Journal of Transport Geography 37 (2014) 10–18

Contents lists available at ScienceDirect

Journal of Transport Geography

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

Flat versus spatially variable tolling: A case study in Fresno, California

Omid M. Rouhani *, Debbie Niemeier

Department of Civil and Environmental Engineering, University of California, One Shields Avenue, Davis, United States

ARTICLE INFO

Keywords:
Road pricing
Variable toll
Flat toll
Tolling profile
Profit maximization
Congestion pricing

ABSTRACT

A number of studies have examined the feasibility of temporal variations in tolls. However, spatial variation in tolls has not received much attention, especially in practice. Spatial variation could effectively reduce congestion and increase profits. To fill this gap, we conduct an empirical application on 3 different road segments using the Fresno, California's transportation planning model. Our modeling results in a number of interesting insights. First, the derived optimal flat toll values are very close to the average variable tolls, but the effects of applying spatially variable tolls on improving total revenues (from 4% to 24%) and total improved travel time (from 18% to 1083%) measures are significant. Second, spatially variable tolls are more effective, but more costly, particularly for arterials, which can be attributed to the higher number of access points for arterials. Third, spatial variations in tolls are more effective for peak hours than for off-peak hours and for social optimization than for profit maximization. Fourth, to improve throughputs for both profit maximization and social optimization, the prevalent tolling pattern along a corridor induces lower final volumes per capacity (V/Cs) (after pricing) at the mainline flow sections and relatively higher final V/Cs at the entrance and exit (boundary) points. Finally, optimal toll patterns are not dependent on vehicle miles traveled (VMTs) or volumes but, rather, are related to targeted V/Cs. Therefore, flow-dependent charges along a corridor should be based on V/Cs rather than on volumes or VMTs.

1. Introduction

A large body of research exists on road/congestion pricing (Arnott et al., 1993; Dial, 1999a,b; Goh, 2002; Rouwendal and Verhoef, 2006; Verhoef, 2007; Anas and Pines, 2013). Congestion pricing is commonly used to reduce road congestion, while road pricing is levied primarily for revenue generation. Taking the fundamental reasons for pricing into account, all the studies related to the road/congestion pricing issue must make an important set of decisions regarding the magnitudes and patterns of the charges under different conditions. Temporal variations of tolls are one of those decisions that have been studied in two major contexts: congestion pricing and road pricing. These studies have tried to answer the question of when and why tolls should be flat (constant) or varied.

In the congestion pricing context, de Palma and Lindsey (2002) and de Palma et al. (2007) simulated the problem of two competing roads and examined a time-based congestion pricing scheme along with maintenance and tolling decisions. Without route choice and elastic demand features, Chu (1999) modeled peak-period commuting along an urban highway to simulate the equilibrium effects of various congestion pricing schedules and showed that flat tolls work better than smoothly varying (temporal) tolls for revenue generation, but smoothly varying tolls are better for congestion reduction and efficiency maximization. However, the results vary significantly based on the congestion level, and some of the examined pricing schedules are neither purely flat nor variable; a hybrid of flat and variable schedules is applied in some cases.

In one of the few practical system-wide congestion pricing studies, Kristoffersson (2013) compared a flat toll to a step toll regime in the toll ring of the Stockholm city network and estimated that the social benefit would be slightly higher for the step toll than for the flat toll, while the average reduction in flow was found to be somewhat larger for the flat toll. The study showed that the flat toll could still be a comparable option to the step toll. However, the step toll was found to be superior in suppressing traffic in peak periods.

Time-based variable pricing has been studied in the profit-oriented road pricing context as well. For instance, de Palma et al. (2008) developed a dynamic simulation model to study the impacts of employing temporally variable tolls versus flat tolls for both publicly and privately-run road systems on a small

* Corresponding author. Tel.: +1 530 204 8576.
E-mail addresses: omrouhani@ucdavis.edu (O.M. Rouhani), dniemeier@ucdavis.edu (D. Niemeier).

http://dx.doi.org/10.1016/j.jtrangeo.2014.03.017
0966-6923/

network. They found that a time-varying toll regime performed better than the public and private step-tolling schemes, and that step tolling led to higher welfare gains but lower profits than flat tolling. Lam (2004) has shown the merits of implementing a temporally variable toll schedule to maintain a specified level of service in the case of value pricing. In the case of distance-based road pricing, May and Milne (2000) have also shown the inferior effectiveness of flat tolls compared with cordon, time-spent-in-congestion, and time-spent-traveling pricing schemes. Burris and Sullivan (2006) analyzed the costs and benefits of variable pricing for two projects: the QuickRide high occupancy/toll (HOT) lanes in Texas and the SR-91 express lanes in California. The study estimated that the benefit/cost ratios were 1.5 and 1.7, respectively. Despite the effectiveness of temporal variation in tolls, the distributional and spill-over effects on other regions might become a problem politically (Larsen, 1995).

Recent studies have examined the application of responsive tolls, which vary in real time as functions of traffic conditions. Yin and Lou (2009) developed two approaches for setting these tolls. Lou et al. (2011) formulated the toll optimization problem to set responsive tolls. The main focus of the aforementioned studies is the variation of tolls in different time periods. One other current stream of research involves analysis of implementing non-linear pricing for tolling (Wang et al., 2011). Here, the applied pricing schemes are functions of road usage, but the use of fixed functions are suboptimal to the use of tolls optimized according to different traffic flow patterns. In fact, the pricing is limited to a function (s) of flow on each road. This function could remain the same for all the road segments. Although the assumption in Wang et al.'s (2011) study is that tolls could vary by demand (or consequently by time), tolls are based on predetermined functions and might not result in optimal road-usage levels.

Less research has been conducted on how tolls could/should be spatially varied to improve traffic performance. Some studies have investigated the system-wide variation in tolls, but not necessarily the spatial patterns of tolls, especially not along a corridor (Fang et al., 2011). For example, a number of researchers have studied second-best pricing schemes extensively, which assume that a subset of links in a network is eligible for tolling (Johansson-Stenman and Sterner, 1998; Zhang and Ge, 2004; Verhoef, 2007). Compared to a first-best pricing based on marginal costs (Yang and Huang, 1998), a second-best pricing is easier to implement and can be more cost/benefit effective. The second-best pricing has usually been applied to smaller networks due to its complexity. These studies typically have not focused on the toll profile, the variability of tolls on a corridor, or even the time variability of tolls.

In one of the rare studies examining spatially variable tolls, Kristoffersson and Engelson (2011) showed that, in the case of the Stockholm toll ring, the social benefit could be increased if the charges at different tolling stations were different from each other, especially for rates imposed for inbound and outbound traffic. In practice, flexible toll rates are only applied temporally and during periods of high traffic volume; examples are I-15 in San Diego (Brownstone et al., 2003; HNTB Corporation, 2006) and the Maryland 200 (MTA, 2013). An exception is Highway 407 in Toronto, Canada, where the toll rate is slightly higher for the regular zone than for the light zone (Toronto 407 ETR, 2013). But this slight difference, which is applied for only two zones, could not capture the potential of a full spatially variable tolling system that applies variable rates on each segment along a road. However, spatial variation could reduce congestion and increase profits effectively just as well as temporal variation.

This research investigates the effects of implementing flat versus variable tolls, both temporally and, more specifically, spatially for recurrent traffic congestion conditions. The effects are examined for both profit-maximization and social-optimal tolls and for peak versus off-peak periods. Specifically, our research goals are to determine spatial patterns of optimal variable pricing along a road corridor, to analyze the effects of implementing variable tolls on profit maximization and social welfare optimization, to determine a reasonable relationship(s) between road usage and optimal toll values, and to evaluate the virtues of a spatial pricing scheme. From these analyses, we can offer recommendations with respect to the trade-offs of using variable versus flat tolling. We use the Fresno, California, road network as our empirical platform and examine three different segments of road corridors using different pricing schemes.

2. Model framework

The bi-level programming concept has been applied to the problems of social welfare optimization and profit maximization (Yin, 2000). The upper-level decision-making process reflects policy makers' (social planners') decisions to choose the toll values that optimize social welfare, as the goal of city officials or policy makers is to maximize total social welfare. As a proxy to social welfare maximization, another objective function could be used, such as total travel time, total fuel consumption, system-wide emissions, total profits from pricing, or a summation of these functions expressed in monetary terms, etc. Here, we used total travel time as the objective function. Following the policy makers' decision, users of the network would make their travel choices (travel or not/mode/route). In the second problem, profit maximization, the upper level reflects road owners' (or even policy makers') desire to maximize profits by setting the proper tolls. As in the first problem, users choose among the streets (paths) or modes resulting from the upper-level decision.

Each of the problems, specifically the lower-level parts of the problem, can be modeled with user equilibrium (UE) (Sheffi, 1984), which can quantify the system-wide effects of any kind of pricing. There are, however, other modeling forms that can also be used (Chu et al., 2012; Mun et al., 2003). In general, many of these other forms are limited to a specific problem or they are unable to reflect the system-wide impacts of any decision.

2.1. Social welfare optimization (system optimal)

The goal of policy makers, city officials, and social planners is to maximize social welfare (SW) by setting proper tolls (decision variable) on a limited number of links. As discussed before, a common objective function is total travel time minimization ($\sum(t_{ij}(x^*_{ij})\cdot x^*_{ij})$) or the sum of travel times on each link multiplied by its equilibrium flows. The travel demand for the link $i-j$ – $x^*_{ij}(\tau)$ – is a function of a vector of tolls, based on the solutions of the lower-level problem. The policy maker's problem can be formulated as follows:

$$(Max\ SW)\underset{\tau_{ij}}{Min}\sum_{(i,j)}(t_{ij}(x^*_{ij}).x^*_{ij}) \tag{1}$$

where x^* is the solution of the user equilibrium (UE) problem with fixed demand in the transportation network based on the following mathematical problem (the solution of the lower-level problem is based on the UE problem (Sheffi, 1984)):

$$(UE)\ \underset{x}{Min}\int_0^{x_{ij}} C_{ij}(u)du \tag{2}$$

$$s.t.:\sum_{p\in p_{ks}} x^{ks}_p = d^{ks},\ \forall(k,s)\in P \tag{3}$$

$$x^{ks}_p \geqslant 0,\ \forall p\in P_{ks},\ \forall(k,s)\in P \tag{4}$$

$$x_{ij} = \sum_{k,s\in P}\ \sum_{p\in P_{ks}} x^{ks}_p\cdot\delta^{ks}_{ij,p},\ \forall(i,j)\in A \tag{5}$$