Congestion Pricing and Capacity Investment in a Gateway-Hinterland Intermodal System

Anming Zhang
Center for Transportation Studies, SAUDER School of Business, The University of British Columbia, Canada

ABSTRACT

This paper examines the impact of road congestion pricing and capacity investment in a gateway-hinterland intermodal system. We find, for a given highway capacity, that when the hinterland region moves from a no-congestion pricing regime to congestion pricing on its road, the traffic going through the gateway would fall if freight trucks have a much higher value of time than non-commercial light vehicles. On the other hand, if road tolls are fixed, then an increase in road capacity stimulates gateway traffic. Our analysis suggests that there is a need for a central (cross-region) body to coordinate decisions concerning capacity investment and congestion pricing for the gateway and hinterland transport facilities. The case of endogenous road tolls, in which road capacity is determined prior to road pricing, is also analyzed.

ACKNOWLEDGEMENT

I would like to thank Yossi Berechman, Andrew Yuen and participants of the International Conference on Gateways and Corridors, Vancouver, B.C., for helpful comments. Partial financial support from the Social Science and Humanities Research Council of Canada (SSHRC) and from Li & Fung Institute for Supply Chain Management and Logistics at the Chinese University of Hong Kong is gratefully acknowledged.

1. INTRODUCTION

The gateway-hinterland intermodal transportation system is commonly observed. In such a system, a gateway, which is usually featured by a major seaport or airport, serves both its own region and its hinterland: Traffic destined for, or originated from, an interior region – the hinterland – passes through the gateway. In the context of seaports, the gateway may be defined as a coastal metropolis with port access to both its hinterland and the rest of the world, which captures a substantial share of total regional and international trade volumes (Berechman, 2007). Consider the cargo flow to the hinterland: Goods from the rest of the world are first shipped to the gateway, and then transported to the hinterland region by trucking, rail or inland waterway. In the case of trucking, transportation will utilize
roads and highways of the hinterland region. For example, international containers are shipped to the Port of New York/New Jersey - the gateway - and then are distributed to its U.S. hinterland by truck (73% of the total container throughput in 2005), waterway (36%) or rail (1%) (Berechman, 2007). Similarly, other major gateways, such as Vancouver, Los Angeles, Long Beach, Rotterdam, Hamburg, Antwerp, Hong Kong, Singapore and Shanghai serve their respective hinterlands by intermodal transportation systems. As argued by van Klink and van den Berg (1998), the gateways are in a unique position to, on the one hand, stimulate intermodal transport and, on the other hand, use the intermodal systems to enlarge their hinterlands. Several studies further argue that hinterland access is also important for the competitiveness of a gateway when it competes with other gateways (e.g., Notteboom, 1997; Kreukels and Wever, 1998; Zhang, 2008).

This paper considers such an intermodal system with a port being situated in the gateway region and a highway (road) in the hinterland region. An obvious characteristic of this intermodal system is that both the gateway’s port and the hinterland’s highway may be prone to congestion. First, it has been widely recognized that congestion is acute at many ports around the world. At the U.S. west-coast ports and the Port of Vancouver, Canada, congestion has become a major problem, owing largely to the strong growth in imports from Asia. The ensuing delays have imposed substantial costs on carriers and shippers. For example, the twelve shipping lines, which accounted for about 70% of trans-Pacific trade, estimated that costs in 2005 rose by 10% because of port congestion (Bloomberg News, December 4, 2005). In Australia, in the first six months of 2005, there was a ship queue of around 50 vessels anchored at any one time for an average of 21 days at Queensland’s Dalrymple Bay Coal Terminal, which cost the coal industry around $2 million per day (Everett, 2006). Moreover, shippers are required to increase inventory due to uncertain delays (Maloni and Jackson, 2005).

As a result of increasing port congestion, governments and port operators have been looking for solutions, which include capacity expansion and congestion pricing. For example, Canada’s Asia-Pacific Gateway and Corridor Initiative will provide funding primarily for expanding capacity at the gateway ports and related facilities in British Columbia. At the Ports of Los Angeles and Long Beach, a ‘traffic mitigation fee’ ($50 per TEU, or $100 for all containers larger than a TEU) is imposed on containers exiting port terminals during peak hours by PierPASS, an organization established by marine terminal operators to reduce congestion. Antwerp, Rotterdam and Hamburg have expanded, or will expand, their port capacities, in response to the port congestion problem due to the substantial increase of imports from China and the new Central Europe members of the EU (Quinn, 2002).

In comparison to the port congestion problem, congestion at connected facilities in the hinterland, such as road, highway, rail and waterway, is less discussed. In a gateway-hinterland intermodal transportation system (sometimes also referred to as a ‘transport supply chain’) however, overall congestion is dictated by the weakest link (or node). Shippers and carriers incur delay costs due

---

1 While the port and highway are considered here, the analysis applies, in principle, to other node-link combinations. A node may be a port, airport, intermodal transfer point, or border custom, whereas a link may be a highway, railway, inland waterway, or bridge.
2 For more information, see www.apgci.gc.ca and www.th.gov.bc.ca/gateway/.
not only to congestion at port facilities, but also to congestion at other parts of the intermodal system. A survey conducted by Maloni and Jackson (2005) suggests that U.S. port managers’ greatest concern in port capacity expansion planning is the capacity constraint imposed by local roads. For instance, the Ports of Los Angeles and Long Beach had to divert a large number of ships to other ports because of truck and rail congestion (Journal of Commerce, August 8, 2005). Similar to the port congestion problem, congestion at roads and highways may be alleviated by capacity expansion and pricing mechanism. Whilst capacity investments have been constantly undertaken for roads and highways around the world, road congestion pricing is still not very common in practice, although it becomes increasingly discussed (debated) in both the academic and policy circles.

These developments naturally raise the question about the nature of interactions between the gateway and hinterland, and their coordination, in congestion pricing and capacity investment. Nevertheless, the issue has yet been adequately addressed in the transportation literature. We could only locate two recent studies that investigate the question, both of which look at the impacts of gateway pricing and capacity expansion on the hinterland’s road system. Berechman (2007) examines empirically the ‘full’ social costs associated with the additional traffic in the New Jersey (hinterland) highway system brought by an expansion of the Port of New York/New Jersey (gateway). He finds that a 6% additional container traffic translates into substantial social costs, ranging from US$663 million to US$1.62 billion per year. (The large range of the estimates is due in large part to different choices of value of time.) Of the full social marginal costs, congestion costs incurred by the carriers and local highway users account for 62%. Berechman argues that the heavy social costs incurred at the hinterland’s highways must be recognized and incorporated into the analysis of gateway port investment plans.

The other study, Yuen, Basso and Zhang (2007), examines the effect of congestion pricing implemented at a gateway (seaport or airport) on its hinterland’s optimal highway tolls, highway congestion and social welfare. They find that if the gateway vertically integrates with its oligopoly carriers, its facility charge will rise, with part of the charge increase being due to congestion pricing. This increase in the gateway charge will in turn result in lower highway tolls, independent of whether the highway is able to price discriminates between the local and transit traffic. But, while the change in the highway’s congestion level is in general ambiguous, the hinterland’s welfare will fall as a result of the increase in the gateway charge. Their analysis thus suggests that the gateway congestion pricing might lead to a gateway facility charge that would be too high for the optimization of the overall gateway-hinterland intermodal system, thereby bringing about the important need for coordination between government agencies and facility operators of the two regions.3

This paper is similar to Yuen, et al. (2007) in that it addresses the interaction and coordination issue in facility pricing in a gateway-hinterland intermodal system. Unlike Yuen, et al., however, the present paper investigates the impact of the hinterland’s highway congestion pricing on the gateway. This is

3 More generally, there is an extensive literature on urban road pricing (see Small and Verheof, 2007, for a literature survey, and a recent study by De Borger, Dunkerley and Proost, 2007, on related issues). The gateway-hinterland pricing and capacity linkage has yet been addressed in this literature, however.
important and relevant. While road congestion tolls are still rare in reality, they become increasingly popular in recent years. In Southern California, four variably-priced express lanes (SR 91 Express Lanes) opened in 1995, in which toll schedule is adjusted every three months based on traffic volumes and patterns observed over the three-month period. Road congestion pricing has also been implemented in San Diego, Singapore, etc. As a further example, Vancouver, British Columbia, has a plan to toll two new bridges, namely, the Port Mann Bridge and the Golden Ears Bridge. However, these highway (and bridge) tolls and capacity expansion projects focus mainly on local congestion, rather than on the congestion problem of the overall gateway-hinterland intermodal system. This is due partly to the existing institutional arrangements: for instance, facilities of such an intermodal transportation system are usually under the control of different stakeholders, including local governments, and port and inland transport facility operators. They may even be under the jurisdictions of different countries. For example, cargos at the Port of Rotterdam are delivered to Italy through a rail corridor with links in the Netherlands, Belgium, Germany, Switzerland and Italy.

Furthermore, different from Yuen, et al. (2007) who look at pricing issues only, the present paper considers both pricing and capacity investment. We find, for given highway capacity, that when the hinterland moves from no-congestion pricing to congestion pricing, while both its local traffic and social welfare improves, the traffic going through the gateway would fall when freight trucks have a much higher value of time than non-commercial light vehicles. On the other hand, for fixed highway tolls, an increase in highway capacity stimulates traffic by both the local road users and gateway cargos. Like Yuen, et al. (2007), our analysis therefore suggests that there is a need for a central (cross-region) body to coordinate decisions concerning capacity investment and congestion pricing for the gateway and hinterland transport facilities. We shall further look at the case of endogenous highway tolls, in which highway capacity is determined prior to highway pricing.

The issue of coordination between different modes in a transport logistics chain has been recognized in the policy arena. However, the focus there is mainly on improving the physical connections among facilities, including, in the context of port-inland transportation, the adequacy of landside connections to ports (e.g., Australian Government, 2005). The European Community recently proposed a standardization and harmonization program concerning intermodal loading units: while the standardized characteristics of containers (e.g., TEU) were usually used for sea mode, swap bodies were usually used for land modes. This program is estimated to provide European industry and transporters with efficiency gains, and a reduction of up to 2% in logistics costs (European Commission, 2004). In contrast, the focus of this paper is on the efficiency implication of coordination among the regions in transport facility pricing and capacity investment. Our main point is that as the gateway and hinterland transport facilities are parts of an intermodal transportation system, their pricing and capacity decisions affect each other’s performance. In particular, we investigate the implications of congestion

---

4 As will be seen below, there are several modelling differences as well, including the more general market structure at the carriers’ level in Yuen, et al. (2007), as opposed to the simplification (perfectly competitive market structure) assumed in the present paper. On the other hand, whilst the linear demand and delay functions are used in Yuen, et al., this paper uses more general demand and delay functions.
pricing and capacity investment of the hinterland’s highways for the gateway’s traffic and welfare.\textsuperscript{5}

The paper is organized as follows: Section 2 sets up the model and analyzes the consumer and carrier problem. Section 3 examines the effects of highway congestion pricing and capacity expansion, and Section 4 contains concluding remarks.

2. THE MODEL

We consider likely the simplest model structure in which our question - what would be the effects of congestion pricing and capacity investment of the hinterland’s highways - can be addressed. There are two regions: a gateway region G, and a hinterland region H. A port is situated in G, and an adjacent highway is situated in H.\textsuperscript{6} Transportation firms (carriers) use both the port and highway to bring X units of a cargo good into region H, where the good is consumed. In addition, the highway is used by region H’s local private cars. The number of these private vehicle trips is denoted by Y.

The highway is congestible; consequently, region H may charge a per-trip congestion toll, denoted t, for the use of its highway. We consider a setting in which region H’s decisions on t and its highway capacity, denoted K, are made prior to the consumers’ and carriers’ decisions.

For given t and K, we now specify the decisions of consumers and carriers. Consider first the demand of H’s local highway users. In addition to the toll, these consumers of highway service may suffer congestion at the highway, and so their full cost is the sum of the toll and the congestion delay cost. The congestion level depends on the traffic volume-capacity ratio. Using \( V \) to denote the total traffic volume at the highway, which, as indicated above, includes both local traffic Y and the traffic associated with cargo transportation, then \( V \) may be expressed as:

\[
V = Y + X
\]  

In (1) the cargo units X are normalized into a ‘private car equivalence’ number. Note that in general, a freight truck movement contributes more to road congestion than a private car movement.\textsuperscript{7}

\textsuperscript{5} To maintain their competitiveness and provide better service to their consumers, various forms of collaboration have become more popular among port operators, shipping companies and inland transportation. For example, many ports use rail connections as a strategic tool to penetrate new markets and retain dominance over existing hinterlands (Debrie, 2004). Major European port operators, such as Eurogate and Hamburger Hafen, have been participating in rail services; major stevedores in Australia, namely, P&O Ports and Patrick Co., have been involved in significant restructuring to control landside chains (Debrie and Gouvernal, 2006). Shipping lines have also engaged in collaborating with inland transportation. For example, three major shipping lines at the Port of Rotterdam - Royal Nedlloyd, Sealand Service and P&O Containers - jointly established the European Rail Shuttle (ERS) in 1994 to link Rotterdam with inland markets. Maserk has been involved in trucking operations between Hong Kong and its southern China hinterland. As to be seen below, the strong interest of operators and carriers in the intermodal collaboration will be abstracted away from our analysis since we focus on the policy coordination between regional governments.

\textsuperscript{6} The gateway and hinterland may be connected by a transportation corridor, which is abstracted away from our modeling consideration here.

\textsuperscript{7} For instance, a ‘private car equivalence’ ratio of 3 or 4 for each freight truck movement is used in Berechman (2007).
Given the specification of $V$, the local highway users thus face the following 'full price':

$$p = D_r(V/K) + t$$

(2)

where $D_r(V/K)$ represents congestion delay cost ($D$ for 'delay') to the local users, which depends, in part, on their value of time. Note that this congestion delay function is a special case of the general delay specification $D_r(V, K)$, in that it satisfies

$$V \frac{\partial D_r}{\partial V} + K \frac{\partial D_r}{\partial K} = 0$$

(3)

and so is homogenous of degree of zero. Given the delay specification, the local users' demand is then given by the full price:

$$P_r(Y) = D_r(V/K) + t$$

(4)

where $P_r(Y)$ denotes the inverse demand function, with $P_r(\cdot) < 0$ (i.e., downward-sloping demand).

Similarly, the inverse demand for the cargo good is written as $P_x(X)$ with $P_x'(\cdot) < 0$. Cargo carriers are assumed to be perfectly competitive; consequently, the cargo price that consumers in region H face is, in equilibrium, equal to its 'full' marginal costs:

$$P_x(X) = C + D_x(V/K) + t$$

(5)

where $C$ is the carriers' constant unit cost, which includes their operating cost as well as any purchase price they pay to the product supplier, and $t$ (highway toll) represents another 'input' cost. Further, $D_x(V/K)$ represents congestion delay cost to the cargo carriers. This function is, same as $D_r(\cdot)$, is homogenous of degree of zero.

We impose the following conventional assumptions on the delay functions:

$$D_r'(\cdot) > 0, \quad D_x'(\cdot) > 0$$

(6)

Inequalities (6) will arise, for given values of time, if congestion delay increases in the traffic volume-capacity ratio. Note that for any given traffic volume-capacity ratio, it is usually $D_x'(\cdot) \neq D_r'(\cdot)$. This is because commercial trucks usually have a higher value of travel timesavings than private cars. For instance, US DOT (2003) recommends a value of travel time between $10.6/hour and $21.2/hour (in 2000 US$) for cars and light trucks, depending on the purpose of the trip (personal, commuting, or 'on-the-clock' travel). In contrast, a study by Levinson and
Smalkoski (2003) finds a mean value of $49.42/hour for commercial vehicle operators in Minnesota, whereas De Jong (2000) finds a range of values from $36/hour to $48/hour for commercial vehicle operators. Given this observation, it is assumed that for any given traffic volume-capacity ratio,

\[ D_X(t) = D_Y(t) \quad (7) \]

Equations (4) and (5) implicitly determine \( Y \) and \( X \) as functions of \( t \) and \( K \):

\[ Y = Y(t, K), \quad X = X(t, K) \quad (8) \]

and so \( V = Y + X \) is also a function of \( t \) and \( K \). The 'shapes' of these functions with respect to \( t \) and \( K \) will be explored in the next section.

3. EFFECT OF HIGHWAY CONGESTION PRICING AND CAPACITY INVESTMENT

3.1. Fixed capacity

Consider first the case where the highway capacity is given. This fixed-capacity assumption may be reasonable, especially in the short run, given that highway capacity expansion is lumpy, costly and irreversible, and usually involves substantial lead times for planning and construction. The fixed-capacity assumption also allows us to focus on the 'pure' effect of highway congestion pricing on the gateway.

More specifically, taking the consumers’ and carriers’ behavior - which is summarized in equations (8) - into account, region H chooses highway toll \( t \) to maximize its social welfare. Here, assuming without loss of generality the highway operating costs to zero, region H maximizes the sum of its toll revenue, \((Y + X)t\), and the region's consumer surplus:

\[ \text{Max } SW = CS + (Y + X)t \quad (9) \]

The consumer surplus (\( CS \)) includes both the surplus from the consumption of cargo \( X \), and the surplus from the consumption of local highway service \( Y \). As is standard in the literature on partial equilibrium industry studies, \( CS \) is assumed to arise from a utility function that can be approximated by the form:

\[ u(Y, X) + \phi \quad (10) \]

where \( Y \) and \( X \) are separable and uncorrelated in the utility function \( u(Y, X) \). Further, \( \phi \) is the expenditure on other goods in region H and its composite price is normalized to unity, with

\[ \frac{\partial u}{\partial Y} = P_Y, \quad \frac{\partial u}{\partial X} = P_X \quad (11) \]
Thus, region H’s consumer surplus can be written as:

\[ CS = u(Y, X) - P_Y Y - P_X X \]  (12)

where \( \phi \) is suppressed for notational simplicity.

Incorporating (12) and (11) then helps to yield the following first-order condition for (9):

\[ t = \frac{(YD'_Y + XD'_X)}{K} \]  (13)

The optimal toll, denoted as \( t^* \), is determined by equation (13). Unfortunately, it is not possible to solve \( t^* \) explicitly with the general specifications of demand and congestion delay cost used here. The pricing rule for the highway given by equation (13) does, however, indicate that \( t^* \) must be strictly positive. This result is not surprising; if there were no congestion at the highways - i.e., \( D'_Y = D'_X = 0 \) - then by (13), \( t^* = 0 \). Thus, the positive toll is the result of highway congestion. It is this congestion toll with which the congestion externality caused by atomistic carriers and drivers is internalized and so the (locally) optimal use of highway can be achieved.

**Proposition 1.** For given highway capacity,

(i) the hinterland region imposes a congestion toll on its highway so as to maximize its social welfare

(ii) When the hinterland moves from no-congestion pricing to congestion pricing, the local highway traffic rises

(iii) When the hinterland moves from no-congestion pricing to congestion pricing, the port traffic at the gateway falls, remains unchanged, and rises if \( D'_Y - D'_X < Kp'_Y \), \( D'_Y - D'_X = Kp'_Y \), and \( D'_Y - D'_X > Kp'_Y \), respectively.

**Proof:** The first part of the proposition has been shown in the text. Thus, when the hinterland moves from no-congestion pricing to congestion pricing at its highway, the highway toll will increase from 0 to a positive toll. It is sufficient, therefore, to show parts (ii) and (iii) by examining the signs of \( Y_i = \frac{\partial Y}{\partial t} \) and \( X_i = \frac{\partial X}{\partial t} \), respectively, with \( X = X(t, K) \) and \( Y = Y(t, K) \) being given by (8). To do this, totally differentiating equations (4) and (5) with respect to \( t \), and using (8), yield:

\[ P'_Y Y_t = D'_Y \frac{Y_t + X_t}{K} + 1, \quad P'_X X_t = D'_X \frac{Y_t + X_t}{K} + 1 \]

Solving the above two equations for \( Y_t \) and \( X_t \) yields:

\[ Y_t = \frac{D'_X - D'_Y - Kp'_X}{-P'_X D'_Y - P'_Y D'_X + Kp'_X p'_Y} \]  (14)
where by downward-sloping demands and conditions (6), the denominator in (14)-(15) is positive. By condition (7) and $p'_{x} < 0$, the numerator of (14) is positive and hence, $Y_{r} > 0$. Further, the sign of $X_{r}$ is the same as that of the numerator of (15), which is negative, zero, and positive if $D'_{Y} - D'_{X} < Kp'_{Y}$, $D'_{Y} - D'_{X} = Kp'_{Y}$, and $D'_{Y} - D'_{X} > Kp'_{Y}$, respectively. Part (iii) of Proposition 1 then follows immediately.

Q.E.D.

Thus, while highway congestion pricing encourages more local traffic, the traffic going through the gateway may fall or rise with highway congestion pricing, depending on the curvatures of the demand and delay functions, as well as the highway traffic and capacity. In particular, for given highway capacity $K$, the impact of highway congestion pricing on the gateway traffic depends on the magnitude of the difference between the carriers' and local users' values of travel timesavings. As noted in (7), freight trucks have a higher value of time than private cars. If such difference were sufficiently high, then an increase in the highway toll would reduce the gateway traffic. On the other hand, if the difference is sufficiently small, then $X$ would rise. These are the conditions that can be checked empirically, such that a specific prediction may be given.

Proposition 1 focuses on the impact of highway congestion tolling on the traffic going through the gateway. What would be the impact on the gateway region's welfare? Given the local monopoly nature of seaports, the profit of the port is likely to increase with its traffic volume $X$. In addition, as $X$ determines the size and economic activity of the gateway, one may argue that there are other benefits for the gateway that increase with $X$. These other benefits may include economic development benefits that are conferred on the gateway region in the form of enhanced employment, expanded output in other sectors, efficient logistics chains, increased tax revenue and higher real estate values in local land. For instance, in the New York/New Jersey gateway, it has been estimated that port-related freight activity contributes annually $18 billion in economic activity and $2.2 billion in tax revenue (Berechman, 2007).

If traffic volume of the gateway can be used as a proxy for its welfare, then Proposition 1 implies that region G's welfare may fall or rise with highway congestion pricing depending, among other things, on the relative magnitudes of carriers' and local users' values of travel time savings. In particular, the difference between the values of time for the cargo carriers and local highway users is large, then congestion tolling in the hinterland - while improving the hinterland region's welfare - would reduce the gateway region's welfare. In this case, the optimal congestion toll for the overall gateway-hinterland intermodal system would be lower than would be if the hinterland region were considered alone.
3.2 Variable capacity

Next we examine the variable capacity case. To focus on the impact of capacity expansion on traffic volumes, we first consider that the highway toll is fixed. The fixed-toll case can arise when congestion pricing is not practiced so that the toll stays at zero. While highway congestion pricing becomes more popular, it is still not common in reality. Lindsey (2007) points out, for example, that in Canada road congestion tolls are non-existent *per se*. Further note that this case can also arise if the capacity and toll decisions are made simultaneously.

**Proposition 2.** For a given highway toll, an increase in highway capacity will increase both the local highway traffic and the port traffic at the gateway.

*Proof:* The proposition follows if the signs of \( Y_k = \partial Y / \partial K \) and \( X_k = \partial X / \partial K \) are both positive, where \( X = X(t, K) \) and \( Y = Y(t, K) \) are given by (8). To do this, totally differentiating equations (4) and (5) with respect to \( t \), and using (8), yield:

\[
P_Y Y_k = D_Y \frac{(Y_k + X_k)K - (Y + X)}{K^2}, \quad P_X X_k = D_X \frac{(Y_k + X_k)K - (Y + X)}{K^2}
\]

Solve the above two equations for \( Y_k \) and \( X_k \):

\[
Y_k = \frac{-(Y + X)P_Y D_Y}{-K P_X D_Y - K P_Y D_X + K^2 P_X P_Y}
\]

\[
X_k = \frac{-(Y + X)P_X D_X}{-K P_X D_Y - K P_Y D_X + K^2 P_X P_Y}
\]

Both are positive given the shapes of the demand and delay functions.

*Q.E.D.*

These effects are as expected: holding highway tolls constant, an increase in highway capacity reduces congestion and hence stimulates traffic. Unlike the generally ambiguous effect of highway congestion pricing on port traffic, here an increase in highway capacity will have an unambiguously positive effect on traffic going through the gateway region. Again, if the gateway’s welfare is positively related to port traffic volume, then the capacity increase in the hinterland would improve the gateway’s welfare. This suggests that in this case, highway capacity would be under-invested for the overall gateway-hinterland intermodal system if the investment decision is made from the perspective of the hinterland region alone.

The above capacity analysis holds highway tolls constant. An alternative formulation of the problem is the case of variable tolls, in which the pricing and capacity decisions are linked with each other. In particular, given the long life, inflexibility and lumpiness of port capacity, it would be sensible to consider a two-stage formulation, with capacity being determined strictly prior to highway tolling. In this two-stage formulation, optimal highway tolls are ‘endogenously’
determined. That is, the optimal toll $t^*$ is now a function of capacity $K$, which is implicitly determined by equation (13). This introduces an additional, indirect channel through which capacity investment affects traffic: that is, investment impacts pricing which in turn impacts traffic. More specifically, the impacts of $K$ on the local highway traffic $Y$ and gateway traffic $X$ are given by, respectively,

$$\frac{dY}{dK} = Y_t \frac{dt^*}{dK} + Y_K$$  \hspace{1cm} (18)$$

$$\frac{dX}{dK} = X_t \frac{dt^*}{dK} + X_K$$  \hspace{1cm} (19)$$

Recall that $Y_t$ and $X_t$ are given by (14) and (15) respectively, whereas $Y_K$ and $X_K$ are given by (16) and (17) respectively. To sign $dY/dK$ and $dX/dK$, we need to sign $t^*_K = dt^* / dK$. The latter is determined by totally differentiating (13) with respect to $K$:

$$t^* + Kt^*_K = D^*_Y \frac{(Y_t t^*_K + Y_K + X_t t^*_K + X_K)K - (Y + X)}{K^2} Y + (Y_t t^*_K + Y_K)D^*_Y +$$

$$D^*_X \frac{(Y_t t^*_K + Y_K + X_t t^*_K + X_K) - (Y + X)}{K^2} X + (X_t t^*_K + X_K)D^*_X$$  \hspace{1cm} (20)$$

As can be seen from equation (20), with the general demand and delay functions, it is not possible to have a closed-form solution for $t^*_K$ and hence to sign $t^*_K$.

With linear delay functions, we can obtain the following expression for $t^*_K$:

$$t^*_K = \frac{D^*_Y Y_K + D^*_X X_K - t^*}{K - D^*_Y Y_t - D^*_X X_t}$$  \hspace{1cm} (21)$$

where $D^*_Y$ and $D^*_X$ are two constant parameters. After some long and tedious manipulations, (21) can be rewritten as

$$t^*_K = \frac{(D^*_Y - D^*_X)(D^*_X P^*_Y Y - D^*_Y P^*_X X) - K P^*_X P^*_Y (D^*_Y Y + D^*_X X)}{(D^*_Y - D^*_X)^2 + K^4 P^*_X P^*_Y}$$  \hspace{1cm} (22)$$

As can be seen, while the denominator of (22) is positive, the numerator can be positive or negative, depending on the relative magnitudes of the carriers’ and local users’ values of time, the shapes of demand functions, and traffic volumes $Y$ and $X$. Here the implications of a capacity increase for highway pricing are in general two-fold: first, the capacity expansion relieves congestion, which implies a lower congestion toll – a negative effect on the price. Second, capacity expansion

---

8 Lindsey (2007) has noted the similar point.
improves highway service, owing to less congestion, which results in a higher willingness-to-pay by the users. This effect will imply a positive effect on the price. Which effect dominates will depend on parameters of the model, as noted above.

**Proposition 3.** With endogenous highway tolls, the impacts of highway capacity expansion on the highway toll and the traffic volumes are in general undermined. For linear delay functions with $D_r' = D_x'$, however, highway capacity expansion will raise the optimal highway toll, and increase both the local highway traffic and the port traffic at the gateway.

*Proof:* The first part of the proposition is discussed in the text. For the second part, when $D_r' = D_x'$, then (22) implies a positive $t_x'$, and the rest of the results then follow by using (18), (19) and Propositions 1 and 2. Q.E.D.

4. **CONCLUDING REMARKS**

This paper has investigated the impact of congestion pricing and capacity investment at the hinterland’s road on both the hinterland and gateway. We found, for a given road capacity, that when the hinterland moves from no-congestion pricing to congestion pricing, while both its local traffic and social welfare improves, the traffic going through the gateway may fall or rise. The latter depends critically on the magnitude of the difference between the freight carriers’ and local users’ values of travel time savings. If freight trucks have a much higher value of time than private commuting cars, then road congestion pricing would reduce the gateway traffic. In this case, if the gateway’s welfare is positively related to port traffic volume, then congestion tolling in the hinterland - while improving the hinterland region’s welfare - will reduce the gateway region’s welfare. Thus, the optimal congestion toll for the overall gateway-hinterland intermodal system would be lower than would be if the toll decision is made with only the hinterland’s interests being considered.

We further showed that for fixed road tolls, an increase in road capacity stimulates traffic by both local road users and gateway cargos. Unlike the generally ambiguous effect of road congestion pricing on port traffic, here an increase in road capacity will have an unambiguously positive effect on traffic going through the gateway. Again, if traffic volume is used as a proxy for the gateway’s welfare, then the capacity increase in the hinterland would improve the gateway’s welfare. This suggests that road capacity would be under-invested for the overall gateway-hinterland intermodal system if the investment decision is made based only on the perspective of the hinterland region. Taken together, our analysis therefore suggested that there may a need for a central (cross-region) body to coordinate decisions concerning capacity investment and congestion pricing for the gateway and hinterland transport facilities.

Finally, if road capacity is determined prior to road pricing, then there is an additional, indirect channel through which capacity investment affects traffic: that is, investment impacts road pricing which in turn will affect traffic. Here the implications of a capacity increase for road pricing are in general two-fold: first, the capacity expansion relieves congestion, which implies a lower congestion toll - a negative effect on the price. Second, capacity expansion improves highway service, owing to less congestion, which would lead to a higher willingness-to-pay
by road users, thus implying a positive effect on the price. Which effect dominates will depend on parameters of the model; consequently, the impacts of highway capacity expansion on optimal highway tolls and the traffic volumes at both the gateway and hinterland are in general undermined. In such cases, empirical investigation is needed in order to ascertain these effects.

REFERENCES

Australian Government (2005), Australia's Export Infrastructure: Report to the Prime Minister by the Exports and Infrastructure Taskforce, Commonwealth of Australia, Canberra


