Regional Science and Urban Economics 42 (2012) 166–176

Contents lists available at SciVerse ScienceDirect

Regional Science and Urban Economics

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

Give or take? Rewards *versus* charges for a congested bottleneck ☆

Jan Rouwendal *,[1], Erik T. Verhoef [1], Jasper Knockaert

Department of Spatial Economics, VU University Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands

ARTICLE INFO

Article history:
Received 2 September 2010
Received in revised form 25 August 2011
Accepted 26 August 2011
Available online 3 September 2011

JEL classification:
R41
R48
D62

Keywords:
Traffic congestion
Road pricing
Subsidies
Rewards
Bottleneck model

ABSTRACT

This paper analyzes the possibilities to relieve traffic congestion using subsidies instead of Pigouvian taxes, as well as revenue-neutral combinations of rewards and taxes ('feebates'). The model considers a Vickrey–ADL model of bottleneck congestion with endogenous scheduling. With inelastic demand, a fine (time-varying) reward is found to be equivalent to a fine toll, and to a continuum of combinations of time-varying tolls and rewards, including fine feebates. When demand is price-sensitive, a reward becomes less attractive from the efficiency viewpoint, because it attracts additional users to the congested bottleneck. As a result, both the second-best optimal fraction of rewarded travelers in the scheme, and the relative efficiency that can be achieved with it, decrease when demand becomes more elastic. Our analytical and simulation results for coarse schemes reveal that a coarse reward is less effective than a coarse feebate, which is itself less effective than a coarse toll. The most efficient coarse system is the step toll, which is also allowed to be positive in the fringes of the peak. Despite the smaller efficiency gains, rewards and feebates may be attractive to use in circumstances where public and political acceptability of tolling is especially low, so that its implementation is unlikely, including the temporary use of price incentives in case of road works and large-scale events.

1. Introduction

Although the idea of road pricing has been around for a long time, and its popularity seems to be on the rise with successful introductions in cities like London, there often still exists fierce resistance against introducing a price for a commodity – access to public roads – that has been freely available for such a long time. Consequently, there are many examples of road pricing proposals that do not survive the planning stage, and that are dropped for political reasons before implementation commences. In cases where temporary changes in supply or demand conditions (*e.g.*, road works, large-scale events) cause severe congestion on a more short-lived basis, the public and hence political acceptability of road pricing may even be smaller than otherwise, notwithstanding the greater potential gains, as it may be perceived to be especially unfair to regular users who are already confronted with longer travel times.

Despite the clear economic case for marginal cost pricing, it may therefore be helpful to explore alternative, more acceptable possibilities to alleviate congestion, other than archetypical pricing solutions. One possibility would be the use of subsidies or rewards, instead of the penalty of tolling. Even though most economists would hasten to point out that positive marginal external costs would call for a positive Pigouvian tax, that subsidies may induce all sorts of perverse incentives, that taxes needed to raise revenues to finance rewards are likely to cause or aggravate distortions elsewhere in the economy, and that a net reward is likely to attract additional users whereas a net reduction is desired, there may be reasons to consider rewards nevertheless. Ultimately, a reward system may be less effective and efficient in combating congestion than a tolling system, but if the latter is unfeasible for political reasons, a more relevant comparison is between a reward system and the absence of any financial incentives. Such considerations have motivated proposals for, for example, combinations of pricing and rationing (Daganzo, 1995), revenue neutral 'credit-based' congestion pricing (Kalmanje and Kockelman, 2005), tradable driving permits (Verhoef et al., 1997), exemptions from paying tolls (Daganzo and Garcia, 2000) and so-called Fast and Intertwined Regular (FAIR) lanes (De Corla-Souza, 2000).

It is the purpose of the present paper to explore the possibilities of rewards in traffic congestion management. We do so by studying various types of rewards in the context of the well-known bottleneck model (Vickrey, 1969; Arnott et al., 1990, 1993). The idea for this paper stems from a Dutch policy experiment called 'Spitsmijden' (Avoiding the Peak), documented in Knockaert et al. (2009) and Bliemer and Van Amelsfort (2010). The purpose of this experiment was to gain insight into the potential of positive financial incentives in the management of peak traffic congestion. To that end, regular users of a given highway (the A12 between Zoetermeer and The Hague) could earn rewards

☆ We thank two anonymous reviewers for helpful comments. Financial support from TRANSUMO is gratefully acknowledged; Verhoef furthermore acknowledges support from the ERC (AdG Grant # 246969 OPTION).

* Corresponding author. Tel.: +31 20 5986093; fax +31 20 5986093.
E-mail address: jrouwendal@feweb.vu.nl (J. Rouwendal).
[1] Affiliated to the Tinbergen Institute, Gustav Mahlerplein 117, 1082 MS Amsterdam, The Netherlands.

0166-0462/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.regsciurbeco.2011.08.011

ranging from 3 to 7 Euros for avoiding the A12 during the morning peak (7:30–9:30 am). Given that participation was voluntary, it is no surprise that relatively strong behavioral impacts from rewarding were found among participants: one might expect a greater probability of participation, the more easily an individual can change behavior and thus earn rewards. According to the automated vehicle observations, the participants reduced their aggregate number of peak hour trips by some 50% when doing so was rewarded. Given that departure time adjustments were, by far, the most popular behavioral response to the incentive (roughly accounting for four-fifths of all adjusted trips), it seems appropriate to use a dynamic model, with endogenous departure time choices, for the present paper.

Having rewards in place when taxes would be first-best optimal makes the policy second-best in nature. In addition, a second and quite different type of second-best considerations will arise in our analyses because we will not only consider smoothly time-varying 'fine' tolls and rewards, but also more realistic 'coarse' systems.[2] These involve a limited number of discrete toll or reward levels during the peak. Our analysis thus fits in a much wider literature on second-best road congestion pricing reviewed in, among many others, Small and Verhoef (2007). Probably closest to our paper are earlier studies that looked into the efficiency of so-called 'coarse tolls' and 'multi-step tolls' in the bottleneck model. Arnott et al. (1990, 1993) were the first to consider this type of problem, and found that the relative efficiency of such measures is above that of 'uniform tolling', which entails a constant toll level throughout the peak, but considerably lower than that of fine tolling. The welfare gains are around 50% for the coarse (single-step) toll, a result that was later confirmed by Chu (1999) in a more elaborate model. Quite intuitively, Laih (1994) showed that the efficiency of step-tolls increases with the number of steps. Xiao et al. (2009) explore coarse tolling with heterogeneous users, and find that this improves the relative efficiency of coarse tolling.

An important aspect of dynamic equilibria with step tolls in the bottleneck model, already identified by Arnott et al. (1990, 1993), is that it entails mass departures in the second part of the peak, since the equilibrium condition of constancy of generalized prices over times requires that a discrete drop in the toll level is matched by a discrete increase in (expected) travel time. The implied mass departures complicate the analytical treatment of step tolls. Lindsey et al. (2010) show that when drivers can halt before passing a toll gantry, in anticipation of a drop in the toll level, there will be no such mass departures in equilibrium. In the present paper, we will ignore this 'braking' behavior, and stick to the original model of coarse tolling in the bottleneck model, as developed by Arnott et al. (1990).

The main difference with earlier studies of step tolling is our focus on rewards. More specifically, we will be looking at rewards, involving non-negative subsidies for all users; 'feebates', involving budget-neutral combinations of taxes and subsidies producing a zero net revenue for the regulator; and, as an important benchmark, tolls, involving non-negative taxes for all users. For all cases to be considered, we assume that tolls and rewards will be faced only by users of the bottleneck, typically in function of the moment of passing it. This means that no rewards will be assumed to be given to individuals who refrain from using the bottleneck altogether, *e.g.* by staying at home or using a public transport option. Although such subsidies would be possible in reality, and would have an impact if demand for the use of the bottleneck is price-sensitive, we dismiss the possibility for two reasons. The practical reason is that such policies would usually be considered to be intolerably expensive in reality, at least if all non-users would qualify for a reward. The analytical reason is that a solid treatment would require the inclusion of a second transport mode in the model. This would harm the transparency of our results when introducing a third type of second-best considerations, in particular if this mode were assumed to be inefficiently priced as is usually the case in reality.

Given that congestion entails an external cost, one would expect that tolls outperform rewards and feebates in terms of efficiency. It is therefore not primarily the ranking of the different policies that we are most interested in, but rather assessing their relative efficiency. Our analysis should thus give insight into the circumstances under which feebates or rewards may offer a worthwhile alternative to tolling in the management of congestion; and if so, how the policy should be designed to maximize its efficiency.

The paper is organized as follows. In the next section we introduce the basic bottleneck model, and discuss the equivalence of fine tolling, fine feebates and fine rewards when demand is completely inelastic. In Section 3 we study the properties of fine schemes in the context of price-sensitive demand. Section 3 thus focuses on the first second-best aspect: the use of rewards where tolls would be optimal. We will find that with rewards in place instead of conventional tolls, it is not optimal to apply the financial incentive throughout the entire peak, but instead it is preferable to reward only a certain a fraction of drivers, so that rewards can be earned only during certain parts of the entire peak period. We thus derive the optimal fractions of rewarded commuters.[3] We will impose that there be a dynamic equilibrium in the sense that all rewarded and non-rewarded drivers should face the same generalized equilibrium price, so that the optimal fraction is realized without making an explicit *ex ante* subdivision between drivers who do or do not potentially qualify for a reward. Next, in Section 4, we move a step closer to practical policy experiments and introduce the other second-best aspect that we are interested in, by analyzing coarse systems where tolls or rewards only change in a few discrete steps during the peak. Finally, Section 5 concludes.

2. The basic bottleneck model and some variants on fine tolling

The analysis in this section uses the basic bottleneck model, in which during the peak a homogeneous group of users of a given size N has to pass a bottleneck with given capacity s. Given that we consider only one single route, the free-flow travel time can be and will be set equal to zero without loss of generality, so that in absence of a queue, a traveler departs from home, passes the bottleneck, and arrives at work all at the same instant. However, a travel delay of $Q(t_d)/s$ is incurred if at the moment t_d of departing from home and joining the queue, its length is $Q(t_d)$. The rate of change of Q when positive is equal to the difference between the departure rate from home, and the capacity of the bottleneck. We consider the following, conventional, cost function; usually associated with Vickrey (1969) and Small (1982):

$$c(t) = \alpha \cdot T(t) + \beta \cdot \max\{0, t^* - t\} + \gamma \cdot \max\{0, t - t^*\}. \tag{1}$$

The Greek characters represent positive parameters, where α denotes the 'value of time'; β indicates the unit shadow cost of schedule delay early, and γ that of schedule delay late. T is the travel delay incurred in the queue before passing the bottleneck, t the arrival time at the bottleneck's exit, and t^* the preferred arrival time, at which schedule delay costs are naturally zero.

[2] We will not be using the terminology of "third-best" or "fourth-best" problems to distinguish schemes with increasing numbers of constraints. It turns out that doing so might easily become confusing rather than helpful. The term "second-best" will therefore be used to denote any policy that seeks to optimize social surplus under at least one binding constraints on the pricing instrument.

[3] This type of optimization was not present in the Spitsmijden experiment mentioned above, where rewards were given only to a limited number of participants. Another difference is that in the experiment, rewards were also given to participants who switched to alternatives such a public transport, whereas in this paper rewards will be assumed to be given only to drivers who pass the bottleneck during the time intervals when a reward applies.