

# STEP TOLLING WITH BOTTLENECK QUEUING CONGESTION<sup>1</sup>

C. Robin Lindsey  
Sauder School of Business  
University of British  
Columbia  
2053 Main Mall  
Vancouver, BC  
Canada V6T 1Z2  
robin.lindsey@sauder.ubc.ca

Vincent A.C. van den Berg<sup>\*</sup>  
Department of Spatial  
Economics  
VU University Amsterdam  
De Boelelaan 1105  
1081 HV Amsterdam  
The Netherlands  
vberg@feweb.vu.nl

Erik T. Verhoef<sup>\*\*</sup>  
Department of Spatial  
Economics  
VU University Amsterdam  
De Boelelaan 1105  
1081 HV Amsterdam  
The Netherlands  
everhoef@feweb.vu.nl

This version: 11/07/12

**Key words:** Congestion pricing, step tolls, bottleneck model, Vickrey model, departure time choice, braking  
**JEL codes:** D62, R41, R48

## **Abstract**

*In most dynamic traffic congestion models, congestion tolls must vary continuously over time to achieve the full optimum. This is also the case in Vickrey's (1969) 'bottleneck model'. To date, the closest approximations of this ideal in practice have so-called 'step tolls', in which the toll takes on different values over discrete time intervals, but is constant within each interval. Given the prevalence of step-tolling schemes they have received surprisingly little attention in the literature. This paper compares two step-toll schemes that have been studied using the bottleneck model by Arnott, de Palma and Lindsey (1990) and Laih (1994). It also proposes a third scheme in which late in the rush hour drivers slow down or stop just before reaching a tolling point, and wait until the toll is lowered from one step to the next step. Such 'braking' behaviour has been observed in practice. Analytical derivations and numerical modelling show that the three tolling schemes have different optimal toll schedules and reduce total social costs by different percentages. These differences persist even in the limit as the number of steps approaches infinity. Braking lowers the welfare gain from tolling by 14% to 21% in the numerical example. Therefore, preventing or limiting braking seems important in designing step-toll systems.*

<sup>\*</sup> Corresponding author.

<sup>\*\*</sup> Affiliated to the Tinbergen Institute, Roetersstraat 31, 1018 WB Amsterdam.

---

<sup>1</sup> This the post-print the article published in Journal of Urban Economics 72(1), see <http://dx.doi.org/10.1016/j.jue.2012.02.001> for the published version or the journal website <http://www.journals.elsevier.com/journal-of-urban-economics/>



## 1. Introduction

In most dynamic traffic congestion models, congestion tolls must vary continuously over time to achieve the full social optimum. This is also true of Vickrey's (1969) 'bottleneck model' of the morning rush hour in which congestion takes the form of queuing at a bottleneck and drivers choose their departure times to minimize their travel cost. Yet no existing road pricing scheme has such sophisticated, continuously time-varying tolls. The closest approximations of this ideal have so-called 'step tolls' in which the toll takes on different values over discrete time intervals, but is constant within each interval.<sup>2</sup> Given the prevalence of step-tolling schemes, they have received surprisingly little attention in the scholarly literature.

Two step-tolling schemes for the bottleneck model have been studied: one by Arnott, de Palma and Lindsey (1990, 1993) and the other by Laih (1994, 2004). For brevity, we will refer to their models and treatments as the 'ADL model' and the 'Laih model' respectively. A fundamental problem of these two models is that they ignore the fact that as the end of the tolling period approaches, drivers have an incentive to stop before reaching the tolling point and wait until the toll ends. Waiting has indeed been observed in practice. From 1975 to 1994, Singapore operated an Area Licensing Scheme in which motorists had to purchase a license to enter the Restricted Zone. Towards the end of the charging period vehicles were seen queuing near gantry points waiting for the charge to end.<sup>3</sup> Singapore introduced Electronic Road Pricing in 1998, and for the first few years tolls were changed in steps in half-hour intervals. Some of the steps were rather large, and motorists were observed slowing down before they

---

<sup>2</sup> Many road-pricing schemes feature step tolls. The London congestion pricing scheme has a fixed toll from 7 am to 6 pm on weekdays (see the scheme's website at [www.tfl.gov.uk/tfl/roadusers/congestioncharge/whereandwhen/](http://www.tfl.gov.uk/tfl/roadusers/congestioncharge/whereandwhen/); accessed 7 January 2012). Singapore's Electronic Road Pricing system features multi-step toll schedules on expressways and arterials. For instance, at the Bugis-Marina Centre (Nicoll Highway), the toll has 8 steps during the weekday morning peak (rates for 8 November 2011 to 5 February 2012, at [www.onemotoring.com.sg/publish/onemotoring/en/on\\_the\\_roads/ERP\\_Rates.html](http://www.onemotoring.com.sg/publish/onemotoring/en/on_the_roads/ERP_Rates.html), accessed on 7 January 2012). The toll scheme in Stockholm has 10 steps between 6:30 am and 6:29 pm ([www.transportstyrelsen.se/en/road/Congestion-tax/Congestion-tax-in-stockholm/Times-amounts/](http://www.transportstyrelsen.se/en/road/Congestion-tax/Congestion-tax-in-stockholm/Times-amounts/), accessed 7 January 2012). And the westbound express-lane on State Route 91 in California has eight steps in the toll during the morning peak on Monday to Thursday with the toll ranging from \$1.30 to \$4.60 (toll schedule effective 1 January 2012, [www.octa.net/91\\_schedules.aspx](http://www.octa.net/91_schedules.aspx), accessed on 7 January 2012).

<sup>3</sup> Png, Olszewski and Menon (1994) show that in 1992 few motorists purchased a license in the last 15 minutes before the morning charging period ended.

reached a toll gantry to take advantage of a toll decrease.<sup>4</sup> In 2003, five-minute graduated rates were introduced between some half-hour periods to deter exactly this behaviour.<sup>5</sup>

Several other examples of waiting can also be mentioned. The Stockholm congestion charging scheme employs step tolls. During the trial in 2006, motorists were seen waiting on the roadside for the toll to drop.<sup>6</sup> In July 2010 congestion pricing was introduced on the San Francisco-Oakland Bay Bridge with a peak-hour toll of \$6 from 5-10 am and 3-7 pm, and a non-peak hour toll of \$4. Lee and Frick (2011) report that motorists have been slowing down and parking illegally on the shoulders to wait for the toll to decrease. According to Linda Lee of the Bay Area Toll Authority this behaviour has been reported to the Metropolitan Transportation Commission at several public meetings, but has not been systematically recorded. Speeding up before a toll increase, in contrast, has not been observed.<sup>7</sup> We explain below why this may be so, and why we therefore do not allow for speeding in the model.

In addition to their neglect of braking, the ADL model and Laih model have other, distinct problems. The ADL model is limited to a single-step toll that is either ‘on’ or ‘off’. Moreover, equilibrium in the model exists only if a mass of drivers departs when the toll comes off, with individual positions in the mass determined randomly. Mass departures are problematic in practice, and empirical evidence for them is scant.

Unlike the ADL model, Laih (1994, 2004) considers multi-step tolling. Laih (2004) also avoids the problem of mass departures by assuming that drivers who choose to pass the bottleneck after a certain tolling period can wait on a set of secondary lanes without impeding other drivers who do pass the bottleneck in that tolling period. However, his analysis is limited in scope because many roads simply do not have separate lanes, or even shoulders,

---

<sup>4</sup> David Boyce recalls seeing motorcycles parked on the shoulder waiting for a toll to drop (personal communication, 12 November 2010).

<sup>5</sup> See Chew (2008) and the Electronic Road Pricing web site at [www.lta.gov.sg/motoring\\_matters/motoring\\_erp.htm](http://www.lta.gov.sg/motoring_matters/motoring_erp.htm) (accessed 14 October 2010). The graduated rates may not have entirely solved the problem. According to Mr. Tan Jin Thong, president of the National Safety Council of Singapore, it is still “common for motorists to use the shoulder to wait for Electronic Road Pricing gantries to cease operation” (as quoted by Chee, 2011). Overtaking using the shoulder is also mentioned in the same article, but not for the purpose of avoiding a toll increase but rather to avoid long waits.

<sup>6</sup> Jonas Eliasson, personal communication, 1 September 2011; Fosgerau (2011) also mentions anecdotes about cars waiting at the roadside until the toll has dropped.

<sup>7</sup> Personal communication, 15 July 2011. The slowing down and stopping also received attention from the press: see Eskenazi (2010) and Mercury News (2010).

where motorists can pull out of the traffic stream. As explained later, without such separate queues, the ADL model would be appropriate, and Laih's analysis would be inconsistent with his parameter-value assumptions.

In this paper we make two contributions. First and foremost, we analyze step-tolling in the bottleneck model under the assumption that drivers do stop and wait for a toll to decrease if the cost of waiting is less than the amount of toll saved. In contrast to the ADL and Laih models, the equilibrium is behaviourally consistent in that it accounts for the possibility that drivers can reduce their travel costs by braking. Unlike in the ADL model, the equilibrium is free of mass departures. And unlike in Laih's formulation, the equilibrium does not require the presence of two sets of traffic lanes. One could object that stopping may be dangerous and poses the risk of a traffic citation or fine. However, stopping does occur in practice, as noted above. Moreover, in lieu of coming to a full and sudden stop, drivers can slow down in the kilometres leading up to a toll point, and time their passage of the toll point just after the toll drops. The analytics of the model are the same whether drivers slow down in this smoother manner, or stop abruptly. The welfare effects are also similar. But empirically, smoothed slowing down is of course much harder to detect than abrupt braking, meaning that the observations from actual toll applications discussed above may in fact underestimate the importance of the problem. To allow for both possibilities, we will refer to the new model as the 'braking model'. A notable feature of equilibrium in the braking model is that the bottleneck is unused while drivers are stopped. We will show to what extent this loss of effective capacity undercuts the efficiency gains from tolling in comparison with the ADL and Laih models.

Our second contribution is to extend the ADL model to multiple toll steps. Doing so generalizes the ADL model, and makes it more empirically relevant since a number of tolling schemes have multiple steps. In the ADL model with multiple toll steps, a mass of drivers departs after each toll reduction. This equilibrium is arguably less problematic than the equilibrium with one step because the numbers of drivers in each mass departure is smaller, and the incentives for drivers to stop before each toll reduction are weaker. After developing the braking model, and extending the ADL model to multiple steps, we compare the equilibria

in the ADL, Laih, and braking models as a function of model parameters and the number of toll steps. We also assess the efficiency loss from step tolling in the braking model due to gaps in usage of the bottleneck. For the base calibration of our numerical model, this loss is around 20% of the gains that the ADL and Laih models predict, suggesting that toll-induced braking is important to take into account in the evaluation of congestion pricing schemes.

While our analysis of step tolling builds significantly on the literature it has some limitations. One limitation is that it ignores drivers' incentives to speed up in order to pass through the tolling point just before a toll increase. At first glance, this might seem to be the logical mirror image of braking before a toll decrease; possibly with a comparable behavioural and economic significance. There are, however, various reasons why we chose to ignore such speeding in our analysis. First, whereas we found multiple examples confirming the relevance of braking in real step-toll applications, we found no reference to toll-induced speeding, suggesting it is empirically less relevant. Second, there is a strong asymmetry in a driver's ability to slow down and to speed up.<sup>8</sup> If a driver slows down, all following drivers can be expected to slow down as well, typically with some delay that increases with the distance between the cars.<sup>9</sup> By contrast, in congested traffic there is not much that a driver can do to make the driver ahead speed up, not in the least place because that driver is similarly constrained. A driver who speeds up regardless will cause a rear-end collision. Third, while it is straightforward for drivers to stop and wait for a toll to drop, it is difficult to time arrival at a tolling point just before the toll increases—especially since it requires synchronisation of personal and system clocks.<sup>10</sup> Finally, and more pragmatically, incorporating speeding into the standard bottleneck model is not straightforward since drivers are assumed to travel at free-flow speed except for waiting in a space-less ('vertical') queue. To model speeding, it would therefore probably be better to use a flow congestion model. Yet doing so would have

---

<sup>8</sup> This asymmetry is clearly reflected in typical car-following models, which specify a driver's acceleration and speed as a function of the distance from and speed of the car in front (the 'leader' in car-following terminology); and not as a function of the distance or speed of the driver behind (see, for example May, 1990). Leaders thus determine followers' behaviour in this model, while a reverse effect is absent.

<sup>9</sup> Hydrodynamic models of traffic congestion feature these delays in response time through the speed of so-called shockwaves (for a review, see Daganzo, 1997).

<sup>10</sup> Jonas Eliasson, personal communication, 1 September 2011.

several drawbacks. First, flow congestion models are less tractable than the bottleneck model and require numerical solution.<sup>11</sup> Second, the bottleneck model is the most widely used dynamic economic model of traffic congestion (Small and Verhoef, 2007). Finally, the step-toll studies on which we build use the bottleneck model.

Another limitation of our analysis is that we consider only toll schedules with the same numbers of steps ‘up’ and steps ‘down’. Still, we do allow the toll levels to vary freely rather than requiring the level of each step-up toll to match the level of the corresponding step-down toll.<sup>12</sup>

The paper is organized as follows. Section 2 begins by presenting the basic bottleneck model, and solving for the no-toll equilibrium and social optimum. It then presents a graphical treatment of the one-step toll in the ADL, Laih and braking models. Section 3 extends the analysis to multi-step tolls and provides a formal derivation of optimal toll schedules for each model. Section 4 provides a brief numerical comparison of the models and assesses the efficiency loss due to braking in the braking model. Section 5 concludes and discusses ideas for future research.

## 2. No-toll equilibrium, first-best optimum, and single-step tolls

### 2.1. No-toll equilibrium and first-best optimum

Both the no-toll equilibrium and the first-best or social optimum of the bottleneck model are described in detail in Arnott, de Palma and Lindsey (1990, 1993, 1998), as well as in various textbooks such as Small and Verhoef (2007). We will therefore be brief, while keeping the exposition self-contained.

---

<sup>11</sup> Chu (1999) analyzes one-step tolling using flow congestion, although he does not consider braking. His one-step toll generally attains a larger fraction of the first-best gain than in the bottleneck model. The fraction depends on the specification of the travel time function and the number of users, whereas in the bottleneck model the fraction is constant as we will show. To numerically solve his model, Chu uses a cobweb procedure. Ge and Stewart (2010) use the cell transmission model which combines flow-congestion on each link and queuing upstream when the inflow rate exceeds link capacity. They consider several toll schedules including step tolls (with a rapid, but finite, rate of increase and decrease in the toll to maintain continuity). Analytical solutions cannot be derived and they use variational inequality methods to obtain numerical solutions. Similar to the bottleneck model, equilibria in their model feature surges in demand before and after tolling is imposed. Unlike in the ADL and Laih model equilibria, during the toll period the inflow rate typically drops below link capacity. Unlike in our paper they ignore the possibility of braking or speeding, and they do not derive optimal tolls but rather compare some pre-specified toll schedules.

<sup>12</sup> Optimal toll schedules in the Laih model and the braking model turn out to be symmetric anyway.

In the bottleneck model there is a continuum of  $N$  identical individuals who travel alone by car from a common origin to a common destination connected by a single road that is subject to bottleneck queuing congestion. Free-flow travel time is normalized to zero without loss of generality. Without a queue, an individual thus departs ‘from home’, passes the bottleneck, and completes her trip by arriving ‘at work’ all at the same moment. If the driver leaves at time  $t$  and there is a queue, she encounters a queuing delay,  $T(t)$ , equal to

$$(1) \quad T(t) = \frac{1}{s} \cdot \left( \int_{t_s}^t r(u) du - s(t - t_s) \right),$$

where  $t_s$  is the time at which the queue began,  $s$  is bottleneck capacity, and  $r(\cdot)$  is the departure rate as a function of time. The driver arrives at time  $t_a = t + T(t)$ .

Travellers incur a unit cost of  $\alpha$  from queuing delay. They also have a preferred arrival time of  $t^*$ , and incur a so-called ‘schedule delay cost’ if they arrive earlier or later. The unit cost of arriving early is  $\beta$ , and the unit cost of arriving late is  $\gamma$ . The ‘generalized cost’ of a trip,  $c$ , is the sum of queuing delay cost and schedule delay cost:

$$(2) \quad c(t) = \alpha \cdot T(t) + \beta \cdot \text{Max}(0, t^* - t_a) + \gamma \cdot \text{Max}(0, t_a - t^*).$$

Let  $\tau(t + T(t))$  denote the toll paid, if any, for an arrival at time  $t + T(t)$ .<sup>13</sup> The generalized price (hereafter referred to as the ‘price’ for brevity) includes the generalized cost and the toll:

$$(3) \quad p(t) = c(t) + \tau(t + T(t)).$$

In equilibrium, all drivers experience the same trip price. During intervals when there is either no toll or a toll that is constant, the queue must evolve so that  $c(t)$  is constant. This requires a departure rate  $s \cdot \alpha / (\alpha - \beta) > s$  for early arrivals, and  $s \cdot \alpha / (\alpha + \gamma) < s$  for late arrivals.

Departures and arrivals take place over a time interval  $[t_s, t_e]$ . Since the bottleneck operates at capacity,

$$(4) \quad t_e = t_s + N / s.$$

Because the first and last drivers do not encounter a queue, their travel times are zero and their schedule delay costs therefore must be equal:

---

<sup>13</sup> The toll schedule can be defined as a function of departure time or arrival time. We define it as a function of the moment that the bottleneck is passed, which is the arrival time at the destination, as this simplifies the discussion of step tolling.



$$(5) \quad \beta \cdot (t^* - t_s) = \gamma \cdot (t_e - t^*).$$

Equations (4) and (5) together yield

$$(6a,b) \quad t_s = t^* - \frac{\gamma}{\beta + \gamma} \cdot \frac{N}{s}, \quad t_e = t^* + \frac{\beta}{\beta + \gamma} \cdot \frac{N}{s}.$$

The equilibrium generalized cost and price in the no-toll (NT) equilibrium are therefore

$$(7) \quad p_{NT} = c_{NT} = \delta \cdot \frac{N}{s} \quad \text{with} \quad \delta \equiv \frac{\beta \cdot \gamma}{\beta + \gamma}.$$

The total private cost (PC) and total social cost (TC) for the  $N$  travellers are equal:

$$(8) \quad PC_{NT} = TC_{NT} = \delta \cdot \frac{N^2}{s}.$$

At  $t_s$  or  $t_e$ , the schedule delay cost is  $\delta \cdot N/s$  and the travel delay cost is 0; for an arrival at  $t^*$ , the schedule delay cost is 0 and the travel delay cost is  $\delta \cdot N/s$ . Because schedule delay cost is linear in arrival time and the generalized cost must be constant over time, for the average user, the schedule delay and travel delay cost both equal  $\frac{1}{2} \cdot \delta \cdot N/s$ . Accordingly, half of the total costs are travel delay costs, and half are schedule delay costs.

Queuing delay is a pure deadweight loss in the bottleneck model: shortening the queue reduces travel time delays without raising schedule delays as long as flow through the bottleneck remains equal to  $s$ . First-best time-varying tolling, also called ‘fine tolling’, therefore eliminates queuing. To achieve this, the toll schedule is chosen to match the cost of travel delay for an arrival at the same moment in the no-toll equilibrium:  $\tau_{FB}(t) = \alpha \cdot Q_{NT}(t + T_{NT}(t))$ . The travel period is still given by eqn. (6a,b). The private cost of a trip remains unchanged, but the social cost of a trip is reduced by half:

$$(9) \quad p_{FB} = \delta \cdot \frac{N}{s}, \quad c_{FB} = \frac{1}{2} \cdot \delta \cdot \frac{N}{s},$$

where FB denotes the first-best optimum. The corresponding total costs are

$$(10) \quad PC_{FB} = \delta \cdot \frac{N^2}{s}, \quad TC_{FB} = \frac{1}{2} \cdot \delta \cdot \frac{N^2}{s}.$$

In the next three subsections we discuss how single-step (coarse) tolling works in the ADL, Laih, and braking models. We defer formal derivations of the equilibria until Section 3 where we provide derivations for multi-step tolls.

## 2.2. Coarse tolling: the ADL model

Arnott, de Palma and Lindsey (1990, 1993) were the first to consider step tolling in the bottleneck model. They focused on a single-step toll, or coarse toll in their terminology. The coarse toll is defined by its level,  $\rho$ , and the times at which it is turned on ( $t^+$ ) and off ( $t^-$ ). The qualitative properties of the optimal coarse toll depend on the relative magnitudes of  $\alpha$  and  $\gamma$ . If  $\alpha \geq \gamma$ , the coarse toll does not affect the timing of arrivals and the solution is straightforward. But if  $\alpha < \gamma$ , as Small (1982) and most later empirical studies have found, the arrival period is affected and the analytics are more complicated.

To see why this is the case, it is helpful to describe in detail the equilibrium with the coarse toll when  $\alpha < \gamma$ . Figure 1 does so by plotting cumulative departures and arrivals against time. The first driver travels at  $t'_s$  before  $t^+$  when the toll is switched on. Departures continue at a rate  $s \cdot \alpha / (\alpha - \beta)$  as in the equilibrium without toll as long as drivers pass the bottleneck before  $t^+$ . The first person to arrive after  $t^+$  and pay the toll must experience a queuing time that is  $\rho / \alpha$  shorter than the queuing time of the person arriving just before  $t^+$  who avoids the toll. At some time,  $t_0$ , departures therefore halt, and the queue starts shrinking. To maximize the reduction in queuing time,  $t^+$  and  $\rho$  are jointly chosen so that the queue reaches zero at  $t^+$ .

Next, there is a period that lasts until the toll is lifted, at  $t^-$ , during which all travellers pay the toll. (The possibility that drivers can brake and wait until the toll is lifted is ignored in the ADL model.) The equilibrium departure rate is  $s \cdot \alpha / (\alpha - \beta)$  for drivers who arrive early, and  $s \cdot \alpha / (\alpha + \gamma)$  for those who arrive late, again the same as in the equilibrium without toll. Parameters  $t^-$  and  $\rho$  are jointly chosen so that the queue reaches zero at  $t^-$ .

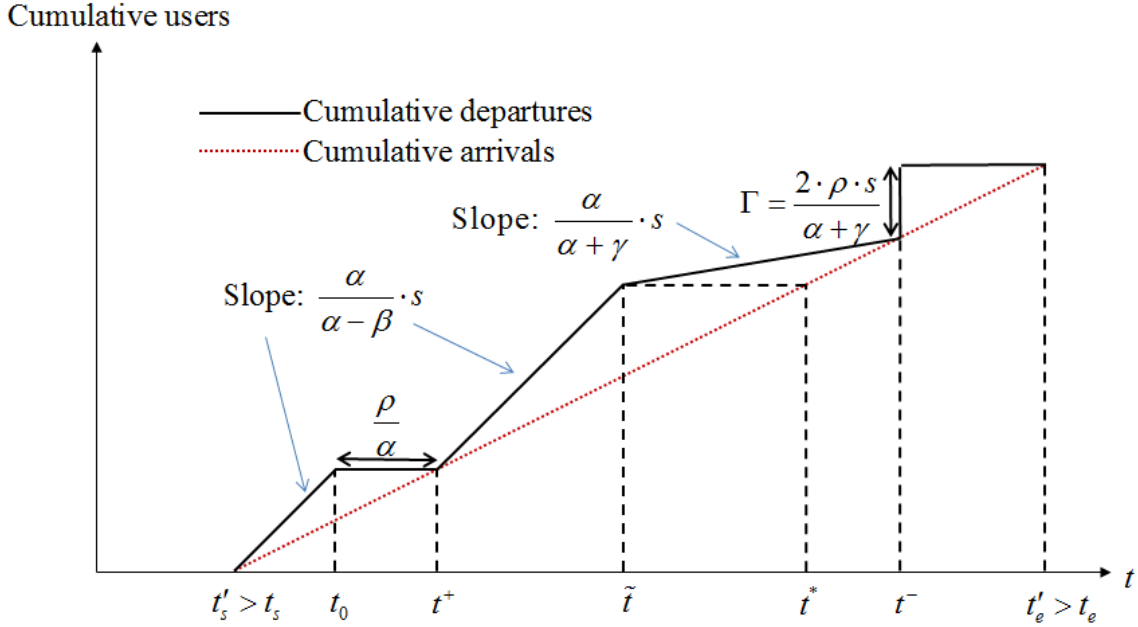


Figure 1. Equilibrium with a coarse toll: The ADL model for  $\alpha < \gamma$

The question now arises how drivers who pass the bottleneck after  $t^-$ , and pay no toll, can be equally well off as the last driver who did pay the toll but faced no queuing delay. Arnott, de Palma and Lindsey (1990) recognized that this requires a mass of individuals to depart at  $t^-$ , with individual positions in the mass determined randomly. The number of drivers in the mass,  $\Gamma$ , is such that the expected generalized cost for someone in the mass equals the deterministic price paid by the last traveller just before  $t^-$ . The unit cost of time spent in the queue is  $\alpha + \gamma$ , and the random amount of time spent in the queue is uniformly distributed between 0 and  $\Gamma/s$ . The expected additional generalized cost for someone in the mass is therefore  $\frac{1}{2} \cdot (\alpha + \gamma) \cdot \Gamma/s$ . Setting this cost equal to  $\rho$  gives  $\Gamma = 2 \cdot \rho \cdot s / (\alpha + \gamma)$ .

We can now explain why the coarse toll delays the start of travel if  $\alpha < \gamma$ . Consider a driver who departs at time  $t^- + \Delta$ , and assume the queue from the mass departure has not yet dissipated so that  $\Delta < \Gamma/s$ . The driver incurs a generalized cost of  $\gamma \cdot (t^- + \Delta - t^*) + (\alpha + \gamma) \cdot (\Gamma - \Delta \cdot s) / s$ . Setting this cost equal to the expected generalized cost of being in the mass one obtains

$$\gamma \cdot (t^- + \Delta - t^*) + (\alpha + \gamma) \cdot \frac{\Gamma - \Delta \cdot s}{s} = \gamma \cdot (t^- - t^*) + \frac{1}{2} \cdot \frac{\Gamma}{s} \cdot (\alpha + \gamma),$$

which simplifies to

$$(11) \quad \Delta \cdot \alpha = \frac{1}{2} \cdot \frac{\Gamma}{s} \cdot (\alpha + \gamma).$$

Equation (11) is consistent with the assumption  $\Delta < \Gamma/s$  only if  $\alpha > \frac{1}{2} \cdot (\alpha + \gamma)$ , or  $\alpha > \gamma$ .

If  $\alpha < \gamma$ , as empirical studies suggest, no one will depart after  $t^-$ , neither when the queue dissipates nor later.<sup>14</sup> The equilibrium depicted in Figure 1 then applies. Because the bottleneck operates at full capacity throughout the peak,<sup>15</sup> the peak has the same duration as the no-toll equilibrium and social optimum. But the timing of the peak,  $[t'_s, t'_e]$ , is now determined not by eqn. (5), but by the condition that the schedule delay cost of the first traveller equals the expected cost of a traveller in the mass:

$$(12) \quad \beta \cdot (t^* - t'_s) = \gamma \cdot (t^- - t^*) + \frac{1}{2} \cdot (\alpha + \gamma) \cdot (t'_e - t^-).$$

The solution works out to<sup>16</sup>

$$(13) \quad t'_s = t^* - \frac{\gamma}{\beta + \gamma} \cdot \frac{N}{s} + \frac{(\gamma - \alpha) \cdot \rho}{(\beta + \gamma) \cdot (\alpha + \gamma)}.$$

Given  $\gamma > \alpha$ , the peak begins later than with no toll. Since the first driver incurs only a schedule delay cost, and the expected price of a trip is the same for everyone, the expected price falls. The drop in price comes despite the fact that the shift in the peak causes total schedule delay costs to rise: the drop in travel delay costs outweighs it.

Now consider the case  $\alpha > \gamma$  which is shown in Figure 2.<sup>17</sup> Equation (11) then implies that  $\Delta < \Gamma/s$ . Contrary to the case with  $\alpha < \gamma$ , there is a final time interval after the mass departure during which drivers depart at rate  $s \cdot \alpha / (\alpha + \gamma)$ . This last cohort of drivers experiences a deterministic trip price, and the last driver in the cohort experiences no queuing

---

<sup>14</sup> This argument applies for any coarse toll, whether it is optimal or not.

<sup>15</sup> We use the term 'peak' for the period between the first and last arrival for ease of reference. As such, the bottleneck model does not have peak and off-peak periods, but rather a continuum of trip times that vary in their attractiveness.

<sup>16</sup> See Arnott, de Palma and Lindsey (1990, eqn. 14b).

<sup>17</sup> Daniel (2009) discusses the equilibrium with multi-step tolls given  $\alpha > \gamma$  in the context of airport congestion.

delay. Therefore, in equilibrium, the schedule delay costs for the first and last drivers are equal, condition (5) applies, and the timing of the peak is given by eqns. (6a,b).

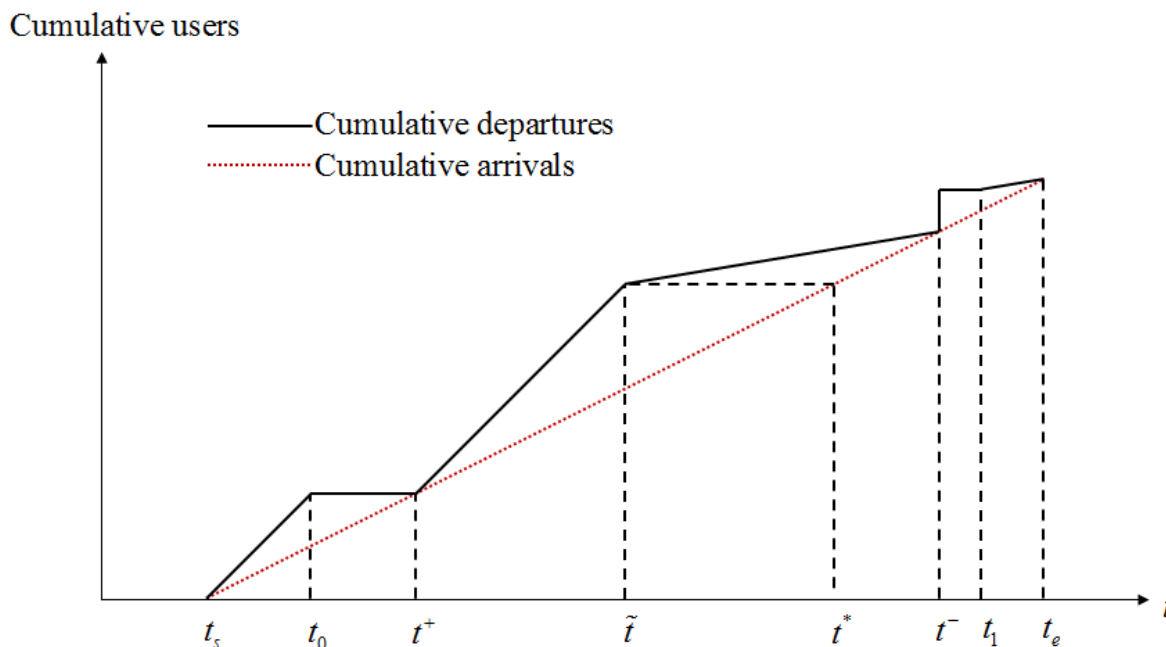


Figure 2. Equilibrium with a coarse toll: the ADL model for  $\alpha > \gamma$

The optimal coarse toll will be derived in Section 3 as a special case of the optimal multi-step toll of the ADL model. We now turn to consider the single-step toll in the Laih model.

### 2.3. Coarse tolling: the Laih model

Laih (1994, 2004) considers multi-step tolls, under the assumption that tolling does not affect the timing of arrivals. In both papers Laih assumes that  $\gamma > \alpha > \beta > 0$ .<sup>18</sup> However, the solution he presents in Laih (1994) is valid only if  $\gamma \leq \alpha$ , unless there are separate queuing facilities. He does not discuss mass departures. Rather, he uses a graphical approach to infer how much queuing time costs are reduced as a function of the number of toll steps. Laih's (2004) solution is valid regardless of the relative magnitudes of  $\gamma$  and  $\alpha$ , because he

<sup>18</sup> Laih (1994) states this assumption as inequality (5) on page 199. Laih (2004) makes the assumption in the text on page 61.

circumvents the problem of mass departures by assuming that users who arrive after  $t^-$ , and do not pay the toll, can wait on a set of secondary lanes or shoulders without impeding other drivers who arrive before  $t^-$  and use the main set of lanes. There are therefore separate queues and no mass departures are needed to obtain equilibrium. And because all travel times are deterministic, the timing of the peak and the equilibrium price are unaffected by tolling.

Figure 3 shows equilibrium in the Laih model with a single-step toll. Travellers who pay the toll arrive between  $t^+$  and  $t^-$ . Between  $t_1$  and  $t^-$  their cumulative departures are shown by the blue dashed line. Travellers who arrive late and do not pay the toll depart between  $t_1$  and  $t^-$ , and arrive between  $t^-$  and  $t_e$ . Their cumulative departures are shown, using a second set of axes, by the green dot-dashed line. Cumulative departures for the two groups combined are identified by the black solid line.

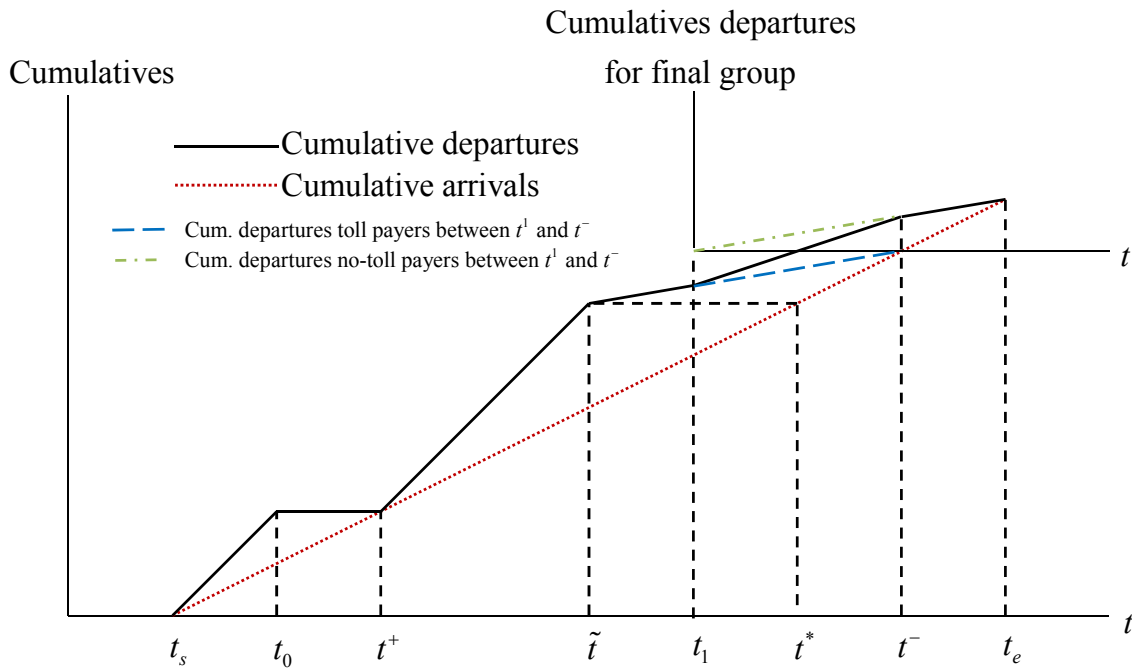


Figure 3. Equilibrium with a coarse toll: the Laih model

2.4. Coarse tolling: the braking model

Both the ADL model and the Laih model overlook that drivers have an incentive to delay reaching the tolling point when the toll is about to drop if the waiting cost they incur is outweighed by the money they will save. The braking model takes this incentive into account. The equilibrium is depicted in Figure 4.

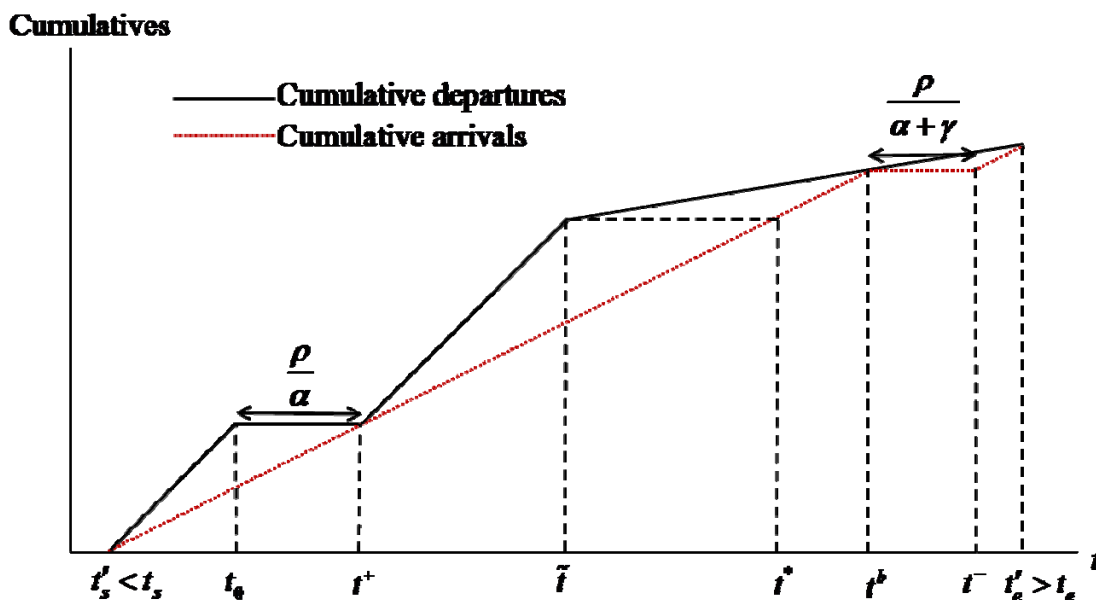


Figure 4. Equilibrium with a coarse toll: the braking model

The equilibrium is qualitatively the same as in the ADL model (cf. Figure 2) up to time  $t^b$  when the last driver pays the toll. The bottleneck is then idle for a period  $t^- - t^b = \rho / (\alpha + \gamma)$ : a queue is building up, but no-one passes the tolling point. Consequently, the arrival period  $[t'_s, t'_e]$  is longer than in the no-toll equilibrium and social optimum. Since all travel times are deterministic, the schedule delay costs of the first and last drivers are equal. Departures start earlier, and end later, than in the no-toll equilibrium and social optimum (i.e.  $t'_s < t_s$  and  $t'_e > t_e$ ), and the equilibrium price is therefore higher. This contrasts with both the ADL model, in which the coarse toll causes private costs to fall, and the Laih model, in which private costs are unchanged by the toll. The cost rises in the braking model because the bottleneck goes unused for a certain time interval before the toll ends.

The braking model avoids the difficulties in the ADL and Laih models. First, it is behaviourally consistent since it accounts for the possibility that drivers can reduce their travel costs by braking. Second, it avoids the problematic nature of mass departures in the ADL model. And third, it avoids the requirement of the Laih model for parallel traffic lanes.

### 3. Multi-step tolls

#### 3.1. Overview

The coarse tolls described in Section 2 are a simple case of multi-step tolls. The same behavioural assumptions and solution concepts apply for tolls with multiple steps in each of the ADL, Laih, and braking models. But the analytics become more complicated, and formulas for the optimal timing and levels of the step tolls in the ADL model are complex and rather opaque.

Before the analytics of the three models are presented, some additional terminology and notation need to be introduced. The toll schedules are assumed to have  $m$  levels with a series of increasing ‘step-up’ tolls while drivers arrive early, and a series of decreasing ‘step-down’ tolls while drivers arrive late. Tolling periods are indexed so that period 1 is the central period spanning a time interval  $(t_1^+, t_1^-)$  that contains  $t^*$ , and with a toll level  $\rho_1$  as shown in Figure 5. Period 2 spans earlier and later time intervals on either side of the central period. Period 2 is bracketed by period 3, and so on through period  $m$ . The early arrival segment of period  $i$  spans period  $[t_i^+, t_{i-1}^+)$  during which the ‘step-up’ toll  $\rho_i^+$  applies. Similarly, the late arrival segment of period  $i$  spans period  $[t_{i-1}^-, t_i^-)$  during which the ‘step-down’ toll  $\rho_i^-$  applies. Since period 1 forms a single connected time interval,  $\rho_1^+ = \rho_1^- = \rho_1$ . The toll is assumed to be zero at the beginning ( $t_s'$ ) and end ( $t_e'$ ) of the travel period; a harmless assumption since demand is perfectly inelastic.



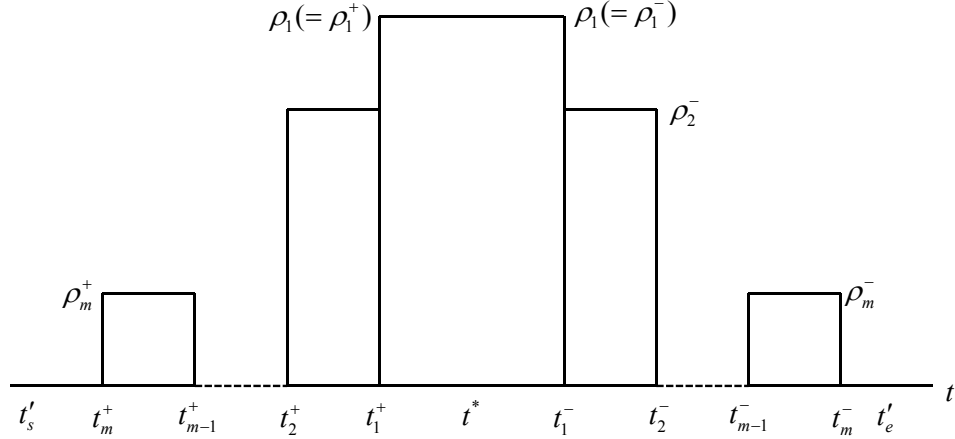


Figure 5. Notation in the multi-step toll models

The efficiency of the optimal  $m$ -step toll of each model can be measured as a fraction of the welfare gain derived from the fine toll by using the index

$$\omega_r = \frac{TC_{NT} - TC_r}{TC_{NT} - TC_{FB}},$$

where  $r$  indexes the model or regime (ADL, Laih, or braking). We begin with the Laih model since it is the easiest to analyze.

### 3.2. Step-tolling in the Laih model

Optimal toll levels and tolling periods for the Laih model are described in Laih (1994, 2004). Appendix A summarizes the derivations. The optimal tolls turn out to be symmetric between early and late segments of each tolling period so that  $\rho_i^+ = \rho_i^- \equiv \rho_i$  for  $i = 1, \dots, m$ , and the  $^+$  and  $^-$  superscripts can be omitted. The  $i$ th step-up toll begins at

$$t_i^+ = t_s + \frac{\rho_i}{\beta},$$

and the  $i$ th step-down toll ends at

$$t_i^- = t_e - \frac{\rho_i}{\gamma}.$$

Optimal toll levels are

$$(14) \quad \rho_1^{Laih} = \frac{m}{m+1} \cdot \delta \cdot \frac{N}{s}, \quad \rho_i^{Laih} = \frac{m+1-i}{m} \cdot \rho_1^{Laih}, \quad i = 2, \dots, m.$$

The central toll,  $\rho_1^{Laih}$ , approaches the peak level of the fine toll,  $\delta \cdot N / s$ , as  $m \rightarrow \infty$ . Because the toll steps up and down in uniform increments, and the time intervals between steps are also uniform, the Laih toll approaches the optimal fine toll as  $m \rightarrow \infty$ . Total social costs are

$$(15) \quad TC_{Laih} = \delta \cdot \frac{N^2}{s} \cdot \left(1 - \frac{1}{2} \cdot \frac{m}{m+1}\right).$$

Using eqns. (8), (10), and (15), it follows that the Laih toll has a relative efficiency of

$$(16) \quad \omega_{Laih} = \frac{m}{1+m}.$$

The Laih toll asymptotically approaches the efficiency of the first-best (fine) toll as  $m \rightarrow \infty$ .

### 3.3. Step-tolling in the ADL model

The ADL model is easy to solve if  $\alpha \geq \gamma$  since it is then equivalent to the Laih model in terms of the toll schedule, private costs, social costs, and relative efficiency. But if  $\alpha < \gamma$ , the ADL model is tedious to solve. In part, this is because the optimal tolls are not symmetric in levels: the  $i$ th step-up toll  $\rho_i^+$  is lower than the corresponding step-down toll  $\rho_i^-$ .<sup>19</sup> Derivations are provided in Appendix B. The optimal tolls have simple recursive formulas:

$$(17) \quad \rho_i^+ = \frac{m+1-i}{m} \cdot \rho_1, \quad i=1, \dots, m, \quad \rho_i^- = (i-2) \cdot \rho_3^- - (i-3) \cdot \rho_2^-, \quad i=2, \dots, m.$$

As in the Laih model, the step-up tolls grow in uniform increments. And the step-down tolls follow a simple recursion relation. However, for  $m > 1$ ,  $\rho_1$ ,  $\rho_2^-$  and  $\rho_3^-$  have complicated expressions.<sup>20</sup> Some noteworthy properties are nevertheless discernible. First, the size of the time shift toward later departures decreases with  $m$ , and approaches zero as  $m \rightarrow \infty$ . Second, the ADL toll approaches the fine toll as  $m \rightarrow \infty$  so that  $\omega_{ADL} \rightarrow 1$ . Third, for  $1 < m < \infty$ , each ADL toll is larger than the corresponding Laih toll. Fourth, for finite  $m$ ,  $\omega_{ADL} > \omega_{Laih}$ . This last inequality is consistent with the fact that the generalized cost of travel decreases with tolling in the ADL model, but not the Laih model.

---

<sup>19</sup> Symmetric tolls are actually even harder to solve than asymmetric tolls and we have been unable to derive closed-form solutions. Imposing symmetry would have no advantage with electronic toll collection technology. And it is unlikely that drivers would gain a lot from toll symmetry in terms of ease of recall or trip planning.

<sup>20</sup> Appendix B provides explicit formulae for  $m=1$  and  $m=2$ .

### 3.4. Step-tolling in the braking model

When multiple toll steps are introduced in the braking model, the bottleneck is idle for a time interval before each step down in the toll. The total time idle turns out to be  $\rho_1 / (\alpha + \gamma)$ , and therefore depends only on the level of the central toll. The number of toll steps only determines the number of intervals into which total idle time is divided. As shown immediately below,  $\rho_1$  is an increasing function of  $m$  so that total idle time actually increases as the toll structure becomes more refined. Hence, with braking the step toll does not asymptotically approach the first-best optimum in the limit  $m \rightarrow \infty$ .

The timing and levels of the optimal step toll in the braking model are derived in Appendix C. The toll levels work out to

$$(18) \quad \rho_1^{Braking} = \frac{m}{m+1} \cdot \delta \cdot \frac{N}{s}, \quad \rho_i^{Braking} = \frac{m+1-i}{m} \cdot \rho_1^{Braking}, \quad i = 2, \dots, m,$$

The toll levels are the same as in eqn. (14) for the Laih model. However, there is a distinction between the time,  $t_i^-$ , when the  $i$ th step-down toll  $\rho_i^-$  is lifted, and the time,  $t_i^b$ , when  $\rho_i^-$  is last paid. Formulae for the  $t_i^+$  and  $t_i^b$  can be written as functions of the beginning or end of the travel period:

$$(19) \quad t_i^+ = t'_s + \frac{\rho_i}{\beta}, \quad i = 1, \dots, m,$$

$$(20) \quad t_i^b = t'_e - \frac{\rho_i}{\gamma}, \quad i = 1, \dots, m.$$

Total social cost is

$$(21) \quad TC_{Braking} = \delta \cdot \frac{N^2}{s} \cdot \left( 1 - \frac{1}{2} \cdot \frac{m}{m+1} \cdot \left( 1 - \frac{\beta \cdot \gamma}{(\beta + \gamma) \cdot (\alpha + \gamma)} \right) \right).$$

The relative efficiency of the braking toll is derived using eqn. (21) and eqn. (10):

$$(22) \quad \omega_{Braking} = \frac{m}{m+1} \cdot \left( 1 - \frac{\beta \cdot \gamma}{(\beta + \gamma) \cdot (\alpha + \gamma)} \right).$$

The relative efficiency of the braking toll can be compared with the efficiency in the Laih and ADL models. The comparison is straightforward for the Laih model. From eqn. (22) and eqn. (16) one has

$$(23) \quad r_{BL} \equiv \frac{\omega_{Braking}}{\omega_{Laih}} = 1 - \frac{\beta \cdot \gamma}{(\beta + \gamma) \cdot (\alpha + \gamma)}.$$

Unless  $\beta = 0$  or  $\gamma = 0$  (in which case travel is costless in all three models because rescheduling to avoid queuing is costless) tolling is less efficient in the braking model than in the Laih model. The ratio,  $r_{BL}$ , is an increasing function of  $\alpha$ .  $r_{BL}$  converges to 1 in the limit  $\alpha \rightarrow \infty$  or  $\gamma \rightarrow \infty$  because total idle time for the bottleneck,  $\rho_1 / (\alpha + \gamma)$ , then approaches zero. Since  $\alpha > \beta$ ,<sup>21</sup> the lower limit is

$$\lim_{\alpha \rightarrow \beta} r_{BL} = 1 - \frac{\beta \gamma}{(\beta + \gamma)^2}.$$

The ratio  $r_{BL}$  reaches a minimum value of 3/4 with  $\beta = \gamma$ . Thus, relative to tolling in the Laih model the efficiency loss from tolling due to braking can be as large as 25%.

Comparison of the braking and ADL models is intractable for general values of  $m$  because the formula for  $\omega_{ADL}$  is very messy. But the comparison is easy for the coarse toll. Setting  $m = 1$  in eqn. (22) one has

$$\omega_{Braking} = \frac{1}{2} \cdot \left( 1 - \frac{\beta \cdot \gamma}{(\beta + \gamma) \cdot (\alpha + \gamma)} \right).$$

For the coarse toll in the ADL model (see Appendix B)

$$\omega_{ADL} = \frac{1}{2} \cdot \left( 1 + \frac{\beta \cdot (\gamma - \alpha)}{(\beta + \gamma) \cdot (\alpha + \gamma)} \right).$$

The relative efficiency of tolling in the two models is therefore

$$r_{BA}|_{m=1} \equiv \frac{\omega_{Braking}}{\omega_{ADL}} = \frac{(\beta + \gamma) \cdot (\alpha + \gamma) - \beta \gamma}{(\beta + \gamma) \cdot (\alpha + \gamma) + \beta \cdot (\gamma - \alpha)}.$$

Since the ADL model is solved here for the case  $\gamma \geq \alpha$ ,  $r_{BA} < 1$  as long as  $\beta > 0$ . Similar to the comparison with the Laih model,  $r_{BA} \rightarrow 1$  in the limit as  $\gamma \rightarrow \infty$ . A minimum value of  $r_{BA} \cong 0.6991$  is reached with  $\beta = \alpha$  and  $\gamma / \alpha = (1 + \sqrt{7}) / 2$ . Relative to tolling in the ADL model, the efficiency loss from tolling due to braking can be as large as 30%.

---

<sup>21</sup> The opposite case,  $\alpha < \beta$ , would imply that early arriving drivers would prefer to extend their trip rather than arrive early.

#### 4. Numerical results

To get a better idea of how much the three tolling models differ quantitatively, and how braking affects the efficiency of tolling, we turn to a brief numerical analysis. We use the same values for the unit cost parameters as Arnott, de Palma and Lindsey (1990):  $\alpha=6.4$ ,  $\beta=3.9$ , and  $\gamma=15.21$ .<sup>22</sup> We set  $N=9000$  and  $s=3600$  vehicles per hour so that travel lasts for  $2\frac{1}{2}$  hours in all regimes except the braking model. Optimal toll schedules were computed (in Mathematica) using the analytical results presented in the appendices, and various numerical consistency checks were performed to make sure that we indeed consider second-best equilibria.

##### 4.1. Tolls and equilibrium prices

Figure 6 shows the three step-tolling schemes for  $m=5$  along with the first-best toll for comparison (the ADL toll is the solid line, the Laih toll is dashed, the braking toll is dot-dashed, and the FB toll is dotted). Consistent with the analytical results in Section 3, toll levels are the same in the braking model and the Laih model, but the timing differs. The early-arrival tolling periods have the same duration, but the late-arrival periods are slightly longer in the braking model. Braking prolongs travel by 18 minutes, which is nearly 12% of the peak duration without braking.<sup>23</sup> In the ADL model the early-arrival tolls differ only slightly from the other two models. The timing of the steps also differs only slightly from the Laih model because of the small departure delay. For late arrivals, the difference in toll levels between the ADL model and the Laih model is larger due to the asymmetry in the ADL toll schedule. The braking toll exceeds the other two tolls as well as the fine toll over most of the travel period (but recall that the braking toll is not paid during intervals when the bottleneck is idle).

---

<sup>22</sup> The results of interest depend only on the ratios  $\beta/\alpha$  and  $\gamma/\alpha$ .

<sup>23</sup> For  $m=1$ , the increase is about 11 minutes, and as  $m \rightarrow \infty$  the increase approaches 21 minutes.

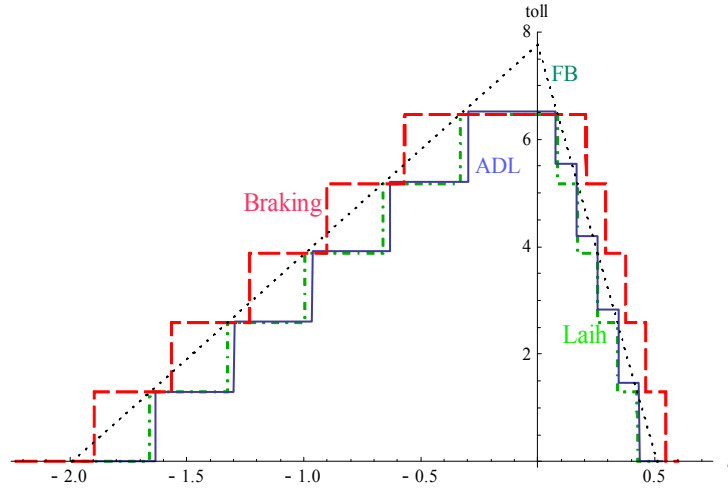


Figure 6. Toll levels with step tolls and first-best (fine) tolls ( $m=5$ )

Figure 7 plots equilibrium prices for the three models against the number of steps. Consistent with the analytical results, the price in the Laih model is independent of the number of steps, and remains equal to the price for the no-toll equilibrium and the social optimum. Also as expected, the price in the ADL model is lower than in the Laih model but approaches it as the number of steps increases. By contrast, in the braking model the price is higher than in the Laih model. Moreover, the difference increases with the number of steps, because the central toll  $\rho_1^{Braking}$  in (18), and therewith total idle time ( $\rho_1^{Braking}/(\alpha + \gamma)$ ), increases. For a single step, the price increase is 7% of the no-toll equilibrium price, for 5 steps it is 12%, and as  $m \rightarrow \infty$  the increase approaches 14%. Somewhat counter-intuitively, therefore, a closer approximation of the continuous toll, through an increase in the number of steps, does not decrease the total idle time in the braking equilibrium, but rather increases it. The reason is that the total idle time depends on the central toll level, but not on the number of steps used to bring the toll back to zero. Every unit reduction brings the same amount of idle time ( $1/(\alpha + \gamma)$ ), irrespective of the size of the step. The level of the central toll, meanwhile, increases with the number of steps because the central toll applies for a shorter period. As a result, the efficiency of step tolling in the braking model will not achieve unity as the number of steps  $m$  approaches infinity, as shown in equation (22) above, as well as in the numerical results to which we now turn.

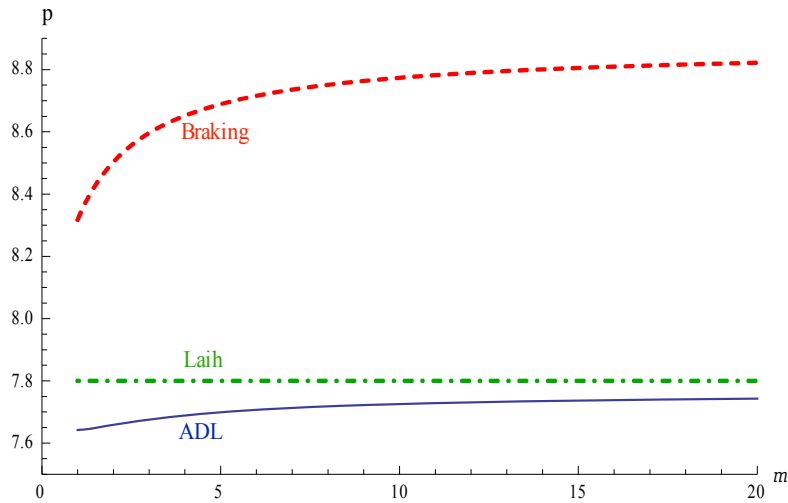


Figure 7. Prices with step tolls as a function of  $m$

#### 4.2. Relative cost reductions

Figure 8 shows the percentage reductions in total social costs for the three models as a function of the number of steps. The highest possible reduction, obtained with first-best pricing, is 50%. As  $m$  increases, the cost reductions for the ADL and Laih models rapidly approach the 50% upper bound. The relative efficiency of the Laih toll is  $m/(1+m)$  so that the cost reduction in percentage terms is  $50 \cdot m/(1+m)$ . The ADL toll does only slightly better. The fact that it does better suggests that if mass departures are observed with step tolls in reality, it need not be welfare enhancing to try and avoid this by opening a parallel queuing facility, *e.g.* a shoulder, even if it were costless, to allow for a Laih equilibrium to replace the ADL equilibrium.

The braking toll produces appreciably smaller cost reductions. As the number of steps approaches infinity, the total cost reduction approaches an upper limit of only 43% rather than 50%, on account of the toll-induced idleness of the bottleneck. As per eqn. (23), the efficiency gain from the braking toll is 14% lower than the gain from the Laih toll regardless of the number of steps. For the ADL toll the difference is slightly larger: 21% with one step, 16% with 5 steps, and asymptotically approaches 14% as  $m \rightarrow \infty$ .

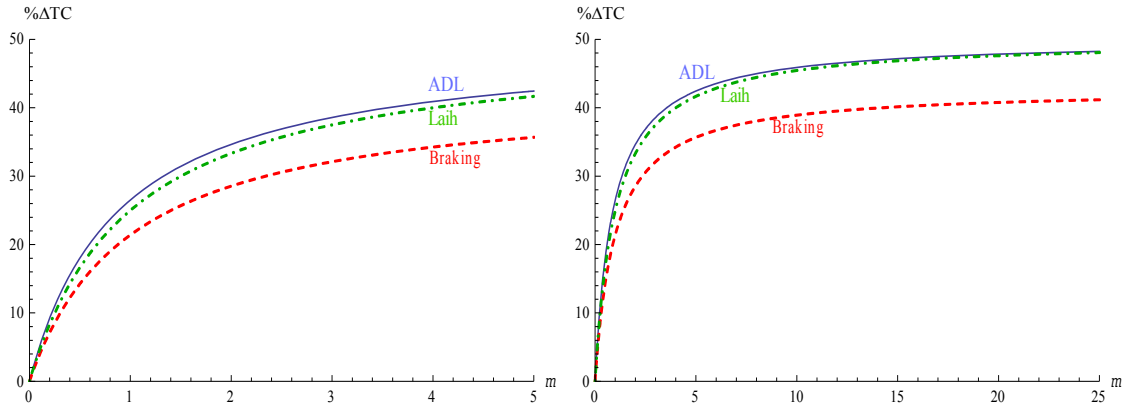


Figure 8. Percentage reduction in social cost as a function of  $m$  (up to  $m=5$  in the left panel, and  $m=25$  in the right panel)

#### 4.3. Sensitivity analysis for relative efficiency gains

The numerical results in Sections 4.1 and 4.2 are derived for a single set of parameter values and it is useful to test the robustness of these results by conducting some sensitivity analysis. Arguably of greatest interest is the efficiency of the braking toll compared to the toll in the other two models, and attention is limited to this here. The relative efficiency of the braking toll and the Laih toll,  $r_{BL}$ , given in eqn. (23), depends only on the ratios  $\beta/\alpha$  and  $\gamma/\alpha$ . Figure 9 plots  $r_{BL}$  over the range  $\beta/\alpha \in [0,1)$  and  $\gamma/\alpha \in [0,10]$ . As noted earlier,  $r_{BL} = 1$  with  $\beta/\alpha = 0$  or  $\gamma/\alpha = 0$ ; and  $r_{BL}$  reaches a minimum value of 0.75 with  $\beta/\alpha = 1$  and  $\gamma/\alpha = 1$ ; see the front corner of Figure 9. Over an appreciable portion of the surface, the braking toll underperforms the Laih toll by 10-15%.

Figure 9. Efficiency gain from braking toll relative to Laih toll



The relative efficiency of the braking toll and the ADL toll,  $r_{BA}$ , depends on  $\beta/\alpha$ ,  $\gamma/\alpha$  and  $m$ . Figure 10 plots  $r_{BA}$  for  $m=1$  in the left panel, and  $m=5$  in the right panel. The two surfaces are very similar. Consistent with the analytical results in Section 3 for the coarse toll,  $r_{BA}$  reaches a minimum value with  $\beta/\alpha=1$  and an intermediate value of  $\gamma/\alpha$  of around 1.8. Under these conditions, a large fraction of travellers arrive late so that the harmful effects of braking are magnified.

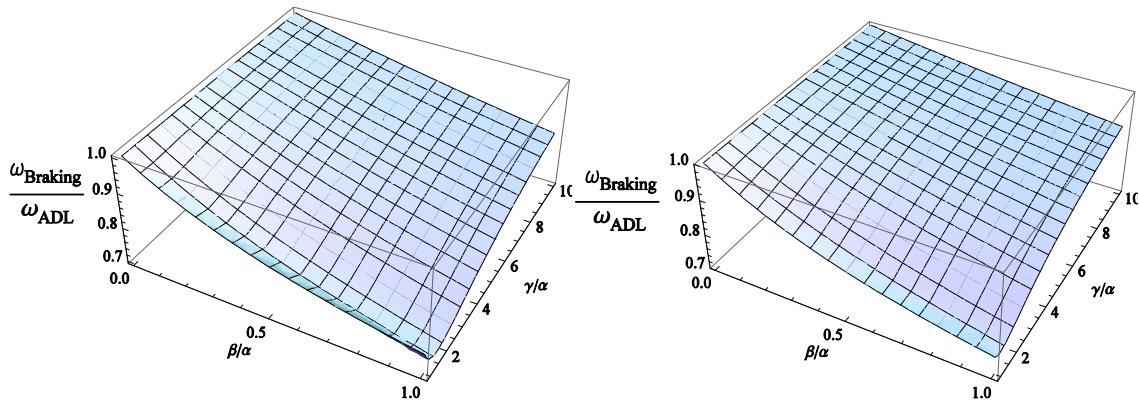


Figure 10. Efficiency gain from braking toll relative to ADL toll ( $m=1$  in the left panel,  $m=5$  in the right panel)

## 5. Conclusions

This paper has reconsidered step tolling in the bottleneck model. Its main contribution is to remedy an inconsistency in earlier studies by Arnott, de Palma and Lindsey (1990, 1993) and Laih (1994, 2004). These studies ignored that when a toll is about to drop, drivers have an incentive to slow down or stop before reaching the tolling point and wait. We develop a ‘braking model’ in which drivers wait if the toll payment saved outweighs the travel time and schedule delay cost incurred by waiting. Unlike in the ADL and Laih models, equilibrium in the braking model is based on internally consistent traveller behaviour. The equilibrium is also consistent with observed cases of waiting on roads with step tolls. An additional advantage of the braking model is that the equilibrium does not require either mass departures—as in the ADL model—or parallel sets of traffic lanes—as in the Laih model.

A second contribution of the paper is to extend the ADL model from a single-step (coarse) toll to multi-step tolls. Comparison of the braking model with the ADL and Laih models reveals that braking can materially reduce the efficiency gains from tolling because the bottleneck is unused while drivers are waiting. The percentage reduction in efficiency gain depends on the unit costs of travel time and schedule delay, and (in the comparison with the ADL model) the number of steps in the toll schedule. In our base case parameterization, the welfare gain of a toll with one step is 21% lower with braking than in the ADL model, and with five steps it is 16% lower; for other parameterizations, the gain can be reduced by up to 30%.

The fact that braking occurs in practice, and can be quite costly, raises the question how it can be prevented. Stopping in the middle of traffic lanes or parking on the shoulder can be deterred by enforcing traffic laws. Slowing down, however, is more difficult to prevent although it can be discouraged by imposing minimum speed limits. In either case, enforcement is costly. The problems associated with braking arise when tolls are levied at specific locations. They could be avoided by replacing location-based toll collection schemes with more sophisticated systems with charges proportional to distance driven.<sup>24</sup>

The analysis in the paper could be extended in various directions. One is to model drivers' incentives to speed in order to reach a tolling point before a toll increase. We argued why toll-induced speeding is likely to be less significant than braking, but it would be worthwhile to investigate this further in the context of a model that does allow for a formal representation of such speeding, such as a car-following or a hydrodynamic model. With heterogeneity of travellers incorporated, overtaking could also occur in a dynamic equilibrium without tolling, and the modelling of speeding would probably be more realistic.

A second possible extension is to account for the inconvenience or risk incurred by drivers in slowing down or stopping. To the extent that these costs are fixed, drivers may choose not to slow down or stop if the toll payment avoided is small. Multi-step tolling schemes with

---

<sup>24</sup> Comprehensive, per-kilometre road-pricing schemes have other advantages, including more complete network coverage and greater spatial/geographical equity.

several, small, toll adjustments may therefore be less vulnerable to braking than coarse schemes with tolls that are only ‘on’ or ‘off’.

In the standard bottleneck model travellers are assumed to care about when they arrive at their destination, but not when they leave the origin. This may be a reasonable approximation for the morning commute, but not for the evening commute or some other types of trips where early departure may be heavily penalized. The bottleneck model with departure-time preferences has been studied by a few authors including Vickrey (1973), Fargier (1983), and de Palma and Lindsey (2002a, 2002b). Fosgerau (2011) considers a single-step toll using the Laih model and a general scheduling utility function that depends on both departure time from the origin and arrival time at the destination. Further analysis of step tolls using models with general trip-timing preferences would be insightful.

### **Acknowledgments**

The authors thank the editor and two anonymous referees for their helpful comments. The usual disclaimer applies. Lindsey would like to thank the Social Sciences and Humanities Research Council of Canada for financial support. Van den Berg and Verhoef gratefully acknowledge support from the ERC (AdG Grant #246969 OPTION). For useful comments we are also grateful to David Boyce, Jonas Eliasson, Andrew Koh, Linda Lee, and the participants at the 57th Annual North American Meetings of the Regional Science Association International.

### **References**

- Arnott, R., de Palma, A., Lindsey, R., 1990. Economics of a bottleneck. *Journal of Urban Economics* 27(1), 111-130.
- Arnott, R., de Palma, A., Lindsey, R., 1993. A structural model of peak-period congestion: a traffic bottleneck with elastic demand. *American Economic Review* 83(1), 161-179.
- Arnott, R., de Palma, A., Lindsey, R., 1998. Recent developments in the bottleneck model. In K.J. Button and E.T. Verhoef (Eds.), *Road Pricing, Traffic Congestion and the Environment: Issues of Efficiency and Social Feasibility*, Aldershot: Edward Elgar, 79-110.
- Chee, K., 2011. Illegal road shoulder use still high. *Asiaone Motoring*, February 9. Retrieved on 15 July 2011, from <http://www.asiaone.com/Motoring/News/Story/A1Story20110209-262506.html>

- Chew, V., 2008. Electronic road pricing: Developments after phase I. National Library Singapore, Singapore Infopedia. Retrieved on 17 August 2010, from [http://infopedia.nl.sg/articles/SIP\\_1386\\_2009-01-05.html](http://infopedia.nl.sg/articles/SIP_1386_2009-01-05.html)
- Chu, X., 1999. Alternative congestion technologies. *Regional Science and Urban Economics* 29(6), 697-722.
- Daganzo, C.F., 1997, *Fundamentals of Transportation and Traffic Operations*. Amsterdam and New York: Pergamon.
- Daniel, J.I., 2009. The deterministic bottleneck model with non-atomistic traffic. Working paper of the Department of Economics, Alfred Lerner College of Business & economics, University of Delaware, no. 2009-08. Retrieved on 25 July 2010, from [www.lerner.udel.edu/sites/default/files/imce/economics/WorkingPapers/2009/UDWP2009-08.pdf](http://www.lerner.udel.edu/sites/default/files/imce/economics/WorkingPapers/2009/UDWP2009-08.pdf)
- De Palma, A., Lindsey, R., 2002a. Comparison of morning and evening commutes in the Vickrey bottleneck model. *Transportation Research Record* 1807, 26-33.
- De Palma, A., Lindsey, R., 2002b. Congestion pricing in the morning and evening peaks: A comparison using the Bottleneck Model. *Proceedings of the 37th Annual Conference of the Canadian Transportation Research Forum: 2002 Transportation Visioning - 2002 and Beyond*. Vancouver, 179-193.
- Eskenazi, J., 2010. Commuters slowing, stopping on freeway to foil Bay Bridge 'Congestion Pricing'. *SF Weekly*, July 1. Retrieved on 15 July 2011, from [http://blogs.sfweekly.com/thesnitch/2010/07/commuters\\_slowing\\_stopping\\_on.php](http://blogs.sfweekly.com/thesnitch/2010/07/commuters_slowing_stopping_on.php)
- Fargier, P.H., 1983. Effects of the choice of departure time on road traffic congestion: Theoretical approach. *Proceedings of the 8th International Symposium on Transportation and Traffic Theory*. Toronto, Canada: University of Toronto Press, 223-263.
- Fosgerau, M., 2011. How a fast lane may replace a congestion toll. *Transportation Research Part B* 45(6), 845-851.
- Ge, Y.E., Stewart, K., 2010. Investigating boundary issues arising from congestion charging in a bottleneck scenario. In C. Tampre, F. Viti and L.H. Immers (Eds.), *New Developments in Transport Planning: Advances in Dynamic Traffic Assignment*. Aldershot: Edward Elgar, 303-326.
- Laih, C.H., 1994. Queuing at a bottleneck with single and multi-step tolls. *Transportation Research Part A* 28(3), 197-208.
- Laih, C.H., 2004. Effects of the optimal step toll scheme on equilibrium commuter behavior. *Applied Economics* 36(1), 59-81.
- Lee, L., Frick, K., 2011. San Francisco-Oakland Bay Bridge Congestion Pricing. In *Congestion pricing projects in the San Francisco Bay Area: I-680 express lanes and Bay Bridge time-of-day pricing*. Retrieved on 15 July 2011, from [http://www.fhwa.dot.gov/ipd/pdfs/revenue/resources/webinars/webinar\\_congestion\\_pricing\\_033011.pdf](http://www.fhwa.dot.gov/ipd/pdfs/revenue/resources/webinars/webinar_congestion_pricing_033011.pdf)
- May, A.D., 1990. *Traffic Flow Fundamentals*. Upper Saddle River, NJ: Prentice-Hall.
- Mercury News, 2010. As congestion pricing kicks in, drivers hit the brakes to save a buck. Posted on 1 July 2010. Retrieved on 12 August 2010, from [http://www.mercurynews.com/news/ci\\_15419447?source=email&nclick\\_check=1](http://www.mercurynews.com/news/ci_15419447?source=email&nclick_check=1)
- Png, G.H., Olszewski, P., Menon, A.P.G., 1994. Estimation of the value of travel time of motorists. *Journal of the Institution of Engineers, Singapore* 34(2), 9-13.
- Small, K.A., 1982. The scheduling of consumer activities: Work trips. *American Economic Review* 72(3), 467-479.
- Small, K.A., Verhoef, E.T., 2007. *The Economics of Urban Transportation*. London: Routledge.

Vickrey, W.S., 1969. Congestion theory and transport investment. *American Economic Review* (Papers and Proceedings) 59(2), 251-260.

Vickrey, W.S., 1973. Pricing, metering, and efficiently using urban transportation facilities. *Highway Research Record* 476, 36-48.

### Appendix A: Derivation of optimal multi-step tolls in the Laih model

At the beginning ( $t_s$ ) and the end ( $t_e$ ) of the travel period in the Laih model there is neither a queue nor a toll. Schedule delay costs at  $t_s$  and  $t_e$  must therefore be equal and eqn. (5) in the text holds. Combined with condition (4) this implies:

$$(A.1) \quad t_s = t^* - \frac{\gamma}{\beta + \gamma} \left( \frac{N}{s} \right),$$

$$(A.2) \quad t_e = t^* + \frac{\beta}{\beta + \gamma} \left( \frac{N}{s} \right).$$

The price,  $p$ , does not change due to step tolling, and total private cost is given by eqn. (8) in the text:

$$(A.3) \quad PC \equiv p \cdot N = \delta \frac{N^2}{s}.$$

In the Laih model there are separate queues during late arrivals for drivers who pay toll  $\rho_i$  and those who pay toll  $\rho_{i+1}$ . The first driver to arrive in each toll period faces no queue. The price paid at time  $t_i^+$  is  $\beta(t^* - t_i^+) + \rho_i$  which must match the price paid at time  $t_s$ ,  $\beta(t^* - t_s)$ . Accordingly, toll  $\rho_i$  starts at time:

$$(A.4) \quad t_i^+ = t_s + \frac{\rho_i}{\beta}, \quad i = 1, \dots, m.$$

Similarly, for late arrivals the first driver to arrive in each toll period faces no queue and incurs a price  $\gamma(t_i^- - t^*) + \rho_i$ . Setting this price equal to the price for the last driver,  $\gamma(t_e - t^*)$ , one obtains:

$$(A.5) \quad t_i^- = t_e - \frac{\rho_i}{\gamma}, \quad i = 1, \dots, m.$$

Total social costs equal total private costs minus toll revenue ( $TR$ ). Let  $TR_i$  denote the toll revenue derived from the combined early and late arrival intervals for toll period  $i$ . Total social costs can then be written

$$(A.6) \quad TC = p \cdot N - TR = PC - TR_1 - \sum_{i=2}^m TR_i.$$

Toll revenues are

(A.7)

$$\begin{aligned}
TR &= TR_1 + \sum_{i=2}^m TR_i = \rho_1 \cdot s (t_1^- - t_1^+) + \left( \sum_{i=2}^m \rho_i s (t_{i-1}^+ - t_i^+) \right) + \left( \sum_{i=2}^m \rho_i s (t_i^- - t_{i-1}^-) \right) \\
&= \rho_1 \cdot s (N/s - \rho_1/\beta - \rho_1/\gamma) + \left( \sum_{i=2}^m s \cdot \rho_i (\rho_{i-1} - \rho_i) / \beta \right) + \left( \sum_{i=2}^m s \cdot \rho_i (\rho_{i-1} - \rho_i) / \gamma \right) \\
&= \rho_1 (N - s\rho_1/\delta) + \left( \frac{s}{\delta} \sum_{i=2}^m \rho_i (\rho_{i-1} - \rho_i) \right).
\end{aligned}$$

Inserting (A.7) into (A.6), and taking derivatives gives the following first-order conditions:

$$(A.8) \quad N - 2 \cdot s (\rho_1 / \delta) + s \cdot \rho_2 / \delta = 0, \quad \text{for } i = 1$$

$$(A.9) \quad \rho_{i-1} - 2\rho_i + \rho_{i+1} = 0, \quad \text{for } i = 2, \dots, m-1$$

$$(A.10) \quad \rho_{m-1} - 2\rho_m = 0, \quad \text{for } i = m$$

Equation (A.10) implies  $\rho_{m-1} = 2\rho_m$ . Solving (A.9) iteratively gives:

$$(A.11) \quad \rho_i = \frac{m+1-i}{m} \rho_1, \quad i = 1, \dots, m.$$

Inserting (A.11) into (A.8) gives:

$$(A.12) \quad \rho_1 = \frac{m}{m+1} \delta \frac{N}{s}.$$

Inserting (A.7), (A.11), and (A.12) into (A.6) implies an equilibrium total social cost of:

$$(A.13) \quad TC = \delta \frac{N^2}{s} \left( 1 - \frac{1}{2} \frac{m}{m+1} \right).$$

## Appendix B: Derivation of optimal multi-step tolls in the ADL model with $\alpha < \gamma$

### B.1. Overview of solution

The solution presented here is based on heuristic reasoning rather than a rigorous derivation of the general structure of the toll. The timing of the step toll is derived on the following assumptions:

- During the early arrival period there is an interval prior to each step up in the toll when no user departs. The queue drops to zero at the instant when the toll steps up.
- During the late arrival period a mass of drivers departs each time the toll steps down. With  $\gamma \geq \alpha$ , the queue drops to zero at the instant when the toll takes its next step down and the

next mass of drivers departs. All drivers who arrive late depart in masses; unlike with  $\gamma < \alpha$  there are no time intervals during which drivers depart at a finite rate.

### B.2. Derivation of solution

The first step is to derive the timing of the toll schedule for given toll levels. For early arrivals the solution is the same for  $\gamma \geq \alpha$  and  $\gamma < \alpha$ . The first user to depart incurs only an early arrival cost and hence pays a price of:

$$(B.1) \quad p(t_s) = \beta(t^* - t_s).$$

The user who departs at  $t_i^+$  immediately after step-up toll  $\rho_i^+$  is imposed incurs an early arrival cost, and pays the toll, but faces no queue. The user's cost is therefore:

$$(B.2) \quad p(t_i^+) = \beta(t^* - t_i^+) + \rho_i^+, \quad i = 1, \dots, m.$$

Time  $t_i^+$  is solved by equating prices in eqn. (B.1) and eqn. (B.2):

$$(B.3) \quad t_i^+ = t_s + \frac{\rho_i^+}{\beta}, \quad i = 1, \dots, m.$$

The user who departs at  $t_i^-$  just before the central-toll interval ends incurs a price:

$$(B.4) \quad p(t_i^-) = \rho_1 + \gamma(t_i^- - t^*).$$

A mass of  $M_i$  users departs at time  $t_i^-$ . These users incur an expected price of

$$(B.5) \quad E \cdot p(t_i^-) = \rho_{i+1}^- + \gamma(t_i^- - t^*) + \frac{1}{2} \frac{M_i}{s} (\alpha + \gamma), \quad i = 1, \dots, m.$$

$M_i$  is solved by equating prices in eqn. (B.4) and eqn. (B.5):

$$(B.6) \quad M_i = \frac{2s}{\alpha + \gamma} (\gamma(t_i^- - t_i^-) + \rho_1 - \rho_{i+1}^-), \quad i = 1, \dots, m.$$

Using eqn. (B.6),

$$(B.7) \quad t_{i+1}^- = t_i^- + \frac{M_i}{s} = \frac{\alpha - \gamma}{\alpha + \gamma} t_i^- + \frac{2\gamma}{\alpha + \gamma} t_1^- + \frac{2}{\alpha + \gamma} (\rho_1 - \rho_{i+1}^-), \quad i = 1, \dots, m-1.$$

Define  $R \equiv (\alpha - \gamma)/(\alpha + \gamma) < 0$ . Applying eqn. (B.7) iteratively, summing the component terms, and simplifying expressions one obtains



$$(B.8) \quad t_i^- = t_1^- + \frac{1}{\gamma}(1 - R^{i-1})\rho_1 - \frac{2}{\alpha + \gamma} \sum_{j=2}^i R^{i-j} \rho_j^-, \quad i = 2, \dots, m.$$

*Equilibrium price as a function of tolls*

The expected price for users in the last mass to depart is:

$$(B.9) \quad E \cdot p(t_m^-) = \gamma(t_m^- - t^*) + \frac{1}{2}(t_e - t_m^-)(\alpha + \gamma).$$

Equate prices in eqn. (B.9) and (B.1), and use  $t_e = t_s + N/s$ :

$$(B.10) \quad \gamma(t_m^- - t^*) + \frac{1}{2}(t_s + N/s - t_m^-)(\alpha + \gamma) = \beta(t^* - t_s).$$

Equate prices in eqn. (B.4) and eqn. (B.1), and solve for  $t_1^-$ :

$$(B.11) \quad t_1^- = \frac{\beta + \gamma}{\gamma} t^* - \frac{\beta t_s + \rho_1^-}{\gamma}.$$

Substitute eqn. (B.11) into eqn. (B.8), set  $i = m$ , substitute the resulting formula for  $t_m^-$  into eqn. (B.10), and solve for  $t_s$ :

$$(B.12) \quad t_s = t^* - \frac{\gamma}{\beta + \gamma} \frac{N}{s} - \frac{R^m}{\beta + \gamma} \left( \rho_1 + \frac{2\gamma}{\alpha + \gamma} R \sum_{j=2}^m R^{-j} \rho_j^- \right).$$

Substitute eqn. (B.12) into eqn. (B.11):

$$(B.13) \quad t_1^- = t^* + \frac{\beta}{\beta + \gamma} \frac{N}{s} + \frac{\beta}{\gamma(\beta + \gamma)} R^m \left( \rho_1 + \frac{2\gamma}{\alpha + \gamma} R \sum_{j=2}^m R^{-j} \rho_j^- \right) - \frac{\rho_1}{\gamma}.$$

Equilibrium travel prices can now be expressed as a function of toll levels alone by substituting eqn. (B.13) into eqn. (B.4):

$$(B.14) \quad p = \frac{\beta\gamma}{\beta + \gamma} \left( \frac{N}{s} + R^m \left( \frac{\rho_1}{\gamma} + \frac{2R}{\alpha + \gamma} \sum_{j=2}^m R^{-j} \rho_j^- \right) \right).$$

*Toll revenues*

The toll revenue generated by the central toll,  $\rho_1$ , is

$$(B.15) \quad TR_1 = \rho_1 s (t_1^- - t_1^+).$$

Substitute eqn. (B.3) with  $i = 1$  and eqn. (B.13) into eqn. (B.15):

$$(B.16) \quad TR_1 = \rho_1 s \left( \frac{N}{s} - \frac{1}{\gamma} \left( \frac{\beta + \gamma}{\beta} - R^m \right) \rho_1 + \frac{2}{\alpha + \gamma} R^{m+1} \sum_{j=2}^m R^{-j} \rho_j^- \right).$$

Revenue from step-up toll  $\rho_i^+$  is:

$$TR_i^+ = \rho_i^+ s (t_{i-1}^+ - t_i^+), \quad i = 2, \dots, m,$$

or, using eqn. (B.3),

$$(B.17) \quad TR_i^+ = \rho_i^+ s \frac{\rho_{i-1}^+ - \rho_i^+}{\beta}, \quad i = 2, \dots, m.$$

Revenue from step-down toll  $\rho_i^-$  is:

$$TR_i^- = \rho_i^- s (t_i^- - t_{i-1}^-), \quad i = 2, \dots, m,$$

or, using eqn. (B.8),

$$(B.18) \quad TR_i^- = \rho_i^- s \left( \frac{(1-R)R^{i-2}}{\gamma} \rho_1 - \frac{2}{\alpha + \gamma} \rho_i^- + \frac{2(1-R)R^{i-1}}{\alpha + \gamma} \sum_{j=2}^{i-1} R^{-j} \rho_j^- \right), \quad i = 2, \dots, m.$$

*Total social costs*

Total social costs are

$$TC = p \cdot N - \sum_{i=1}^m TR_i = p \cdot N - \left( TR_1 + \sum_{i=2}^m TR_i^+ + \sum_{i=2}^m TR_i^- \right).$$

Using eqns. (B.6), (B.17) and (B.18),  $TC$  can be written as:

(B.19)

$$\begin{aligned} TC &= \delta \frac{N^2}{s} + \left( \frac{\delta N R^m}{\gamma} - N \right) \rho_1 + \frac{s}{\gamma} \left( \frac{\beta + \gamma}{\beta} - R^m \right) \rho_1^2 \\ &+ \frac{2}{\alpha + \gamma} R^{m+1} (\delta N - s \rho_1) \sum_{i=2}^m R^{-i} \rho_i^- - \frac{s}{\beta} \sum_{i=2}^m \rho_i^+ (\rho_{i-1}^+ - \rho_i^+) + \frac{2s}{\alpha + \gamma} \sum_{i=2}^m R^{-i} (\rho_i^-)^2 \\ &- \frac{(1-R)s}{\gamma} R^{-2} \rho_1 \sum_{i=2}^m R^i \rho_i^- - \frac{2(1-R)R^{-1}}{\alpha + \gamma} s \sum_{i=2}^m R^i \rho_i^- \sum_{j=2}^{i-1} R^{-j} \rho_j^-. \end{aligned}$$

*First-order conditions for tolls*

The first-order conditions for optimal tolls are derived by differentiating eqn. (B.19). The first-order condition for  $\rho_1$  is:

$$(B.20) \quad \left( 1 + \frac{\gamma}{\beta} - R^m \right) \left( 2\rho_1 - \delta \frac{N}{s} \right) = (1-R)R^{-2} \sum_{i=2}^m R^i \rho_i^- + (1-R)R^{m+1} \sum_{i=2}^m R^{-i} \rho_i^- + \frac{\gamma}{\beta} \rho_2^+.$$

The first-order conditions for the step-up tolls,  $\rho_i^+$ , are

$$(B.21) \quad \begin{aligned} \rho_{i-1}^+ - 2\rho_i^+ + \rho_{i+1}^+ &= 0, & i = 2, \dots, m-1, \\ \rho_{m-1}^+ - 2\rho_m^+ &= 0, & \text{for } i = m. \end{aligned}$$

Eqn. (B.21b) yields

$$(B.22) \quad \rho_m^+ = \frac{1}{2} \rho_{m-1}^+.$$

The remaining tolls can be derived using the recursive equation:

$$(B.23) \quad \rho_i^+ = \frac{m+1-i}{m+2-i} \rho_{i-1}^+, \quad i = 2, \dots, m.$$

Eqn. (B.23) is proved by induction. Given eqn. (B.22) it holds for  $i = m$ . Assume it holds for general  $i$ . For  $i-1$ , eqn. (B.21a) and eqn. (B.23) imply

$$\rho_{i-2}^+ - 2\rho_{i-1}^+ + \rho_i^+ = \rho_{i-2}^+ - 2\rho_{i-1}^+ + \frac{m+1-i}{m+2-i} \rho_{i-1}^+ = \rho_{i-2}^+ - \frac{m+3-i}{m+2-i} \rho_{i-1}^+ = 0,$$

or

$$\rho_{i-1}^+ = \frac{m+2-i}{m+3-i} \rho_{i-2}^+.$$

Eqn. (B.23) therefore holds for  $i-1$   $\square$ .

Applying eqn. (B.23) iteratively one obtains, finally,

$$(B.24) \quad \rho_i^+ = \frac{m+1-i}{m} \rho_1^+, \quad i = 1, \dots, m.$$

Setting  $i=2$  in eqn. (B.24), substituting the expression for  $\rho_2^+$  into eqn. (B.20), and collecting terms gives

$$(B.25) \quad \begin{aligned} \left( \frac{m+1}{m} \frac{\gamma}{\beta} + 2(1-R^m) \right) \rho_1 - (1-R) R^{-2} \sum_{i=2}^m R^i \rho_i^- + (1-R) R^{m+1} \sum_{i=2}^m R^{-i} \rho_i^- \\ = \left( 1 + \frac{\gamma}{\beta} - R^m \right) \delta \frac{N}{s}. \end{aligned}$$

The first-order conditions for the step-down tolls,  $\rho_i^-$ , are

$$(B.26) \quad \begin{aligned} (R^{m+1-i} + R^{i-2}) \rho_1 - 2\rho_i^- + (1-R) \left( R^{i-1} \sum_{j=2}^{i-1} R^{-j} \rho_j^- + R^{-1-i} \sum_{j=i+1}^m R^j \rho_j^- \right) \\ = R^{m+1-i} \delta \frac{N}{s}, \quad i = 2, \dots, m-1, \end{aligned}$$

and for  $i = m$ ,

$$(B.27) \quad (R + R^{m-2})\rho_1 - 2\rho_m^- + (1-R)R^{m-1}\sum_{j=2}^{m-1}R^{-j}\rho_j^- = R\delta\frac{N}{s}.$$

(Note that setting  $i = m$  in eqn. (B.26) gives eqn. (B.27).) Replace  $i$  with  $i-1$  in eqn. (B.26), multiply by  $R$ , and subtract the result from eqn. (B.26) to get

$$(B.28) \quad R^{m+1-i}(1-R)\rho_1 - (2-R)\rho_i^- + \rho_{i-1}^- + (1-R)^2R^{-i}\sum_{j=i+1}^mR^j\rho_j^- = R^{m+1-i}(1-R)\delta\frac{N}{s}.$$

Replace  $i$  with  $i+1$  in eqn. (B.28), multiply by  $R$ , and subtract the result from eqn. (B.28) to get

$$\rho_{i-1}^- - 2\rho_i^- + \rho_{i+1}^- = 0, \quad i = 3, \dots, m-1.$$

Rewrite these equations as

$$(B.29) \quad \rho_i^- = 2\rho_{i-1}^- - \rho_{i-2}^-, \quad i = 4, \dots, m.$$

We now prove by induction that

$$(B.30) \quad \rho_i^- = (i-2)\rho_3^- - (i-3)\rho_2^-, \quad i = 2, \dots, m.$$

Eqn. (B.30) holds trivially for  $i=2$  and  $i=3$ , and by eqn. (B.29) it also holds for  $i=4$ .

Assume it holds for general  $i$ . For  $i+1$ , eqn. (B.29) gives

$$\begin{aligned} \rho_{i+1}^- &= 2\rho_i^- - \rho_{i-1}^- = 2((i-2)\rho_3^- - (i-3)\rho_2^-) - ((i-3)\rho_3^- - (i-4)\rho_2^-) \\ &= (i-1)\rho_3^- - (i-2)\rho_2^-. \end{aligned}$$

This concludes the proof.  $\square$

Substitute eqn. (B.30) into eqn. (B.27) to get

$$(B.31) \quad \begin{aligned} &(R + R^{m-2})\rho_1 - 2((m-2)\rho_3^- - (m-3)\rho_2^-) \\ &+ (1-R)R^{m-1}\sum_{j=2}^{m-1}R^{-j}((j-2)\rho_3^- - (j-3)\rho_2^-) = R\delta\frac{N}{s}. \end{aligned}$$

Working out the summation terms in (B.31), and rearranging components, yields a linear equation in  $\rho_1$ ,  $\rho_2^-$  and  $\rho_3^-$ :

$$(B.32) \quad \begin{aligned} &(1-R)(R + R^{m-2})\rho_1 + (m-2 - (m-3)R - 2R^{m-2} + R^{m-1})\rho_2 \\ &+ (1-m + (m-2)R + R^{m-2})\rho_3^- = R(1-R)\delta N/s. \end{aligned}$$

Substitute eqn. (B.30) into eqn. (B.25)

$$(B.33) \quad \left( \frac{m+1}{m} \frac{\gamma}{\beta} + 2(1-R^m) \right) \rho_1 - (1-R) R^{-2} \sum_{i=2}^m R^i \left( (i-2) \rho_3^- - (i-3) \rho_2^- \right) \\ - (1-R) R^{m+1} \sum_{i=2}^m R^{-i} \left( (i-2) \rho_3^- - (i-3) \rho_2^- \right) = \left( 1 + \frac{\gamma}{\beta} - R^m \right) \delta \frac{N}{s}.$$

Combining the summation terms in (B.33), and rearranging components, yields a second linear equation in  $\rho_1$ ,  $\rho_2^-$  and  $\rho_3^-$ :

$$(B.34) \quad \left( \frac{m+1}{m} \frac{\gamma}{\beta} + 2(1-R^m) \right) \rho_1 + (-1 + (m-2)R - (m-2)R^{m-1} + R^m) \rho_2^- \\ - R(m-1)(1-R^{m-2}) \rho_3^- = \left( 1 + \frac{\gamma}{\beta} - R^m \right) \delta \frac{N}{s}.$$

Substitute eqn. (B.30) into eqn. (B.26) with  $i=2$ :

$$(B.35) \quad (1+R^{m-1}) \rho_1 - 2\rho_2^- + (1-R) \rho_3^- + (1-R) R^{-3} \sum_{j=4}^m R^j \left( (j-2) \rho_3^- - (j-3) \rho_2^- \right) = R^{m-1} \delta \frac{N}{s}.$$

Combining the summation terms in (B.35), and rearranging components, yields a third, nonlinear, equation in  $\rho_1$ ,  $\rho_2^-$  and  $\rho_3^-$ :

$$(B.36) \quad (1+R^{m-1}) \rho_1 - 2\rho_2^- + (1-R) \rho_3^- \\ + (1-R) \left[ R \frac{1-R^{m-3}}{1-R} (3\rho_2^- - 2\rho_3^-) + \frac{R(4-3R-(1+m)R^{m-3}+mR^{m-2})}{(1-R)^2} \right] (\rho_3^- - \rho_2^-) \\ = R^{m-1} \delta \frac{N}{s}.$$

### Optimal tolls

Equations (B.32), (B.34) and (B.36) can be solved simultaneously for  $\rho_1$ ,  $\rho_2^-$  and  $\rho_3^-$ . The solutions are homogeneous of degree one in parameters  $(\alpha, \beta, \gamma)$ . Introducing new variables  $b \equiv \beta/\alpha$  and  $c \equiv \gamma/\alpha$  one obtains:

$$(B.37) \quad \rho_1 = \delta \frac{N}{s} \frac{m(1+2b+c(1+m+(2m-1)(2b+c)+(m-1)c^2))}{1+m+4bm+c((m+1)^2+4bm^2+c(m-1-c+2m^2+m^2c))},$$

$$(B.38) \quad \rho_2 = \delta \frac{N}{s} \frac{m-1+c(c^2+c-1-m+2m^2+8m(m-1)b+2m^2c^2-3c^2m-5mc+4m^2c)}{2(1+m+4bm+c((m+1)^2+4bm^2+c(m-1-c+2m^2+m^2c)))},$$

$$(B.39) \quad \rho_3 = \delta \frac{N}{s} \frac{m-1+c(2m^2-3m-3+8m(m-2)b+(4m^2-9m+1)c+(2m^2-5m+3)c^2)}{2(1+m+4bm+c((m+1)^2+4bm^2+c(m-1-c+2m^2+m^2c)))}.$$

The remaining tolls,  $\rho_4^-, \dots, \rho_m^-$ , can then be solved by repeated application of eqn. (B.30).

The formula for the equilibrium cost,  $TC$ , is cumbersome.

The one-step toll has a rather simple solution:

$$\rho_1 = \delta \frac{N}{2s},$$

$$TC = \delta \frac{N^2}{s} \left( 1 - \frac{\gamma(\alpha + 2\beta + \gamma)}{4(\alpha + \gamma)(\beta + \gamma)} \right).$$

However, the two-step toll is already rather complex:

$$\rho_1 = \frac{2}{3} \delta \frac{N}{s} \left( 1 + \frac{2\alpha\beta(\gamma - \alpha)}{3\alpha^3 + 3\gamma^3 + \alpha^2(8\beta + 9\gamma) + \alpha\gamma(16\beta + 9\gamma)} \right),$$

$$\rho_2^- = \rho_1 \frac{\alpha^3 + 5\alpha^2\gamma + 3\gamma^3 + \alpha\gamma(16\beta + 7\gamma)}{4(\alpha^3 + \gamma^3 + 3\alpha\gamma(2\beta + \gamma) + \alpha^2(2\beta + 3\gamma))},$$

$$TC = \delta \frac{N^2}{s} \left( 1 - \frac{\gamma(2\alpha^3 + \gamma^2(3\beta + 2\gamma) + \alpha^2(11\beta + 6\gamma) + 2\alpha(8\beta^2 + 9\beta\gamma + 3\gamma^2))}{2(\beta + \gamma)(3\alpha^3 + 3\gamma^3 + \alpha^2(8\beta + 9\gamma) + \alpha\gamma(16\beta + 9\gamma))} \right).$$

### B.3. Properties of the solution

(a) Limiting behaviour as  $\gamma \rightarrow \alpha$

Equation (B.37) can be rewritten

$$\rho_1 = \delta \frac{N}{s} \frac{m(4m(1+b) + (c-1)(8m-4+2b(2m-1) + (c-1)(4m-2) + (c-1)^2))}{4m(1+b)(m+1) + (c-1)(4(2m^2 + bm^2 + m-1) + (c-1)(5m^2 + m-4) + (c-1)^2(m^2-1))}.$$

In the limit  $\gamma \rightarrow \alpha$ ,  $c \rightarrow 1$  and  $\rho_1 \rightarrow m/(m+1)$ . Similarly, it is readily shown that

$\rho_2 \rightarrow (m-1)/(m+1)$  and  $\rho_3 \rightarrow (m-2)/(m+1)$ . Using eqn. (B.30) for  $i > 3$ , one finds

$\rho_i \rightarrow (m+1-i)/(m+1)$ .

*(b) Sizes of toll steps*

Given eqn. (B.24) the size of the step between consecutive step-up tolls is constant. The size of the step between consecutive step-down tolls is constant for tolls  $\rho_2^-$  through  $\rho_m^-$ . To see this, apply eqn. (B.30) for  $i$  and  $i-1$ :

$$\rho_i^- = (i-2)\rho_3^- - (i-3)\rho_2^-,$$

$$\rho_{i-1}^- = (i-3)\rho_3^- - (i-4)\rho_2^-.$$

The difference is

$$\rho_i^- - \rho_{i-1}^- = \rho_3^- - \rho_2^-, \quad i = 4, \dots, m.$$

The step down from  $\rho_1$  to  $\rho_2^-$  is smaller than the other toll steps. Using eqns. ((B.37), (B.38) and (B.39) one obtains after some algebra:

$$\rho_1 - \rho_2^- - (\rho_2^- - \rho_3^-)^s = (c-1)((c+1)(m(c+1) - (c-1)) + 4mb) > 0.$$

*(c) Asymptotic behaviour as the number of toll steps increases***Value of the peak toll**

Taking the limit  $m \rightarrow \infty$ , all terms in the numerator and denominator of eqn. (B.37) drop out except those involving  $m^2$ , and  $\lim_{m \rightarrow \infty} \rho_1 = \delta N / s$ . The peak-period toll thus approaches the peak value of the continuous-time (fine) toll in the limit as  $m \rightarrow \infty$ . Since all steps between consecutive step-up tolls are constant, and all steps between consecutive step-down tolls are constant except for  $\rho_1 - \rho_2^-$  (which is smaller than the other down-steps), the optimal  $m$ -step toll converges to the optimal fine toll in the limit as  $m \rightarrow \infty$ . By continuity, the efficiency gain from the optimal  $m$ -step toll converges to the efficiency gain from the optimal fine toll as well.

**Timing of departures**

Using equation (B.12) it is straightforward but tedious to show that the start time of the peak,  $t_s$ , is a monotonically decreasing function of  $m$ . In the limit as  $m \rightarrow \infty$ , the timing of departures converges to the timing with no toll. From eqn. (B.12) the limiting value of  $t_s$  is:

$$\lim_{m \rightarrow \infty} t_s = t^* - \frac{\gamma}{\beta + \gamma} \frac{N}{s} - \frac{\rho_1}{\beta + \gamma} \lim_{m \rightarrow \infty} \underbrace{R^m}_{=0} - \frac{2\gamma R}{(\alpha + \gamma)(\beta + \gamma)} \lim_{m \rightarrow \infty} \underbrace{\sum_{j=2}^m R^{m-j} \rho_j^-}_{Z}.$$

Term  $Z$  can be written  $Z = \sum_{k=0}^{m-2} R^k \rho_{m-k}^-$ . Since the  $\rho_i^-$  are monotonically decreasing with  $i$ , and have uniform decrements after  $i=2$ , for any given  $k$ :

$$Z < (1 + R + \dots + R^k) \rho_{m-k}^- + R^{k+1} (1 + R + \dots + R^{m-k-3}) \rho_2^- = \underbrace{\frac{1-R^{k+1}}{1-R}}_{(i)} \underbrace{\rho_{m-k}^-}_{(ii)} + \underbrace{R^{k+1}}_{(iii)} \underbrace{\frac{1-R^{m-k-2}}{1-R}}_{(iv)} \underbrace{\rho_2^-}_{(v)}.$$

Now set  $k = \lceil \sqrt{m} \rceil$ : the smallest integer not less than  $\sqrt{m}$ . In the limit  $m \rightarrow \infty$ , terms (i), (iv) and (v) approach finite positive values. Terms (ii) and (iii) converge to zero. Hence  $\lim_{m \rightarrow \infty} Z = 0$ ,

$$\text{and } \lim_{m \rightarrow \infty} t_s = t^* - \frac{\gamma}{\beta + \gamma} \frac{N}{s}. \quad \square$$

### Comparison of ADL tolls with Laih tolls

As noted in the text, if  $m > 1$  and  $\gamma > \alpha$ , each ADL toll is larger than the corresponding Laih toll. Formally,

- (a)  $\rho_1^{ADL} > \rho_1^{Laih}$ ,
- (b)  $\rho_i^{ADL+} > \rho_i^{Laih}$  for  $i = 2, \dots, m$ .
- (c)  $\rho_i^{ADL-} > \rho_i^{Laih}$  for  $i = 2, \dots, m$ .

#### Proof of (a):

Toll  $\rho_1^{Laih}$  is given in eqn. (14), and toll  $\rho_1^{ADL}$  is given in eqn. (B.37). Taking the difference one obtains:

$$\rho_1^{ADL} - \rho_1^{Laih} = (m-1)(c-1)$$

where  $c \equiv \gamma / \alpha$ .

#### Proof of (b):

From eqn. (14),  $\rho_i^{Laih} = \frac{m+1-i}{m} \cdot \rho_1^{Laih}$ ,  $i = 2, \dots, m$ , and from eqn. (17),

$$\rho_i^{ADL+} = \frac{m+1-i}{m} \cdot \rho_1^{ADL}, \quad i = 2, \dots, m. \text{ Result (b) then follows immediately from result (a).}$$

#### Proof of (c):

From eqn. (B.30),  $\rho_i^- = (i-2)\rho_3^- - (i-3)\rho_2^-$  for  $i = 2, \dots, m$ . Using eqn. (B.38), eqn. (B.39) and eqn. (B.14) one obtains after some algebraic manipulations:

$$\rho_i^{ADL-} - \rho_i^{Laih} = (c-1) \left[ (m^2 - 1)(c^2 - 1) + 2((m+1)(1+c) + 4mb)(m+1-i) \right]$$

where  $b \equiv \beta / \alpha$ . Given  $m > 1$  and  $c > 1$ , every term inside the square brackets is strictly positive.



### Appendix C: Derivation of the optimal multi-step tolls in the braking model

In the braking model, the last driver to arrive during the  $i$ th toll period arrives at  $t_i^b$ , and faces no queue. The next driver waits before the tolling point for a period  $\Delta t_i \equiv t_i^- - t_i^b$ , until the toll is lowered at  $t_i^-$ . In equilibrium, the price for the driver who arrives at  $t_i^b$  and pays a toll of  $\rho_i$  equals the price for the first driver who waits and pays  $\rho_{i-1}$  instead. Therefore,  $\Delta t_i = (\rho_{i-1} - \rho_i) / (\alpha + \gamma)$ . The cumulative amount of time that bottleneck capacity is unused equals:

$$(C.1) \quad \Delta t = \sum_{i=1}^n \Delta t_i = \frac{\rho_1}{\alpha + \gamma}.$$

The total slack time,  $\Delta t$ , does not depend directly on  $m$ . Yet, because  $\rho_1$  increases with  $m$ ,  $\Delta t$  also increases. At the start ( $t'_s$ ) and end ( $t'_e$ ) of the peak, travel delays are zero, and the schedule delay costs must be equal. All users must pass the bottleneck during the time that it is used,  $t'_e - t'_s - \Delta t$ . Hence:

$$(C.2) \quad t_s = t^* - \frac{\gamma}{\beta + \gamma} \left( \frac{N}{s} + \Delta t \right) = t^* - \frac{\gamma}{\beta + \gamma} \left( \frac{N}{s} + \frac{\rho_1}{\alpha + \gamma} \right),$$

$$(C.3) \quad t_e = t^* + \frac{\beta}{\beta + \gamma} \left( \frac{N}{s} + \Delta t \right) = t^* + \frac{\beta}{\beta + \gamma} \left( \frac{N}{s} + \frac{\rho_1}{\alpha + \gamma} \right).$$

The first driver to pay the step-up toll  $\rho_i$  faces no queue. The price paid at time  $t_i^+$  is  $\beta(t^* - t_i^+) + \rho_i$  which must match the price at time  $t'_s$ ,  $\beta(t^* - t'_s)$ . The toll therefore starts at time:

$$(C.4) \quad t_i^+ = t'_s + \frac{\rho_i}{\beta}, \quad i = 1, \dots, m.$$

Similarly, the last driver to pay the step-down toll  $\rho_i$  arrives at time:

$$(C.5) \quad t_i^b = t'_e - \frac{\rho_i}{\gamma}, \quad i = 1, \dots, m.$$

Total social costs equal total private costs minus toll revenue:

$$(C.6) \quad TC = PC - TR = PC - TR_1 - \sum_{i=2}^m TR_i.$$

The component terms are:

$$(C.7) \quad PC = \delta N \left( \frac{N}{s} + \Delta t \right) = \delta N \left( \frac{N}{s} + \frac{\rho_1}{\alpha + \gamma} \right).$$

$$(C.8) \quad TR_1 = s\rho_1(t_1^- - t_1^+) = s\rho_1\left(\frac{N}{s} + \frac{\rho_1}{\alpha + \gamma} - \frac{\rho_1}{\beta} - \frac{\rho_1}{\gamma}\right) = s\rho_1\left(\frac{N}{s} + \frac{\rho_1}{\alpha + \gamma} - \frac{\rho_1}{\delta}\right).$$

$$(C.9) \quad \begin{aligned} TR_i &= s\rho_i\left[\left(t_{i-1}^+ - t_i^+\right) + \left(t_i^b - t_{i-1}^b - \Delta t_{i-1}\right)\right] \\ &= s\rho_i\left[\left(\frac{\rho_{i-1} - \rho_i}{\beta}\right) + \left(\frac{\rho_{i-1} - \rho_i}{\gamma} - \frac{\rho_{i-1} - \rho_i}{\alpha + \gamma}\right)\right] \quad i = 2, \dots, m. \\ &= s\rho_i(\rho_{i-1} - \rho_i)\left(\frac{1}{\delta} - \frac{1}{\alpha + \gamma}\right) \end{aligned}$$

The first-order conditions for minimizing  $TC$  with respect to the tolls are determined by inserting eqns. (C.2) through (C.5) into (C.1), and taking derivatives:

$$(C.10) \quad \begin{aligned} \frac{\delta}{\alpha + \gamma}N - N + s\rho_1\left(\frac{2}{\alpha + \gamma} - \frac{2}{\delta}\right) - s\rho_2\left(\frac{1}{\delta} - \frac{1}{\alpha + \gamma}\right) & \quad \text{for } i = 1, \\ \rho_{i-1} - 2\rho_i + \rho_{i+1} = 0, & \quad \text{for } i = 2, \dots, m-1, \\ \rho_{m-1} - 2\rho_m = 0 & \quad \text{for } i = m. \end{aligned}$$

Eqn. (C.10) for  $i = 1, \dots, m-1$  gives the same first-order conditions as in the Laih model. Hence:

$$(C.11) \quad \rho_i = \frac{m+1-i}{m}\rho_1, \quad i = 1, \dots, m.$$

Inserting (C.11) for  $\rho_2$  into (C.10a) gives:

$$(C.12) \quad \rho_1 = \frac{m}{m+1}\delta\frac{N}{s},$$

which is the same formula as in the Laih model. The formulas are also the same for the time at which the first driver arrives during the  $i$ th early toll ( $t_i^+$ ), and the time when the last driver arrives during the  $i$ th late toll period ( $t_i^-$  in the Laih model, and  $t_i^b$  in the braking model). However, the timing of the peak differs because bottleneck capacity is unused for a time  $\Delta t_i$  after  $t_i^b$ . Inserting eqns. (C.11) and (C.12) for the tolls into the total-cost function gives:

$$(C.13) \quad TC_{Braking} = \delta\frac{N^2}{s}\left(1 - \frac{1}{2}\frac{m}{m+1}\left(1 - \frac{\beta\gamma}{(\beta + \gamma)(\alpha + \gamma)}\right)\right),$$

$$(C.14) \quad \omega_{Braking} = \frac{TC_{NT} - TC_{Braking}}{TC_{NT} - TC_{FB}} = \frac{m}{1+m}\left(1 - \frac{\beta\gamma}{(\beta + \gamma)(\alpha + \gamma)}\right).$$