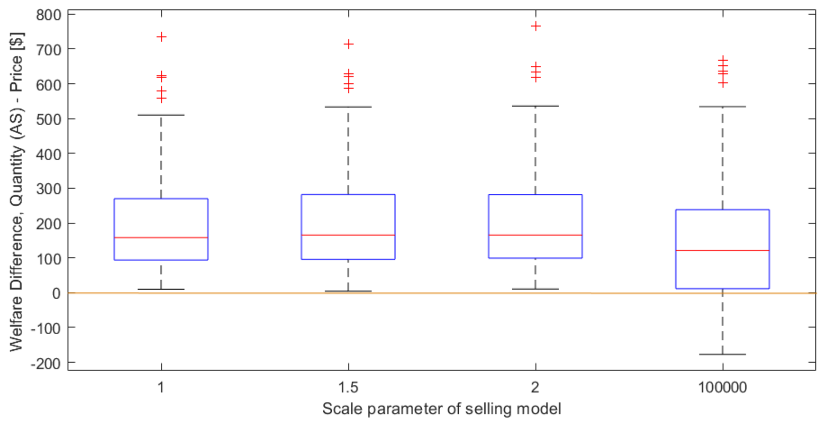


Figure 8: Welfare Differences (adaptive token supply): Effect of selling model

in the box plot in Figure 8. As shown in Section 5.2.3 this has the effect of forcing the token supply to be equal across time periods, reducing the efficiency of the quantity control system in a manner that is not redressed even with the adaptive token supply. This once again highlights the importance of market design in the efficiency of the quantity control system.

6. Conclusions

This paper develops a methodology to compare price control (tolling in dollars) and quantity control (tolling in tokens) instruments in the context of a within-day setting with departure time choice. In the quantity control system, users are provided an initial endowment of tokens by the regulator and incur a token charge to travel in a specific time period. Tokens can be bought and sold within a marketplace at a price determined by a market clearing mechanism in each time period. A key feature of the market model is that the selling decisions of users are explicitly considered.

Numerical experiments across a wide range of scenarios with demand uncertainty yield the following key insights. First, when the tolls (in dollars and tokens) can be adapted from day to day, the two instruments are equivalent. Second, when the token supply is fixed across days or states of nature and the tolls (in dollars and tokens) are non-adaptive, the quantity control instrument is superior in welfare terms when congestion effects are more severe, i.e. steep congestion functions (realistic BPR models), high demand levels and high day-to-day variability. In these scenarios, the optimal network usage is relatively similar across states whereas the optimal toll in dollar amounts varies significantly across states. Third, non-rational selling behavior, which has the effect of equalizing token supply across time intervals leads to a deterioration in the performance of the quantity instrument. However, in general the token system is robust (in welfare terms) with respect to selling behavior in the market. Fourth, when the token supply can be adapted from day-to-day, the quantity instrument is superior in all scenarios where selling behavior is rational. Finally, when toll revenues in the price instrument are equally redistributed (typically difficult to implement in practice) and tokens (in the quantity instrument) are equally distributed, tolling in tokens is marginally more equitable in scenarios where congestion effects are more severe. These findings make a potential case for quantity control.

Several points are however noteworthy. First, income effects and second order effects on the use of toll revenues are not considered. Second, transaction costs associated with the trading of permits, the process of finding a buyer or seller, negotiating a price, etc. are ignored. These are likely to affect the overall welfare of the quantity control system (see Nie (2012)). However, as noted by Brands et al. (2020), transaction costs may be minimized through suitable market designs. For instance, they make use of a price setting intermediary with whom users trade, and point out that this can significantly reduce transaction and negotiation costs compared to designs that include consumer to consumer trading (and over existing designs such as Dutch and English auctions, sealed-bid auctions and Vickerey auction markets). Third, the public acceptability of tradable permits is not necessarily guaranteed and will depend on the initial allocation of permits and the extent of volatility in the permit market.

There are several avenues of further research including the use of more realistic network and congestion models, the consideration of both departure time and route/mode choice, and the inclusion of income effects and transaction costs.

Acknowledgements

André de Palma was partially supported by the NSFC-JPI UE joint research project "MAAT" (project no. 18356856).

References

Akamatsu, T., Wada, K., 2017. Tradable network permits: A new scheme for the most efficient use of network capacity. Transportation Research Part C: Emerging Technologies 79, 178–195.

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

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

Cirillo, C., Axhausen, K.W., 2006. Evidence on the distribution of values of travel time savings from a six-week diary. Transportation Research Part A: policy and practice 40, 444–457.

Delle Site, P., de Palma, A., Kilani, K., et al., 2021. Consumers' welfare and compensating variation: survey and mode choice application. Technical Report. THEMA (THéorie Economique, Modélisation et Applications), CY Cergy Paris Université.

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

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. Environmental and Resource Economics 10, 63–99.

Grant-Muller, S., Xu, M., 2014. The role of tradable credit schemes in road traffic congestion management. Transport Reviews 34, 128–149.

He, F., Yin, Y., Shirmohammadi, N., Nie, Y.M., 2013. Tradable credit schemes on networks with mixed equilibrium behaviors. Transportation Research Part B: Methodological 57, 47–65.

Hess, S., Bierlaire, M., Polak, J.W., 2005. Estimation of value of travel-time savings using mixed logit models. Transportation Research Part A: Policy and Practice 39, 221–236.

Jaensirisak, S., Wardman, M., May, A., 2005. Explaining variations in public acceptability of road pricing schemes. Journal of Transport Economics and Policy (JTEP) 39, 127–154.

Kalmanje, S., Kockelman, K.M., 2004. Credit-based congestion pricing: Travel, land value, and welfare impacts. Transportation Research Record 1864, 45–53.

Kockelman, K.M., Kalmanje, S., 2005. Credit-based congestion pricing: a policy proposal and the public's response. Transportation Research Part A: Policy and Practice 39, 671–690.

Laffont, J.J., 1977. More on prices vs. quantities. The Review of Economic Studies 44, 177–182.

Liu, Y., Nie, Y.M., 2017. A credit-based congestion management scheme in general two-mode networks with multiclass users. Networks and Spatial Economics 17, 681–711.

Nie, Y., 2012. Transaction costs and tradable mobility permits. Transportation Research Part B, 46, 189–203.

Nie, Y.M., 2015. A new tradable credit scheme for the morning commute problem. Networks and Spatial Economics 15, 719–741.

Nie, Y.M., Yin, Y., 2013. Managing rush hour travel choices with tradable credit scheme. Transportation Research Part B: Methodological 50, 1–19.

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

de Palma, A., Lindsey, R., 2020. Tradable permit schemes for congestible facilities with uncertain supply and demand. Economics of Transportation 21, 100149.

de Palma, A., Proost, S., Seshadri, R., Ben-Akiva, M., 2018. Congestion tolling-dollars versus tokens: A comparative analysis. Transportation Research Part B: Methodological 108, 261–280.

Prato, C.G., Rasmussen, T.K., Nielsen, O.A., 2014. Estimating value of congestion and of reliability from observation of route choice behavior of car drivers. Transportation Research Record 2412, 20–27.

Raux, C., 2004. The use of transferable permits in transport policy. Transportation Research Part D: Transport and Environment 9, 185–197.

Rezaeinia, P., de Palma, A., Lindsey, R., 2021. Traffic Congestion Control: Tradable Permits vs Road Tolls. Technical Report. THEMA (Théorie Economique, Modélisation et Applications), CY Cergy Paris Université.

Santos, G., Verhoef, E., 2011. Road congestion pricing, in: A Handbook of Transport Economics. Edward Elgar Publishing.

Shirmohammadi, N., Zangui, M., Yin, Y., Nie, Y., 2013. Analysis and design of tradable credit schemes under uncertainty. Transportation Research Record 2333, 27–36.

Tian, L.J., Yang, H., Huang, H.J., 2013. Tradable credit schemes for managing bottleneck congestion and modal split with heterogeneous users. Transportation Research Part E: Logistics and Transportation Review 54, 1–13.

Trannoy, A., 2011. Equity dimensions of transport policies. Handbook of Transport Economics, Edward Elgar, Cheltenham, UK , 625–649.

Verhoef, E., Nijkamp, P., Rietveld, P., 1997. Tradeable permits: their potential in the regulation of road transport externalities. Environment and Planning B: Planning and Design 24, 527–548.

Wang, G., Gao, Z., Xu, M., 2014a. An mpec formulation and its cutting constraint algorithm for continuous network design problem with multi-user classes. Applied Mathematical Modelling 38, 1846–1858.

Wang, G., Gao, Z., Xu, M., Sun, H., 2014b. Models and a relaxation algorithm for continuous network design problem with a tradable credit scheme and equity constraints. Computers & Operations Research 41, 252–261.

Wang, X., Yang, H., Zhu, D., Li, C., 2012. Tradable travel credits for congestion management with heterogeneous users. Transportation Research Part E: Logistics and Transportation Review 48, 426–437.

Weitzman, M.L., 1974. Prices vs. quantities. The review of economic studies 41, 477–491.

Wu, D., Yin, Y., Lawphongpanich, S., Yang, H., 2012. Design of more equitable congestion pricing and tradable credit schemes for multimodal transportation networks. Transportation Research Part B: Methodological 46, 1273–1287.

Xiao, F., Qian, Z.S., Zhang, H.M., 2013. Managing bottleneck congestion with tradable credits. Transportation Research Part B: Methodological 56, 1–14.

Yang, H., Wang, X., 2011. Managing network mobility with tradable credits. Transportation Research Part B: Methodological 45, 580–594.

Zhu, D.L., Yang, H., Li, C.M., Wang, X.L., 2015. Properties of the multiclass traffic network equilibria under a tradable credit scheme. Transportation Science 49, 519–534.

Appendix A: Glossary

Notation

Symbol	Description
B_i^n	Alternative specific benefit for individual n in time interval T_i $(i = 0 \dots 3)$
C_i	Capacity in time interval T_i $(i = 1 \dots 3)$
D_i	Demand for tokens in time interval $T_i, i \in \tilde{I}$
I	Set of time intervals $(T_i, i = 0 \dots 3)$
$\tilde{I}$	Subset of time intervals T_i where $\delta_i > M$
M	Token endowment per traveler
M^{sk}	Token endowment per traveler on day sk $(k = 1, 2)$
N	Number of travelers
p_i	Token market price in time interval T_i $(i = 1 \dots 3)$
$\bar{p}_i$	Perceived token market price for time interval T_i $(i = 1 \dots 3)$
p_i^{sk}	Token market price in time interval T_i $(i = 1 \dots 3)$ on day sk $(k = 1, 2)$
Q_i	Probability of selling in time interval $T_i, i \in \tilde{I}$
sk	State of nature or day $(k = 1, 2)$
S_i	Supply of tokens in time interval $T_i, i \in \tilde{I}$
t_i^{FF}	Free flow travel time in interval T_i $(i = 1 \dots 3)$
t_i	Congested travel time in interval T_i $(i = 1 \dots 3)$
T_i	Time interval i $(i = 0 \dots 3)$
$U^n(T_i)$	Utility of time interval i for individual n
$V^n(T_i)$	Systematic utility of time interval i for individual n
$V_{sk}^n(T_i^{sk})$	Systematic utility of time interval i on day sk for individual n
X_i	Flow in time interval T_i $(i = 1 \dots 3)$
X_i^{sk}	Flow in time interval T_i $(i = 1 \dots 3)$ on day sk $(k = 1, 2)$
α^n	Value of time of individual n
α_i	BPR function parameter for interval T_i $(i = 1 \dots 3)$
β_i	BPR function parameter for interval T_i $(i = 1 \dots 3)$
δ_i	Toll in tokens for time interval T_i $(i = 1 \dots 3)$
ϵ_i	Error term in utility for time interval T_i
μ_n	Scale parameter of individual n (mobility decision)
$\bar{\mu}$	Scale parameter of selling model
Ω_P	Optimum welfare of price instrument
Ω_Q	Optimum welfare of quantity instrument
τ_i	Toll in dollars for time interval T_i $(i = 1 \dots 3)$