



HAL
open science

Cost Recovery from Congestion Tolls with Long-run Uncertainty

Robin Lindsey, André de Palma

► **To cite this version:**

Robin Lindsey, André de Palma. Cost Recovery from Congestion Tolls with Long-run Uncertainty. 2013. hal-00784299

HAL Id: hal-00784299

<https://hal.science/hal-00784299>

Preprint submitted on 4 Feb 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Cost Recovery from Congestion Tolls with Long-run Uncertainty

Robin Lindsey
Sauder School of Business
University of British Columbia
2053 Main Mall, Vancouver, British Columbia
CANADA V6T 1Z2
Robin.Lindsey@sauder.ubc.ca
Voice: 1 (604) 822 3323; FAX: 1 (604) 822 9574

André de Palma
Economics and Management Department
Ecole Normale Supérieure de Cachan
61, avenue du Président Wilson
94235 Cachan Cedex
and Ecole Polytechnique
France
andre.depalma@ens-cachan.fr
Voice: +33 6 63 64 43 20 ; Fax : + 33 1 47 40 24 60

February 1, 2013

Key words: Congestion pricing; cost recovery; road capacity; cost uncertainty; demand uncertainty; irreversible investment

JEL codes: D62, H21, R41, R42, R48

Abstract

According to the seminal Cost Recovery Theorem the revenues from congestion tolls pay for the capacity costs of an optimal-sized facility if capacity is perfectly divisible, and if user costs and capacity costs have constant scale economies. This paper extends the Theorem to long-run uncertainty about investment costs, user costs, and demand. It proves that if constant scale economies hold at all times and in all states, and if the toll can be varied freely over time and by state, then expected discounted toll revenues cover expected discounted investment costs over a facility's lifetime. If the marginal cost of investment is constant and investment is reversible, then expected cost recovery is also achieved for each investment. Cost recovery is quite sensitive to estimated initial demand, and moderately sensitive to the estimated growth rate of demand. Natural variability in demand can result in substantial surpluses or deficits over a facility's lifetime.

1 INTRODUCTION

Herbert Mohring made a number of landmark contributions to transportation economics over his long career. One of his greatest achievements is the self-financing or cost recovery¹ theorem (Mohring and Harwitz, 1962). The theorem states that the revenues from optimal congestion tolls pay for the capacity costs of an optimal-sized facility if capacity is perfectly divisible, and if user costs and capacity costs both have constant scale economies. The theorem is of interest for at least two reasons. First, it establishes that pricing a facility at marginal social cost to support efficient usage may be compatible with pricing the facility at average cost to finance it.² Second, the theorem is appealing from a normative standpoint because it shows that efficient pricing is consistent with the user-pay principle: there is no need to subsidize users, and users do not have to pay more than costs for the services they consume. The lack of need for a subsidy is especially attractive given the chronic shortage of funds for investment in, and operations of, public facilities.

Mohring and Harwitz derived their theorem for a deterministic environment. Yet uncertainty is practically important for many transportation and other facilities. In the case of roads, demand and capacity fluctuate unpredictably from day to day due to weather, accidents, unplanned road maintenance, and so on. Lindsey (2009) shows that the theorem continues to hold in the face of short-run uncertainty if two additional assumptions are imposed: individuals learn supply and demand conditions before deciding whether to use a facility, and the congestion toll is varied *responsively* to maintain efficient usage levels.

This paper is concerned with uncertainty about investment costs, user costs, and demand over a facility's lifetime rather than with short-run capacity and demand fluctuations. The theoretical treatment is general, but the focus is on roads. The costs and time required to build, expand and rehabilitate a road are uncertain. Costs can rise because of changes in technical specifications, new construction methods, demands from municipalities for better network connections and so

¹ The terms "self-financing" and "cost recovery" will be used interchangeably in this paper.

² Some prominent economists at the time, including Beckmann et al. (1956) and Nelson (1962), had expressed doubts that the goals of efficient usage and cost recovery could be reconciled. Indeed, the tension between the two goals dates back to Jules Dupuit and Arthur Pigou. For a historical review see Lindsey (2006).

on (Nijkamp and Ubbels, 1999; Berechman, 2009). Major cost overruns and delays are common for toll roads. In a large international survey, Flyvbjerg et al. (2003) found an average cost escalation of 20.4% for road projects, and 33.8% for bridges and tunnels.³

Road operations and maintenance costs are also unpredictable. Input costs (e.g. labour, fuel, and material) can vary significantly over time. Natural disasters such as earthquakes and hurricanes cause extensive damage. Climate change affects the frequency and severity of extreme weather, flooding, frost heave, and so on. Traffic volumes are also a major source of uncertainty. In another large international survey, Flyvbjerg et al. (2006) found that, for half of road projects, actual traffic deviated from forecasted traffic by more than $\pm 20\%$.⁴ Traffic volumes are affected by a host of unpredictable factors: project completion time, economic growth rates, fuel prices, land-use developments, construction of competing or complementary roads, environmental concerns that curb automobile usage, changing preferences with respect to housing and mode choice, and so on.⁵

Despite improvements in data collection and econometric methods, forecasts have not become more accurate over time (Flyvbjerg et al., 2006; Transportation Research Board, 2006). Optimistic demand projections tend to be the norm for toll road projects. Bain (2009) identifies several reasons: lower-than-expected travel time savings; over-estimation of drivers' values of time⁶ and corresponding willingness to pay tolls; and errors in designing complex tolling schemes in which tolls vary by vehicle type, section of road, and time of day.

Technology is a third factor that can affect cost recovery over a road's lifetime. Traffic management system techniques such as ramp metering help to regulate demand. Incident Management Systems reduce the duration of traffic incidents. Advanced Traveler Information Systems notify motorists about traffic conditions. Road vehicles are becoming smaller, smarter, and safer. Vehicle collision avoidance systems, lane-departure warning systems, driver fatigue monitoring systems, heads-up displays, and improved braking systems are reducing the

³ Other studies of cost overruns include Odeck (2004) and Berechman and Chen (2011).

⁴ Other studies of bad forecasts include Prozzi et al. (2009) and Williams-Derry (2011).

⁵ Lindsey (2012) discusses the future evolution of road travel demand.

⁶ See Hensher and Goodwin (2004).

probability of accidents that contribute to congestion. By increasing effective road capacity, and managing demand, these technologies help to improve the utilization of roadways.⁷

A final influence on capacity and cost recovery is flexible road capacity design. The capacity of existing roads can be increased or decreased by re-striping lanes, allowing vehicles to use shoulders during peak periods, changing speed limits, introducing or eliminating features to accommodate public transit and/or bicycles, and so on (Ng and Small, 2012). The appropriate date at which to make these adjustments depends on traffic volumes, ITS technology and vehicle designs, and is therefore unpredictable.

This paper addresses two major questions about cost recovery. First and foremost, does the Mohring-Harwitz cost recovery theorem extend to long-run uncertainty in some well-defined sense? Second, how likely are large surpluses or deficits due either to errors and biases in investment decisions or to natural variability in demand and other factors over a facility's lifetime? Section 2 reviews the theoretical literature. Section 3 sets out the model. Section 4 present two versions of a cost recovery theorem with long-run uncertainty. Section 5 examines the prevalence of surpluses and deficits, and Section 6 concludes. Most of the algebraic analysis is relegated to the Appendix.

2 THEORETICAL BACKGROUND

The seminal result on cost recovery is due to Mohring and Harwitz (1962). Stated formally:

Mohring-Harwitz Cost Recovery Theorem⁸ Assume: (a) the user cost function is homogeneous of degree zero in usage and capacity⁹; (b) capacity is perfectly divisible and (c) capacity is supplied at unit cost elasticity. Then toll revenues pay for the capacity costs of an optimal-sized facility.¹⁰

⁷ As evidence, capacity values in the US Highway Capacity Manual have increased over time (Elefteriadou, 2004).

⁸ Mohring and Harwitz (1962, pp. 84-86).

⁹ With homogeneity of degree zero, the cost of usage does not change if usage and capacity are scaled proportionally.

¹⁰ The conditions for self-financing are actually less restrictive than stated because capacity can be defined in such a way that condition (c) is always satisfied. A more general version of the theorem states that the ratio of toll revenues to capacity costs equals the local elasticity of the capacity cost function.

The Mohring-Harwitz Cost Recovery Theorem (henceforth CRT) has been extended in various directions and is relatively robust.¹¹ The main requirements are that assumptions (a), (b) and (c) continue to hold, and tolls be flexible enough to price usage at marginal cost.

A few studies have addressed aspects of optimal investment and congestion pricing under uncertainty without considering cost recovery per se. DeVany and Saving (1980) analyze capacity and toll decisions for competing, profit-maximizing highways when demand fluctuates randomly over time. In the steady-state equilibrium of their model (which is allocatively efficient) firms earn zero expected profits so that their expected toll revenues cover their capital costs. However, this result is due to competition and does not depend on the properties of the user cost and capacity cost functions.¹²

Kraus (1982) assesses the effect of demand uncertainty on optimal highway capacity. He assumes that demand remains constant over time, but its level is unknown when capacity is built. More specifically, the planner is uncertain about a parameter of the representative individual's utility function, where utility depends on travel and a composite good. The planner maximizes the representative individual's utility subject to covering capacity costs with expected revenues from congestion tolls and a head tax. Uncertainty induces the planner to increase capacity. Depending on parameter values, the congestion toll can be higher or lower with uncertainty although reductions seem more likely. With higher capacity and lower tolls cost recovery falls, although Kraus does not explore this implication.

Two more recent papers form the basis of our model and analysis: one by Lindsey (2009) on cost recovery with short-run uncertainty, and one by Arnott and Kraus (1998) on long-run cost recovery with no uncertainty. Lindsey (2009) assumes that demand and capacity fluctuate randomly according to a stationary probability distribution. His main result is reproduced here with assumptions additional to those of the Mohring-Harwitz CRT italicized for ease of reference:

¹¹ See de Palma and Lindsey (2007) for a review.

¹² Indeed, to obtain a limit on firm size DeVany and Saving (1980) assume that firms experience decreasing returns to scale.

Cost recovery with short-run uncertainty.¹³ Assume: (a) the user cost function is homogeneous of degree zero in usage and capacity; (b) *design* capacity is perfectly divisible; (c) *design* capacity is supplied at unit cost elasticity; (d) *in any state realized capacity is proportional to design capacity and (e) prospective users know the state and the toll is fully flexible.* Then *expected* toll revenues pay for the capacity costs of an optimal-sized facility.

The CRT with short-run uncertainty differs from the Mohring-Harwitz CRT in three respects. First, design capacity must be distinguished from realized capacity because realized capacity is random. Second, users must know the state so that they can make informed decisions on whether to use the facility. Third, the toll must be fully flexible so that first-best (i.e., marginal cost) pricing can be implemented in all states. As Lindsey (2009) remarks, if demand and capacity shocks are statistically independent then the Law of Large Numbers applies, and congestion toll revenues converge toward capacity costs so that cost recovery is achieved in practice as well as in expectation.

Arnott and Kraus (1998) undertake a comprehensive analysis of long-run cost recovery under non-stationary demand conditions when the facility manager has perfect foresight. They work through one series of models in which capacity is added continuously in time, and another series of models in which capacity is added intermittently. They show that, given assumptions that parallel those of the Mohring-Harwitz CRT, self-financing usually holds in present-value terms if prices are flexible over time.¹⁴ Anticipating the main results of this paper they also conjecture that with uncertainty self-financing holds in expected PDV terms if prices are not only flexible over time, but also can be conditioned on the state.

In this paper we adopt Arnott and Kraus's most general formulation of capacity investment technology.¹⁵ The facility is built from scratch and expanded intermittently as demand grows. Their result (which they do not formally state) is summarized here as:

¹³ Lindsey (2009, Proposition 1).

¹⁴ The one exception is when the size of investments is constrained.

¹⁵ Arnott and Kraus (1998, Section 4.4).

Cost recovery in the long run.¹⁶ Assume: (a) the user cost function is homogeneous of degree zero in usage and capacity; (b) capacity is perfectly divisible and (c) *capacity investment functions are homogeneous of degree one in investment and capacity*. Then *discounted toll revenues pay for the capacity costs of an optimal-sized facility over its lifetime*.

The Arnott-Kraus CRT differs from the Mohring-Harwitz CRT in two respects. First, a distinction is made between capacity and investment since capacity is not all built at once. The assumption of homogeneity degree one of capacity costs is replaced by the more general condition that the cost of investment is homogeneous of degree one in investment and capacity. Second, cost recovery is achieved in PDV terms over the facility's lifetime rather than "all at once" as in the static Mohring-Harwitz CRT. Toll revenues do not, in general, pay for investment costs over each investment cycle.

We now turn to our model which extends the Arnott-Kraus model to uncertainty, and also allows for irreversibility of investment.¹⁷

3 THE MODEL

The model is adapted from Arnott and Kraus (1998, Section 4.4). There is a single, isolated facility which agents decide whether to use. Agents are identical except for their willingness to pay for usage. Demand and supply conditions depend on time (t) and the state (w). The probability distribution of states can vary autonomously over time, and the distribution of states perceived by the facility manager changes as uncertainty is resolved over time. Aggregate demand at time t in state w , N_{tw} , is a function of the full price or generalized cost of usage, p_{tw} : $N_{tw}(p_{tw})$. The full price is the sum of the user cost and the toll. The user cost at time t in state w is a function $C_{tw}(N_{tw}, s_{tw})$ of usage (N_{tw}) and capacity (s_{tw}). Given a toll τ_{tw} at time t in state w , the full price is

¹⁶ Arnott and Kraus (1998, pp. 176-178).

¹⁷ Arnott and Kraus (1998) include irreversibility in one of their models with continuous capacity investment (Section 3.3), but not in the intermittent investment model we use.

$$p_{tw} = C_{tw}(N_{tw}, s_{tw}) + \tau_{tw}. \quad (1)$$

Initial design capacity, \hat{s}_0 , is zero. Investments in capacity are made at fixed times T_i , $i=1,2,\dots,M$, with $T_1=0$. The number of investments (M) may be finite or infinite. The facility manager chooses the amount of investment, I_i , at time T_i , for all i . Depending on the nature of capacity, investment may be reversible (I_i free) or irreversible ($I_i \geq 0$). Capacity does not depreciate so that design capacity at time T_i , \hat{s}_i , is the sum of investments up to and including investment i : $\hat{s}_i = \sum_{j=1}^i I_j$. Design capacity is interpreted to be the maximum capacity available under ideal conditions. Since the capacity available at time t in state w , s_{tw} , may be less than design capacity, $s_{tw} \leq \hat{s}_i$ for $t \in [T_i, T_{i+1})$. The cost of investment is specified by a function $A_i(I_i, \hat{s}_i)$ which is increasing with respect to I_i , and increasing or decreasing with respect to \hat{s}_i . Investment costs can evolve from investment cycle to cycle (hence the subscript i on A) due to technology change or other factors.¹⁸ To allow for cost overruns it is assumed that the $A_i(\cdot)$ are stochastic, and the uncertainty is not resolved until after the investments are made.

The facility manager is assumed to be risk neutral and maximizes welfare as measured by expected PDV social surplus when the first investment is made at $T_1=0$. Social surplus equals consumers' surplus plus toll revenue minus investment costs. At time t in state w , consumers' surplus is $\int_{p_{tw}} N_{tw}(p) dp$, and toll revenue is $\tau_{tw} N_{tw}(p_{tw})$. To allow for a time-varying discount rate, the discount factor at time t is written $e^{-\Gamma_t}$ where $\Gamma_t \equiv \int_{u=0}^t r(u) du$. Let E^i denote the expectations operator at time T_i , and let $f_t^i(w)$ denote the probability density perceived at time T_i of states at time t .¹⁹ Welfare can then be written:

¹⁸ Arnott and Kraus (1998) allow the investment function to depend on the time elapsed since the previous investment, $T_i - T_{i-1}$. This possibility is ignored here since the timing of investments is exogenous, and any such dependence would not affect results of interest.

¹⁹ Density $f_t^i(w)$ depends on random events from 0 to T_i , and hence is a random variable from the perspective of T_i if $t > T_i$.

$$\begin{aligned}
W = E^1 \left\{ \sum_{i=1}^M \int_{t=T_i}^{T_{i+1}} e^{-\Gamma t} \int_w \left(\int_{p_{tw}} N_{tw}(p) dp + \tau_{tw} N_{tw}(p_{tw}) \right) f_t^i(w) dw dt \right\} \\
- E^1 \left\{ \sum_{i=1}^M e^{-\Gamma T_i} E^i \{ A_i(I_i, \hat{s}_i) \} \right\}.
\end{aligned} \tag{2}$$

The capacity accumulation constraints are

$$\hat{s}_i \leq \sum_{j=1}^i I_j, \quad i=1 \dots M \quad [\phi_i], \tag{3}$$

where the ϕ_i are multipliers. To allow for the possibility that investment is irreversible,

nonnegativity constraints on investment are also imposed:

$$I_i \geq 0, \quad i=1 \dots M \quad [\mu_i], \tag{4}$$

where the μ_i are multipliers. If investment is reversible, the multipliers μ_i are zero by default.

Given eqns. (2)-(4), the Lagrangian is²⁰

$$\begin{aligned}
\mathfrak{L} = E^1 \left\{ \sum_{i=1}^M \int_{t=T_i}^{T_{i+1}} e^{-\Gamma t} \int_w \left(\int_{p_{tw}} N_{tw}(p) dp + \tau_{tw} N_{tw}(p_{tw}) \right) f_t^i(w) dw dt \right\} \\
- E^1 \left\{ \sum_{i=1}^M e^{-\Gamma T_i} E^i \{ A_i(I_i, \hat{s}_i) \} \right\} + E^1 \left\{ \sum_{i=1}^M e^{-\Gamma T_i} E^i \left\{ \phi_i \left(\sum_{j=1}^i I_j - \hat{s}_i \right) \right\} \right\} \\
+ E^1 \left\{ \sum_{i=1}^M e^{-\Gamma T_i} E^i \{ \mu_i I_i \} \right\}.
\end{aligned} \tag{5}$$

The Lagrangian in eqn. (5) is maximized by choosing tolls $\{\tau_{tw}\}$ and investments $\{I_i\}$.

4 THE COST RECOVERY THEOREM

Because the toll is assumed to be freely variable, the optimal toll at each time and in each state is given by the classical formula:

$$\tau_{tw} = \frac{\partial C_{tw}}{\partial N_{tw}} N_{tw}, \quad \forall t, w. \tag{6}$$

Toll revenue at time t in state w is therefore $R_{tw} = \tau_{tw} N_{tw} = (\partial C_{tw} / \partial N_{tw}) N_{tw}^2$, and expected PDV

toll revenues over the full time horizon are

²⁰ The multipliers ϕ_i and μ_i are discounted in eqn. (5) to facilitate interpretation and manipulation of first-order conditions.

$$E^1 \{R\} = E^1 \left\{ \sum_{i=1}^M \int_{t=T_i}^{T_{i+1}} e^{-\Gamma_t} \int_w \frac{\partial C_{tw}}{\partial N_{tw}} N_{tw}^2 f_t^i(w) dw dt \right\}. \quad (7)$$

The optimal investment rules are derived in the appendix. If the user cost functions $C_{tw}(N_{tw}, s_{tw})$ are homogeneous of degree zero, the optimality conditions yield

$$E^1 \left\{ \sum_{i=1}^M \int_{t=T_i}^{T_{i+1}} e^{-\Gamma_t} \int_w \frac{\partial C_{tw}}{\partial N_{tw}} N_{tw}^2 \varepsilon_{tw}^i f_t^i(w) dw dt \right\} = E^1 \left\{ \sum_{i=1}^M e^{-\Gamma_{T_i}} \left(E^i \left\{ \frac{\partial A_t}{\partial \hat{s}_i} \hat{s}_i \right\} + E^i \left\{ \frac{\partial A_t}{\partial I_i} I_i \right\} \right) \right\}, \quad (8)$$

where $\varepsilon_{tw}^i \equiv (\partial s_{tw} / \partial \hat{s}_i)(\hat{s}_i / s_{tw})$ is the elasticity of realized capacity at time t in state w with respect to design capacity during the time interval $t \in [T_i, T_{i+1})$. The left-hand side of (8) matches expected PDV toll revenues in eqn. (7) if $\varepsilon_{tw}^i = 1, \forall t, w, i$. If the investment cost functions $A_t(\cdot)$ are homogeneous of degree one in I_i and \hat{s}_i , the right-hand side of (8) reduces to

$E^1 \left\{ \sum_{i=1}^M e^{-\Gamma_{T_i}} E^i \{A_t(I_i, \hat{s}_i)\} \right\}$ which is expected PDV investment costs. Expected PDV toll revenues then match expected PDV investment costs, and cost recovery is achieved. This result is formalized in:

Theorem 1 (Cost recovery with long-run uncertainty-1). Assume: (a) *at any date and in any state* the user cost function is homogeneous of degree zero in usage and capacity; (b) capacity is perfectly divisible; (c) capacity investment functions are homogeneous of degree one in investment and capacity; (d) *at any date and in any state realized capacity is proportional to design capacity* and (e) *prospective users know the state and the toll is fully flexible*. Then *expected* discounted toll revenues pay for *expected* discounted investment costs of an optimal-sized facility over its lifetime.

Theorem 1 generalizes the Arnott-Kraus CRT with perfect foresight as well as the CRT with short-run uncertainty in Lindsey (2009). The intuition for Theorem 1 is similar to the intuition in Lindsey (2009): random fluctuations in construction costs, capacity availability, and demand are equivalent in expected terms to systematic fluctuations with the same frequency distribution. Expected revenues are then the same too. The three assumptions in Theorem 1 that are additional to the assumptions required for the Arnott-Kraus CRT are italicized for ease of reference. First, the user cost functions must be homogeneous of degree zero at all dates and in all states. Second,

proportionality of actual and realized capacity must hold at all times and in all states. Third, the toll must be flexible at all times and in all states.

Two features of Theorem 1 deserve emphasis. First, unlike with the Arnott-Kraus CRT, costs are recovered only in expected terms. Because construction costs, capacity, and demand are random, a surplus or deficit may be realized in practice. Second, expected cost recovery is achieved only over the facility's lifetime. Expected toll revenues do not necessarily cover investment costs over each investment cycle. Surpluses may be anticipated over some cycles, and deficits over others. To assure expected cost recovery over each investment cycle, two additional assumptions are required: (1) the marginal cost of investment must be constant and independent of design capacity and (2) investment must be reversible. Given these additional assumptions, a stronger variant of Theorem 1 obtains:

Theorem 2 (Cost recovery with long-run uncertainty-2): Let assumptions (a)-(e) of Theorem 1 hold. Further assume: (f) the marginal cost of investment is constant and does not depend on design capacity and (g) investment is reversible. Then expected discounted toll revenues pay for expected discounted investment costs of an optimal-sized facility *over each investment cycle*.

Theorem 2 is proved in the appendix. Although the assumptions required for Theorems 1 and 2 are more stringent than for the other cost recovery theorems, they still admit a wide range of possibilities. Construction costs can rise or fall, and construction technology can change as long as the investment cost function retains constant scale economies. Demand can increase or decrease, the discount rate can change, and the probability distribution of states can evolve over time. Travel costs can change due to advances in ITS technology, vehicle design, and road design. Theorems 1 and 2 also admit cost-recovery with short-run uncertainty if the toll is responsive to short-run capacity and demand shocks such as traffic accidents. If no such short-run shocks ever occur, toll flexibility is only required with respect to slower-evolving conditions such as seasonal or business-cycle-related demand fluctuations.

Empirical evidence on the validity of the assumptions underlying the Mohring-Harwitz CRT is limited and mixed.²¹ An additional consideration relevant for Theorems 1 and 2 is whether the

²¹ See Small and Verhoef (2007, Section 5.1) and de Palma and Lindsey (2007, Section 4).

frequency and severity of cost overruns varies with project scale. As Jahren and Asha (1990) point out, one might expect cost overruns to be more prevalent for large projects since they are generally more complex. However, because the stakes are also higher with large projects, more care may be exercised in the bidding and planning process, and better managers may be assigned to the projects. In their large international survey, Flyvbjerg et al. (2003) found no statistically significant effect of estimated project cost on cost escalation. By contrast, Odeck (2004) found for a sample of Norwegian road construction projects that cost overruns were more prevalent for smaller projects. Given the limited and inconclusive evidence on cost overruns, we take assumption (c) in Theorems 1 and 2 to be a reasonable working hypothesis.

5 DEVIATIONS FROM COST RECOVERY

Theorems 1 and 2 establish that, under plausible assumptions, *expected* PDV toll revenues cover *expected* PDV construction costs. Yet the theorems are silent about the degree to which costs will actually be recovered once uncertainty has been resolved. Both surpluses and deficits are possible, and experience with roads and other transportation infrastructure projects shows that large departures from cost recovery do occur. Deficits are particularly damaging for toll road projects that are financed through non-recourse loans since they may lead to bankruptcy (Welde and Odeck, 2011). Surpluses are also problematic if they result from decisions to build roads that are too small and fail to provide adequate congestion relief.²²

In this section we examine deviations from cost recovery. We begin by considering cost overruns and errors in capacity choice decisions. We then consider deviations from cost recovery due to the inherent unpredictability of demand. Finally, we explore the effects of investment irreversibility. Expected cost recovery over individual investment cycles is guaranteed when investment is reversible (cf Theorem 2) but not when it is irreversible (cf Theorem 1).

²² According to Flyvbjerg et al. (2006), this has been the case for a number of road projects in the UK.

5.1 Investment mistakes

5.1.1 Construction cost overruns

The implications for cost recovery of cost overruns are fairly obvious and warrant only brief discussion. In their survey, Flyvbjerg et al. (2003) found that actual costs exceeded forecast costs by an average of 20.4% for roads and 33.8% for bridges and tunnels. These figures might suggest that on average a fraction $1/1.204 = 0.831$ of costs will be recovered for roads, and a fraction $1/1.338 = 0.747$ of costs will be recovered for bridges and tunnels. However, this disregards the possibility that costs were deliberately underestimated, possibly by contractors or political supporters of the projects, in the hope of getting them approved. It also neglects that operations, maintenance, and rehabilitation contribute to lifetime project costs. Thus, construction cost overruns alone are unlikely to drive average cost recovery below about 75%.

5.1.2 Errors in capacity choice

A number of parameters besides construction costs affect cost recovery through their effect on the choice of design capacity. The effects of these parameters are explored in the next subsection using a parametric version of the general model in Section 4. But it is instructive first to consider the effects of design capacity errors directly using a simple, static approach. Let $p(N)$ be the inverse demand curve and assume that the unit cost of investment is a constant, k . Write the user cost function as $C(N, s)$ and assume that capacity is always fully available so that $s = \hat{s}$. The cost recovery ratio is then $\rho = R/(k\hat{s}) = R/(ks)$, and the elasticity of cost recovery with respect to capacity is $\varepsilon_s^\rho = (d\rho/ds)(s/\rho)$. This elasticity can be written $\varepsilon_s^\rho = \varepsilon_s^R - 1 = (dR/ds)(s/R) - 1$, where ε_s^R is the elasticity of toll revenue with respect to capacity. As shown in the Appendix,

$$\varepsilon_s^R = \frac{p_N C_{Ns} N + (2C_N + C_{NN} N) C_s}{-p_N C_s + (2C_N + C_{NN} N) C_s}. \quad (9)$$

The elasticity of toll revenue depends on the slope of the inverse demand curve and the first and second derivatives of the user cost function.

With perfectly elastic demand ($p_N = 0$), eqn. (9) simplifies to $\varepsilon_s^R = 1$. Demand then varies proportionally with capacity and the toll does not change. Consequently, toll revenue varies proportionally with capacity, $\varepsilon_s^P = 0$, and cost recovery is unaffected by mistakes in choosing design capacity.

In the opposite limiting case of perfectly inelastic demand ($p_N = -\infty$), eqn. (9) reduces to

$$\varepsilon_s^R = -\frac{C_{Ns}N}{C_s} = \frac{\partial C_N}{\partial s} \frac{s}{C_N} < 0. \quad (10)$$

An increase in capacity causes revenues to fall because congestion is reduced, the toll drops, and usage does not change. According to most empirical traffic speed-volume curves, travel speeds are quite sensitive to volume-capacity ratios when congestion is severe. Under these conditions, the elasticity $(\partial C_N / \partial s)(s / C_N)$ in eqn. (10) is large in magnitude. The elasticity ε_s^R is then negative and large in magnitude, and ε_s^P is smaller yet. Errors in choosing capacity will therefore affect cost recovery severely.

5.1.3 Errors in parameter estimates

To assess the effects of errors in estimating parameter values we now consider a specific instance of the general model. For simplicity and clarity, attention is focused on the first investment cycle over the time interval $[T_1, T_2] = [0, T]$. The cost of investment is $kI_1 = k\hat{s}$ where k is a constant.²³ The interest rate, r , is assumed to be constant over time.

Demand at time t in state w is given by a constant-price-elasticity function $N_{tw} = n_{tw}P_{tw}^{-\eta}$ where n_{tw} is demand "intensity" and $\eta > 0$ is the price elasticity. Demand intensity evolves according to Geometric Brownian Motion following the stochastic differential equation $dn_t = gn_t dt + \sigma n_t dW_t$, where g is the mean growth rate or drift parameter, σ is the standard

²³ To account for the fact that initial capacity may remain useful after time T , the unit cost of capacity can be deflated. If the cost of capacity is the same for the second investment, the effective cost of capacity for the first investment is $k(1 - e^{-rT})$.

deviation, and W_t is a Wiener process.²⁴ Demand intensity when the road is built, n_0 , is assumed to be known.

User cost is given by the conventional Bureau of Public Roads formula $C(N_{nw}, s_{nw}) = d_t (N_{nw}/s_{nw})^\chi$.²⁵ The coefficient d_t can change over time for several reasons. It will rise if the value of travel time increases. It will fall as advances in ITS technology, vehicle design, and road design smooth traffic flow and permit higher speeds to be safely maintained. To capture these and other influences, d_t is assumed to decline slowly along an exponential path $d_t = d_0 e^{-\xi t}$ where parameter ξ describes the rate of technological progress net of any increase in value of time.²⁶

Capacity is assumed to be fully available at all times in all states so that uncertainty arises only on the demand side. Given these assumptions, optimal design capacity works out to (see the appendix)

$$\hat{s} = n_0 \left(\frac{\chi}{k} \right)^{\frac{1}{\lambda}} (1 + \chi)^{-\eta} d_0^{\frac{1-\eta}{1+\chi}} \left((e^{hT} - 1) / h \right)^{\frac{1}{\lambda}}, \quad (11)$$

where $\lambda \equiv (1 + \chi) / (1 + \chi\eta)$ and $h \equiv \lambda (\xi\chi\eta + g + (\lambda - 1)\sigma^2 / 2) - (r + \xi\chi)$. Design capacity is proportional to initial demand intensity. This follows from the assumption that capacity can be built at a constant unit cost, and a property of Geometric Brownian Motion that demand at all future dates is proportional to initial demand.

Expected cumulative usage, $E\{U\}$, is

²⁴ Geometric Brownian Motion is frequently assumed in the options and option value literature (Dixit and Pindyck, 1994) as well as in transportation studies (e.g., Rose, 1998; Nagae and Akamatsu, 2006; Saphores and Boarnet, 2006; Doan and Patel, 2010; Pimentel et al., 2010). Marathe and Ryan (2005) find empirical support for Geometric Brownian Motion in the case of usage of two established services (electric power consumption and airline passenger enplanements), but not for two emergent services (cellphones and the Internet).

²⁵ The constant term in the BPR formula is omitted to permit analytical solutions. The user cost function therefore specifies the variable or congestion-dependent component of travel cost while omitting the free-flow component.

²⁶ Parameter χ is included in the exponent of the formula for d_t so that d_t can be interpreted as a multiplicative capacity expansion factor. In principle, d_t could be treated like n_t as a random variable.

$$E\{U\} \equiv E\left\{\int_{t=0}^T N_{tw} dt\right\} = n_0^{\frac{1}{1+\chi\eta}} \left((1+\chi)d_0\hat{s}^{-\chi}\right)^{\frac{-\eta}{1+\chi\eta}} (e^{mT} - 1) / m, \quad (12)$$

where $m \equiv (g + \xi\chi\eta)/(1 + \chi\eta) - \chi\eta\sigma^2 / (2(1 + \chi\eta)^2)$. $E\{U\}$ has an elasticity with respect to design capacity of $\varepsilon_s^U = \chi\eta/(1 + \chi\eta) \in (0,1)$. Building a larger road results in greater usage because congestion is reduced. However, as in the static model, usage increases proportionally less than capacity except in the limiting case where demand is perfectly elastic (i.e., $\eta = \infty$). Expected PDV toll revenues work out to

$$E\{R\} = n_0^\lambda \chi(1 + \chi)^{-\eta\lambda} d_0^{\frac{1-\eta}{1+\chi\eta}} (\hat{s})^{-\chi(1-\eta\lambda)} (e^{hT} - 1) / h. \quad (13)$$

The elasticity of $E\{R\}$ with respect to design capacity is $\varepsilon_s^R = -\chi(1-\eta)/(1 + \chi\eta)$. Toll revenue increases or decreases depending on whether demand is elastic or inelastic (i.e., $\eta > 1$ or $\eta < 1$). The elasticity of cost recovery, ρ , with respect to design capacity is $\varepsilon_s^\rho = \varepsilon_s^R - 1 = -(1 + \chi)/(1 + \chi\eta) < 0$. Consistent with the static model, building a larger road reduces cost recovery except in the limiting case $\eta = \infty$. By substituting eqn. (11) for \hat{s} into eqn. (13), it is straightforward to show that expected PDV revenues equal capacity costs: $E\{R\} = k\hat{s}$.

To gain further insights we proceed numerically. We parameterize the model for a three-lane road with a design capacity of $\hat{s} = 6,000$ vehicles per hour. Daily peak-period usage in year 0 is assumed to be $N_0 = 12,000$ so that the peak period lasts for two hours. In year 0 the generalized cost (net of free-flow cost) is $p_0 = \$7.50$, and the toll is $\tau_0 = \$5.00$. Parameter values that support this equilibrium are $\chi = 2$, $\eta = 0.25$, $d_0 = 0.625$, $n_0 = 19,858$, and $k = 131.93$. The duration of the first investment cycle is set to $T = 15$, and the discount rate is set at $r = 0.05$.

It is difficult to judge a reasonable value for parameter ξ . There is also little published information on which to base values of g and σ . Suitable values will vary with rates of national and regional economic growth, rate of growth in automobile ownership, and other factors.²⁷ For exploratory purposes we set the net rate of technological progress to $\xi = 0.01$, the mean annual

²⁷ Prozzi et al. (2009) review forecast and realized traffic growth rates for selected toll roads in North America.

growth rate of demand to $g = 0.02$, and the standard deviation of the annual growth rate to $\sigma = 0.05$. Alternative values of ξ , g , and σ are considered in the sensitivity analysis below.

With these base-case parameters, mean cumulative traffic volume over the 15-year time horizon is 74.5 million. Cumulative expected PDV toll revenues are \$288.9 million, which match construction costs *as per* Theorem 1 and Theorem 2.²⁸

To assess how parameter estimates affect design capacity and expected cost recovery, we alter parameter values one at a time while assuming that the true parameter values remain equal to their base-case values. To do so we use eqn. (11) to compute design capacity with the modified parameter values, and then use eqns. (12) and (13) to compute expected cumulative usage and toll revenues with the true parameter values. Given $\chi = 2$ and $\eta = 0.25$, the elasticity formulas given above yield $\varepsilon_s^U = 1/3$ and $\varepsilon_s^P = -2$. Expected cumulative usage is a slowly increasing function of design capacity, and cost recovery is a sharply decreasing function.

The results of the sensitivity analysis are reported in Table 1. Each quantity is stated as a percentage of the value that would obtain with correctly specified parameters. The first panel of Table 1 shows the effects of misestimating unit construction costs. Estimated parameter values are marked with a \sim (tilde) to distinguish them from the true values. Thus, $\tilde{k} = (5/6)k$ indicates that the unit cost is underestimated by one sixth. In this case, the road is overbuilt by 9.5%. Expected usage increases, but only by a little over 3%.²⁹ The largest effect is on cost recovery which changes in proportion to parameter k . Overestimating construction costs by 1/5 has a similar but opposite effect on capacity, cumulative usage, and cost recovery.

Panel 2 of Table 1 shows that errors in estimating initial demand intensity have a pronounced effect on expected cost recovery. If demand intensity is underestimated by 1/3 (which is not uncommon in practice) the road is built to 2/3 of its optimal capacity. Construction costs are reduced commensurately. The toll, meanwhile, rises considerably because of the nonlinear dependence of congestion delay on capacity. The end result is that over the 15-year period the road generates an expected revenue surplus equal to 125% of its cost. By contrast, if demand is

²⁸ Total construction costs are $365 \hat{k}_s = 365 * 131.93 * 6,000 = \288.9 million.

²⁹ The relative magnitude of these changes is consistent with the point elasticity $\varepsilon_s^U = 1/3$.

overestimated by 1/2 (also not unrealistic), expected toll revenues cover less than half construction costs.

Panel 3 assesses the effects of the mean growth rate in demand. As expected, underestimation of growth results in a smaller road capacity, lower cumulative usage, and higher expected cost recovery. If growth is disregarded, tolls recover on average nearly one third more than costs. By contrast, if growth is overestimated by 2 percentage points, expected cost recovery falls 26% below costs.

Panel 4 shows that misestimating variability in the growth of demand has little effect on design capacity, cumulative usage, or cost recovery. Similarly, panel 5 shows that varying the demand elasticity up or down by a factor of two has only small effects. The rate of technological progress does not have a major influence either (cf. panel 6). Misestimating the rate of progress by ± 1 percentage point has about the same effect as misestimating the mean growth rate of demand by about ± 0.5 percentage points in the opposite direction.

5.2 Probability distribution of revenues and cost recovery

The previous subsection examined deviations from expected cost recovery due to errors in estimating parameters that affect construction costs, demand, or congestion. We now turn to the question of how much revenues and cost recovery for an optimally designed road can differ from their expected values due to the inherent variability of demand.

One simple way to address this question is to consider the probability distribution of revenues at a given date. Doing so gives only a partial view of how much cumulative revenue is generated over a road's full lifetime. However, it has the advantage that closed-form analytical solutions exist. Table 2 presents summary statistics using the base-case parameter values for aggregates of interest in year 10 (i.e., two thirds of the way through the 15-year time horizon).

Uncertainty in the model is driven by stochastic growth in demand intensity. By year 10, expected intensity has increased to $n_0 e^{gt} = (19,858) e^{(0.02)10} = 24,255$. The coefficient of variation (CV) of demand intensity at $t=10$ is about 0.16. The CV for traffic volume is smaller because congestion has a dampening effect on usage. The CV for the toll is larger because traffic volume has a nonlinear effect on congestion delay, and traffic volume and delay vary in tandem. Revenue has the largest CV of all because it is the product of traffic volume and the toll which also vary in tandem. Expected daily revenues in year 10 are \$83,005. This is some 38% higher

than the daily revenues of \$60,000 generated just after the road is built. Nevertheless, because of the large variability in demand there is a 0.19 probability that actual revenues in year 10 are below their starting level.

To derive the probability distribution of cumulative PDV revenues over the full time horizon it is necessary to use numerical methods. This was done by solving the stochastic differential equation for n_t using small (monthly) time increments over the 15-year horizon, and taking a random draw for the Wiener increment, dW_t , at each monthly step. The process was repeated 5,000 times³⁰ to generate the frequency distribution shown in Figure 1. The distribution does not follow a lognormal or other standard distribution. It has a coefficient of variation of 0.225, a coefficient of skewness of 0.869, and a coefficient of kurtosis of 4.41. The distribution is therefore positively skewed. Various calculations can be done with the distribution. For example, revenues at the first quartile of the distribution are 84.1% of mean revenues which implies that they would cover 84.1% of PDV costs. With probability 0.112, cumulative PDV revenues are less than 3/4 of PDV costs. Similarly, with probability 0.073, cumulative PDV revenues exceed 4/3 of PDV costs. Such statistics provide a sense of how far revenues may deviate from achieving exact cost recovery.

5.3 Irreversible investment

According to Theorem 2, if the marginal cost of investment is constant and does not depend on design capacity, and if investment is also reversible, then expected cost recovery holds over each investment cycle. But if investment is irreversible, expected cost recovery is achieved only over the facility's full lifetime, and an expected surplus or deficit can accrue over individual investment cycles. The pattern of expected surpluses and deficits from cycle to cycle will depend on various factors: changes in construction costs and technology, the evolution of macroeconomic conditions, changes in travel behaviour and land usage patterns, and so on.

A comprehensive analysis of surpluses and deficits with irreversible investment would be tedious, and general results are likely to be elusive. However, a clear pattern can be identified in the case of two investments (i.e., $M=2$). Initial capacity is built at time $T_1 = 0$. Capacity can be

³⁰ The simulation took nearly 4 hours using Maple version 8 on a laptop with an Intel 2.53 GHz processor and 6 GB of RAM. Very similar results were obtained with 2,500 iterations.

added at time T_2 , but capacity built at T_1 cannot be liquidated. The facility ends operation at time T_3 , which may be finite or infinite.

In this setting there is an incentive to restrain initial investment. If demand does not grow as quickly as anticipated over the first cycle, capacity may exceed the optimal level for the second cycle and the unwanted capacity cannot be profitably liquidated. By contrast, if demand grows quickly over the first cycle, the shortage can be rectified at time T_2 by investing in lots of additional capacity.³¹ Since investment is held back at T_1 , an expected surplus accrues in the first cycle, and since any shortage of investment at T_1 is made up at T_2 , an expected deficit is incurred in the second cycle. As per Theorem 1, the expected deficit in the second cycle is equal in magnitude to the expected surplus for the first cycle.

The size of the expected surplus and deficit can be explored numerically with a two-period version of the specific model used in subsections 5.1.3 and 5.2.³² We assume $T_1 = 0$, $T_2 = 15$, and $T_3 = 30$, and use the same base-case parameter values as before. Capacity is assumed to be fully salvaged at T_3 so that the effective cost of the initial investment is $k(1 - e^{-rT_3})$. With the base-case parameter values, optimal investment at T_1 is 6,000. On average, a further 1,514 units are invested at T_2 .

The results are presented in Table 3. With the base-case parameter values, no investment occurs at T_2 with probability 0.137. Nevertheless, period 1 investment is held back by a mere half percent relative to the optimum with reversible investment. On average, a surplus of about 2% is earned over the first investment cycle. Similar results are obtained with other parameter values. Irreversibility has a slightly larger effect on cost recovery with a lower mean growth rate in demand (lower g) and a higher variability in the growth rate (higher σ). Varying parameters η , ξ , χ , and r has little effect. Overall, the example suggests that irreversibility has a rather modest effect on cost recovery. Further analysis using empirically calibrated examples are needed to assess the robustness of this finding.

³¹ The incentive to hold back investment in period 1 is analogous to the incentive to postpone investment in the option value literature.

³² The analytical solution, which is rather tedious, is outlined in the appendix.

6 CONCLUSIONS

The Mohring and Harwitz (1962) cost recovery theorem is a seminal contribution to transportation economics as well as microeconomic theory. It establishes that if capacity is perfectly divisible, and if user costs and capacity costs both have constant scale economies, then the revenues from optimal congestion charges pay for the capacity costs of an optimal-sized facility.

This paper investigates whether the cost recovery theorem extends to long-run uncertainty about investment costs, user costs, and demand over the lifetime of a facility. The main conclusion (Theorem 1) is that if constant scale economies continue to hold in expectation at all times and in all states, and if the toll can be varied freely over time and by state, then expected PDV toll revenues cover expected PDV investment costs over a facility's lifetime. Furthermore, if the marginal cost of investment is constant, and investment is reversible, then expected cost recovery is achieved over each investment cycle (Theorem 2). The two theorems generalize results derived for long-run cost recovery with perfect foresight by Arnott and Kraus (1998), and for cost recovery with short-run uncertainty by Lindsey (2009).

The paper also investigates the size of surpluses or deficits that can arise either from errors in estimating key parameters or from natural variability of demand. Using a parametric model it is shown that cost recovery is quite sensitive to estimated initial demand, and moderately sensitive to the growth rate of demand. Even with no investment mistakes, natural variability in demand can result in substantial surpluses or deficits over a facility's lifetime.

Several caveats are in order. First, further research is clearly warranted using more detailed and accurate information on the causes and magnitudes of the various uncertainties that affect cost recovery. Second, the analysis of deviations from cost recovery was conducted mainly with a specific parametric model. The sensitivity analysis was also limited to variation of parameters one at a time. In practice, demand and revenue forecasts often make inaccurate assumptions about multiple parameters at once (Transportation Research Board, 2006). Monte Carlo analysis can be used in an attempt to deal with multiple sources of uncertainty, although it can be difficult to determine the degree of correlation between the variables.

Third, the demand curve has been treated as exogenous in the sense that it does not depend on the size of investments. In practice, usage of a new facility often builds gradually to a steady-state level or steady growth rate (Transportation Research Board, 2006). The so-called "ramp-

up" period for toll roads can last for several years. Demand grows gradually because drivers learn about the new facility with a delay, or are reluctant to try it. The speed at which demand grows depends on experience. Experience is a function of usage, usage is a function of (expected) costs, and costs depend on capacity. It may be possible to accelerate growth in demand by building a larger facility. Setting low, introductory tolls is another common method of gaining market share for a new product. However, such strategies depart from the pure welfare-maximizing objective that underlies the cost recovery theorem.

Fourth, we have ignored uncertainty about delays in project completion which arise due to difficulties in securing a right of way, interest group objections, funding problems, unforeseen engineering and construction difficulties, and so on (Berechman, 2009). The time required to complete a project affects both PDV construction costs and PDV benefits. PDV costs may be reduced if construction is extended over a longer time period, but benefits will also be reduced since usage of the infrastructure will be delayed.

Fifth, we have ignored external costs such as air pollution and noise caused during construction. Measures that are taken to reduce these costs will increase total construction costs, but they do not affect cost recovery unless they affect scale economies in construction costs.

Finally, we have assumed that users are infinitesimal in the sense that they have no market power. This is generally a realistic assumption for roads, but not for airports where individual airlines sometimes account for a substantial fraction of overall traffic. When users have market power the cost recovery theorem usually breaks down (see Zhang and Czerny (2012) for a literature review) although it can be restored if it is recognized that large users have an incentive to contribute to capacity (Verhoef, 2012).

7 ROLE OF THE FUNDING SOURCE

For financial support Lindsey would like to thank the Social Sciences and Humanities Research Council of Canada and the UBC Sauder School of Business new faculty start-up grant. Neither SSHRC nor the Sauder School was involved in any specific aspects of this paper or in the decision to submit the paper for publication.

8 ACKNOWLEDGMENTS

For helpful comments we are grateful to seminar participants at the Kuhmo-Nectar Conference on Transport Economics (2010), the Annual Conference of the Canadian Transportation Research Forum (2011), the Annual Canadian Economic Association Conference (2011), and the Sauder School of Business, UBC (2011). Any errors are our responsibility. A preliminary and incomplete version of this paper was presented at the 46th Annual Conference of the Canadian Transportation Research Forum (de Palma and Lindsey, 2011). That version excludes the theoretical development in Section 3, as well as parts of Section 4.

9 NOTATIONAL GLOSSARY

- $A_i(\)$ investment cost function
- $C_{tw}(\)$ user cost function at time t in state w
- E^i expectations operator at time T_i
- $f_t^i(w)$ probability density perceived at time T_i of state w at time t .
- h composite parameter (specific dynamic model)
- I_i capacity investment at time T_i
- k unit cost of investment (specific dynamic model)
- m composite parameter (specific dynamic model)
- M number of investments (possibly infinite)
- $N_{tw}(\)$ demand function at time t in state w
- p_{tw} full price or generalized cost of usage at time t in state w
- $p(N)$ inverse demand function (static model)
- $r(t)$ instantaneous discount rate at time t
- t time
- R toll revenue
- \hat{s}_i design capacity at time T_i
- s_{tw} capacity realized at time t in state w
- T_i time of investment i
- W welfare

Greek characters

- Γ_t Cumulative discount rate at time t
- ε_{tw}^i elasticity of realized capacity s_{tw} with respect to design capacity \hat{s}_i for $t \in [T_i, T_{i+1})$
- ε_s^R elasticity of toll revenue with respect to design capacity
- ε_s^ρ elasticity of cost recovery ratio with respect to design capacity

η	demand elasticity (specific dynamic model)
λ	composite parameter (specific dynamic model)
μ_i	multiplier for capacity accumulation constraint at time T_i
ξ	rate of technological progress (specific dynamic model)
ρ	fraction of capacity cost recovered by toll revenue
τ_{tw}	toll at time t in state w
ϕ_i	multiplier for nonnegative investment constraint at time T_i
w	state
\mathfrak{L}	Lagrangian

10 APPENDIX

10.1 Optimal investment rules for $M = \infty$

Given the Lagrangian in eqn. (5), the first-order condition for I_i is

$$-e^{-\Gamma T_i} E^i \frac{\partial A_i}{\partial I_i} + E^i \sum_{j=i}^M e^{-\Gamma T_j} \phi_j + e^{-\Gamma T_i} \mu_i = 0, \quad (\text{A1})$$

and the first-order condition for \hat{s}_i is

$$\int_{t=T_i}^{T_{i+1}} e^{-\Gamma t} \int_w \left(\left[-N_{tw}(p_{tw}) + \tau_{tw} \frac{\partial N_{tw}}{\partial p_{tw}} \right] \frac{\partial p_{tw}}{\partial s_{tw}} \frac{\partial s_{tw}}{\partial \hat{s}_i} \right) f_t^i(w) dw dt$$

$$-e^{-\Gamma T_i} \left(E^i \frac{\partial A_i}{\partial \hat{s}_i} + \phi_i \right) = 0, \quad i = 1 \dots M - 1. \quad (\text{A2})$$

From eqn. (1) in the text, $\frac{\partial p_{tw}}{\partial s_{tw}} = \frac{\partial C_{tw}}{\partial N_{tw}} \frac{\partial N_{tw}}{\partial p_{tw}} \frac{\partial p_{tw}}{\partial s_{tw}} + \frac{\partial C_{tw}}{\partial s_{tw}}$ which resolves to

$$\frac{\partial p_{tw}}{\partial s_{tw}} = \frac{\partial C_{tw} / \partial s_{tw}}{1 - (\partial C_{tw} / \partial N_{tw})(\partial N_{tw} / \partial p_{tw})}. \quad (\text{A3})$$

Substituting eqns. (6) and (A3) into (A2) gives

$$\int_{t=T_i}^{T_{i+1}} e^{-\Gamma_i} \int_w \left[-N_{tw} (p_{tw}) + \tau_{tw} \frac{\partial N_{tw}}{\partial p_{tw}} \right] \frac{\partial C_{tw} / \partial s_{tw}}{1 - (\partial C_{tw} / \partial N_{tw})(\partial N_{tw} / \partial p_{tw})} \frac{\partial s_{tw}}{\partial \hat{s}_i} f_t^i(w) dw dt - e^{-\Gamma_{T_i}} \left(E^i \frac{\partial A_i}{\partial \hat{s}_i} + \phi_i \right) = 0. \quad (\text{A4})$$

If $C_{tw}(N_{tw}, s_{tw})$ is homogeneous of degree zero, $(\partial C_{tw} / \partial s_{tw}) s_{tw} = -(\partial C_{tw} / \partial N_{tw}) N_{tw}$. Introducing the elasticities $\varepsilon_{tw}^i \equiv (\partial s_{tw} / \partial \hat{s}_i)(\hat{s}_i / s_{tw})$, $\forall t, w, i$, one has

$$\frac{\partial C_{tw}}{\partial s_{tw}} \frac{\partial s_{tw}}{\partial \hat{s}_i} = \frac{\partial C_{tw}}{\partial s_{tw}} \frac{s_{tw}}{\hat{s}_i} \varepsilon_{tw}^i = -\frac{\partial C_{tw}}{\partial N_{tw}} \frac{N_{tw}}{\hat{s}_i} \varepsilon_{tw}^i. \quad (\text{A5})$$

Using eqns. (6) and (A5), and multiplying through by \hat{s}_i , eqn. (A4) simplifies to

$$\int_{t=T_i}^{T_{i+1}} e^{-\Gamma_i} \int_w \frac{\partial C_{tw}}{\partial N_{tw}} N_{tw}^2 \varepsilon_{tw}^i f_t^i(w) dw dt - e^{-\Gamma_{T_i}} \left(E^i \frac{\partial A_i}{\partial \hat{s}_i} + \phi_i \right) \hat{s}_i = 0. \quad (\text{A6})$$

Next, subtract equation (A1) for investment $i+1$ from equation (A1) for investment i to get:

$$\begin{aligned} & -e^{-\Gamma_{T_{i+1}}} E^{i+1} \frac{\partial A_{i+1}}{\partial I_{i+1}} - e^{-\Gamma_{T_i}} E^i \frac{\partial A_i}{\partial I_i} + e^{-\Gamma_{T_i}} \phi_i + (E^i - E^{i+1}) \sum_{j=i+1}^M e^{-\Gamma_{T_j}} \phi_j, \\ & + e^{-\Gamma_{T_i}} \mu_i - e^{-\Gamma_{T_{i+1}}} \mu_{i+1} = 0, \quad 1 = 1 \dots M - 1 \end{aligned} \quad (\text{A7})$$

or

$$\begin{aligned} \phi_i &= E^i \frac{\partial A_i}{\partial I_i} - e^{-\Gamma_{T_{i+1}} + \Gamma_{T_i}} E^{i+1} \frac{\partial A_{i+1}}{\partial I_{i+1}} + e^{\Gamma_{T_i}} (E^{i+1} - E^i) \sum_{j=i+1}^M e^{-\Gamma_{T_j}} \phi_j \\ & + e^{-\Gamma_{T_{i+1}} + \Gamma_{T_i}} \mu_{i+1} - \mu_i = 0, \quad i = 1 \dots M - 1. \end{aligned} \quad (\text{A8})$$

Substitute (A8) into (A6) to get

$$\begin{aligned} & \int_{t=T_i}^{T_{i+1}} e^{-\Gamma_i} \int_w \frac{\partial C_{tw}}{\partial N_{tw}} N_{tw}^2 \varepsilon_{tw}^i f_t^i(w) dw dt \\ & = e^{-\Gamma_{T_i}} \left(E^i \frac{\partial A_i}{\partial \hat{s}_i} + E^i \frac{\partial A_i}{\partial I_i} - e^{-\Gamma_{T_{i+1}} + \Gamma_{T_i}} E^{i+1} \frac{\partial A_{i+1}}{\partial I_{i+1}} + e^{\Gamma_{T_i}} (E^{i+1} - E^i) \sum_{j=i+1}^M e^{-\Gamma_{T_j}} \phi_j \right) \hat{s}_i \\ & \quad \left(+ e^{-\Gamma_{T_{i+1}} + \Gamma_{T_i}} E^{i+1} - \mu_i \right) \end{aligned} \quad (\text{A9})$$

Sum (A9) over all i , and take expectations at T_1 to get

$$\begin{aligned}
& E^1 \left\{ \sum_{i=1}^M \int_{t=T_i}^{T_{i+1}} e^{-\Gamma_i} \int_w \frac{\partial C_{tw}}{\partial N_{tw}} N_{tw}^2 \mathcal{E}_{tw}^i f_t^i(w) dw dt \right\} \\
& = E^1 \left\{ \sum_{i=1}^M \left[e^{-\Gamma_{T_i}} \left(E^i \frac{\partial A_i}{\partial \hat{s}_i} + E^i \frac{\partial A_i}{\partial I_i} \right) \hat{s}_i - e^{-\Gamma_{T_{i+1}}} E^{i+1} \frac{\partial A_{i+1}}{\partial I_{i+1}} \hat{s}_i + \left((E^{i+1} - E^i) \sum_{j=i+1}^M e^{-\Gamma_{T_j}} \phi_j \right) \hat{s}_i \right] \right. \\
& \quad \left. + \left(e^{-\Gamma_{T_{i+1}}} \mu_{i+1} - e^{-\Gamma_{T_i}} \mu_i \right) \hat{s}_i \right\}. \tag{A10}
\end{aligned}$$

The right-hand-side of (A10) can be rewritten

$$\begin{aligned}
& E^1 \left\{ \sum_{i=1}^M e^{-\Gamma_{T_i}} \left(E^i \frac{\partial A_i}{\partial \hat{s}_i} + E^i \frac{\partial A_i}{\partial I_i} \right) \hat{s}_i \right\} - E^1 \left\{ \sum_{i=2}^M e^{-\Gamma_{T_i}} E^i \frac{\partial A_i}{\partial I_i} (\hat{s}_i - I_i) \right\} \\
& + E^1 \left\{ \sum_{i=1}^M (E^{i+1} - E^i) \sum_{j=i+1}^M e^{-\Gamma_{T_j}} \phi_j \hat{s}_i \right\} + E^1 \left\{ \sum_{i=1}^M \left(e^{-\Gamma_{T_{i+1}}} \mu_{i+1} - e^{-\Gamma_{T_i}} \mu_i \right) \hat{s}_i \right\} \\
& = E^1 \left\{ \sum_{i=1}^M e^{-\Gamma_{T_i}} E^i \frac{\partial A_i}{\partial \hat{s}_i} \hat{s}_i \right\} + e^{-\Gamma_{T_1}} E^1 \frac{\partial A_1}{\partial I_1} \hat{s}_1 + E^1 \left\{ \sum_{i=2}^M e^{-\Gamma_{T_i}} E^i \frac{\partial A_i}{\partial I_i} I_i \right\} \\
& + E^1 \left\{ \sum_{i=1}^M (E^{i+1} - E^i) \sum_{j=i+1}^M e^{-\Gamma_{T_j}} \phi_j \hat{s}_i \right\} + E^1 \left\{ \sum_{i=1}^M \left(e^{-\Gamma_{T_{i+1}}} \mu_{i+1} - e^{-\Gamma_{T_i}} \mu_i \right) \hat{s}_i \right\} \\
& = E^1 \left\{ \sum_{i=1}^M e^{-\Gamma_{T_i}} \left(E^i \frac{\partial A_i}{\partial \hat{s}_i} \hat{s}_i + E^i \frac{\partial A_i}{\partial I_i} I_i \right) \right\} + E^1 \left\{ \sum_{i=1}^M (E^{i+1} - E^i) \sum_{j=i+1}^M e^{-\Gamma_{T_j}} \phi_j \hat{s}_i \right\} \\
& + E^1 \left\{ \sum_{i=1}^M \left(e^{-\Gamma_{T_{i+1}}} \mu_{i+1} - e^{-\Gamma_{T_i}} \mu_i \right) \hat{s}_i \right\}. \tag{A11}
\end{aligned}$$

By the law of iterated expectations, the second term in (A11) is zero. The third term can be written $E^1 \left\{ \sum_{i=2}^{M+1} e^{-\Gamma_{T_i}} \mu_i \hat{s}_{i-1} - \sum_{i=1}^M e^{-\Gamma_{T_i}} \mu_i (\hat{s}_{i-1} + I_i) \right\} = -e^{-\Gamma_{T_1}} \mu_1 \hat{s}_0 = 0$, where the first equality follows from the complementary slackness $\mu_i I_i = 0, \forall i$, and the second equality follows from $\hat{s}_0 = 0$. Equation (A10) therefore simplifies to

$$E^1 \left\{ \sum_{i=1}^M \int_{t=T_i}^{T_{i+1}} e^{-\Gamma_i} \int_w \frac{\partial C_{tw}}{\partial N_{tw}} N_{tw}^2 \mathcal{E}_{tw}^i f_t^i(w) dw dt \right\} = E^1 \left\{ \sum_{i=1}^M e^{-\Gamma_{T_i}} \left(E^i \frac{\partial A_i}{\partial \hat{s}_i} \hat{s}_i + E^i \frac{\partial A_i}{\partial I_i} I_i \right) \right\}, \tag{A12}$$

which is recorded as eqn. (8) in the text.

10.2 Optimal investment rules for $M < \infty$

If $M < \infty$, eqn. (A8) cannot be derived from eqn. (A1) for the last investment at T_M . However, eqn. (A1) can be solved directly to give

$$\phi_M = E^M \frac{\partial A_M}{\partial I_M} - \mu_M. \tag{A13}$$

Substituting (A13) into (A6), gives

$$\int_{t=T_M}^{T_{M+1}} e^{-\Gamma_t} \int_w \frac{\partial C_{tw}}{\partial N_{tw}} N_{tw}^2 \varepsilon_{tw}^M f_t^M(w) dw dt = e^{-\Gamma_{T_M}} \left(E^M \frac{\partial A_M}{\partial \hat{s}_M} + E^M \frac{\partial A_M}{\partial I_M} - \mu_M \right) \hat{s}_M. \quad (\text{A14})$$

Equation (A9) still applies for investments $i=1 \dots M-1$. Summing (A9) for $i=1 \dots M-1$, adding (A14), and taking expectations at T_1 gives a counterpart to (A10):

$$E^1 \left\{ \sum_{i=1}^M \int_{t=T_i}^{T_{i+1}} e^{-\Gamma_t} \int_w \frac{\partial C_{tw}}{\partial N_{tw}} N_{tw}^2 \varepsilon_{tw}^i f_t^i(w) dw dt \right\} \\ = E^1 \left\{ \sum_{i=1}^{M-1} \left[e^{-\Gamma_{T_i}} \left(E^i \frac{\partial A_i}{\partial \hat{s}_i} + E^i \frac{\partial A_i}{\partial I_i} \right) \hat{s}_i - e^{-\Gamma_{T_{i+1}}} E^{i+1} \frac{\partial A_{i+1}}{\partial I_{i+1}} \hat{s}_i + \left((E^{i+1} - E^i) \sum_{j=i+1}^M e^{-\Gamma_{T_j}} \phi_j \right) \hat{s}_i \right] \right. \\ \left. + \left(e^{-\Gamma_{T_{M+1}}} \mu_{M+1} - e^{-\Gamma_{T_1}} \mu_1 \right) \hat{s}_1 + e^{-\Gamma_{T_M}} \left(E^M \frac{\partial A_M}{\partial \hat{s}_M} + E^M \frac{\partial A_M}{\partial I_M} - \mu_M \right) \hat{s}_M \right\}.$$

Following the same steps as for $M = \infty$ yields eqn. (A12) again.

10.3 Proof of Theorem 2

If $A_i(\cdot)$ is independent of \hat{s}_i , $\partial A_i / \partial \hat{s}_i = 0$ for all i . The marginal cost of investment for cycle i is $\partial A_i / \partial I_i = k_i$ for some constant k_i . If investment is reversible, all multipliers μ_i are zero. With these simplifications, eqn. (A9) reduces to

$$\int_{t=T_i}^{T_{i+1}} e^{-\Gamma_t} \int_w \frac{\partial C_{tw}}{\partial N_{tw}} N_{tw}^2 \varepsilon_{tw}^i f_t^i(w) dw dt = \left(e^{-\Gamma_{T_i}} k_i - e^{-\Gamma_{T_{i+1}}} k_{i+1} \right) \hat{s}_i. \quad (\text{A15})$$

The right-hand side of (A15) is the PDV cost of capital over investment cycle i . With $\varepsilon_{tw}^i = 1$, $\forall t, w, i$, the left-hand side of (A15) is expected PDV toll revenue for investment cycle i . Toll revenues and investment costs are therefore equal in expected present discounted value.

If $M < \infty$, eqn. (A14) for the last investment simplifies, with $\varepsilon_{tw}^M = 1$, to

$$\int_{t=T_M}^{T_{M+1}} e^{-\Gamma_t} \int_w \frac{\partial C_{tw}}{\partial N_{tw}} N_{tw}^2 f_t^M(w) dw dt = e^{-\Gamma_{T_M}} k_M \hat{s}_M. \quad (\text{A16})$$

The left-hand side of eqn. (A16) is expected PDV toll revenue for investment cycle M , and the right-hand side is the PDV cost of the last investment.

10.4 Analytics of the cost recovery ratio in the static model

Toll revenue is $R = \tau N = C_N N^2$, hence

$$\frac{dR}{ds} = C_{Ns}N^2 + (2C_N + C_{NN}N)N \frac{dN}{ds}. \quad (\text{A17})$$

From the formula for generalized cost, $p = C + \tau = C + C_N N$, one obtains

$$\frac{dN}{ds} = \frac{C_s + C_{Ns}N}{p_N - (2C_N + C_{NN}N)}. \quad (\text{A18})$$

Substituting (A18) into (A17) gives

$$\begin{aligned} \frac{dR}{ds} &= \frac{p_N C_{Ns} N^2 + (2C_N + C_{NN}N) C_s N}{p_N - (2C_N + C_{NN}N)}, \text{ and} \\ \varepsilon_s^R &= \frac{dR}{ds} \frac{s}{R} = \frac{p_N C_{Ns} N + (2C_N + C_{NN}N) C_s}{p_N - (2C_N + C_{NN}N)} \frac{s}{C_N N}. \end{aligned} \quad (\text{A19})$$

From homogeneity degree zero of the user cost function, $C_N N = -C_s s$, and (A19) can be written

$$\varepsilon_s^R = \frac{p_N C_{Ns} N + (2C_N + C_{NN}N) C_s}{-p_N C_s + (2C_N + C_{NN}N) C_s}. \quad (\text{A20})$$

With perfectly inelastic demand, eqn. (A20) reduces to

$$\varepsilon_s^R = -\frac{C_{Ns}N}{C_s} = \frac{C_{Ns}s}{C_N} = \frac{\partial C_N}{\partial s} \frac{s}{C_N},$$

where the second equality follows by using the relation $C_N N = -C_s s$ again.

10.5 Example with single investment cycle

Capacity is deterministic in the example so that $s_w = \hat{s}_1$. To simplify notation, \hat{s}_1 is written \hat{s} .

Since demand intensity, n_t , is the only random variable, the subscript w will be suppressed; N_t , τ_t and p_t are understood to be random variables that depend on n_t .

Optimal capacity

Optimal capacity can be solved using eqn. (A14) by setting $M = 1$, $T_1 = 0$, $T_2 = T$, $\varepsilon_w^1 = 1$,

$A_1(I_1, \hat{s}_1) = kI_1 = k\hat{s}$, $\mu_1 = 0$, $C_t(N_t, s_t) = d_t(N_t/\hat{s})^\chi$, and $N_t = n_t p_t^{-\eta}$. Equation (A14) becomes

$$\int_{t=0}^T e^{-\rho t} \int_{n_t} d_t \chi N_t^{1+\chi} \hat{s}^{-\chi} f_t(n_t) dn_t dt = k\hat{s}. \quad (\text{A21})$$

Using $d_t = d_0 e^{-\xi \chi t}$, (A21) can be rearranged as a formula for optimal capacity:

$$\hat{s} = \left(\frac{d_0 \chi}{k} \int_{t=0}^T e^{-(r+\xi\chi)t} \left(\int_{n_t} N_t^{1+\chi} f_t(n_t) dn_t \right) dt \right)^{\frac{1}{1+\chi}}. \quad (\text{A22})$$

The optimal toll is

$$\tau_t = \frac{\partial C_t}{\partial N_t} N_t = d_t \chi \left(\frac{N_t}{\hat{s}} \right)^\chi,$$

and the generalized cost of a trip is

$$p_t = C_t(N_t, s_t) + \tau_t = (1 + \chi) d_t \chi \left(\frac{N_t}{\hat{s}} \right)^\chi.$$

Hence

$$N_t = n_t p_t^{-\eta} = n_t (1 + \chi)^{-\eta} \chi d_t^{-\eta} \left(\frac{N_t}{\hat{s}} \right)^{-\chi\eta},$$

which yields

$$N_t = n_t^{\frac{1}{1+\chi\eta}} \left((1 + \chi) d_0 \hat{s}^{-\chi} e^{-\xi\chi t} \right)^{\frac{-\eta}{1+\chi\eta}}. \quad (\text{A23})$$

The inner integral in eqn. (A22) is therefore:

$$\begin{aligned} \int_{n_t} N_t^{1+\chi} f_t(n_t) dn_t &= \int_{n_t} n_t^{\frac{1+\chi}{1+\chi\eta}} \left((1 + \chi) d_0 \hat{s}^{-\chi} \right)^{\frac{-\eta(1+\chi)}{1+\chi\eta}} e^{\frac{\xi\chi\eta(1+\chi)}{1+\chi\eta} t} f_t(n_t) dn_t, \\ &= \left((1 + \chi) d_0 \hat{s}^{-\chi} \right)^{\frac{-\eta(1+\chi)}{1+\chi\eta}} e^{\frac{\xi\chi\eta(1+\chi)}{1+\chi\eta} t} \int_{n_t} n_t^{\frac{1+\chi}{1+\chi\eta}} f_t(n_t) dn_t. \end{aligned} \quad (\text{A24})$$

Demand intensity evolves according to Geometric Brownian Motion and has the pdf

$$f_t(n) = \frac{1}{\sqrt{2\pi n\sigma}\sqrt{t}} \exp \left\{ - \frac{\left[\ln(n/n_0) + \left(\frac{\sigma^2}{2} - g \right) t \right]^2}{2\sigma^2 t} \right\}. \quad (\text{A25})$$

Using the composite variable $\lambda \equiv (1 + \chi)/(1 + \chi\eta)$, the integral in (A24) works out to

$$\int_{n_t} n_t^\lambda f_t(n_t) dn_t = n_0^\lambda \exp \left\{ \lambda \left(g + (\lambda - 1) \frac{\sigma^2}{2} \right) t \right\}. \quad (\text{A26})$$

Substitution of (A26) into (A24), and (A24) into (A22) yields eqn. (11) in the text:

$$\hat{s} = n_0 \left(\frac{\chi}{k} \right)^{\frac{1}{\lambda}} (1 + \chi)^{-\eta} d_0^{\frac{1+\chi}{1+\chi\eta}} \left((e^{hT} - 1)/h \right)^{\frac{1}{\lambda}}, \quad (\text{A27})$$

where $h \equiv \lambda(\xi\chi\eta + g + (\lambda - 1)\sigma^2 / 2) - (r + \xi\chi)$.

Expected toll revenues

Expected toll revenues are given by the left-hand side of eqn. (A21). Equation (13) in the text is obtained by using eqns. (A24), (A26), and (A27):

$$E\{R\} = n_0^\lambda \chi (1 + \chi)^{-\eta\lambda} d_0^{\frac{1-\eta}{1+\chi\eta}} (\hat{s})^{-\chi(1-\eta\lambda)} (e^{hT} - 1) / h. \quad (\text{A28})$$

Expected cumulative usage

Usage at time t is given by eqn. (A23). Expected usage at time t is

$$\begin{aligned} E\{N_t\} &= \left((1 + \chi) d_0 \hat{s}^{-\chi} e^{-\xi\chi t} \right)^{\frac{-\eta}{1+\chi\eta}} \int_{n_t} \frac{1}{n_t^{1+\chi\eta}} f_t(n_t) dn_t \\ &= n_0^{\frac{1}{1+\chi\eta}} \left((1 + \chi) d_0 \hat{s}^{-\chi} e^{-\xi\chi t} \right)^{\frac{-\eta}{1+\chi\eta}} e^{\frac{1}{1+\chi\eta} \left(g - \frac{\chi\eta}{2} \frac{\sigma^2}{1+\chi\eta} \right) t}. \end{aligned}$$

Expected cumulative usage is eqn. (12) in the text:

$$E\{U\} = \left\{ \int_{t=0}^T E\{N_t\} dt \right\} = n_0^{\frac{1}{1+\chi\eta}} \left((1 + \chi) d_0 \hat{s}^{-\chi} \right)^{\frac{-\eta}{1+\chi\eta}} (e^{mT} - 1) / m,$$

where $m \equiv (g + \xi\chi\eta) / (1 + \chi\eta) - \chi\eta\sigma^2 / (2(1 + \chi\eta)^2)$.

10.6 Two-period investment with irreversibility

Optimal investment at T_2 depends on the demand intensity realized at T_2 , n_{T_2} , or n_2 for brevity.

If n_2 exceeds a threshold level, $I_2 > 0$ and $\mu_2 = 0$. If n_2 is below the threshold, $I_2 = 0$ and

$\mu_2 > 0$. The threshold is an increasing function of \hat{s}_1 , and accordingly will be written $\bar{n}_2(\hat{s}_1)$. If

$I_2 > 0$, the optimal \hat{s}_2 is given by eqn. (A27) with n_2 in place of n_0 , $d_0 e^{-\xi\chi T_2}$ in place of d_0 , and

$T = T_3 - T_2$:

$$\hat{s}_2 = n_2 \left(\frac{\chi}{k(1 - e^{-r(T_3 - T_2)})} \right)^{\frac{1}{\lambda}} (1 + \chi)^{-\eta} d_0^{\frac{1-\eta}{1+\chi}} e^{-\xi\chi \frac{1-\eta}{1+\chi} T_2} \left(\frac{e^{h(T_3 - T_2)} - 1}{h} \right)^{\frac{1}{\lambda}}. \quad (\text{A29})$$

(The threshold value $\bar{n}_2(\hat{s}_1)$ is solved by replacing \hat{s}_2 in (A29) with \hat{s}_1 , and solving for n_2 .)

Optimal investment at T_1 for the general model is given by eqn. (A9) with $i=1$. Using the specific

investment cost functions $A_1(I_1, \hat{s}_1) = kI_1$ and $A_2(I_2, \hat{s}_2) = kI_2$, and retaining the general notation for the moment, eqn. (A9) simplifies to

$$\int_{t=T_1}^{T_2} e^{-\Gamma_t} \int_w \frac{\partial C_{nw}}{\partial N_{nw}} N_{nw}^2 \mathcal{E}_{nw}^1 f_t^1(w) dw dt = e^{-\Gamma_{T_1}} \left(k - e^{-\Gamma_{T_2} + \Gamma_{T_1}} k + e^{\Gamma_{T_1}} (E^2 - E^1) e^{-\Gamma_{T_2}} \phi_2 \right) \hat{s}_1. \quad (\text{A30})$$

Since $I_1 > 0$, $\mu_1 = 0$. From eqn. (A13) with $M = 2$, $\phi_2 = k - \mu_2$. Using these formulas, (A30) simplifies to

$$\int_{t=T_1}^{T_2} e^{-\Gamma_t} \int_w \frac{\partial C_{nw}}{\partial N_{nw}} N_{nw}^2 \mathcal{E}_{nw}^1 f_t^1(w) dw dt = \left(e^{-\Gamma_{T_1}} k - e^{-\Gamma_{T_2}} E^1 (k - \mu_2) \right) \hat{s}_1. \quad (\text{A31})$$

Substituting $M=2$ into eqn. (A14) yields

$$\int_{t=T_2}^{T_3} e^{-\Gamma_t} \int_w \frac{\partial C_{nw}}{\partial N_{nw}} N_{nw}^2 \mathcal{E}_{nw}^2 f_t^2(w) dw dt = e^{-\Gamma_{T_2}} (k - \mu_2) \hat{s}_2, \text{ or}$$

$$e^{-\Gamma_{T_2}} (k - \mu_2) = \hat{s}_2^{-1} \int_{t=T_2}^{T_3} e^{-\Gamma_t} \int_w \frac{\partial C_{nw}}{\partial N_{nw}} N_{nw}^2 \mathcal{E}_{nw}^2 f_t^2(w) dw dt. \quad (\text{A32})$$

Substituting (A32) into (A31) gives

$$\hat{s}_1^{-1} \int_{t=T_1}^{T_2} e^{-\Gamma_t} \int_w \frac{\partial C_{nw}}{\partial N_{nw}} N_{nw}^2 \mathcal{E}_{nw}^1 f_t^1(w) dw dt + E^1 \left\{ \hat{s}_2^{-1} \int_{t=T_2}^{T_3} e^{-\Gamma_t} \int_w \frac{\partial C_{nw}}{\partial N_{nw}} N_{nw}^2 \mathcal{E}_{nw}^2 f_t^2(w) dw dt \right\} = e^{-\Gamma_{T_1}} k. \quad (\text{A33})$$

The second term in brackets in (A33) is random from the perspective of T_1 when \hat{s}_1 is chosen. If $n_2 \leq \bar{n}_2(\hat{s}_1)$, $\hat{s}_2 = \hat{s}_1$, and if $n_2 > \bar{n}_2(\hat{s}_1)$, \hat{s}_2 is given by eqn. (A29). The probability density of n_2 is given by (A25) with $t = T_2$ and n replaced by n_2 :

$$f_{T_2}(n_2) = \frac{1}{\sqrt{2\pi n_2 \sigma \sqrt{T_2}}} \exp \left\{ - \frac{\left[\ln(n_2/n_0) + \left(\frac{\sigma^2}{2} - g \right) T_2 \right]^2}{2\sigma^2 T_2} \right\}.$$

Let $f_t(n|n_2)$ denote the pdf of n for $t > T_2$ given $n = n_2$ at time T_2 . Substituting in the functional forms of the specific model, eqn. (A33) can be written

$$\begin{aligned}
& \hat{s}_1^{-1} \int_{t=0}^{T_2} e^{-rt} \left(\int_n d(t) \chi \hat{s}_1^{-\lambda} N_t^{1+\lambda} f_t(n) dn \right) dt \\
& + \hat{s}_1^{-1} \int_{n_2=0}^{\bar{n}_2(\hat{s}_1)} f_{T_2}(n_2) \left\{ \int_{t=T_2}^{T_3} e^{-rt} \left(\int_n d(t) \chi \hat{s}_1^{-\lambda} N_t^{1+\lambda} f_t(n | n_2) dn \right) dt \right\} dn_2 \\
& + \int_{n_2=\bar{n}_2(\hat{s}_1)}^{\infty} f_{T_2}(n_2) (\hat{s}_2(n_2))^{-1} \left\{ \int_{t=T_2}^{T_3} e^{-rt} \left(\int_n d(t) \chi (\hat{s}_2(n_2))^{-\lambda} N_t^{1+\lambda} f_t(n | n_2) dn \right) dt \right\} dn_2 \\
& = k(1 - e^{-rT_3}).
\end{aligned} \tag{A34}$$

Making the remaining substitutions of the specific model, (A34) becomes

$$\begin{aligned}
& \chi(1 + \chi)^{-\eta\lambda} d_0^{1-\eta\lambda} \hat{s}_1^{-\lambda} \left(n_0^\lambda \frac{e^{hT_2} - 1}{h} + e^{-\lambda \left(s + (\lambda-1) \frac{\sigma^2}{2} \right) T_2} \frac{e^{hT_3} - e^{hT_2}}{h} \int_{n_2=0}^{\bar{n}_2(\hat{s}_1)} f_2(n_2) n_2^\lambda dn_2 \right) \\
& + k(e^{-rT_2} - e^{-rT_3}) (1 - F_{T_2}(\bar{n}_2(\hat{s}_1))) = k(1 - e^{-rT_3}),
\end{aligned} \tag{A35}$$

where $F_{T_2}(n_2)$ is the cumulative distribution corresponding to $f_{T_2}(n_2)$. \hat{s}_1 appears three times in eqn. (A35) and has to be solved numerically.

11 REFERENCES

- Arnott, R., Kraus, M., 1998. Self-financing of congestible facilities in a growing economy. In: Pines, D., Sadka, E., Zilcha, I. (Eds.), *Topics in Public Economics: Theoretical and Applied Analysis*, Cambridge: Cambridge University Press, pp. 161-184.
- Bain, R., 2009. Error and optimism bias in toll road traffic forecasts. *Transportation* 36(5), 469-482.
- Beckmann, M., McGuire, C.B., Winsten, C.B., 1956. *Studies in the Economics of Transportation*. New Haven, CT: Yale University Press.
- Berechman, J., 2009. *The Evaluation of Transportation Investment Projects*. Routledge, New York.
- Berechman, J., Chen, L., 2011. Incorporating risk of cost overruns into transportation capital projects decision-making. *Journal of Transport Economics and Policy* 45(1), 83-103.
- de Palma, A., Lindsey, R., 2007. Transport user charges and cost recovery. In: de Palma, A., Lindsey, R., Proost, S. (Eds.), *Investment and the Use of Tax and Toll Revenues in the Transport Sector*, *Research in Transportation Economics*, Vol. 19, Elsevier, Amsterdam, pp. 29-58.
- de Palma, A., Lindsey, R., 2011. Cost recovery from road tolls with long-run uncertainty. *Proceedings of the 46th Annual Conference of the Canadian Transportation Research Forum: Transportation and Logistics Trends and Policies: Successes and Failures*, Gatineau, May 29 – June 1, 513-527.
- De Vany, A., Saving, T.R., 1980. Competition and highway pricing for stochastic traffic. *Journal of Business* 53(1), 45-60.
- Dixit, A., Pindyck, R.S., 1994. *Investment Under Uncertainty*. Princeton University Press, New Jersey.
- Doan, P., Patel, K., 2010. Toll road investment under uncertainty. March (<http://www.chaire-eppp.org/files_chaire/Toll_road_investment_under_uncertainty_-_Phuong_Doan_Paris_Sorbonne_2010.pdf> [February 2, 2012])
- Elefteriadou, L., 2004. Highway capacity. In: Kutz, M. (Ed.), *Handbook of Transportation Engineering*. McGraw-Hill, New York (Chapter 8).
- Flyvbjerg, B., Skamris Holm, M.K., Buhl, S.L., 2003. How common and how large are cost overruns in transport infrastructure projects? *Transport Reviews* 23(1), 71–88.
- Flyvbjerg, B., Skamris Holm, M.K., Buhl S.L., 2006. Inaccuracy in traffic forecasts. *Transport Reviews* 26(1), 1-24.

Hensher, D.A., Goodwin, P., 2004. Using values of travel time savings for toll roads: avoiding some common errors. *Transport Policy* 11(2), 171-181.

Jahren, C.T., Asha, A.M., 1990. Predictors of cost-overrun rates. *Journal of Construction Engineering and Management* 116(3), 548-552.

Kraus, M., 1982. Highway pricing and capacity choice under uncertain demand. *Journal of Urban Economics* 12, 122-128.

Lindsey, R., 2006. Do economists reach a conclusion on highway pricing?: The intellectual history of an idea. *Econ Journal Watch* (<http://www.econjournalwatch.org>) 3(2), 292-379.

Lindsey, R., 2009. Cost recovery from congestion tolls with random capacity and demand. *Journal of Urban Economics* 66, 16–24.

Lindsey, R., 2012. Road pricing and investment. *Economics of Transportation* 1(1-2), 49–63.

Marathe, R., Ryan, S., 2005. On the validity of the Geometric Brownian Motion assumption, *The Engineering Economist* 50(2), 159-192.

Mohring, H., Harwitz, M., 1962. *Highway Benefits: An Analytical Framework*. Northwestern University Press, Evanston, Illinois.

Nagae, T., Akamatsu, T., 2006. Dynamic revenue management of a toll road project under transportation demand uncertainty. *Networks and Spatial Economics* 6, 345–357.

Nelson, J. C., 1962. The pricing of highway, waterway, and airway facilities. *American Economic Review (Papers and Proceedings)* 52(2), 426-432.

Ng, C-F., Small, K.A., 2012. Tradeoffs among free-flow speed, capacity, cost, and environmental footprint in highway design. *Transportation*. DOI: 10.1007/s11116-012-9395-8.

Nijkamp, P., Ubbels, B., 1999. How reliable are estimates of Infrastructure Cost? A comparative analysis. *International Journal of Transport Economics* 26(1), 23-53.

Odeck, J., 2004. Cost overruns in road construction — What are their sizes and determinants. *Transport Policy* 11, 43-53.

Pimentel, P.M., Azevedo Pereira, J., Couto, G., 2012. High-speed rail transport valuation. *The European Journal of Finance* 18(2), 167-183.

Prozzi, J., Flanagan, K., Loftus-Otway, L., Porterfield, B., Persad, K., Prozzi, J.A., Walton, C.M., 2009. Actual vs. forecasted toll usage: a case study review. July 2008; revised August 2009, Center for Transportation Research, The University of Texas at Austin (<http://www.utexas.edu/research/ctr/pdf_reports/0_6044_1.pdf > [February 11, 2011]).

- Rose, S., 1998. Valuation of interacting real options in a tollroad infrastructure project. *The Quarterly Review of Economics and Finance* 38(4), 711-723.
- Saphores, J.-D., Boarnet, M.G., 2006. Uncertainty and the timing of an urban congestion relief investment - The no-land case. *Journal of Urban Economics* 59(2), 189-208.
- Small, K.A., Verhoef, E.T., 2007. *The Economics of Urban Transportation*. Routledge, London.
- Transportation Research Board, 2006. National Cooperative Highway Research Program (NCHRP) Synthesis 364: Estimating Toll Road Demand and Revenue (<http://www.trb.org/news/blurb_detail.asp?id=7570> [January 27, 2013]).
- Verhoef, E.T., 2012. Cost recovery of congested infrastructure under market power. Tinbergen Institute Discussion Paper TI 2012-064/3 (<<http://ssrn.com/abstract=2101680>> [July 19, 2012])
- Welde, M., Odeck, J., 2011. Do planners get it right? The accuracy of travel demand forecasting in Norway. *EJTIR* 11(1), 80-95.
- Williams-Derry, C., 2011. Toll avoidance and transportation funding: official estimates frequently overestimate traffic and revenue for toll roads (<<http://www.sightline.org/research/sprawl/toll-avoidance-and-transportation-funding>> [January 31, 2013])
- Zhang, A., and Czerny, A.I., 2012. Airports and airlines economics and policy: an interpretive review of recent research. *Economics of Transportation* 1(1-2), 15–34.

Table 1 : Sensitivity to parameter errors

Misspecified parameter	Design Capacity [%]	Expected cumulative usage [%]	Expected cost recovery rate [%]
(1) Unit construction cost			
$\tilde{k} = (5/6)k$	109.5	103.1	83.3
$\tilde{k} = (6/5)k$	91.3	97.0	120.0
(2) Initial demand			
$\tilde{n}_0 = (2/3)n_0$	66.7	87.4	225
$\tilde{n}_0 = (3/2)n_0$	150.0	114.5	44.4
(3) Mean growth rate in demand (base case $g = 0.02$)			
$\tilde{g} = 0$	87.3	95.6	131.3
$\tilde{g} = 0.01$	93.3	97.7	115.0
$\tilde{g} = 0.03$	107.6	102.5	86.3
$\tilde{g} = 0.04$	116.3	105.2	73.9
(4) Standard deviation of demand growth (base case $\tilde{\sigma} = 0.05$)			
$\tilde{\sigma} = 0$	99.1	99.7	101.8
$\tilde{\sigma} = 0.1$	102.7	100.9	94.7
(5) Demand elasticity (base case $\tilde{\eta} = 0.25$)			
$\tilde{\eta} = 0.125$	101.7	100.6	96.6
$\tilde{\eta} = 0.5$	96.9	99.0	106.4
(6) Rate of technological progress (base case $\tilde{\xi} = 0.01$)			
$\tilde{\xi} = 0$	103.7	101.2	93.0
$\tilde{\xi} = 0.02$	96.5	98.8	107.3

Source: Authors' calculation

Table 2: Variability in daily traffic, toll and toll revenue in year 10

Variable	Mean	Standard deviation	Coefficient of variation
Demand intensity	24,255	3,859	0.159
Traffic volume	14,129	1,489	0.105
Toll level	\$5.75	\$1.23	0.213
Toll revenue	\$83,005	\$26,841	0.323

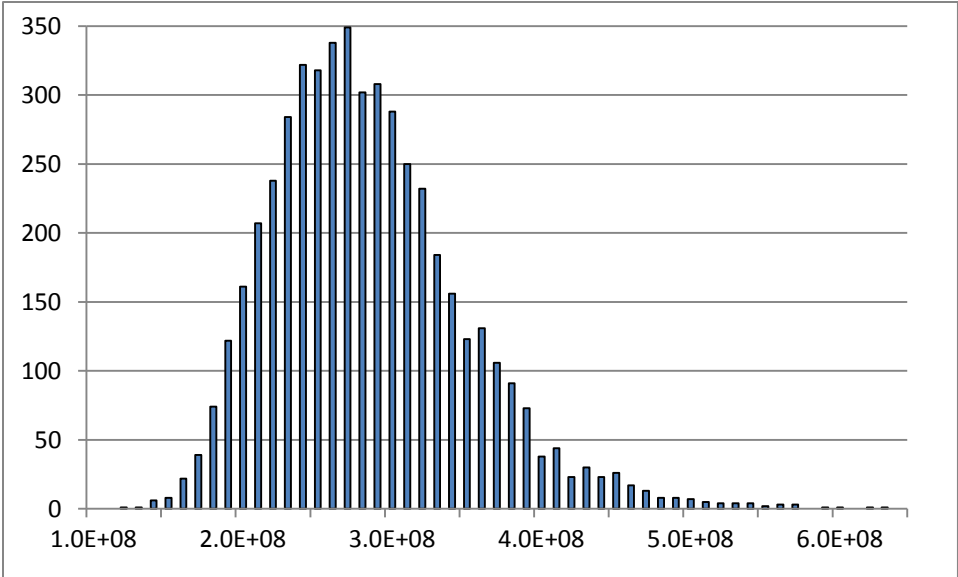
Source: Authors' calculation

Table 3 : Two-period investment example with irreversible investment in period 1

Parameter Variation	Investment in period 1 relative to reversible case [%]	Prob. of no period 2 investment	Expected cost recovery rate in period 1 [%]
(1) Base case	99.46	0.137	101.09
(2) Mean growth rate in demand (base case $g = 0.02$)			
$g = 0.01$	98.27	0.352	103.56
$g = 0.03$	99.90	0.033	100.21
(3) Standard deviation of demand growth (base case $\tilde{\sigma} = 0.05$)			
$\sigma = 0.1$	97.40	0.325	105.40
(4) Demand elasticity (base case $\tilde{\eta} = 0.25$)			
$\eta = 0.125$	99.41	0.151	101.42
$\eta = 0.5$	99.56	0.112	100.66
(5) Rate of technological progress (base case $\xi = 0.01$)			
$\xi = 0$	99.75	0.071	100.50
$\xi = 0.02$	98.98	0.232	102.08
(6) Curvature of user cost function (base case $\chi = 2$)			
$\chi = 1$	99.54	0.118	100.75
$\chi = 4$	99.40	0.154	101.50
(7) Discount rate (base case $r = 0.05$)			
$r = 0.03$	99.29	0.135	101.44
$r = 0.07$	99.60	0.139	100.81

Source: Authors' calculation

Figure 1: Probability density for cumulative present value revenues (base-case parameter values, 5,000 simulations)



Source: Author's construction