Transportation Infrastructure Investments, Pricing and Gateway Competition: Policy Considerations

Robin Lindsey
Department of Economics, University of Alberta, Canada

ABSTRACT

To alleviate congestion and facilitate Asia-Pacific trade the federal and western provincial governments have launched several major transportation infrastructure investments. This paper addresses some public policy issues regarding investments of this type. Does the private sector have incentives to make efficient investments and to impose efficient congestion charges? What about local governments? Are subsidies warranted either for investments or for ongoing operations and maintenance of infrastructure? Should governments pursue strategic trade objectives? Do exist institutions, regulations and practices conducive to making good decisions? Some dangers of making wrong investment decisions are identified.

1. INTRODUCTION

Rapid growth in Asia-Pacific trade is straining transportation infrastructure in Canada. Congestion delays are occurring not only at the Port of Vancouver, but also on railway and highway links to and from the Port and at intermodal transfer points. These delays impede not only Canadian imports and exports, but also transshipments to and from the U.S., freight shipments within Canada, and passenger transportation – especially on highways and bridges in the Lower Mainland of British Columbia.

In response to the need for new and upgraded infrastructure the federal government launched the Asia-Pacific Gateway and Corridor Initiative (APGCI) in October 2006 by granting $591 million for investment and other policy measures to enhance transportation infrastructure in western Canada. The 2007 federal budget added to APGCI another $410 million over seven years, as well as money for a national fund for gateways and border crossings and an expansion of the Gas Tax Fund transfer to municipalities. And the B.C. government has introduced a parallel Asia Pacific Initiative to promote trade and investment.

As a result of these initiatives several major transportation infrastructure investments are underway including new roads and bridges in the B.C. Lower Mainland, capacity expansion at the Ports of Vancouver and Prince Rupert, railway corridor enhancements and various projects in the Prairie Provinces. Some of the new and upgraded facilities will benefit local traffic. But the primary goal is to
boost international trade by increasing throughput capacity and reducing congestion delays. Indeed, the federal initiative sets a target to increase Canada’s share of trade flows.\(^1\) Only a small proportion of container traffic to and from the western U.S. currently moves through B.C. ports. Given the rapid growth in Asia-Pacific trade, and severe congestion at U.S. seaports and on U.S. rail lines and highways, the U.S. market is seen as a growth opportunity.

This paper is organized around five general public-policy questions about gateways and corridors\(^2\) that encompass not only the APGCI, but also initiatives elsewhere in Canada and other countries:

1. Is usage of infrastructure priced efficiently for the various transportation modes?
2. Does the private sector have incentives to impose efficient user charges?
3. Are government subsidies warranted for infrastructure investments?
4. Are government subsidies warranted for ongoing operations and maintenance of infrastructure?
5. Are existing institutions, regulations and practices conducive to making good decisions?

The paper tackles these and related questions from a theoretical perspective that draws heavily on transportation economics. Attention is focused on freight transportation although much of the analysis applies to passenger transportation as well. Section 2 begins by discussing the links between pricing and investment decisions. It then reviews the economics of congestion pricing by public and private sector infrastructure managers (Questions 1 and 2). Section 3 considers infrastructure investments in an environment with competing and complementary facilities and addresses Question 3. Section 4 reviews the theory and empirical evidence on infrastructure cost recovery and the case for operating subsidies (Question 4). Section 5 considers some of the dangers of making bad investment decisions and touches on Question 5. Section 6 concludes.

2. INFRASTRUCTURE PRICING DECISIONS

2.1. Interdependence of pricing and investment decisions

The demand for transportation depends on the generalized cost or full price of transportation. For freight the full price comprises fuel costs, intermodal transfer charges, infrastructure user charges, taxes, and the cost of time spent in transit — which includes wages of the driver, conductor or pilot, and the cost of interest, physical depreciation and technological depreciation of the shipment. A change in any component of the full price will affect the volume of transportation and hence the benefit from expanding or upgrading infrastructure. It is theoretically

\(^1\) It seeks to “boost Canada’s commerce with the Asia-Pacific region; increase the Gateway’s share of North America bound container imports from Asia; [and] improve the efficiency and reliability of the Gateway for Canadian and North American exports.” (Transport Canada 2006, 3).

\(^2\) Gateways can be defined as nodes in an intermodal transportation network that have particular topological or geographical features (Transport Canada 2006; Parsons et al. 2006). Corridors are major links in the network that connect one or more gateways. This paper treats transportation infrastructure generically without distinguishing — unless otherwise indicated — between gateways and corridors.
ambiguous whether the benefit is an increasing or decreasing function of the full price level. But the benefit can be quite sensitive to prices. Given the long life, inflexibility and lumpiness of transportation infrastructure, adding or removing capacity in the appropriate amounts is difficult or impossible. Investment decisions should therefore be taken with consideration not only for projected growth rates in economic activity that affect freight transport demand, but also changes in prices and tax rates as well as changes in the structure of prices and taxes.

With the notable exception of roads infrastructure user charges are levied on all transport modes in Canada. Congestion pricing is not yet standard practice for any mode of transportation in Canada. But recent developments suggest that this may change (see the Appendix for a brief discussion). Changes in taxes are also likely. Airports and seaports in Canada currently receive less favourable tax treatment than do their counterparts in the U.S., and federal ground rents paid by major airports are contentious. There is pressure to reduce these charges although the March 2007 federal budget made no changes. There is also growing support for environmental charges on transportation (as well as other sectors of the economy) in the form of fuel taxes or carbon taxes. As such changes in user charges and taxes occur they will affect the usage of transport infrastructure and the benefits derived from it.

Just as prices affect efficient investment levels, so do investments affect the full price of transportation. By shifting out the supply curve of infrastructure services, infrastructure investment influences user costs through changes in congestion delay, vehicle wear and tear, accident hazards and so on. Whether the infrastructure is managed by the public or private sector, investment is also likely to change user charges. If investment entails capacity expansion that relieves congestion, user charges may fall. But if investment consists of building more direct or faster routes, or improving reliability, demand will rise and higher charges may be forthcoming. Whatever the case, it is generically true that pricing and investment decisions are interdependent and should therefore be considered jointly rather than in isolation.

---

A series of theoretical studies have examined whether second-best infrastructure capacity when usage is underpriced is greater or smaller than first-best capacity. There are two opposing effects. With inefficiently low prices usage is excessive and expanding capacity leads to a further increase in usage that contributes to the distortion. However, more traffic benefits from any congestion relief. Little can be said about the net effect without specific assumptions (Arnott and Yan 2000).

For example, in the absence of congestion tolls capacity expansion can induce an increase in usage to the point where congestion delays are no shorter than before the expansion and the potential benefits of the investment are completely dissipated. This is an extreme manifestation of latent demand that materializes when supply is expanded. Empirical estimates of latent demand vary. According to Noland and Lewison (2002) the elasticity of vehicle-km traveled with respect to road network capacity ranges from 0.2 to 0.6 in the short run, and 0.6 to 1.0 in the long run. However, a recent study by the Conference Board of Canada (2006) found no statistically significant increase in long-run travel in Canada in response to road investments.

Infrastructure management and service provision are separated for air and maritime transportation, and airport and seaport charges are paid by airlines and shippers that provide service on behalf of final consumers. (Airport improvement fees are an exception since they are paid directly by passengers.) For rail transportation the major players – Canadian National and Canadian Pacific – own the infrastructure, and user charges are bundled with freight rates. But infrastructure costs for all three modes are passed on to customers and consequently affect the full price that shippers pay.
2.2. Efficient pricing, and pricing by private and local public infrastructure managers

The levels and structure of user charges for infrastructure depend on who manages the infrastructure, the position of the infrastructure in the transportation network or supply chain, and other factors. This subsection summarizes the major cases with a view to assessing the efficiency of prices and the implications of price setting for investment decisions. The focus is on congestion pricing since congestion is the most costly external cost of transportation (at least for road transportation) and it is the primary motivation for the APGCI investments.

2.2.1. Efficient pricing

For infrastructure to be used efficiently the marginal benefit from usage must balance the marginal costs incurred including user costs, infrastructure costs and external social costs. This calls for marginal social cost pricing. If there are no distortions in other markets so that first-best conditions apply, user charges should be set equal to resource costs (i.e. the cost of inputs used to provide the service) plus a Pigouvian tax to internalize any external costs. If other markets are distorted – as they typically are in practice – efficiency calls for second-best pricing. The rules for second-best pricing differ for mispriced substitutes and mispriced complements. If a substitute is underpriced (e.g. a congested, toll-free highway), then the second-best price is set below the first-best price in order to limit overuse of the substitute. If the substitute is overpriced, then the second-best price is set above the first-best level to encourage usage of the substitute. For complements the rules are reversed: for an underpriced complement the second-best price is above the first-best price, and for an overpriced complement the second-best price is below it. In practice both mispriced substitutes and mispriced complements are likely to be present, and the second-best price can be difficult to determine.

2.2.2. Pricing with private-sector management

If infrastructure is managed by an unregulated profit-maximizing entity, prices will generally diverge from efficient prices for two reasons. First, the manager will add a markup above marginal cost in proportion to its market power. Second, the manager will disregard some of the external costs of transportation. It is important to distinguish here between external public costs that are borne by society at large such as emissions and noise, and external user costs that are external to the individual user but are borne by users collectively such as congestion, infrastructure damage and a portion of accident costs. External public costs will not be internalized by private infrastructure managers. Nor will they be internalized by users or by independent service providers (e.g. trucking companies using public highways or air freight companies operating at airports). But private infrastructure managers do have an incentive to account for external user costs since these costs affect users’ willingness to pay to utilize the facility. In particular, by constraining usage in order to reduce congestion delays a manager provides a more valuable service and can charge a higher price for it.
The efficiency of private-sector congestion tolls depends on the market power of the manager. A private manager with no market power will charge the first-best congestion toll (Mohring 1985). But a manager with market power sets a congestion toll that is too high because the profit-maximizing markup is a percentage of the generalized cost, which includes external costs. The cost of congestion is therefore reflected not only in the congestion toll but also in the markup. The distortion is magnified if there are two or more private infrastructure managers in the supply chain—each of whom levies a congestion toll. This is an example of double marginalization—familiar in industrial organization theory—whereby upstream and downstream firms each add a markup to their prices while disregarding the loss of business they inflict on each other.

The double marginalization problem also arises when individual users of the infrastructure account for an appreciable fraction of aggregate traffic. An airline that operates multiple flights from a given airport is a good example. The airline incurs costs on itself in the sense that the airline’s individual customers and flights impose congestion delays on the airline’s other customers and flights. The airline has an incentive to internalize this self-imposed cost by (implicitly) levying a charge on its own flights that varies in proportion to its share of total airport traffic. There is empirical evidence that airlines do just this (Brueckner 2002). Similar logic applies to large shipping lines, although this does not appear to have been formally studied in the literature. In such cases the second-best congestion charge for a public infrastructure manager is reduced by the amount of the self-congestion charge since imposing a full charge would reduce traffic volume below the optimal level. But a private manager will impose the full congestion charge plus a markup. Charging twice for the same congestion delays is inefficient, and the inefficiency is exacerbated in proportion to the market power of each party. The cumulative effect can be especially severe if there are multiple links in the supply chain such as interlining flights or intermodal freight shipments that traverse a series of congested nodes and links (Basso 2005).

2.2.3. Pricing with local public-sector management

In the review of efficient pricing in Section 2.2.1 it is implicitly assumed that all users are residents in the sense that their welfare is given equal weight with operating costs and revenues from user charges. In practice, most facilities are used by nonresidents too. Foreign airlines use domestic airports, foreign shipping lines using domestic seaports, and out-of-province trucking companies use provincial highways. In such cases a public-sector facility manager can be said to be “local”.

Depending on a local facility manager’s mandate, it may treat nonresident users as a source of profit rather than as constituents of the jurisdiction. If so, the manager’s objectives will be intermediate between those of a global welfare-
maximizing public manager and a pure profit-maximizing private manager. If the manager can price discriminate, it will generally want to charge nonresidents a higher price than residents.\(^9\) But the prices for both groups will include the full external cost that they impose on users collectively. In particular, the price for residents will include a congestion charge for the costs imposed on nonresident users because the willingness to pay of nonresidents varies dollar for dollar with the congestion costs they incur.\(^10\) In this respect, a local public-sector manager’s behavior is the same as that of a private-sector manager’s, and the same observations apply regarding the exercise of market power and the markup included in the congestion charge.

### 2.2.4. Pricing with substitutes and complements

Most transportation facilities face competition from competing facilities while also benefiting from complementary relationships with facilities in the same supply chain. The price chosen by the manager of one facility depends on the prices and capacities of other facilities. The sign of this functional relationship – which is generally the same for public and private sector facilities – is shown in Table 1 for the case of two facilities, \(A\) and \(B\).\(^11\)

In the case of competing or substitute facilities (first row of Table 1) an increase in the price charged at Facility \(A\) (e.g. because its costs have increased) will boost demand for and congestion at \(B\), and induce the manager of \(B\) to raise its price.

<table>
<thead>
<tr>
<th>Relationship between (A) and (B):</th>
<th>Effect on price of (B) of:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Price of (A)</td>
</tr>
<tr>
<td>Substitutes</td>
<td>+</td>
</tr>
<tr>
<td>Complements</td>
<td>-</td>
</tr>
</tbody>
</table>

**Table 1: Typical comparative statics properties of prices**

By contrast, an increase in the capacity of \(A\) will alleviate delays at \(A\), encourage users to switch from \(B\) to \(A\), and induce the manager of \(B\) to reduce its price. An increase in the capacity of \(B\) has a theoretically ambiguous effect. It reduces congestion at \(B\), and hence the congestion toll component of the price at \(B\), but it

\(^9\) There are exceptions. As Barter (2006) points out, Singapore charges goods vehicles that enter Singapore from Malaysia an entry fee of only $S10 per month. This is a tiny fraction of the $S20 charge per weekday imposed on foreign-registered cars, and it is also small compared to the average monthly electronic road pricing charges paid by residents of Singapore. According to Barter, Singapore imposes the low goods vehicle charge to encourage shippers to transship freight through Singapore’s ports rather than use competing ports in Johor.

\(^10\) This symmetry reflects the reciprocal nature of congestion.

\(^11\) Table 1, as well as Table 2 in Section 3, is constructed from results derived in Tirole (1988, §4.2), De Palma and Lindsey (2000), De Borger, Proost and Van Dender (2005), De Borger and van Dender (2006) and De Borger, Dunkerley and Proost (2006). The results are labeled “typical” because it is possible to construct examples in which signs are reversed.
also increases the equilibrium market share and market power of \( B \). The net effect
can go either way.\(^{12}\)

In the case of complements the signs of the functional relationships are
generally opposite to those for substitutes. If Facility \( A \) raises its price or reduces
its capacity, demand for \( B \) falls and \( B \) will reduce its price. If \( B \) increases its
capacity it will choose to reduce its price since congestion will fall and – unlike the
case of substitutes – it does not gain business from \( A \) since the two facilities serve
the same customers.

The interplay between pricing and capacity reviewed in this section is clearly
important when considering infrastructure investment decisions. This is the subject
of the next section.

3. INFRASTRUCTURE INVESTMENT DECISIONS

3.1. Incentives for efficient investment

The investment incentives for public and private sector infrastructure managers
are similar as far as efficiency. For any given level of traffic it is optimal to
minimize the sum of (annualized) infrastructure capacity costs and user costs.\(^{13}\) An
increase in capacity is therefore warranted if it reduces user costs by more than
the cost of the investment. A private manager can recoup the investment cost in
greater revenues, and a public manager can recoup the investment cost in the sum
of enhanced revenues and consumers’ surplus. In general, efficient capacity levels
vary inversely with prices because higher prices depress the volume of traffic that
will benefit from investment. To the extent that private-sector prices exceed first-
best prices, private investment will be below first-best investment levels. In
principle, this bias could be addressed with an investment subsidy, but a more
direct intervention is to regulate prices or facilitate competition by encouraging
the creation of alternative infrastructure facilities.

3.2. Strategic investment considerations

Since infrastructure investments are long lasting and typically irreversible, they
can influence the behavior of other players in ways that can either benefit or harm
the investor. Investments are therefore “strategic” in the sense that the term is
used in dynamic game theory. Strategic effects can matter for both public and
private-sector infrastructure managers, and as explained in the next two
subsections they can play out through changes in the pricing and investment
decisions of other players.

---

\(^{12}\) In De Borger and Van Dender’s (2006) private-duopoly model the profit-maximizing price of a firm is
independent of its capacity. As they note, this result derives from their assumption that the demand curve is
linear.

\(^{13}\) See Small and Verhoef (2007, §6.1).
3.2.1. Strategic effects on prices

The strategic effects of investments on prices have been analyzed in a number of theoretical studies.\textsuperscript{14} The prototypical model is a two-player, two-stage game in which two infrastructure managers, $A$ and $B$, simultaneously choose investments (or capacities) in stage 1, and simultaneously choose prices in stage 2. The results are qualitatively similar for private and local public-sector managers and can be summarized by reference to the middle column of Table 1 in Section 2.

In the case of substitutes an increase in the capacity of $A$ will induce $B$ to cut its price. Because a price cut at a competing facility is undesirable for either a private or a local public manager, the manager has a strategic motive to invest in less capacity than it otherwise would. Holding back capacity in stage 1 of the game reduces the intensity of price competition in stage 2. This reinforces the bias towards underinvestment in capacity due to the exercise of market power.

In the case of complements an increase in the capacity of $A$ will induce $B$ to raise its price. This is disadvantageous to $A$ because the facilities are complementary. Thus – as is true for substitutes – the manager of $A$ has a strategic motive to hold back capacity. Similarly, $B$ has a strategic incentive to hold back on its capacity. Naturally, holding back capacity is mutually harmful since the facilities are complements. In practice it may be possible to resolve the conflict of interest by merging the two facilities or through some investment agreement.

3.2.2. Strategic effects on investments

Investments can have strategic effects on other players' investments only if the investment – or at least a credible commitment to invest – is made before the other players can move. Such a situation can be modeled by extending the two-stage game to a three-stage game in which first $A$ invests, then $B$ invests, and finally $A$ and $B$ simultaneously choose prices. The typical effect of $A$'s capacity choice on $B$'s capacity choice is shown in Table 2.

<table>
<thead>
<tr>
<th>Substitutes</th>
<th>Complements</th>
</tr>
</thead>
<tbody>
<tr>
<td>-</td>
<td>+</td>
</tr>
</tbody>
</table>

Table 2: Typical comparative statics properties of capacities

If the facilities are substitutes, $A$ has a strategic incentive to “overinvest” in capacity since this will deter $B$ from investing as much. If the facilities are complements, $A$ again has an incentive to overinvest because this will encourage $B$ to invest more. The strategic incentives for complements and substitutes are thus qualitatively the same even though the mechanics differ. Intriguingly, the strategic incentives for influencing investment are the opposite of the strategic incentives for influencing prices identified above. The timing of decisions is therefore crucial since a decision to overinvest in the hope of influencing other players’ investments

\textsuperscript{14} See the references in footnote 11.
will be counterproductive if the investment is made too late, and merely impacts adversely on the final price-setting stage of the game.\textsuperscript{15}

A few examples of alleged strategic overinvestment by governments have been cited in the literature. According to Phang (2003, p.29) Singapore has invested in airport capacity “far in advance of actual need” in order to discourage competing investments and maintain the dominant position of its hub airport in Southeast Asia. Similarly, Singapore built its mass rapid transit (MRT) system instead of a cheaper bus system in the belief that the MRT would be more effective at attracting high value-added investments and would deter other countries in the area from making similar investments. Barter (2006) remarks that strategic motives have occasionally overridden Singapore’s normally cautious approach to investments based on cost-benefit analysis.

Other documented examples of strategic investments or “destructive competition” are regional lobbying for airports in Thailand and competition for outside investment between municipalities in China (World Bank, 2005). As for Canada it has been suggested that a more aggressive investment strategy should be implemented at the Port of Prince Rupert in order to preempt competing port development plans in Mexico (Pitts 2006).

3.3. Arguments for and against infrastructure investment aid

Having reviewed the efficiency and strategic motivations for infrastructure investment, we now consider the merits and drawbacks of investment aid — either in the form of government aid to private-sector investments, or federal aid to provincial or other “local” governments. Some efficiency and public good considerations are addressed in this section.\textsuperscript{16} The case for strategic investment aid is taken up in Section 3.4.

3.3.1. Arguments for investment aid

A common, but controversial, argument for aiding private-sector investments is that the private sector is biased against large and risky projects. Transportation infrastructure projects such as major bridges and port expansions are large by most standards and they are also risky for several reasons. They require substantial lead times for planning and construction, which exposes them to the risk of cost escalation\textsuperscript{17} and changes in environmental and other appraisal criteria. They are long-lived, irreversible and typically have no alternative uses. Demand for freight transportation is volatile because it is a derived demand and depends on global macroeconomic conditions, exchange rates, commodity prices and other factors. Infrastructure capacity is subject to short-term disruptions due to equipment breakdowns, strikes or bad weather, as well as major shocks such as

\textsuperscript{\footnotesize 15}As Daltung et al. (1987) show for a case study of the Norwegian government and the Caribbean cruise shipping industry, subsidization of home-country firms may be disadvantageous unless foreign competitors reduce their capacities in response.

\textsuperscript{\footnotesize 16}Other arguments for government intervention based on market failures are reviewed in Boardman and Vining (2007).

\textsuperscript{\footnotesize 17}A recent example is the Mackenzie Valley gas pipeline project. On March 12, 2007, Imperial Oil and other project proponents submitted a revised cost estimate of $16.2 billion: some 110 percent higher than the original (2004) estimate of $7.7 billion.
natural disasters or terrorist attacks that have more severe and long-lasting effects.\textsuperscript{18}

There has been relatively little analysis – at least in the economics literature – of the implications of risks on optimal facility capacity. Optimal design capacity is determined by a tradeoff between the costs of holding slack capacity during off-peak and/or undisrupted periods, and the benefits of having reserve capacity available at times when demand is unusually high or a portion of capacity is out of commission. There is some presumption that optimal design capacity is increased by uncertainty.\textsuperscript{19} However, there is recent evidence that carriers and other private-sector stakeholders in North American freight transportation are dissuaded by uncertainty from investing in capital investments to upgrade their infrastructure (Ortiz et al. 2006).

According to public economic theory, aid to investments – whether public or private – may be warranted if they generate positive spillover effects that cannot be captured by the investor. This argument applies to provincial toll-free highways that are used by out-of-province residents and trucking firms. It also applies at a supranational level to investments that benefit trading partners by reducing the costs of trade (De Mooij et al 2005; Cárcamo-Díaz 2005; Cárcamo-Díaz and Goddard 2007). In theory, a donor country can itself benefit by helping trading partners improve their own infrastructure (Mun and Nakagama 2006; Fukuyama 2006) although the level of aid will be Pareto suboptimal.

One positive spillover effect that is often suggested as a reason to subsidize infrastructure is regional development. In the past, regional development was a priority of transportation policy in Canada, and explicit or implicit subsidies were widespread. But federal policy is now more commercially oriented. Regional development continues to be part of the explicit or implicit mandate for seaports in the U.S. and continental Europe. Yet it has been argued that the employment benefits from seaports have declined in importance with the shift toward capital-intensive operations. Jobs may be generated outside seaports (Verbeke 2007), but any employment gains may come at the expense of losses for other projects and/or other transportation modes (Sichelschmidt 1999). And gateways are increasingly disconnected from the cities in which they reside (Hall 2007).

Agglomeration economies are a related and more recently recognized potential benefit from infrastructure investments. As elaborated in the New Economic Geography literature agglomeration economies can derive from local demand for traded goods, thick labor markets, technology spillovers, international status and other positive feedback effects. As Baldwin et al. (2003) explain, infrastructure projects may have little effect on the location of economic activities below a threshold level, but a very large impact when a critical mass is reached. However, it is also possible that new infrastructure can promote dispersion rather than agglomeration. Moreover, some funding mechanisms such as the U.S. Highway Trust Fund tend to favor investments in suburban areas rather than city centers (Anderson and Lakshmanan 2004). And investments that are designed to


\textsuperscript{19}Kraus (1982) and d’Ouville and McDonald (1990) examined the implications of demand uncertainty for highway design and concluded that optimal capacity is increased. See also Colledge (2007).
promote development in low-density or remote regions may work against agglomeration economies.

### 3.3.2. Arguments against investment aid

As discussed in Section 3.2 speedy investment may have a strategic benefit, either by deterring a rival investment or by encouraging a complementary investment. However, speed has drawbacks. It may divert resources from regular operations and maintenance. U.S. railroads, for example, have been responding to the rapid growth in Asia-Pacific trade by expanding capacity on major corridors while deferring maintenance elsewhere (Ortiz et al. 2006). Rapid investment may raise the costs of other infrastructure projects. And it may contribute to general overheating of the economy. Overheating has been a concern in China and other Asian countries that have invested especially rapidly in infrastructure (World Bank 2005). It is also a source of unease in Alberta — in large part because of the economic boom caused by high oil and gas prices.

A second and widely recognized drawback of offering generous investment aid is that it can create moral hazard on the part of potential recipients. Availability of regional funds for cross-national transportation infrastructure such as the Initiative for the Integration of Regional Infrastructure in South America (IIRSA) can induce member countries to overstate the positive spillover effects of their proposed projects in order to secure funding (Cárcamo-Díaz and Goddard 2007). And if the funds have to be allocated quickly, the set of feasible recipients may be limited to projects already in advanced stages of planning that will probably be carried out anyway. According to Sichelschmidt (1999) this has been the case for some projects in the European Union funded Trans-European Network (TEN) program.

A third downside of investment aid is that it will be largely wasted unless the facility has natural advantages or achieves sufficient economies of scale to be competitive. In particular major seaports with good access to markets have significant economic advantages that make demand for port services price inelastic (Corbett et al. 2007) and entry for new seaports difficult (Medda and Carbonarao 2007).

### 3.4. Strategic government intervention

Section 3.2 considered a three-stage game in which Facility A can invest before Facility B and influence B’s investment choice in an advantageous way. Facility A can also gain an advantage if the first move of the game is taken by a government in the form of a subsidy that induces A to invest more aggressively.\(^\text{20}\) As Brander (1995) explains, a government subsidy helps a domestic firm not only directly by reducing its costs, but also strategically by enabling it to commit to a higher output.\(^\text{21}\) Brander also notes (p.1409) “the optimal subsidy is increasing in the

---

\(^\text{20}\) For this strategy to work A’s aggressive stance must of course be anticipated by B.

\(^\text{21}\) Firms can also be induced to expand outputs by altering their goals. As Ircha (2001, 212) observes, most U.S. ports view regional development as their major role. This often leads them (with the aid of state and local government subsidies) to set low prices and provide enhanced levels of service. He calls for similar incentives and aid for Canadian ports.
relative cost advantage of the domestic firm.” Provided the subsidy level is chosen judiciously, the cost of the subsidy to taxpayers is more than offset by an increase in the firm’s profits.

Government aid can take forms other than subsidies such as grants, loans or loan guarantees, tax breaks, accelerated depreciation allowances, and provision of assets at below costs (European Commission 2001, p.10). Such support is credible insofar as it requires only a one-time commitment since the durability of infrastructure assures that the effects of the aid will be long-lived as long as the infrastructure is adequately maintained. In the case of the Ports of Vancouver and Prince Rupert support may also exploit the “relative cost advantage” Brander mentions since the ports have shorter ocean line-haul routes to Asia than U.S. or Mexican ports. Prince Rupert also has the advantage of being congestion-free. The CN rail line linking the port to the interior is also congestion free, and trains can maintain high average speeds because of low grades.  

One danger in pursuing strategic investment aid is that the appropriate form of intervention is sensitive to the nature of competition. In the classical Boeing-Airbus example it is assumed that the two aircraft manufacturers compete in quantities. If competition actually plays out with prices rather than quantities, a tax rather than a subsidy is called for — as indeed is clear from the analysis of strategic price effects in Section 3.2.1.

Another complication — noted in Section 2.2 — is that facilities have both substitutes and complements, and this creates conflicting strategic incentives. Indeed, it may not always be clear whether two given facilities are substitutes or complements. The Ports of Vancouver and Prince Rupert would appear to be substitutes, but transshipment between them is possible and one port may provide backup capacity for the other in case of prolonged congestion or capacity disruptions. According to McMillan (2006), Canada’s two Tier 1 ports – Vancouver and Halifax – are complementary in the transcontinental supply chain, and he argues that regional ports that feed traffic to them are complementary as well. However, competition (at least for funding) between Halifax and regional ports in Atlantic Canada is discernible in discussions of the Atlantic Canada Gateway initiative (Horibe 2007).

A third drawback of strategic investment aid — also noted by Brander (1995, p.1411) — is that the benefits are diluted if recipients compete with each other rather than with competitors in other jurisdictions. More generally the potential benefits may be insufficient to warrant intervention. A final caveat is that if all countries engage in subsidization, they may all end up worse off (the Prisoners’ Dilemma).  

---

* The potential cost advantage of shipping containers through the new Prince Rupert terminal to the Twin Cities region in the U.S. is assessed in Stewart et al. (2007).
* As Brander (1995, 1417) remarks: “The Bertrand model is not necessarily any less plausible than the Cournot model as an approximation to actual conduct. Because it is hard to know in practice which of the two models (if either) is appropriate in a given case … even finding the sign or direction of the optimal policy might be difficult.”
* A case in point may be the widespread state aid to seaports in northern Europe.
4. COST RECOVERY AND THE CASE FOR OPERATING SUBSIDIES

Many publicly operated transportation facilities are obliged to be self-financing, including commercial airports in Canada and Canada Port Authorities. This raises the question whether public pricing and investment decisions such as analyzed in Sections 2 and 3 are compatible with cost recovery. It is of normative interest to know whether efficient (i.e. marginal social cost) pricing is consistent with the user-pays principle. It is also practically important since if a deficit would result under efficient pricing, either the facility must be subsidized or prices must be adjusted to close the gap – preferably in a way that sacrifices the least in terms of efficiency and competitive advantage.

4.1. Cost recovery and returns to scale

The classical result on cost recovery is the Cost Recovery Theorem due to Mohring and Harwitz (1962). The theorem states that the revenues from marginal social cost pricing of a congestible facility just suffice to pay for construction and operating costs if three conditions hold: (a) capacity is perfectly divisible and supplied at constant marginal cost; (b) user costs are homogeneous of degree zero; and (c) capacity is optimal. The empirical evidence on the three conditions varies by mode. Given the dearth of direct user charges for highways, and the vertically integrated structure of railways, airports and seaports are the most relevant for discussion.

(a) Infrastructure costs and capacity indivisibilities: Empirical studies find that airport infrastructure enjoys positive scale economies for traffic volumes up to 3-4 million passengers/year, and approximately constant economies beyond that level (Morrison 1983; Doganis 1992; Graham 2003; Forsyth 2005). For seaports there appear to be positive scale economies at least for certain components of infrastructure (Haralambides and Veenstra 2002; Strandenes 2004). The significance of capacity indivisibilities is difficult to judge. Some elements of capacity come in discrete lumps. For airports this is more the case for runways than for terminals and other landside facilities. Within limits, capacity can be varied without altering physical infrastructure by implementing more advanced air traffic control, Intelligent Transportation Systems (ITS) technology, more frequent maintenance and so on. However, some minimum level of infrastructure is required to handle any traffic, and airports and seaports with low traffic volumes are bound to run a deficit.

One input with an imperfectly elastic supply for many airports and seaports is land. To the extent that expanding capacity requires more land and drives up its cost, cost economies of scale will be smaller than returns to scale. However, if variable factor costs are properly accounted for the cost recovery theorem still holds (Small 1999).

---

* This section draws on De Palma and Lindsey (2004, 2007).
* User costs are homogeneous of degree zero if the cost of usage does not change if total usage and capacity are increased or decreased by equal proportions. Conditions (a) and (b) of the theorem effectively require constant economies of scale in infrastructure provision and infrastructure usage.
* This assumes that capacity is built once and for all, and that the purchaser of land cannot price discriminate.
(b) **Infrastructure usage:** Economies of scale in usage prevail at both airports and seaports for three reasons. First, there are economies of traffic density from increased service frequency as traffic volume grows. These economies are well documented for airports (Caves et al. 1984; Brueckner and Spiller 1994). For maritime transportation service frequency is less important for bulk freight – which is transported under contract – than for liner shipping which operates according to a schedule.\(^2\) Second, economies from massed reserves exist due to randomness in the arrival times of vessels and the time required to load and unload them. These economies are prevalent at airports (Morrison 1983) as well as seaports (Haralambides and Veenstra 2002) and they have grown at seaports as container ships increase in size. Third, to minimize the total time spent in transit for freight, maritime carriers prefer to load exports at their last port of call and to discharge imports at their first port of call (Baird 2002). Customers of seaports that handle lots of freight therefore enjoy an advantage in terms of shorter shipping times.

(c) **Optimality of capacity:** Given capacity indivisibilities, and long lead times for capacity expansion, infrastructure capacity is unlikely to be exactly optimal except at infrequent points in time. A schedule of infrequent and large investments will result in alternating periods of surplus and deficit. Cost recovery in present-value terms does not require that capacity additions be made at an optimal time. But it does require that the size of capacity additions be optimal conditional on the timing (Arnott and Kraus 1998), which will generally not be possible if investments can be made only in lumpy amounts. Other than for Oum and Zhang’s (1990) analysis for airports there appear to have been no empirical studies of whether a surplus or deficit is likely to result when capacity is indivisible.

The Cost Recovery Theorem has been generalized in various directions. While it is relatively robust, there are several reasons why the theorem and its extensions are unlikely to hold in practice. Four of these reasons warrant brief mention here.

1. **Unpriced congestion elsewhere on the network:** Unless usage is priced efficiently on the whole network, prices should be set according to second-best rules. As noted in Section 2.2.1, if unpriced facilities are substitutes then prices should be set below marginal cost. This will result in a deficit, *ceteris paribus.* If the unpriced facilities are complements, then prices should be set above marginal cost, which will result in a surplus.

2. **Self-internalization of congestion by large users:** As discussed in Section 3.1, users that account for an appreciable fraction of traffic has an incentive to internalize the congestion costs they impose on themselves. Infrastructure congestion charges should be reduced commensurately, and this will result in a deficit, *ceteris paribus.*

3. **Imperfect capital markets:** Cost recovery in present-value terms holds only if the operator can borrow freely to finance investments. This is not the case for Canada Port Authorities because the Canada Marine Act imposes stringent borrowing limits, and prohibits federal subsidies.

---

\(^2\) Achieving these economies is a driving force behind the growth of global shipping alliances.
4. **Strategic investment incentives**: Strategic investment incentives encourage overinvestment or underinvestment. This violates the definition of optimal capacity underlying the Cost Recovery Theorem.

4.2. **The case for subsidy**

The evidence reviewed above indicates that there are constant or moderately positive scale economies for airport and seaport infrastructure. Scale economies in usage are prevalent for both modes. The picture is muddied by capacity indivisibilities, treatment of land costs, strategic investment incentives and other considerations. But on balance it appears likely that a deficit will result.

If cost recovery is required — as is the case for commercial airports and Canada Port Authorities — then additional revenues must be raised by implementing Ramsey pricing rules or some other form of price discrimination. Airport price structures typically feature substantial price discrimination, and cross-subsidization between aeronautical and non-aeronautical (i.e. commercial) operations is feasible (Forsyth 2006). Seaports, however, do not have access to significant concession revenues from passenger traffic. Under the Canada Marine Act, Canada Port Authorities are constrained from engaging in commercial activities such as recreational marinas and landside developments that could yield additional revenues. And they can receive no federal subsidies for operations. In contrast, most U.S. ports do receive subsidies and they have power to levy taxes (Fawcett 2007).

Given current legislation in Canada, and the fiscal advantages of U.S. ports, there are arguments for subsidizing Canadian seaports. However, they need to be weighed against the usual drawbacks of subsidization. Subsidies burden taxpayers. Some of the benefits from subsidies may accrue to foreign shippers rather than to residents. Subsidy would tilt the playing field against other modes of freight transportation. And subsidy weakens incentives for cost efficiency — in part because users have less incentive to encourage productivity improvements (Asteris 2006). Transport policy in Canada has evolved towards commercialization, and the case for renewed aid may be difficult to make.

5. **DANGERS OF MAKING WRONG INVESTMENT DECISIONS**

Making good transport infrastructure investment decisions is inherently difficult given the complexity of supply chains for freight transportation, the durable and inflexible nature of infrastructure, the interdependence of pricing and investment decisions, strategic considerations and so on. Many factors can contribute to making mistakes. Discussion here is limited to two factors that apply generally to major infrastructure investments: faulty cost-benefit analysis and lack of coordination between decision makers.
5.1. Faulty cost-benefit analysis

Cost-benefit analysis is a well-established approach to project evaluation (Boardman et al. 2006) but it is easily misused. Mackie and Preston (1998) list twenty-one errors in transportation project appraisal that use cost-benefit methods. Attention is limited here to four errors that are especially relevant to the APGCI and similar initiatives.

1. **Inappropriate choice of study area**: The study area is the region within which the impacts of a project are evaluated. Time and budget constraints may oblige analysts to select a limited study area, although as far as standing (i.e. whose welfare counts) it is usual to tally costs and benefits nationally. Moreover, to assess properly a major gateway project that can affect freight markets globally, it may be necessary to extend the study area outside national boundaries. Defining the study area too narrowly may leave out important network effects, and introduce significant bias in a project’s estimated benefits and/or costs (Van Exel et al. 2002).

2. **Bad demand forecasts**: Transportation demand depends on exchange rates, commodity prices and other factors that are difficult to forecast accurately even over short time horizons. The general view is that rapid growth in Asian trade will be sustained for some time, although this outlook is not universal (Horibe 2007; Rodrigue 2007). But demand for a particular gateway or corridor depends not only on global demand but also on investments in competing facilities, developments in Open Skies Agreements, etc., which are also difficult to predict.

3. **Errors in assessing strategic responses**: As Section 3.2 explains, investments can have strategic effects on competitors, supply chain partners and other actors that play out through changes in prices, investments, entry/exit decisions and so on. These responses, too, are difficult to predict. A case in point is the failure of the Eurotunnel consortium to forecast the strength of the competitive response from ferry companies (Mackie and Preston 1998) or the degree to which low-cost airlines would penetrate the medium-distance passenger transportation market in Europe.

4. **Appraisal optimism**: Mackie and Preston (1998) conclude that ‘appraisal optimism’ is the greatest danger in project evaluation. As they explain

“Appraisal optimism happens because the information contained in the appraisal tends to be owned by scheme promoters who have obvious incentives to bias the appraisal – deliberately or unwittingly .... This is a particularly acute problem if the scheme is in the public rather than private sector, since the normal commercial checks and balances on excessive optimism do not apply.” (Mackie and Preston 1998, p.6)

Appraisal optimism is presumably a contributing factor to the prevalence of white elephants in urban transit systems, airports and other transportation infrastructure (Flyvbjerg et al. 2003). However, as discussed next, a lack of

---

29 It is also often bypassed in favour of other, less appropriate, techniques (Forsyth 2007).
coordination in formulating and implementing infrastructure projects may create an opposing bias against adequate investment as well as against infrastructure maintenance.

5.2. Lack of coordination

The likelihood of making bad investment and related decisions is increased not only by conflicts or misunderstandings between countries, but also within countries by lack of coordination between different transportation modes, between governments and private-sector stakeholders, and between government agencies. Two quotes from Canadian sources highlight these problems:

“In a region so dependent upon trade, there is little coordination in western Canada of port terminal and inland road and rail infrastructure development. [The result is] over or under-utilized infrastructure, missed opportunities and the potential for duplication of investments. An inability to balance supply and demand creates negative and lasting perceptions among users who experience service problems and/or increased costs.” (British Columbia Ministry 2005, p.11)

“In many cases, federal and provincial responsibilities overlap and the various regulatory regimes differ. Project developers find themselves dealing with several agencies that make no effort to coordinate...” (Asia Pacific Foundation of Canada 2006, p.25)

Coordination problems for investment tend to be greater the greater the interdependence in the players’ payoffs (e.g. because of network effects) and the larger the number of stages involved in the investment. Coordination also tends to be more difficult the larger the set of decision makers and stakeholders who are involved. This set can include firms at various stages of the supply chain; national, provincial and municipal governments; ministries with mandates for transportation, infrastructure, finance and the environment; indigenous communities, lobby groups and so on. At a provincial level the APGCI mainly concerns B.C. and Alberta. For the Atlantic Gateway initiative four Maritime provinces have stakes. And for the Windsor-Detroit Corridor project – which includes a new bridge and an access road linking it to Highway 401 – the parties involved include the federal government, the Ontario provincial government, the state of Michigan, and firms on both sides of the border engaged as public-private partnership participants.30

A frequent source of problems for infrastructure projects are inconsistencies in investment and budgeting decisions. When the economy is strong, financing for major infrastructure projects is often approved. But if the budget is in deficit, projects are vulnerable to cutbacks or cancellation because this provides large (albeit one-off) savings without a loss of current benefits (World Bank 2005). Separation between planning and financing decisions also contributes to inadequate expenditure on operations and maintenance. One reason is that maintenance costs are often underestimated (Cárcamo-Díaz and Goddard 2007, p.56). Another is a lack of long-term guaranteed funding for maintenance. Yet

---

30 A further complication is that there are two competing proposals: one by Cen Tra Inc. – which owns and operates the Ambassador Bridge – to build a new span adjacent to the existing bridge, and another proposal by the Detroit River Tunnel Partnership to use an existing rail tunnel for trucks (Jang 2007).
another is that a spurt of investment eventually leads to a large stock of aging infrastructure in need of maintenance, and consequently fewer funds left over for further investment.

Integrated transportation policy has long been seen as a way to overcome coordination problems (May et al. 2005; Van de Velde 2005). TransLink – the Greater Vancouver Transportation Authority – has served as a role model for coordination and cooperation between local governments in the planning and funding of roads and public transportation. However, in 2007 an independent review panel issued a report recommending a reorganization of TransLink on the grounds that it was not meeting its mandate adequately (TransLink Governance Review Panel B.C. 2007). At the federal level Transport Canada has a Policy Group devoted to strategic policy spanning all transportation modes. But – for better or worse – there is no equivalent institution to the EU to coordinate decisions between Canada and Asia, the U.S. or other parts of the world.

6. CONCLUSIONS

Major transportation infrastructure investment programs such as the Asia-Pacific Gateway and Corridor Initiative are inherently risky. Investing too little and too slowly may result in a failure to capitalize on opportunities, or even a loss of business that is very difficult to regain. But investing too much may lead to substantial deficits if traffic does not increase enough to recoup investment costs.\textsuperscript{31} Appropriate investment levels are especially tricky to ascertain if the marginal returns to investment are (locally) increasing in the amount spent because of agglomeration economies or advantageous strategic responses from other players. But determining the threshold level is difficult – particularly since it can take some time for the effects of investment to be realized.

Given the general and wide-ranging perspective taken in this paper, it is out of place to offer specific policy recommendations for the Asia-Pacific Gateway and Corridor Initiative, the Atlantic Canada Gateway initiative or other programs. Nevertheless, two directions for change seem logical to pursue. One is to place greater reliance on direct user charges. Road pricing in the form of congestion tolls is the most obvious candidate, but airport congestion pricing also deserves consideration. The other recommendation is to relax constraints under the Canada Marine Act that prohibit Canada Port Authorities from engaging in commercial activities and that limit the amount they can borrow to fund investments.

The case for providing investment aid – either to the private sector or to lower-tier governments – is difficult to judge. It requires a careful weighing of the potential merits – such as positive spillover effects from agglomeration economies and preemption of rival investments, against the potential drawbacks – such as overheating the economy and burdening taxpayers with white elephants. As a rule, governments should not try to pick winners.

\textsuperscript{31} Under the conditions of the Cost Recovery Theorem a given percentage error in choosing capacity can result in a much larger percentage financial surplus or deficit (De Palma and Lindsey 2004).
7. ACKNOWLEDGMENTS

For helpful comments I am grateful to participants at the Asia Pacific Gateway and Corridor Round Table in Calgary, March 29, 2007, and the International Conference on Gateways and Corridors in Vancouver, May 2, 2007. Special thanks are due to Steve Shepherdson for suggestions on how to focus the paper.

8. REFERENCES

Boardman, A.E. and A.R. Vining (2007), “Can P3s contribute to the upgrade of Canada’s Asia-Pacific trade infrastructure?”, presentation at the International Conference on Gateways and Corridors, Vancouver, May 4

Brueckner, J.K. (2002), "Airport congestion when carriers have market power", American Economic Review 92(5), 1357-1375


Holguín-Veras, J. (2005), “The truths, the myths, and the possible in freight road pricing in congested urban areas", working paper, Department of Civil and Environmental Engineering, Rensselaer Polytechnic Institute
Verbeke, A. (2007), “Strategic port and gateway development: European experiences”, presentation at the Asia Pacific Gateway and Corridor Round Table, Calgary, March 29

9. APPENDIX: CONGESTION PRICING IN PRACTICE

Seaports: Congestion at seaports is manifest not only in vessels waiting in harbours to load and unload, but also at intermodal transfer points. Access to many seaports is still rationed by “first-come-first-serve” queue discipline rather than by pricing in any form (Strandenes 2004). However, congestion pricing has been implemented at some ports. Marine terminal operators at the Ports of Los Angeles and Long Beach implemented it as part of the PierPASS program in July 2005. 32 Shipping lines also charge shippers higher rates at times of peak demand. 33 For example, in March 2007, major shipping lines sending containers to the Port of Vancouver introduced ‘congestion surcharges’ that varied with the size of container, whether it was transported by truck or rail within Canada, and its destination (Stueck 2007).

As far as seaport charges generally, many Canadian fees are levied on a service or per-use basis 34 whereas U.S. fees are typically one-time or annual fees (HLB Decision Economics Inc 2001). In this respect, practice is closer to marginal social cost pricing in Canada than in the U.S.

Airports: Airport congestion pricing has been advocated at least since Levine (1969). Four airports in the U.S. have slot controls, but no airport in North America has yet implemented congestion fees. Nevertheless, airport congestion pricing is an area of active research 35 and the idea seems to be catching on.

Highways: While numerous toll roads are operating around the world, only a handful of tolled facilities in North America employ tolls for the purpose of congestion pricing. The tolls on Highway 407 do vary slightly by time of day, but the highway is privately operated and congestion pricing is not explicitly mandated in the concession agreement. Nevertheless, despite longstanding opposition to road pricing it is slowly gaining credence and numerous toll road projects are being studied or underway. British Columbia currently plans to toll two bridges: the existing Port Mann Bridge after it is twinned, and the new Golden Ears Bridge.

---

32 See http://www.pierpass.org/. The goal of the PierPASS program is to reduce congestion and improve local air quality by shifting deliveries to off-peak hours. A “Traffic Mitigation Fee” is charged for most cargo movements during peak hours. Cargo owners have responded by shifting cargo pickup and delivery to nights and weekends.
33 I am grateful to Trevor Heaver for making this point.
34 Depending on the type of charge it can be levied per tonne, per square meter, per day, per hour, per metre (of dock) per day, etc.
35 See, for example, Brueckner (2002), Basso (2005), Zhang and Zhang (2006) and Basso and Zhang (2007).
which is scheduled to open in 2009. There is some support to extend tolling to all bridges in the Vancouver region, and road pricing in some form is also being put forward for Toronto and Montréal.\textsuperscript{36} There appears to be limited scope to alter truck movements via pricing (Holguín-Veras 2005). But congestion pricing can still benefit truckers by reducing non-commercial vehicle traffic.

\textsuperscript{36} The prospects of toll roads in Canada are assessed in Lindsey (2007).