Confidential

Pricing Congestion in Sydney

Discussion Paper

Prepared for Infrastructure New South Wales

April 2012

smart infrastructure facility
UNIVERSITY OF WOLLONGONG

ACIL Tasman
Economics Policy Strategy

ACIL Tasman Pty Ltd

ABN 68 102 652 148
Internet www.aciltasman.com.au

Melbourne (Head Office)

Level 4, 114 William Street
Melbourne VIC 3000

Telephone (+61 3) 9604 4400
Facsimile (+61 3) 9604 4455
Email melbourne@aciltasman.com.au

Brisbane

Level 15, 127 Creek Street
Brisbane QLD 4000
GPO Box 32
Brisbane QLD 4001

Telephone (+61 7) 3009 8700
Facsimile (+61 7) 3009 8799
Email brisbane@aciltasman.com.au

Canberra

Level 1, 33 Ainslie Place
Canberra City ACT 2600
GPO Box 1322
Canberra ACT 2601

Telephone (+61 2) 6103 8200
Facsimile (+61 2) 6103 8233
Email canberra@aciltasman.com.au

Perth

Centa Building C2, 118 Railway Street
West Perth WA 6005

Telephone (+61 8) 9449 9600
Facsimile (+61 8) 9322 3955
Email perth@aciltasman.com.au

Sydney

PO Box 1554
Double Bay NSW 1360

Telephone (+61 2) 9389 7842
Facsimile (+61 2) 8080 8142
Email sydney@aciltasman.com.au

For information on this report

Please contact:

Ken Willett, ACIL Tasman
Telephone (07) 30098702
Mobile 0412045455
Email k.willett@aciltasman.com.au

Contributing team members

Henry Ergas, SMART
Infrastructure Facility
University of Wollongong

David Greig, ACIL
Tasman

Contents

List of boxes

No table of figures entries found.

Executive Summary

Sydney's Congestion is Economically Costly

Traffic congestion was worsened relentlessly in Sydney for decades. Sydney's congestion problem is now serious and very costly.

Congestion costs include loss of time, schedule disruptions and unpredictability, and higher stress, fuel-use, vehicle emissions, and risk of crashes. However, congestion-alleviation also entails costs. Therefore, some congestion is economically justifiable, but current levels are not.

The Bureau of Infrastructure Transport and Regional Economics prepared order of magnitude estimates of the economically unjustifiable component of costs of congestion in Sydney. This is the deadweight loss or economic inefficiency element. The estimates included $2.4 billion in 1995, $3.5 billion in 2005, and $5.4 billion in 2012.

Nature and Causes of Congestion

Traffic in Sydney has marked peaks linked to travel to and from work and education sites. Traffic has grown with population, economic activity, incomes, falling real vehicle prices, and rising vehicle quality. Because provision of arterial road capacity has not kept pace with peak demand growth, excess demand for arterial road space has resulted.

Excess demand is usually short-lived when goods and services are supplied and accessed via the market system, because prices rise, inducing adjustments to quantity demanded, thereby rationing scarce supply in the short-term, and encouraging supply of additional quantity over time. However, the price mechanism has been used on only a few arterial roads in Sydney, leaving most of the arterial network unpriced. Moreover, the limited pricing has not been applied in ways that would ration scarce supply of road space at peak times and locations. In the context of growing demand for arterial road space and the absence of rationing of scarce road space through pricing, excess demand has increased relentlessly.

Queuing has rationed scarce road space. This queuing is better known as traffic congestion. Queues have spread as they have lengthened and disrupted traffic flows elsewhere on the arterial network. Queues have persisted longer as road-users have adjusted to lengthening and spreading queues and growing traffic volumes. This has meant broadening peak periods.

Not pricing road networks makes economic sense when road-users can be added without impeding others. It is economically inappropriate when road-users impose costs on others with impunity, as occurs in congested conditions.

While road-users adjust to costs of congestion that they experience individually, they do not adjust their behaviour in response to costs they impose upon others. The congestion costs that an additional road-user imposes on others on a congested network (marginal external costs of congestion) are large relative to the congestion costs borne by the extra road-user itself, because extra costs are imposed on all of the very large number of users of the network. The result is over-use of the arterial road network at peak times and locations and too much congestion from a social perspective. Put another way, resources are used inefficiently.

This is typically regarded as a form of market failure. However, it could reasonably be referred to as policy failure, because governments control the road network, changes to it, and pricing (or non-pricing) of access to parts of it. Typically, governments have not established or allowed development of a market for road space. They have also chosen not to simulate the efficient operation of a market by charging a road-use a price equivalent to marginal external costs of congestion (congestion pricing), which together with costs borne by each road-user would mean the effective access price paid by each road-user equalled the marginal social cost of road-use.

Past Anti-Congestion Policy Measures

Almost everywhere, not just in New South Wales, governments have persisted in excluding congestion pricing from their changing mixes of anti-congestion measures, despite increasing urging from economists over the past 60 years to apply this policy instrument. These changing policy mixes have typically failed to stop congestion from worsening in medium-sized and large cities around the world. So, failed anti-congestion strategies are the norm.

Until the late-1970s, governments typically saw road building as the solution to congestion. However, high costs and assumed futility because of traffic attraction by new capacity ("induced traffic" or "induced demand") led to changes to anti-congestion strategies. Governments increasingly switched resources from roads to public transport, cycling and walking facilities, and operating subsidies for public transport. In some cases, governments re-allocated some pre-existing road lanes from general purpose use to access by buses and other multi-passenger vehicles, which effectively added to public transport subsidies. Many governments buttressed these policy changes with measures such as higher on-street and off-street parking charges, information programs regarding public transport services, and promotion of car-pooling arrangements.

All of these policy instruments were meant to reduce demand for road space and increase demand for alternatives to road-use by single occupancy vehicles. Transport planners typically described some or all of these policy instruments as "demand management measures". They considered them to be substitutes for congestion pricing.

These "demand management measures" failed to stop the inexorable worsening of congestion, even though the major measure, subsidised public transport, involved 100 per cent capital subsidies and operating subsidies in excess of 75 per cent of operating costs. Indeed, costs of all of these "demand management measures" have been found to be high relative to numbers of passengers attracted from single-occupancy vehicles. An important oversight by proponents of these measures is that they are just as likely as increases in road capacity to be undermined by "induced traffic". Another neglected problem is that public transport subsidies have facilitated inefficient operating arrangements (x-inefficiency).

Public transport, cycling and walking have often been described as "sustainable transport", because use of these modes by commuters reduces congestion and emissions caused by cars. However, the fiscal unsustainability of an ineffective strategy of trying to reduce congestion to acceptable levels through heavy subsidies has been overlooked.

Many governments also took steps to change urban land regulation policies to try to increase urban densities, at least in and around major activity centres and major public transport hubs and routes. They hoped that this would encourage greater use of transport modes other than cars, and improve the viability of public transport. However, these actions have not reduced congestion and may have increased it. They have overridden consumer preferences and distorted relative prices of land and capital, inducing substitution of capital for land. The result has been resource misallocation.

Because massive public transport subsidies, other "demand management" policies, and regulated increases in urban density have made little impact on congestion, some governments, notably those in New South Wales, Victoria and Queensland re-considered their policies of restraint on provision of general purpose arterial road capacity, particularly in the case of by-pass or orbital roads. Toll roads (typically involving public private partnerships) were often preferred to provision of free-access arterials, because of the high costs of urban arterial road provision in the context of fiscal stress associated with high costs of maintaining public transport subsidies.

While governments have claimed that toll roads would help alleviate congestion, these roads typically have been priced simply to recover full costs (including a reasonable rate of return on capital). Such pricing is incompatible with congestion-alleviation, because full cost recovery is possible only if tolls are set to toll-off sufficient potential users to ensure a wide difference in quality of service between tolled and free-access facilities. Pricing of new roads to alleviate congestion would require low and possibly negative prices.

Economically Sensible Anti-Congestion Policy Reform

Current anti-congestion policy-mixes for Sydney, as for most other major metropolitan areas, are ineffective and economically inefficient. Economically sensible reform would substitute congestion pricing for heavy public transport subsidies, parking levies/supply restraints, and tolling of new roads.

Ideally, prices under a congestion pricing regime would reflect marginal external costs of congestion -- the difference between congestion costs caused and borne by each road-user. Prices would vary over time, across the network, and between vehicle-types. Prices would be highest in the busiest periods and locations, and for the largest vehicles. Zero prices would apply in free-flow conditions.

This "internalisation" of marginal external costs of congestion would induce changes to travel modes, routes and times, reducing traffic at peak times and locations. Delays, stress, fuel and emissions would be cut and transport facilities would be better utilised.

Congestion pricing would ensure "induced traffic" effects did not undermine benefits of new road, public transport, cycling and walking facilities, and information programs on urban transport options. Therefore, it would increase benefits from these initiatives. Meanwhile, these infrastructure and "demand management measures" would help pricing to induce changes in peak-period travel behaviour.

Congestion pricing is primarily a policy instrument for alleviation of congestion in an efficient way. It is very different concept to applying tolls to new roads to recover their full costs or to existing roads to raise money for further investments in urban transport infrastructure or some other purpose. Unlike cost-recovery tolling of new road segments in dispersed locations, or tolling of existing roads to raise money, congestion pricing would improve efficiency of use of metropolitan road and public transport networks.

Of course, congestion pricing yields revenue as a by-product of its primary function. Moreover, there is reasonable evidence to suggest that under plausible assumptions, a well-designed congestion and road damage pricing system could provide enough revenue to cover full costs of providing and maintaining a metropolitan urban arterial road network.

Parking levies and supply restrictions have sometimes been proposed as a simplified form of congestion pricing. However, these measures would not address the contribution to congestion of through-traffic, commercial vehicles, and the length, route and timing of trips. In contrast, a well-designed congestion pricing system would do so.

Pricing of crowded roads would improve bus fuel economy, trip times, and service reliability. It would increase demand for bus and rail services, allowing higher service-frequency and route-density, which would attract still more passengers. Induced increases in residential and commercial densities around public transport corridors and destinations would reinforce these trends. A cycle of increasing demand for services and declining unit social costs of public transport-use would occur.

Congestion pricing should be accompanied by a restructuring of public transport fares. Congestion pricing raises effective prices of using single-occupancy vehicles in peak times and locations relative to effective prices of travel at other times and routes, and by other transport modes, including public transport. Therefore, continuation of subsidies to public transport to change relative prices of car and public transport-use would be redundant. Moreover, the reduced cost structure of public transport would have to be factored into fares. They should also be adjusted to manage passenger congestion and allow for broader, flatter peak periods. The various effects of congestion pricing should improve public transport's viability, reducing subsidy requirements.

It is extremely important to note that congestion pricing is an essential element of an economically efficient anti-congestion package for Sydney, but it is not sufficient. It must be complemented by increases in road capacity – particularly debottlenecking and by-pass investments – and increases in public transport capacity, but not public transport subsidies. Capacity increases are required for efficient congestion alleviation beyond the short-term future.

"Simplified" Congestion Pricing

The economic ideal of network-wide time- and location-variable congestion pricing has not yet been adopted anywhere. However, "simplified" schemes have been implemented in Singapore, London and Stockholm.

These "simplified" schemes have been based on charging for crossing a cordon around a city centre (Singapore and Stockholm) within peak periods or being on a public road within a cordon for a period before, during and after normal business hours (London). The London area pricing scheme involves a flat charge in the charging period and a zero charge at other times. The Singapore and Stockholm cordon pricing regimes apply variable charges. In Stockholm, charges vary according to times of crossing the single cordon. In Singapore, charges vary between zones comprising the cordoned area, between roads linking with the cordon zones, and between roads and cordon zones. Congestion charges in Singapore also vary between times of the day, and times of the year. Singapore prices are also adjusted every three months to maintain levels of road service.

The Singapore scheme is supported by dedicated short range communication (dsrc) technology. The London and Stockholm schemes are supported by automatic number plate recognition (anpr) technology. Because these technologies have already been successfully deployed to support cordon and area pricing regimes, it could reasonably be expected that they could be used to support cordon or area pricing in Sydney without difficulty.

Although existing cordon and area pricing regimes fall well-short of the ideal of network-wide variable congestion pricing, the considerable evidence available indicates that they have substantially alleviated congestion and improved the efficiency of resource allocation. Moreover, the anti-congestion and resource allocation performance of each could be improved by fine-tuning price differentiation. In the case of London and to a lesser extent Stockholm, important gains could be made by elimination of price exemptions and discounts.

The London scheme has much higher operating costs relative to revenue than the others, not only because of substantial exemptions and discounts, but also because it has more than 11 times the number of detection points. One reason it has more detection points is that it is an area pricing scheme and therefore presences within the cordon have to be detected, not just movements across the cordon. The major cost factor is that the London scheme has more than 9.5 times the number of boundary detection points. This is a result of differences in natural and built topographies of these cities.

Sydney, like London, lacks topographical features that would allow a relatively small number of detection points for a cordon or cordons. In addition, Sydney is characterised by considerable dispersion of economic activity and therefore employment, combined with a radial public transport system focussed on the central business district. This would not be compatible with congestion pricing based on a single cordon or multiple adjacent, non-concentric cordons. It would require multiple dispersed non-concentric cordons and probably pricing of some links to those cordons. This would substantially raise the cost of the system.

More importantly, dealing with Sydney's characteristics would greatly increase the complexity and difficulty of designing a cordon-based congestion pricing system. This would increase risk of shifting congestion between locations. Considerable quantitative analysis with a sophisticated model specified to include pricing responses would be required, along with considerable additional quantitative and qualitative analysis, to provide a foundation for design of a "simplified" congestion pricing regime for Sydney.

A danger is that placing too much emphasis on "simplicity" might inadvertently lead to complexity and undesirable inefficiency because of congestion shifting issues. These would have to be dealt with via various modifications both initially and as problems become apparent.

Network-Wide Variable Congestion Pricing in Sydney

Because of Sydney's particular circumstances, the *Ministerial Inquiry into Sustainable Transport in New South Wales*, which reported in December 2003, proposed a network-wide variable congestion pricing system (potentially in combination with road damage pricing) for "Greater Sydney" within 5-10 years. It explained that this would allow more efficient finer differentiation of pricing than cordon-based schemes, and could be extended to include pricing of other external costs, particularly road damage by heavy vehicles.

The Ministerial Inquiry envisaged that this pricing regime would be supported by dsrc or anpr technology. However, it acknowledged that a "vast" number of detection points would be required to avoid circumnavigation of these points, and observed that this problem would limit the arterial roads that could be included in the pricing regime until it could be overcome. The Ministerial Inquiry did not suggest how this problem might be resolved.

Two potential solutions would be to work on, or wait for others to undertake the task of adapting global position system technologies to support network-wide variable congestion pricing. The former would involve much higher risk for the NSW Government than the latter. In the meantime, the time consuming and expensive task of upgrading or otherwise providing complements to congestion pricing could be undertaken.

Another option could be to phase-in congestion pricing via variable cordon-based pricing, involving geographically dispersed cordons, plus variable pricing of selected links with cordoned areas. This would be supported by established anpr or dsrc road pricing technologies. The phasing-in would follow provision of infrastructure complements to pricing. Later, coverage would be expanded through application of global position system technology to support network-wide variable congestion pricing.

A precedent for an evolutionary approach to congestion pricing has been provided by Singapore. It has progressively increased the degree of sophistication of its scheme over time. Meanwhile, roads and public transport have been progressively improved, with major upgrades occurring before significant changes to the spatial configuration of the scheme. However, an evolutionary approach based on cordon pricing would be much more difficult in Sydney than Singapore, because of particularly disadvantageous circumstances in Sydney.

Another option with an evolutionary element would be to start with a "no queue" pricing regime involving time-variant prices on specific parts of the road network to eliminate queuing on and around those network locations and segments. Charges would target pre-determined speeds as in Singapore or levels of service as for "dynamically priced" HOT or tolled express lanes in the United States. A case may be made for such an arrangement on the basis of a compromise between economic, computational, and acceptability considerations. The roll-out of "no queue" charging would be complemented by improvements to roads and public transport. Later, when global positioning system technology has been adapted and successfully deployed elsewhere to support network-wide variable congestion pricing, a transition could be made to such a scheme.

Determination of Prices

The task of determining appropriate prices for congestion pricing is very difficult. It is likely to be even more difficult for "simplified" schemes than network-wide schemes, because in the former case, account has to be taken of the effects on unpriced parts of the network outside the scope of the scheme.

These difficulties suggest a "no queue" approach to pricing initially because:

- it is much less computationally demanding
- implementation and refinement can be evolutionary
- it would approach efficient pricing as it is extended across the network
- significant efficiency gains could be achieved even on individual links
- the target of queue elimination is easily defined, visible, verifiable, easily explained, and likely to be readily accepted.

The difficulty of determining prices is not a valid reason to eschew price variability over time and across locations. Variability can still be used to provide incentives to change travel times, routes and modes to alleviate congestion and improve the efficiency of resource allocation.

Existing Tolled Facilities

Implementation of a congestion pricing regime in metropolitan Sydney would be complicated by existing arrangements with private operators of tolled roads. These cover important parts of the road network, particularly the majority of the road segments that Transport for NSW has described as the "110-kilometre Sydney orbital network". Tolling arrangements on these roads are not consistent across tolled facilities. Moreover, the tolls were not designed to alleviate congestion.

To implement congestion pricing in Sydney, it would be necessary to deal with existing arrangements with private operators. An appropriate approach would be to determine the congestion pricing regime as if the tolling arrangements did not exist, and then re-negotiate the agreements before implementation of congestion pricing. Profit-neutral "shadow tolls" could be paid to the private operators by government at rates that took into account the effects of variable congestion prices (including zero prices) borne by users of those road segments.

Acceptability

The major obstacle to congestion pricing around the world has been acceptability concerns. However, there are solid reasons to believe that acceptability issues could be addressed by astutely recycling congestion pricing revenues to fund complementary network improvements satisfying rigorous social-cost benefit analyses.

Infrastructure could be put in place in time for the application of congestion pricing by borrowing funds, with debt servicing capacity supported by resources yielded by congestion pricing, public transport pricing reforms, and consequent smaller subsidy requirements for public transport. This timing would be important for acceptability of congestion pricing.

Road and public transport pricing reforms would underpin the effectiveness, economic efficiency and fiscal sustainability of a package of complementary urban transport infrastructure investments. Meanwhile, these investments would support congestion pricing in efficiently alleviating congestion.

The Commonwealth Government could improve acceptability of congestion pricing by cutting fuel tax to make room for congestion pricing. This would also reduce tax-induced inefficiencies and inefficiencies associated with vertical intergovernmental fiscal imbalance in the Australian federal system of government.

Competent design, implementation and community engagement in respect of such a package could take a few years. Planning should start immediately to so that no time is wasted in deploying an effective and efficient congestion alleviation package for metropolitan Sydney.

1 Introduction

A traffic congestion crisis is looming in metropolitan Sydney. Congestion has been worsening inexorably. It is a matter of great concern to road-users, businesses and government.

Traffic congestion afflicts metropolitan areas around the world. It is widely perceived as a normal, but greatly disliked part of life in medium-sized and large cities. Commuters complain about it in office discussions and social gatherings, newspapers run stories about it, and even the world's most popular singing group referred to the problem in one of its most popular songs (see Box 1).

Box 1 Congestion Spoils a Beautiful Day

"You're out of luck

And the reason you had to care

The traffic is stuck

And you're not moving anywhere."

Source: Paul Hewson (Bono), U2 (2000), *Beautiful Day*.

The ubiquitousness of the problem indicates that failed anti-congestion strategies are the norm. While international fashion in anti-congestion strategies adopted by governments has changed over the past few decades, the trend of worsening congestion has not.

Until the mid- to late-1970s, governments typically saw road building as the solution to congestion. Thereafter, most governments transitioned to the view that such activity is futile and too costly. This change in perception was linked to a new orthodoxy that "you can't build your way out of congestion". It was based on a theory that suggested more road capacity attracted sufficient additional road usage to cancel out the congestion-alleviating effect of costly road capacity increases.

In this context, governments increasingly switched resources from roads to provision of public transport, cycling and walking facilities, and operating subsidies for public transport. In some cases, governments re-allocated some pre-existing road lanes from general purpose use to access by buses and other multi-passenger vehicles. Many governments buttressed these policy changes with measures such as higher on-street and off-street parking charges, information programs regarding public transport services, and promotion of car-pooling arrangements. All of these policy instruments were meant to reduce demand for road space and increase demand for alternatives to single occupancy vehicles. Transport planners typically described some or all of these policy instruments as "demand management" measures.

Many governments also took steps to change urban land regulation policies to try to increase urban densities, at least in and around major activity centres and major public transport hubs and routes.

They hoped that this would encourage greater use of transport modes other than cars, and improve the viability of public transport.

However, massive public transport subsidies, other "demand management" policies and land regulation changes to increase urban density have made little impact on congestion. Indeed, congestion continues to worsen in most of the world's major cities.

Consequently, some governments, notably those in New South Wales, Victoria and Queensland re-considered their policies of constraint on the amount of arterial road space available for general purpose use. In particular, doubts arose regarding the appropriateness of not providing by-pass or ring roads. However, the new preferred method of funding provision of new arterial capacity to alleviate congestion was tolling, often involving public private partnerships. While governments claimed that toll roads would help alleviate congestion, in most cases, these roads were priced simply to recover costs including a reasonable rate of return on capital, rather than to facilitate congestion alleviation.

With very few exceptions, governments have eschewed pricing congested parts of an urban network (congestion pricing) to alleviate congestion. The notable exceptions have been government action to implement highly simplified versions of congestion pricing in Singapore (1975), London (2003) and Stockholm (2007).

In contrast to the policy stances of most governments, there has been strong support for congestion pricing among economists for more than 50 years, particularly over the past 10-15 years. During the latter period, economists have been joined by increasing numbers of transport engineers and urban planners.

This advocacy has not swayed most politicians, who perceive congestion pricing to be political poison. Typically, politicians have relegated congestion pricing to the "too hard basket".

For example, a previous New South Wales Government shelved a proposal by the *Ministerial Inquiry into Sustainable Transport in New South Wales* headed by Tom Parry (2003) to apply a network-wide, variable congestion pricing regime (potentially in combination with road damage pricing) in "Greater Sydney"[1] within 5-10 years, to be accompanied by rationalisation of existing taxation of motorists. In December 2006, a Council of Australian Governments Competition and Regulation Working Group (2006) observed that congestion pricing stood out as the most effective option for alleviating congestion and improving the efficiency and productivity of the transport network, at least when delivered as part of a complete package of complementary measures. In 2009, the Victorian Government ceased investigation of congestion pricing in Melbourne. This investigation had been prompted by a report by the Victorian Competition and Efficiency Commission (2006), which favoured congestion pricing and recommended further investigation of its application in Melbourne. In 2010, the Queensland Government halted a departmental investigation of congestion pricing in Brisbane after its existence became public knowledge.

Overseas examples included:

---

[1] "Greater Sydney" was defined by the Ministerial Inquiry to include Sydney, Wollongong, the Blue Mountains, the Central Coast, Newcastle, and parts of the Hunter, Southern Highlands and Shoalhaven regions.

- abandonment of proposals to apply congestion pricing in Edinburgh and Manchester following referenda in 2005 and 2008, respectively
- the political scuttling in 2008 of congestion pricing proposals for New York
- inaction in the United Kingdom following an on line petition against congestion pricing in 2007, despite strong support for congestion pricing from a major transport study released in December 2006 (Eddington, others, 2006a, 2006b).

The authors of the United Kingdom transport study had argued that a well-designed, network-wide, differential, congestion-targeted road pricing scheme was the “stand-out” strategy for alleviating congestion. They envisaged that congestion pricing would be part of a “sophisticated policy mix”, involving environmental pricing, proper pricing across all modes of transport, and sustained targeted infrastructure investment. They said such a regime had the potential to deliver substantial benefits for the economy, road users, taxpayers, and the environment.

The Henry Tax Review recommended that governments in Australia should analyse the application of variable congestion pricing to the road network (Henry, others, 2009, Part 2, p. 377):

> “Recommendation 61: Governments should analyse the potential network-wide benefits and costs of introducing variable congestion pricing on existing tolled roads (or lanes), and consider extending existing technology across heavily congested parts of the network. Beyond that, new technologies may further enable wider application of road pricing if proven cost-effective. In general, congestion charges should apply to all registered vehicles using congested roads. The use of revenues should be transparent to the community and subject to further institutional reform.”

In a discussion paper on a long-term transport plan for New South Wales, Transport for New South Wales (2012, pp. 57, 91) indicated the Government's willingness to at least consider pricing arterial roads to manage congestion. However, congestion pricing was only a minor topic of discussion in the report, covering about two-thirds of a page in two locations in text spanning 90 pages.

Infrastructure New South Wales commissioned ACIL Tasman and SMART Infrastructure Facility, University of Wollongong to prepare this paper to highlight issues to be considered in deciding whether or not to undertake any further investigation of congestion pricing in Sydney, and issues that would need to be addressed in a detailed investigation. This paper discusses congestion pricing as a stand-alone policy instrument and, more appropriately, as part of a package of complementary anti-congestion measures. The discussion includes comparison of:

- congestion pricing with alternative measures
- alternative forms of congestion pricing
- anti-congestion policy packages with and without congestion pricing.

2 Congestion's Nature, Costs and Causes

Traffic congestion occurs when and where there is excess demand for, or deficient supply of road space. Congestion is queuing for scarce road space. Queuing serves as a rationing mechanism.

Congestion involves costs. These costs include longer travel times, reduced predictability (increased unreliability) of travel times, more trip rescheduling, greater frustration and stress, higher fuel costs, and greater noxious and greenhouse gases emissions than occur in free-flow conditions. Congestion also increases the risk of crashes, but severity of crashes tends to be less in low-speed, stop-start driving conditions.

Government intervention is economically justified if resources are misallocated because of market and policy failures, and the benefits of intervention exceed the costs. The form and extent of intervention chosen from the available options should be those that yield the greatest surplus of benefits over costs.

Each additional entity accessing a congested road network imposes costs on others using the network without paying compensation. These costs vary according to vehicle type, as well as time and location. Such costs are known as external costs in the economics literature. The difference between costs resulting from the presence of an additional entity (social marginal costs) and the costs borne by that entity and each other entity are the marginal external costs of congestion. They are large relative to the congestion costs borne by the entity itself (indicated by average congestion cost), because an additional user of a congested network imposes extra costs on every other user, and in aggregate, the extra costs are large. While the extra cost imposed on each road-user in congested conditions by an additional road-user is small, the number of road-users is high relative to capacity of the network in those circumstances. Therefore, aggregate effects caused by an additional road-user are large.

When each entity makes decisions on road-use, it considers only costs and benefits that it experiences. Each entity ignores the additional congestion costs it imposes on others. Because governments typically have not required road-users to confront the costs they impose on others, the urban road system is over-used at peak times and locations. There is too much congestion from a social perspective. Consequently, resources are used inefficiently.

This is typically regarded as a form of market failure. However, it could reasonably be referred to as policy failure, because governments control the road network, changes to it, and pricing (or non-pricing) of access to parts of it. Typically, governments have not established or allowed development of a market for road space. They have also chosen not to simulate the efficient operation of a market by charging a road-use a price equivalent to marginal external costs of congestion, which together with costs borne by each road-user would mean the effective access price paid by each road-user equalled the marginal social cost of road-use.

Intervention to correct this market failure involves costs as well as benefits. Action that reduced congestion so much that marginal costs of intervention exceeded marginal benefits would have gone too far.

It should not be presumed that the existence of congestion and associated costs warrants government intervention to eliminate congestion. That would be warranted only in circumstances in which the benefits exceed the costs. The benefits of moving to zero congestion at particular locations and in specific periods would have to be weighed against the costs of excluding sufficient road-users (by pricing or regulation) or adding enough capacity. Elimination of congestion is most likely to be consistent with an efficient allocation of resources when congestion is attributable to a bottleneck or choke point. In other circumstances, some amount of congestion at various locations at peak times typically would be efficient.

David Cosgrove and David Gargett (2007, p. 107) of Bureau of Transport and Regional Economics prepared "order of magnitude" estimates of the costs of congestion in excess of those that would be consistent with an efficient allocation of resources in each of Australia's capital cities. These estimates indicated the economic inefficiency costs or deadweight losses to Australian society for congestion. They estimated that deadweight losses from congestion in Australian capital cities totalled $9.4 billion in 2005, comprising $3.5 billion in trip delay and variability costs for private road-users, $3.6 billion in trip delay and variability costs for business road-users, $1.2 billion of extra vehicle operating costs, and $1.1 billion of extra pollution costs. Sydney's share of the Australian total deadweight losses from congestion was estimated to be about $3.5 billion.

Cosgrove and Gargett (2007, p. 106) estimated that deadweight losses from congestion in Australian capital cities represented about 50 per cent (range, 30 to 55 per cent) of total congestion costs. External costs of congestion were about 70 per cent (range, 60 to 75 per cent) of total congestion costs.

Cosgrove and Gargett (2007) also forecast growth of deadweight losses from congestion out to 2020. They estimated deadweight losses from congestion in Sydney of $5.4 billion in 2012 and $7.8 billion in 2020. This was based on "business as usual" forecasts of transport sector activity, and on an assumption of no improvement in congestion management, including a continuation of growth of supply or road space that is significantly below demand growth.

Traditionally, governments have taken responsibility for provision, maintenance and operation of the road network, and have allowed free access to roads funded by various taxes. The economic rationale is that an un-congested road displays a characteristic of a pure public good, “non-rival consumption”. This means allowing an extra user does not detract from benefits enjoyed by others or involve other costs. Consequently, the social cost of adding each extra road-user (marginal social cost) is approximately zero, while each additional user collects benefits (marginal social benefit is positive).

Roads do not have the other characteristic of a pure public good, “non-excludability”. It is feasible to use prices to exclude some potential users of a road, although a system to exclude non-paying cars involves costs. However, excluding an extra user through pricing in excess of marginal social cost, which approximates to zero, would be economically inefficient, because the benefits gained by the willing user would exceed social marginal cost of allowing access.

However, the case for not charging for road-use is undermined when excess demand emerges at zero price. Then, external costs of congestion arise and roads become more like private goods and less like public goods, with access by one detracting from the benefits enjoyed by others.

Similarly, the case for not charging for road-use is weakened when road-users impose costs on other vehicle owners and taxpayers by causing pavement damage (another type of external costs) or on taxpayers through government decisions to construct stronger pavements to limit damage. Heavy vehicles are the major cause of road damage. The extent of damage caused rises exponentially with weight per axle.

Excess demand can grow at any price if demand growth outstrips supply of additional road space and price adjustments are not allowed. Excess demand can decline (eventually turning into excess supply) if supply of road space grows faster than demand. Growing excess demand has been the norm in Sydney and other large and medium-sized cities around the world. Therefore, worsening traffic congestion has been a global phenomenon for many decades.

Growth of demand for road space has been influenced by several factors. Population growth is just one relevant factor.

Falling real car prices, declining fuel consumption, and improving comfort, quality, and reliability of cars have added to road usage. Consumers have shown strong preferences for the comfort, time-saving, convenience, flexibility, reliability, privacy, and freedom from harassment attributes of single-occupant vehicles over buses, trains and other vehicle-sharing. Consequently, demand for cars and hence road-space has risen with incomes and value of time.

Cars often offer a substantial time saving advantage over public transport because of the fixed time penalty of the order of 15-20 minutes associated with public transport, regardless of the distance travelled. Passengers must move to a pick-up point, wait for a bus or train, which may not run to the timetable, and then walk from the drop-off point to the destination (Car drivers may also have to walk from parking stations to their workplaces). In addition, public transport services have multiple pick-up and drop-off points, which increase the time penalty associated with public transport. Even "express services" incur a time penalty because of this, albeit not as large as for all-stops services. Congestion has reduced some of the time saving advantage of cars over trains, but buses are affected by congestion, like cars, except to the extent that they operate on dedicated lanes or busways.

Heavy subsidies for public transport have resulted in lower fares and more services for public transport. This would have reduced demand for road space for cars.

In Australia, despite heavy subsidies, public transport’s share of passenger transport in metropolitan areas remained relatively static from 1977 to 2005, following a strong declining trend over the previous 30 years. In Sydney, public transport’s market share remained around 13.5-15.5 per cent, substantially higher than in other Australian metropolitan areas (Cosgrove, Gargett, 2007, pp. 24-39). After 2005, public transport's market share rose significantly in Melbourne, Brisbane and Perth, but not in Sydney and other Australian metropolitan areas. Annual public transport trips per capita rose in the former group of cities from the mid-1990s, but not the latter group, including Sydney, where public transport trips per capita continued falling (Cosgrove, 2011, p. 7).

The reasons for these changes, including differences between metropolitan areas, are not clear. Possible explanations could include:

- a strong rise in automotive fuel prices from 2005 that reached an historical high in 2008, and recovered from a slump in 2009 to current high levels
- a precautionary increase in saving following the 2008 global financial crisis and subsequent persistence of considerable economic uncertainty
- differences in centralisation or decentralisation of employment opportunities in metropolitan areas, associated with variations in congestion patterns.

As a result of preference and service value considerations, historically, car numbers and car travel have increased faster than population, and per capita travel has increased faster than per capita income. However, the differential rates of growth of car numbers and car travel slowed as ownership levels edged towards saturation (Cosgrove, Gargett, 2007, pp. 24, 28-29). Since 2008, it appears that kilometres travelled per capita in Sydney have fallen slightly. In some other Australian metropolitan areas, the decline has been more significant. This appears to be related to the influence of fuel prices and precautionary saving in the context of the global financial crisis and subsequent persistence of economic uncertainty (Gargett, 2012, pp. 44-46, 71).

Road freight vehicle activity has added to demand for urban road space. Such activity has been rising faster than population, car use and gross domestic product (Cosgrove, Gargett, 2007). This appears to be attributable to the growing complexity of metropolitan economies and increasing integration of regional, national and international economies (Downs, 2004, p. 38).

Rising population, increasing car-based mobility, preferences for detached houses on large land lots, unpriced roads, and land-use regulations have yielded sprawling, low-density suburbia. The same factors have induced decentralisation of employment and activity centres. Low-density residential and activity areas have, in turn, induced more car-use. These interdependent occurrences have tended to undermine the viability of public transport. This has meant poor service availability, which has reinforced the trend to car-use, particularly for cross-town trips. But, low-density, decentralised land-uses have tended to disperse car-use, meaning lower overall vehicle density. While greater car-use and less public transport patronage have added to congestion, lower vehicle density has worked the other way.

The decentralisation of employment and business centres potentially could reduce pressure on roads in and around the central business district (cbd) and other dense activity centres. However, this does not apply to the extent that cars and trucks have to use roads serving dense commercial/employment centres, like the central business district, to access decentralised centres.

The demand for road space is heavily concentrated in early morning and late afternoon peak-periods, and on routes servicing major activity centres, because of the desire of business enterprises, governments and educational institutions to have participants together during much the same hours of the day to take advantage of perceived benefits of interaction and supervision on a face to face basis. Many arterial roads that are heavily congested in peak periods, have substantial excess capacity at other times. However, peaks have been spreading.

The supply of arterial road space has lagged growth in demand in Sydney for decades. There has also been a trend to provision of tolled arterial road space under public private partnership arrangements.

The potential benefits of this new capacity in alleviation of congestion have been undermined by tolling-off effects in the context of zero pricing of other parts of the road network (apart from the Sydney Harbour Bridge).

The arterial road supply deficit appears to have been linked with an ideological preference among transport and urban planners for government support for public transport over roads, the rising fiscal burden of public transport subsidies, high costs of providing new arterial road capacity in metropolitan areas, allocation of low priority to infrastructure investment, the Commonwealth Government's view that traffic congestion is the responsibility of state governments, and political aversion to pricing of use of pre-existing components of the metropolitan arterial road network.

Indirect road-use charges, such as fuel taxes and vehicle registration fees, do not sufficiently ration access to crowded roads to reduce congestion to anything close to the socially optimal level. There is only a vague, remote link between fuel purchase decisions and choices between driving times and routes and travel modes. Other vehicle taxes are independent of time and location of vehicle-use.

Therefore, growth of demand for arterial road space has outstripped growth of supply at close to a nominal zero price in the untolled road network. Therefore, scarce road space has been rationed by queues that have grown in number and length and persisted for growing periods of time. These circumstances have been aptly described by Nobel Laureate in Economics, Gary Becker, and seminal contributor to the economics literature on congestion pricing, Sir Alan Walters, as quoted in Box 2.

Box 2 **Sir Alan Walters and Nobel Laureate Gary Becker on Causes of Congestion**

Alan Walters:

"While it may be efficient to have some congestion on the highway, the fact that the motorist is not required to pay for the congestion he causes will induce too many motorists to use the road and there will be too much congestion. These conditions are probably typical of large conurbations throughout the world. Rarely do user charges reflect adequately the congestion in large cities – traffic jams and snail like speeds are the consequence. These are the wastes of user charges that are too low."

Gary Becker:

"An iron law of economics states that demand always expands beyond supply of free goods to cause congestion and queues. Drivers caught in traffic jams on freeways in and around major cities of the world regularly run afoul of this law"

Source: Walters (1968), pp. 11-12; Becker (1998), p. 26.

3 Concept of Congestion Pricing

3.1 Origins of the Concept

The origins of the congestion pricing concept can be traced back to Arthur Pigou's (1920) classic, *The Economics of Welfare*. Pigou proposed that charges be applied to internalise external costs in general, and he specifically suggested such charges to deal with congested roads.

Pigou's general proposal to levy charges to internalise external costs triggered much interest among economists. However, Pigou's discussion of congestion pricing received little attention for more than 30 years, apart from Frank Knight's (1924) critique and extension of it. In response to Knight's critique, Pigou deleted discussion of congestion pricing from the second and later editions of *The Economics of Welfare*.

In the early 1950s, interest in the concept of congestion pricing was revived, with papers by James Buchanan (1952), Alan Walters (1954) and William Vickrey (1955). Milton Friedman and Daniel Boorstin (1952) also wrote about congestion pricing, but their piece was not published until 1996.

From the late 1950s, congestion pricing became a focus of serious investigation, triggered by major seminal contributions from Vickrey (1959, 1963, 1969), Walters (1961) and Mohring and Harwitz (1962). Indeed, in the early 1960s, the United Kingdom Government commissioned a study (Smeed, Walters, Roth and others 1964) on the topic. Later, Alan Walters (1969) and Gabriel Roth (1967) published separate large monographs on congestion pricing.

Alan Walters (1961) and William Vickrey (1963) developed a basic model of traffic congestion based on the relationship between speeds and flows established by traffic engineers. However, this static model did not allow for re-scheduling of trips between peak and shoulder periods, the common phenomenon of congestion associated with bottlenecks or choke points, and the appearance and disappearance of hypercongestion (traffic flows as well as speeds decline as more cars are added) linked to the speed-flow relationship and bottlenecks. To analyse these issues, Vickrey (1969) formulated a dynamic model, which became known as the bottleneck model. Later, this model was developed by Richard Arnott, Andre de Palma and Robin Lindsey (1990, 1993, 1994, 1998). These models provide insights in respect of different types of congestion and the effects of congestion pricing.

Economists have persistently argued that congestion pricing is the key to efficient alleviation of congestion, but policy makers have been reluctant to apply it. Historically, the main obstacles have been concerns about practicality, fairness, and political acceptability.

However, technological advances have greatly diminished practicality issues to such an extent that reasonably sophisticated congestion charging regimes have been established in Singapore, London and Stockholm. While these schemes fall well short of the ideal, ongoing improvements in technology would allow implementation of more sophisticated regimes in future. A brief outline of available technologies has been provided in Appendix A.

Now, the main impediment to application of some form of congestion pricing in other large cities is concern about re-distributional effects. Recently, however, attitudes have been changing because of poor performance of alternatives to congestion pricing, and the widely acknowledged success of simplified congestion pricing and complementary measures in Singapore, London, and Stockholm in reducing congestion.

Substantial literature has accumulated on the idealised concept of congestion pricing and simplified versions of it. For overviews, see Button (2004), Arnott, Rave and Schöb (2005), Lindsey (2006), Small and Verhoef (2007), Parry (2009), and Fosgerau and Van Dender (2010).

3.2 Basic Theoretical Principles

The idea underlying congestion pricing is to price the costs that road-users impose on others without payment of compensation (external costs) when they use the road network at busy locations and times. These external costs include time delays, schedule delays, greater trip time variability, reduced productivity, extra fuel, effects of more emissions, and possibly additional crash costs.

Ideally, prices would reflect marginal external congestion costs (differences between short-run marginal social cost and marginal private cost), and apply on a network-wide, variable basis. They would be zero in free-flow conditions and increase with the degree of congestion, which varies with time, location, and various events, such as sporting activities, bad weather and crashes. Prices would also vary with the footprint and acceleration characteristics of vehicles. Differential lane pricing could be incorporated to allow for heterogeneity in values of time. Such a differential congestion pricing regime could be extended to price marginal external costs of road damage by also varying prices having regard to road-type and axle weight of vehicles.

Ideal congestion prices would make road-users confront congestion costs they impose on others. By internalising these costs, road-users would be induced to change their travel behaviour. In the short-term, changes in behaviour could include alterations to travel times, routes and modes. It could also include alterations to the amount of travel undertaken. For example, it could induce more frequent working at home and videoconferences instead of travel for face to face meetings. In the longer term, congestion pricing could induce changes to workplace and residential locations.

Congestion pricing is an application of the concept of pricing in accordance with short-run social marginal cost to achieve an efficient allocation of resources. This was noted, by Sir Alan Walters, a seminal contributor to the economic literature on congestion pricing (see Box 3).

Box 3 **Social Marginal Cost Pricing of Roads**

"The theory of marginal cost pricing suggests that taxes (congestion charges) be levied to reduce demand until traffic flow is at a level where private unit cost (tax) is equal to marginal social cost."

Source: Walters (1961), p. 680.

Ideal congestion pricing would be accompanied in the short-term by a restructuring of public transport fares to a social marginal cost basis. This would mean roads and public transport are priced on a consistent basis. At present, public transport fares are priced below social marginal cost in an attempt to alleviate congestion by changing relative prices of road-use and public transport to reduce demand for road space. Such pricing would not be appropriate in the context of congestion pricing, which increases relative prices of road-use.

Ideal congestion and public transport pricing would reduce traffic congestion and public transport passenger congestion to the optimal level (typically not zero) in the short-term with existing infrastructure (in the absence of countervailing market and policy failures). Reducing congestion to the optimal level in the long-term requires ideal pricing plus efficient provision of urban transport infrastructure (roads and public transport). Both too much congestion (under-pricing or under-investment) and too little congestion (overpricing or overinvestment) mean resources are allocated inefficiently.

The purpose and structure of congestion prices differ greatly from those of typical tolls. The basic purpose of congestion pricing is to alleviate congestion, and ideally, reduce it to the optimal level. The ideal structure would involve prices that vary with congestion across the urban road network and over time. In contrast, conventional tolls are applied to discrete facilities, not the network, to pay for their provision. Typically, they do not vary with congestion or any determinant of it, such as time of day.

The ideal form of congestion pricing has not yet been applied anywhere. Only highly simplified versions have been implemented, and in very few cities. Existing simplified schemes have been discussed in sub-section 3.4.

3.3 Interacting Market and Policy Imperfections

Application of congestion pricing to address inefficiencies in resource allocation associated with too much congestion must be considered in the context of the many other causes of misallocation of resources in the New South Wales and Australian economies. Some of these sources of resource misallocation interact with others. In addition, policy instruments targeting specific inefficiencies may also affect the extent of other inefficiencies. The importance of interactions such as these was originally pointed out in a more general economic policy context by the first Nobel Laureate in economics, Jan Tinbergen (1952), and independently by Bent Hansen (1955).

There is widespread support for the view that spatial agglomeration of economic activity provides benefits to producers and consumers through non-market interaction for which no compensation is paid. These external benefits are believed to be a dominant cause of agglomeration and its productivity effects. Because these benefits are unpriced, they are undersupplied in an otherwise efficient economy. However, agglomeration is accompanied by the inefficiency of congestion as people travel to interact. If congestion is priced to internalise external costs of congestion, the interaction declines and therefore external benefits of interaction are reduced (Arnott, 2007, 2005; Fosgerau, Van Dender, 2010).

Richard Arnott (2007) argued that if government is constrained from subsidising agglomeration, but not constrained from applying congestion pricing, a "second-best" policy would be appropriate. This would involve congestion prices substantially lower than marginal external costs of congestion.

Mogens Fosgerau and Kurt Van Dender (2010) agreed with Arnott that external benefits of agglomeration are not well understood, but argued that agglomeration policies could be separated from anti-congestion policy. However, they noted that it was valid to warn that effects of productivity of a city should be considered when congestion pricing was being formulated.

Congestion exacerbates external costs of vehicle emissions. This occurs because congestion involves too many car trips from a social perspective, and because fuel consumption is much higher in very low speed and stop-start conditions, than in free-flow conditions. Congestion pricing with complementary measures reduces these emissions along with congestion.

Failure to address external costs of vehicle emissions through pricing increases the extent of congestion, because it means driving is socially excessive. Trying to reduce emissions through regulations that lower fuel consumption is likely to increase congestion because lower fuel consumption reduces costs of driving, encouraging extra car-use (the rebound effect).[2]

Reducing congestion to the optimal level requires emissions pricing, as well as congestion pricing and complementary measures. Similarly, cutting emissions to the optimal level requires congestion pricing with complementary measures, as well as emissions pricing.

Congestion increases the likelihood of crashes, but tends to reduce the severity of the larger number of crashes. Failure to tax vehicles or price vehicle and third party insurance arrangements on a distance-linked basis, means too much driving which adds to congestion. Both congestion and crash pricing are required to reduce congestion and crashes to socially optimal levels.

Congestion exacerbates the tendency of taxes on labour income and fuel tax to discourage labour supply, because congestion reduces nett returns to working.[3] Indeed, an econometric study by Kent Hymel (2009) found that congestion dampens employment growth and this effect increases with congestion. The effect of a 10 per cent increase in congestion in 10 large United States ranged between 2.5 per cent and 4.7 per cent reductions in employment growth.

Because taxes on labour income tend to reduce labour supply, they also tend to lower the level of congestion. Congestion pricing would reduce congestion further. Less congestion would encourage labour supply by reducing costs of working. On the other hand, congestion pricing would tend to discourage labour supply by adding to costs of working. The effects of congestion pricing are expected to outweigh the effects of less congestion on labour supply for lower income earners. In cases of bottleneck congestion or hypercongestion, this would be less common, because there would be a lower

---

[2] For a discussion of the rebound effect, see Hymel, Small and Van Dender (2010).

[3] The nett labour supply effect is result of effects working in opposite directions. While some may leave the labour market or work less because of the reduction of nett returns to working, others may work more to make up for the reduction in their nett returns to work. With fuel tax, there is an additional complication. While fuel tax tends to deter labour supply by increasing the cost of going to work, it tends to increase labour supply to the extent that fuel-use is a complement to leisure.

income threshold below which the effect of congestion pricing outweighs the effect of less congestion (see sub-section 3.6).

However, nett adverse effects of congestion pricing on labour supply could be at least offset by judicious (socially productive) recycling or use of nett revenue (including public transport subsidy savings). This nett revenue could be as large as the welfare gains from correcting external costs of hypercongestion or bottleneck congestion, and several times larger than the smaller welfare gains from addressing external costs of other congestion (see sub-section 3.6). Judicious recycling could include reductions in distortionary taxes or funding of government projects providing comparable efficiency gains (Parry, 2009, pp. 467, 474-475).

For example, modelling by Ian Parry and Antonio Bento (2001) indicated that the nett effect of pricing of congestion and recycling of revenues to reduce taxes on labour would be an increase in labour supply. Welfare gains from this labour supply effect were estimated to be at least as large as gains from reduction of congestion. A caveat is that the analysis of Parry and Bento related to the United States. Marginal effective tax rates on labour income for lower earners are higher in Australia than the United States. Moreover, because of interactions between the welfare and tax systems, there are likely to be significant numbers of secondary income earners close to the extensive margin. This means that when faced with an increase in travel costs, they could leave the labour force. So, with labour supply more distorted to begin with, the resource misallocation costs of distorting it further could be materially greater than those in the United States.

Public transport subsidies tend to encourage labour supply by reducing the cost of going to and from work and education sites. To some extent, this would offset the adverse effects of labour income taxes on labour supply. However, public transport subsidies have to be funded and the marginal cost of public funds is substantial,[4] because of resource allocation inefficiencies (deadweight losses) caused by taxation, which may include adverse labour supply effects (Small, Verhoef, 2007, p. 158). Of course, the cost could be kept down by restricting the subsidies to low income people through a means-tested voucher system.

If congestion pricing revenue was used to finance retention of some public transport subsidies in this form, it would be appropriate to reduce congestion prices, because both measures change relative prices. Then, replacement revenue for lower prices and revenue allocated to subsidies would have to be found at the marginal cost of public funds to finance infrastructure to complement congestion pricing. It would be more direct and efficient to recycle congestion price revenue to address labour

---

[4] Kleven and Kreiner (2003) estimated the marginal cost of public funds in OECD countries. For Australia, they estimated that the marginal cost of public funds would be around 1.28 for a proportional tax reform and about 1.72 for a progressive tax reform. The numbers after the decimal points indicate the deadweight loss associated with raising an extra dollar of revenue. The figures for Australia are similar to those for Canada, significantly higher than the for the United States and Japan, but substantially less than for economies in north-western Europe.

Separate estimates of the marginal cost of public funds in Australia were made for a range of taxes to assist the Henry Tax Review. The estimates included 1.24 for labour income tax, 1.4 for corporate income tax, 1.08 for GST, 1.15 for fuel tax, 1.08 for land tax, 1.34 for conveyancing stamp duties, 1.41 for payroll tax, and 1.7 for royalties and crude oil excise (KPMG Econtech, 2010).

income tax effects directly or to finance infrastructure to complement congestion pricing in alleviating congestion.

Because fuel tax discourages driving to some extent, it tends to reduce external costs of congestion, crashes and emissions. However, except for the case of greenhouse gas emissions, it is a crude device for internalising these external costs, particularly external costs of congestion. Fuel bills do not signal when and where fuel is consumed. Congestion pricing is a far superior anti-congestion device.

Fuel tax applies in full to fuel used in light vehicles providing transport services to business. Therefore, it taxes an intermediate input, which impedes production efficiency. Cutting fuel tax would address that source of production inefficiency. Congestion pricing would further improve production efficiency by internalising external costs of congestion borne and generated by business.

The nett effects on labour supply of using congestion pricing revenue to cut fuel tax are not clear. Congestion pricing would tend to reduce labour supply, at least from lower income groups. Cutting fuel tax would tend to increase it by lowering the cost of going to work and reducing tax on fuel as an intermediate input to production that is complementary to labour. However, cutting fuel tax would reduce labour supply to the extent that fuel-use is a complement to leisure.

A problem with cuts in fuel tax and labour income tax is that these taxes are controlled by the Commonwealth Government, not state governments (and to some extent) local governments, which have responsibility for roads and public transport. A way of dealing with this problem has been discussed in sub-section 6.4.

3.4 Simplified Versions of Congestion Pricing

Various simplified versions of congestion pricing have been discussed below. They include cordon-based pricing regimes in Singapore, London and Stockholm, and some other arrangements that have been described as simplified forms of congestion pricing.

In the case of the London and Stockholm schemes, a considerable amount of public commentary has been focussed on the ratios of costs to revenues. These ratios provide useful crude indications of the technical and administrative efficiency of management arrangements for congestion pricing schemes and the degree of generosity of exemptions and discounts. However, beyond that, their usefulness is doubtful.

The purpose of congestion pricing is to alleviate congestion, not to raise revenue. An assessment should consider benefits from alleviation of congestion, other by-product benefits, and all resource costs, not just those incurred by government. Government revenue is a transfer from a social perspective, not a benefit *per se*. Government revenue from congestion pricing is beneficial to society only to the extent that it displaces revenue involving greater deadweight losses or it funds programs with a benefit/cost ratio sufficiently greater than one to cover deadweight losses associated with poor design of the congestion charging system. If costs to government are to be used as a basis for comparison, they should be weighed against social benefits less other social costs.

3.4.1 Evolution of Singapore's Multi-Segment Cordon Pricing Scheme

Outline of Scheme

Singapore led the world in applying a simplified form of congestion charging in June 1975. A more sophisticated, form of congestion pricing still operates in Singapore, although it is still simpler than the ideal of network-wide variable congestion pricing..

Charges on vehicle usage under Singapore’s evolving congestion charging regime have been accompanied throughout by substantial imposts on vehicle ownership that have also changed over time. In general, imposts on vehicle ownership have declined with increases in the scope and level of congestion charges.

The initial, simple, paper-based, area pricing system was known as the Area Licensing Scheme (ALS). The ALS required the purchase and display of daily paper licences prior to accessing a Restricted Zone (RZ) during morning peak periods. Vehicles could move within or leave the RZ without extra payment. Enforcement officers made visual checks at control points on boundaries of the RZ.

The ALS was modified on several occasions to increase price differentiation or extend its coverage. An important extension was the introduction of a paper-based Road Pricing Scheme on selected expressways in 1994.

In 1998, the paper-based regime was replaced by a more sophisticated congestion pricing system labelled Electronic Road Pricing (ERP). It is a cordon pricing regime that imposes charges when the cordon is crossed at specified times. ERP has evolved into a multi-zone cordon charging system.

Initially, charges were set lower under ERP than ALS, but traffic fell by 15 per cent overall and 16 per cent in the morning peak, because of improved price discrimination. Also, the original ERP involved a RZ in respect of inner-Singapore, plus pricing on selected limited-access expressways and arterials. There were control points on the cordons and relevant major roads.

Charges varied with time. Charges applied only on weekdays from 7.30 am to 7.00 pm for the Restricted Zone, from 7.30 am to 9.30 am on selected express and major arterial roads. Prices increase and decrease in half-hourly steps before and after the peak, respectively. There was no charge at other times on weekdays and at any time on weekends.

Two adjacent pricing cordons in the RZ were created in 2005 to facilitate additional price differentiation. One cordon encompassed an area that is primarily a shopping precinct. The other related to an area dominated by offices. The price differentiation addressed different traffic peaks and patterns for the two precincts, including through-traffic problems in the shopping precinct. Charges for through-traffic via two gantries went up, while charges for vehicles travelling to the shops were reduced, except that pricing was applied to traffic stopping at the shopping precinct on Saturdays.

After August 2005, charges applied from 6.00 pm to 8.00 pm for one of the selected expressways northbound from the city. After congestion reappeared on this expressway, another gantry was erected closer to the city in November 2007. It operates from 5.30 pm to 10.30 pm on weekdays and catches traffic exiting before the gantry that was established in 2005.

In 2008, the office precinct was divided into two zones with a new pricing cordon along the Singapore River to deal with intra-city traffic. Charges applied for crossings in either direction, but only in the evening peak period from 6pm to 8pm, the period in which congestion had become a problem.

Over time, gantry numbers have been expanded. This has occurred on the outer cordon, as well as to segment the Restricted Zone. After the 2008 adjustments to the cordon, only 15 gantries were in place to detect traffic movements across the cordon.

Charges vary with location. They are higher on expressways than for crossing the Restricted Zone cordon.

The charging schedule is reviewed quarterly and for school holidays to achieve targeted speed ranges. These are 45-65 km/hour for expressways and 20-30 km/hour for arterial roads.

Charges vary with vehicle types, based on allocated passenger car unit (pcu) equivalents, which are proxy measures of the dynamic road space occupied by a moving vehicle. Pcu allocations for various vehicle types are: cars, taxis and light goods vehicles, one pcu; heavy goods vehicles, 1.5 pcu; buses, 2 pcu; motorcycles, 0.5 pcu. The only vehicles exempt from charges are emergency vehicles.

ERP uses dedicated short-range communication (dsrc) involving in-vehicle transponders and stored-value smart cards plus gantries with radio antennae for detection and cameras to record vehicles lacking smart cards or sufficient funds.

Proceeds of congestion pricing are not earmarked for complementary transport projects or tax cuts. All transport projects require economic justification before funds are allocated from a central pool.

The government has made considerable effort to publicise the point that the ERP system is a traffic management tool, not a mechanism for collecting revenue. To reinforce this, the government has reduced up-front taxes on vehicles and annual licence fees, whenever major changes have been made to the ERP scheme. When changes to ERP were made in 2008, the cut in vehicle taxes in aggregate dollar terms was about 57 per cent more than the expected additional congestion charge revenue (Chin, 2010, pp. 11, 12).

Also, ERP revisions in 2008 were accompanied by increases in public transport capacity. This was done to provide viable public transport alternatives to those choosing not to drive following changes to ERP.

Outcomes

When the ALS scheme was first applied in Singapore in 1975, it cut traffic to the restricted zone by 73 per cent and average speeds rose from 19 to 36 km/hour. The pricing was based on judgment, rather than any calculation of external costs of congestion. It is widely accepted that pricing was well above the optimal level. One estimate was that the initial pricing was about 50 per cent too high. Despite that, the daily licence fee was raised from S$3 to S$4 after 6 months. However, later adjustments moved the price closer to the estimated optimal level (Santos, Li, Koh, 2004).

When ERP replaced ALS in 1998, charges were lowered, but traffic volumes fell by 15 per cent overall and by 16 per cent in the morning peak-period, because of better price discrimination, including elimination of unlimited entries to the cordoned area for a single payment. The much greater price variability under the ERP also reduced the boundary problem of motorists rushing to beat the start or waiting for the end of pricing for the day. However, it did not eliminate it completely.

Subsequent adjustments to ERP to address re-emergence of congestion in some locations produced positive outcomes. This highlighted the importance of being prepared to modify the design of a cordon pricing regime to correct deficiencies in the system and to address new circumstances.

Introduction of a separate cordon for the shopping precinct and differentiated pricing for that cordon in 2005 cut the proportion of through-traffic from 35 per cent to 20 per cent, with slight increases in trips to the shopping precinct on weekdays and Saturdays and significant increases on Sundays. Traffic on streets in the shopping precinct was reduced by 14 to 36 per cent during priced periods on weekdays and by 19 to 34 per cent on Saturdays. The reduction in traffic in the shopping precinct derived mainly from the cut in the amount of through-traffic (Chin, 2010, pp. 9-10).

In 2005, the erection of a gantry on a major expressway carrying traffic north from the city, and charging via that gantry from 6 pm to 8 pm on weekdays reduced traffic by 25 per cent initially. Introduction of a second gantry closer to the city in November 2007 in response to reappearance of congestion on this expressway, and charging via operation of the gantry from 5.30 pm to 10.30 pm on weekdays reduced traffic by about 20 per cent (Chin, 2010, pp. 8-9).

The division of the business precinct cordon into two cordon segments in 2008, and charging for crossing the cordon between segments in either direction from 6 pm to 8 pm reduced intra-city traffic across this line by 28 to 37 per cent during the evening peak period (Chin, 2010). A subsequent drop in the charge rates because of higher speeds brought down the top of the range of the reduction in intra-city traffic to 30 per cent (Chin, 2010, pp. 10-11).

In Singapore, train services are run by private operators under long-term licensing arrangements without any operating and maintenance subsidy from the government (Menon, 2006, p. 139). The government continues to make unsubsidised improvements to public transport infrastructure and services in a system that is profitable (Santos, Li, Koh, 2004, p. 231). This is quite different to other developed economies, where subsidies are very large. Singapore’s multi-zone cordon pricing regime has contributed to the superior financial performance of Singapore's train system. It has triggered a “virtuous cycle” of higher demand for, and improved economic performance of public transport, and better services (Santos, Li, Koh, 2004, p. 231).

The costs of running the Singapore multi-zone cordon pricing regime are about 21 per cent of revenues. This is slightly lower percentage than for the congestion pricing schemes in Stockholm (22.4 per cent) and much less than for the London scheme (60 per cent).

Charges are not set to reflect marginal external cost. Therefore, they depart from a key principle of optimal congestion pricing. Therefore, the economic efficiency gains from the Singapore multi-zone cordon pricing regime would be lower than could be expected under an ideal congestion pricing scheme.

Although the Singapore cordon pricing scheme appears to have been successful at alleviating congestion, it has not been subjected to a benefit-cost analysis for which results have been published (Santos, 2010; Anas, Lindsey, 2011, p. 73).

3.4.2 Area Pricing – London

Outline of Scheme

The London Congestion Charging Scheme (LCCS) came into effect on 17 February 2003. The LCCS is an area pricing system, applying a flat rate daily charge to vehicles driven or parked on public roads on workdays between 7.00 am and 6.00 pm in central London. Payment of the daily charge provides an entitlement to unlimited travel, to, from and within the charging zone during the charging period.

Cameras were located at 174 access points to the charging zone. Another 24 cameras were located within the charging zone to detect vehicles driven or parked on public roads.

Originally, the daily charge was £5. It was raised to £8 on 4 July 2005 and to £10 on 4 January 2011. If payments were made by an automated payment system, Congestion Charging Auto Pay, which was introduced on 4 January 2011, the daily charge was £9. These charges were not based on estimates of marginal external congestion cost.

On 19 February 2007, the charging zone was extended westward, with the area approximately doubling. The western extension was removed on 4 January 2011. This had been promised during an election campaign by the successor to the Lord Mayor who had championed the LCCS and initiated the western extension of the charging zone.

The original finishing time for charging was 6.30 pm. This was adjusted to 6.00 pm on 19 February 2007.

The congestion charge does not apply on the surrounding boundary route. During the operation of the western extension, the charge continued not to apply to use of the road that formed the western boundary of the original charging zone.

A discount of 90 per cent was granted to vehicles registered to people residing within the charging zone and a few areas outside it, subject to an annual registration fee of £10. Exemptions or discounts of 100 per cent were granted to a wide range of vehicles, including motor cycles, bicycles, buses, London taxis, vehicles used by the disabled, emergency vehicles and alternative fuel vehicles cutting emissions at least 40 per cent below Euro 4 standards.

If the charge is not paid in advance or on the day of travel in the charging zone or from 4 January 2011 through the automated payment system, penalties apply. Vehicle registration numbers are submitted on payment and entered into a database. Video cameras capture images of vehicles entering and within the zone. Automatic number plate recognition (anpr) technology identifies registration numbers with an accuracy rate of 90 per cent. These are electronically compared with the data base of paid and exempt vehicles. If a match is found, images and vehicle details are discarded. If there is no match, the number is referred for manual checking, and if warranted, the issue of penalty notices.

About £50 million of unpaid congestion charge fines has not been paid by embassies and diplomatic staff in London, including representatives of the United States, Russia, and Germany. They have argued the congestion charge is a tax and therefore they are exempted from paying it under the Vienna Convention. Both Transport for London and the U.K. Government have insisted that the congestion charge is a price not a tax.

Application of the LCCS was preceded by extensive public consultation over a period of three years. To pre-empt concern about the charge being regressive and to support switching to public transport, significant expansion of bus capacity was provided by the time the LCCS was implemented, and most nett revenue was earmarked for public transport.

In late-2006, the Lord Mayor of London at that time, Ken Livingston, proposed to give more weight to the emissions abatement by-product of the LCCS by applying, from October 2008, a complex set of congestion charge discounts and surcharges based on potential CO_2 and noxious emissions. In July 2008, the subsequent Lord Mayor, Boris Johnson, announced that the emissions-related charging structure would not be implemented. Reasons for this decision included the observation that the revised charging regime would encourage travel by smaller vehicles, resulting in more congestion and pollution. However, on 4 January 2011, a "Greener Vehicle Discount Scheme" was introduced, providing a 100 per cent discount of the congestion charge for cars that emit less than 100 g/km of CO_2 and meet the Euro 5 standard for air quality.

Outcomes

In London, implementation of the LCCS delivered an 18 per cent reduction in vehicles entering the charging zone, and a cut of 15 per cent in vehicles circulating within the charging zone during charging hours. Consistent with this, traffic fell on radial roads approaching the zone. There was no systematic evidence of significantly increased traffic outside charging hours or in the area surrounding the charging zone.

Application of the LCCS yielded a 30 per cent reduction in congestion, defined as the difference between the average network travel rate and the free flow network travel rate in minutes per vehicle kilometre. Although traffic increased on the unpriced ring-road forming the boundary of the scheme, congestion fell marginally because of better traffic management arrangements.

Traffic volume and speed change effects of LCCS reduced carbon dioxide emissions in the charging zone and on the boundary ring-road by 15.7 per cent and 4.7 per cent, respectively. Emissions of nitrogen oxides and particulate matter fell by 7.9 per cent and 6.3 per cent, respectively, in the charging zone, and fell by 0.2 per cent and rose by 2.8 per cent, respectively, on the boundary ring-road.

The initial degree of congestion relief has been broadly maintained since the LCCS was introduced. This is after allowing for interventions having long-term capacity-reduction effects for general traffic, such as reallocation of road space to bus lanes, signalling changes, “public realm” schemes, and temporary capacity-reduction events, such as road disruption for maintenance and works in respect of utilities. The capacity-reduction activities removed some of the congestion and emissions reduction gains from the LCCS.

Traffic entering the western extension of the congestion charging zone (from February 2007) fell by 10 to 15 per cent compared to 2006. Congestion declined by 20 to 25 per cent. Traffic on the western boundary road rose by up to 5 per cent, as expected. Traffic entering the original zone increased by up to 4 per cent, but circulating traffic fell.

The LCCS has been widely criticised because of the high costs of running the scheme relative to revenue. In London, these costs account for 60 per cent of revenue, compared to around 21 and 22.4 per cent in Singapore and Stockholm, respectively. The high cost of running the LCCS has been attributed to:

- the establishment of 174 boundary checkpoints compared to 15 and 18 for Singapore and Stockholm, plus 29 cameras inside the cordon (because of area pricing) compared to none in Singapore and Stockholm (as cordon pricing applies)
- the choice of technology and administrative systems
- an unfavourable deal with the private administrator of the system
- numerous exemptions, and 90 per cent discounts for residents of the charging zone and a few areas outside of it.

While the proportion of revenue required to cover costs provides an indication of management costs being too high and exemptions and discounts being too generous, it is not an appropriate measure of the efficiency of a congestion pricing system, as explained above. Nevertheless, the efficiency of the scheme has been undermined by (Button, 2008, p. 213):

- exemptions and discounts that undermine the congestion-alleviation potential of the scheme
- a charge that does not vary with the degree of congestion and is not related to marginal external costs of congestion
- technical and administrative arrangements that unnecessarily waste resources
- allocation of nett revenues and road space to increase subsidies to public transport that was already heavily subsidised, resulting in deadweight losses of the kind associated with any crosssubsidy
- x-inefficiency[5] in the way that bus lanes and other public transport infrastructure are supplied and increased, involving decisions clearly not based on any commercial criteria.

Despite these deficiencies, a social benefit-cost analysis of the LCCS, excluding the western extension, by Transport for London (2007) yielded benefit/cost ratios of around 2 and 2.5, with charges of £5 and £8 per day, respectively.

However, an earlier “tentative economic appraisal” of the London area pricing scheme by Prud'homme and Bocarejo (2005) indicated social costs significantly exceeded social benefits. They concluded, “The London congestion charge, which is a great technical and political success, seems to be an economic failure.” However, Mackie (2005) and Raux (2005) were critical of the methodology of Prud'homme and Bocarejo on several counts, and argued that flaws in their methodology were responsible for the negative result they achieved.

---

[5] X-inefficiency refers to poor performance in transformation of inputs into outputs.

Later economic assessments of the LCCS, prior to the western extension, by Leape (2006), and Santos (2008a,b,c) were positive overall. However, Santos (2008a,b,c) and Kay (2007) doubted that the western extension (since abandoned) of the London charging zone would yield nett economic benefits. They noted that while most travel in the original charging zone originated elsewhere, travel within the western zone generally originated within that area or came from or through the original charging zone. Because of the flat price and discount arrangements, the congestion-alleviation effect and revenue raising capacity of the extension would be less than for the original scheme.

There is little doubt that the LCCS could have provided higher nett benefits with variable pricing targeting marginal external costs of congestion. Similarly, efficiency gains could be increased by substantially reducing the wide range of 90 per cent and 100 per cent exemptions.

The exemptions for "green" vehicles indicate unnecessary jumbling of objectives and poor targeting of policy instruments to objectives. Congestion charging's primary purpose is to reduce congestion, but it reduces emissions as a by-product, because fuel consumption is much higher in congested conditions than free flow conditions. If it desired to target reduction of emissions, the sensible approach would be to deploy another policy instrument for that purpose. Trying to modify one instrument to achieve two targets instead undermines its ability to perform well in pursuit of both. These points were emphasised by Nobel Laureate Jan Tinbergen (1952) and Bent Hansen (1955) in the early 1950s.

Despite the various flaws of the LCCS, eminent transport economist Kenneth Button (2008, p.214) concluded:

> "London as well as a few other locations has provided a practical example of how a crude monetary pricing regime is more efficient at allocating resources than the use of queues."

3.4.3 Cordon Pricing - Stockholm

Outline of Scheme

Stockholm adopted congestion pricing in the form of cordon pricing regime in August 2007 following a successful trial from January-August 2006 and a referendum in September 2006.

In the three years prior to the trial, an intense and aggressive scaremongering campaign was conducted by the Stockholm Chamber of Commerce, the Automobile Association and a morning daily newspaper, Dagens Nyheter. The community attitude was generally negative during that period, but became more positive as the trial proceeded.

In the referendum, 51.3 per cent of voters in Stockholm city voted in favour of applying congestion pricing permanently. In the remainder of the metropolitan area, only 39.8 per cent voted in favour of continuing with congestion pricing. Although the government changed at an election conducted at the same time as the referendum, the new Swedish parliament supported permanent implementation of congestion charging.

In response to questions put to voters in Stockholm city (not the whole metropolitan area) regarding allocation of revenues, the original plan to earmark most of the nett revenue for public transport was abandoned. Instead, nett revenue has been hypothecated for road investment, starting with provision of a ring-road around Stockholm.

Congestion charges apply when vehicles pass 18 control points at all entrances/exits to the designated congestion area, a roughly circular area with a 3 km radius, including the entire central business district (cbd). Most of the cordoned area is on 14 islands. Bridges connect islands to other islands and the remainder of Stockholm.

Charges vary depending on entry and exit times with no charge from 6.30 pm to 6.29 am and on weekends. Maximum charges of SEK20 ($A2.80) apply from 7.30 am to 8.30 am and 4.00 pm to 5.30 pm. Other charges are SEK15 from 7.00 am to 7.29 am and 3.30 pm to 3.59 pm, and SEK10 from 6.30 am to 6.59 am and 6.00 pm to 6.29 pm. The maximum charge in a day was capped at SEK60 ($A8.40).

Exemptions from congestion charges apply to a wide range of road-users. They include vehicles registered abroad, emergency vehicles, transportation service vehicles, large buses, taxis, motorcycles, one vehicle used by each holder of a disabled person's parking permit, and vehicles fuelled by ethanol, bio-gas or electricity.

There are two free-of-charge openings for through traffic. One is a north-south highway. The other is for vehicles passing directly from an island (Lidingö) to the east of the inner city, through the toll cordon and beyond, within 30 minutes.

During the trial in Stockholm, overlapping technologies were used. They were dedicated short range communication (dsrc) and automatic number plate recognition (anpr) technologies. The dsrc technology involved on-board transponder tags that communicate with receivers at control points. In addition, vehicles passing through control points were photographed and optical character recognition software was used to identify licence plates of vehicles without transponder tags to provide evidence to support enforcement. A sophisticated recognition system used algorithms to make a second attempt at identifying number plates that are unclear because of light, weather or poor camera angles. When the Stockholm cordon pricing system was re-introduced following the referendum, the dsrc technology was phased out, because the trial showed that the anpr technology worked well and subsequent tweaking further improved its effectiveness. By 2008, the identification achieved by anpr, with a small amount of manual support, was regularly between 95 and 99 per cent.

Substantial improvements in public transport were put in place 4 months in advance of the trial. These included 18 new express bus routes and provision of 18 per cent more park-and-ride spaces. An effective communications strategy was implemented.

Outcomes

The pricing scheme cut traffic across the cordon by about 22 per cent relative to pre-pricing levels during the period of the trial and following reintroduction of cordon pricing permanently. Traffic in major inner city streets declined by about 15 per cent. The decline within the cordon was less than over the cordon, because vehicles moving entirely within the cordon were not charged and road space was freed up by the reduction in traffic volumes across the cordon. Traffic volumes declined at locations far out from the cordon. Traffic on link roads on the city's outskirts did not increase (Eliasson, others, 2009, pp. 242-243).

Subsequently, traffic across the cordon increased gradually, as a result of growing population and car ownership, declining real cordon charges through inflation, and a steadily increasing proportion of exempted alternative fuel cars among those crossing the cordon (Eliasson, 2008, p. 399).

The reduction in congestion and improvement in travel time reliability were significantly stronger than the reduction in car traffic. Delay times fell by one-third during the morning peak period and by one-half during the afternoon/evening peak period. The duration of the highest decile of the travel time distribution fell to one-third or less of pre-charging levels for some categories of roads, such as arterials during the afternoon/evening peak period (Eliasson, others, 2009, p. 244).

Greenhouse gas emissions from road vehicles were cut by around 14 per cent within the cordon, and by 2-3 per cent in the metropolitan area following introduction and re-introduction of cordon charging. Noxious emissions were reduced by about 8.5 to 14 per cent, depending on the pollutant within the cordon, and by about 5 per cent in the metropolitan area. Emissions reductions were constrained by greater use of old buses in the accompanying extension of bus services (Eliasson, others, 2009, p. 245).

Substantial public transport improvements were completed 4 months prior to the application of cordon price, but passengers did not rise significantly until cordon pricing applied (Anas, Lindsey, 2010, p. 75).

With hindsight, the costs of implementing the Stockholm cordon pricing system were substantially higher than necessary. At least three factors contributed to this.[6]

First, costs were driven up by the government's determination to ensure the system worked properly from the start without technical and administrative gliches that would have undermined the system's acceptability. Risk-proofing the system was central to the government's procurement approach and therefore it paid for several costly back-up and fail-safe mechanisms (Hamilton, 2010).

Second, costs were raised because of issues regarding the legal status of the cordon charges. The charges started as "environmental charges" when tender processes commenced, but legally became taxes after selection of the prime contractor. The implications of the change only became clear after the contract was signed. Two hundred changes had to be made to the system and some parts had to be re-built, with consequent increases in the cost of the system, tentatively estimated to be 15 to 25 per cent (Hamilton, 2010).

Third, the Lidingö Island exemption meant that the effective identification level requirements of the anpr system had to be higher than nominal, formal requirements because people travelling between the island and locations outside the cordon area had to cross the cordon twice. It was estimated by the National Road Administration that this could have raised the cost of establishing the system by a "couple of hundred million SEK ($A28 million)" (Hamilton, 2010).

---

[6] For a detailed discussion of these and other cost issues with the Stockholm system, see Hamilton (2010).

A series of articles in Dagens Nyheter newspaper in 2008 criticised the Stockholm congestion charging system on the grounds that costs of the scheme were unnecessarily high and took about 50 per cent of revenue. The first point has been discussed above. The second point is inaccurate as explained below.

In any event, as explained above, the system should be assessed by comparing social benefits to social costs, not by comparing government revenue from the system to costs to government. The observations by Carl Hamilton (2010) on this matter in Box 4 are apposite.

Box 4 **Comparing congestion charging revenue with system costs an irrelevant measurement**

"The criticisms raised against the (Stockholm congestion pricing) project by Dagens Nyheter, the Automobile Association, and the Chamber of Commerce fail in two important aspects. First, they do not compare the cost of establishing the system to the value it generates, in terms of reduced traffic congestion, only in terms of the revenue collected, which is an irrelevant measurement. Congestion charging is by no means the most effective way to collect tax revenue. Rather, tax revenue is a positive by-product, while congestion relief is the major value generated. Second, they fail to recognise the extraordinary performance in getting a fully functional system in place in time for launch –- politically, administratively, commercially, organisationally, and technically. .... the Stockholm congestion charging was put in place quickly enough and at sufficient quality to swing opinion in time for the referendum."

Source: Hamilton (2010), p. 19.

Jan Jansson (2008) reviewed, reconciled and adjusted early (2006) benefit-cost analyses based on the Stockholm congestion charging trial undertaken by consulting firm, Transek, and Prud'homme and Kopp (2006), and later adjustments by SIKA, a government agency responsible for coordinating use and development of transport benefit-cost analysis in Sweden. Jansson's adjustments indicated a nett benefit from congestion charging of about $A11 million per year, and a benefit/cost ratio of 1.19.

A subsequent benefit-cost analysis of the trial undertaken by Jonas Eliasson (2009) extrapolated benefits and costs forward and provided an estimate of nett social benefit of SEK650 million ($A91.6 million) per year and a nett present value of SEK6.3 billion ($A882 million), with constant yearly benefits, and SEK7.6 billion ($1,064 million) if benefits grow at 1.5 per cent per year with traffic growth. The ratio of the nett present value of all benefits and costs except operating and start-up costs to the present value of start-up and operating costs was 2.5 in the former case and 2.6 in the latter case. Eliasson suggested that benefits included in the analysis were conservatively low, and included costs were conservatively high.

One important difference between the Jansson and Eliasson analyses was the estimated cost of running the Stockholm congestion charging scheme. Jansson (2008, p. 182) assumed the annual cost of investment and operation of charging system was about $A64 million per year, based on figures provided by SIKA in 2006. Eliasson (2009, p. 470) assumed ongoing investments and maintenance is estimated to be around SEK220 million ($A30.8 million) per year, using estimates provided by the National Road Administration (NRA), which is responsible for the Stockholm congestion system. The latter figure is around 27 per cent of annual revenues from congestion charges. Carl Hamilton (2010, pp. 7-8) pointed out that the cost of running the system had been lower than estimated by NRA and was expected to be about SEK180 million ($A25.2 million per year) from 2010. This suggests a ratio of running costs to revenue of 22.4 per cent, which is similar to the ratio for Singapore's multi-zone cordon pricing system.

Another major difference between the analyses was Jansson's exclusion and Eliasson's inclusion of a premium on nett revenue from congestion charging in recognition of the inefficiency costs or deadweight losses associated with government revenue from distortionary taxes. Eliasson followed the established practice in Swedish benefit-cost analyses of including this premium. It amounted to SEK182 million ($A25.5 million) per year.

Jansson (2008, p. 185) argued against inclusion of this benefit on the grounds that it would compensate for leaving out of the benefit-cost analysis the adverse effect of congestion charging on labour supply (the tax interaction effect discussed in sub-section 3.3 above). However, he acknowledged that increased public transport supply, which accompanied cordon charging, could work in the opposite direction.

If the tax interaction effect should be accounted for in some way, it would also seem appropriate to take into account offsetting revenue recycling effects too. It would be more correct to account for all three effects than to follow Jansson (2008) and leave them all out, or follow Eliasson and include one, but not the other two.

Eliasson (2008, pp. 478, 479) found that the value of time and travel reliability gains from Stockholm's congestion charging system represented about 70 per cent of revenue from charges. Eliasson (2008) and Anas and Lindsey (2010) pointed out that this was very high compared to most theoretical or model-based studies. Typically, these are speed-flow models, which do not allow for re-scheduling of trips within the peak and shoulder periods and do not take into account queuing caused by bottlenecks or choke points in the road network. If these phenomena are allowed for (as in the bottleneck model of congestion originated by Vickrey, 1969, and developed by Arnott, de Palma and Lindsey, 1990, 1993, 1994, 1998), estimates of welfare gains from congestion pricing rise in absolute terms and relative to government revenue (Parry, 2009, p. 467). This matter has been discussed further in sub-section 3.6 below.

3.4.4 HOT Lanes and Tolled Express Lanes

Outline of Schemes

In the U.S.A., for many years, there was a trend towards establishment of high occupancy vehicle (HOV) lanes, which effectively subsidised buses and car pooling. This was followed by a trend towards conversion of underutilised HOV lanes to high occupancy toll (HOT) lanes, and provision of new tolled express lanes adjacent to existing unpriced freeway lanes. On these facilities, low occupancy vehicles pay tolls, while HOVs have free access.

Variable pricing has been adopted on most of these facilities. In some cases, prices vary during the day according to a published schedule. The schedule is adjusted periodically, typically every three months. In other cases, prices are adjusted every few minutes according to traffic conditions beforehand ("dynamic pricing"). Typically, indicative price schedules are also published as a guide to potential users of the "dynamically priced" facilities.

Variable pricing of HOT lane and tolled express lane facilities as described above has often been labelled "value pricing". This term is meant to reflect the provision of additional road-use value through pricing.

Often, the HOT and tolled express lane schemes have been described as simplified forms of congestion pricing, even in cases when variable pricing has not been applied. Also, they have also often been depicted as productive steps towards cordon pricing, area pricing or more sophisticated congestion pricing regimes, and as arrangements that would facilitate public acceptance of actual congestion pricing systems in future.

These arrangements are quite distinct from the tolling concept adopted in Australia. Here, governments have provided new tolled urban road links (usually under public private partnership arrangements), typically ring and by-pass roads, rather than addition of tolled lanes to otherwise free-access freeways. In addition, unlike the approach in U.S.A., tolls in Australia usually do not vary with time of day or with congestion.

Outcomes

It is doubtful that tolling of ad hoc road links, as in the case of widely scattered HOT lanes and tolled express lanes could reasonably be regarded as enabling acceptance of future cordon pricing, area pricing or more sophisticated congestion pricing regimes. Tolling of such ad hoc road links adjacent to close substitutes with zero prices, in the context of zero pricing on most of the remainder of network, is very different to cordon pricing, area pricing, and network-wide variable pricing, for which the closest substitutes are unpriced cycling, walking and off-peak use of roads, and priced public transport. The HOT and tolled express lane concepts are narrow in scope, involve conversion of HOV lanes or provision of new lanes, and could cover costs only if too much congestion is maintained on the network. In contrast, congestion pricing, particularly network-wide variable congestion pricing, is broad in scope – ideally applying when and wherever there is congestion on existing and new facilities – alleviates congestion efficiently, and reduces pressure for new facilities. Another major difference is that congestion pricing requires astute application of revenues to avoid distributional and acceptability concerns, while HOT and tolled express lanes avoid these concerns by applying prices only to facilities not previously accessible by low-occupancy vehicles.

Economic modelling of HOV and HOT/tolled express lane arrangements has shown that they are greatly inferior to network-wide variable congestion pricing. HOV lanes in particular were found to yield substantial economic efficiency losses, while network-wide variable congestion pricing yielded substantial gains (Harvey, 2005; Small, Winston and Yan, 2006), so long as those prices are set to efficient levels. Tolling of ad hoc road segments and lanes, even if priced for the purpose of alleviating congestion, could, at best, yield less than one quarter of the efficiency gains of network-wide variable congestion pricing. Prices set on ad hoc road segments and lanes to alleviate congestion would range from negative to very low. Therefore, pricing of such roads to recover costs is in conflict with congestion alleviation (Parry, 2002; Verhoef, Small, 2004; Small, Verhoef, 2007; Downs 2004).

However, HOT lane and priced express lane arrangements can provide significant welfare gains relative to HOV lanes and no extra lanes. The heterogeneity of road-users is an important contributor to these gains. Road-users with different values of time are able to decide between priced and unpriced lanes having regard to the charge for using the priced lane and different traffic volumes and speeds in priced and unpriced lanes (Parry, 2009, p. 471).

Nevertheless, the application of "dynamic pricing" for HOT lanes and tolled express lanes has provided important insights for congestion pricing.

First, it has highlighted the significance of heterogeneity of road-users for formulation of prices for a network. Time-, location-, and lane-differentiated prices are appropriate responses to differences in users' values of time. However, various simulation studies have found that the efficiency gains from differentiated pricing of lanes may not be large (about 10 per cent) in the context of network-wide variable congestion pricing (Verhoef, Small, 2004; Parry, 2002, 2009).

Second, the evidence available indicates that drivers have adapted easily and rationally to "dynamic pricing". There is no evidence that drivers have experienced more difficulty with "dynamic pricing" than with pricing that involves time- and location-based pricing schedules that are adjusted at pre-determined intervals. In turn, there is no evidence of more adjustment problems for drivers with scheduled varying prices than with flat tolls.

3.4.5 Parking Levies and Restraints

Outline of Schemes

In many major cities, governments have used price, tax and supply restraint measures in busy locations to push up the effective price of parking to discourage driving in those locations and on access roads to them. Typically the purpose of these measures is to reduce traffic congestion. However, these measures have also been promoted as devices to reduce greenhouse gas and local pollutants.

If prices of on-street and off-street parking in a congested area are raised by direction or via a tax impost by more than is necessary to ration supply, driving that terminates and originates at that location would be discouraged. The Sydney, Melbourne and Perth parking levies are examples of tax measures ostensibly aimed at reducing congestion.

The effective price of parking is also raised by artificially restricting supply of parking spaces. The restriction could take the form of direct reduction of on-street parking spaces, regulated restriction of off-street parking space numbers under urban planning regimes, or limitation of time spent in parking spaces. The effective price increase could take the form of an explicit increase in price or more "cruising for parking", meaning drivers expend fuel and time. Again, driving that terminates in the parking-restricted area is discouraged.

Parking measures aimed at congestion-alleviation are typically used as "second best" alternatives or surrogates for congestion pricing. The concept is that higher effective parking prices will induce drivers to switch from single occupancy vehicles to car-pooling, public transport, cycling, walking and working from home. Such parking measures may be politically favoured over congestion pricing, because of concerns about redistribution and consequent acceptability issues with the latter.

Governments sometimes apply both parking measures and public transport subsidies, as anti-congestion measures. Like parking measures, subsidising public transport is a "second-best" alternative to congestion pricing.

Outcomes

Parking measures reduce car travel to the central business district and other major activity centres subjected to effective increases in parking prices. This occurs through switching to alternative transport modes or alternative activities. Whether or not travel times are changed would depend on the structure of parking measures. However, measures that achieve desired time-switching would be difficult to design and administer.

By raising the effective price of driving to major activity centres, parking measures lower the relative price of public transport, inducing greater patronage. Consequently, parking measures allow exploitation of economies of scale/density/frequency with lower public transport subsidies or attract more patronage from cars to public transport, potentially without additional subsidies. Parking taxes (but not supply restrictions) also provide revenue that can be recycled to upgrade infrastructure.

Parking measures do not deal with through-traffic and tend to induce more of it by freeing-up space on radial roads. The extent of the "induced traffic" effect would be greater, the larger is the proportion of traffic that previously terminated and originated in the area subject to parking measures. The "induced traffic" phenomenon has been discussed in detail in section 4.

Public transport subsidies also are much less effective in discouraging cross city driving than driving to the central business district. They also cause an by attracting car-users from radial roads and freeing-up space there. This adds to the "induced traffic" effect on radial roads arising from parking measures (see sub-section 4.2.1).

Parking measures have a greater relative effect on short drives to the central business district than long drives, because the effective increase in parking price is a higher proportion of the full cost of short drives than long drives. Parking measures affect short-trip drivers more than long-trip drivers, because the effective increase in parking price is a higher proportion of the full cost of short drives than long drives. Therefore, parking measures do not adequately deal with congestion caused by terminating traffic, particularly congestion <u>not</u> close to the parking-priced/supply restricted area. Also, sub-optimal sprawl may be encouraged.

Subsidised public transport typically provides better services for commuters close to the central business district than those with long commutes. In this respect, effects of parking measures and public transport subsidies tend to reinforce each other in discouraging short range driving, rather than long range driving to the central business district.

Supply restrictions that encourage "cruising for parking" waste time and fuel. It is not clear whether or not this effect offsets the effect on congestion of discouraging driving to the area.

Parking measures would not address the contribution of buses and commercial vehicles delivering and collecting people and goods. Buses and trucks contribute to congestion not only because of their presence, but also because of their size, acceleration characteristics, and loading/unloading activities.

It would be impractical to devise parking measures that for all parking spaces in busy locations discriminated to reflect timing of entry and extent, traffic conditions at those times, and location of the origin, destination and route of the driver.

Parking measures favour through-drivers over drivers terminating and originating in the area subject to such measures and favour longer distance over shorter distance commuters. These factors would be reflected in relative property values.

Regulatory measures to restrict the supply of parking spaces impose imposts on drivers by stealth. They provide effective transfers from drivers to owners of pre-existing parking spaces. They do not provide revenue for government to recycle.

Policy makers' preference for parking measures over congestion pricing on acceptability grounds neglects the clear superiority of the latter as a congestion-alleviation measure. It also neglects the greater revenue-raising and revenue-recycling opportunities presented by congestion pricing. Astute use of congestion pricing's superior revenues would facilitate acceptability as discussed below.

Parking prices (not cruising) should be increased sufficiently to clear the market, but beyond that, parking measures are clearly inferior to well-designed congestion pricing as an anti-congestion instrument and as a complement to direct anti-emissions measures. Supply restraint measures that cause cruising for parking are least desirable.

3.5 Theory Versus Practice

3.5.1 Technology up with Theory

About 50 years ago, when William Vickrey (1959) formulated a proposed congestion pricing for Washington, and Reuben Smeed, Alan Walters, Gabriel Roth and others (1964) recommended implementation of congestion pricing in the United Kingdom, it was basically a theoretical concept that would have been practical to implement only in crude form or at very high cost. Now, technology has caught up with the theory, and technological advances continue unabated.

A brief description and discussion of technologies currently in use to support congestion pricing, and existing technologies that have been identified as having considerable potential to support more sophisticated congestion pricing regimes have been provided at Appendix A. A key point is that mature technologies are available that could be adapted and deployed to support a highly sophisticated version of congestion pricing in practice. An important caveat is that these steps have not yet been undertaken in practice.

Of course, greater technical sophistication yielding precise information that changes promptly with changing conditions can mean greater cost, although technological advances are continually reducing these costs. In addition, the capacity to cover whole cities with economies of scale through application of technologies such as global position systems would reduce costs. Also, the finer differentiation of pricing according to time, location and the conditions that is provided by such technologies eliminates costs of dealing with inadvertent congestion shifting problems and reduces costs of adjusting the pricing regime to changes in circumstances.

3.5.2 Coping with Information

An important issue is the point that usefulness of greater sophistication may be limited by the capacity of drivers to absorb price information that varies quickly with changing conditions. Therefore, trade-offs may be necessary between economic gains from greater sophistication, and economic costs of technologies and information overload for drivers.

This point was made long ago by Smeed, Walters, Roth and others (1964). A toned-down warning regarding this matter was provided 5 years ago by Kenneth Small and Erik Verhoef (2007) in the preeminent text book on urban transport economics.

These views are not universally held. Jonas Eliasson (2010) argued that road-users had shown that they were capable of coping and adapting to apparently complex road pricing arrangements. Peter Bonsall and others (2007) argued, based on research for the U.K. Department for Transport, that while people prefer simple to complex pricing structures, they are able to respond sensibly to quite complex price structures provided they are clear and logical. They pointed out that complex price signals are most likely to be blunted when the price structure lacks an obvious logic or has not been fully explained or communicated to road-users.

Quotes reflecting these contrasting views are presented in Box 5.

Box 5 **Efficiency Versus Simplicity in Congestion Pricing Design**

Smeed, Walters, Roth, others:

"The most efficient price system might appear to be one in which price varied with cost on every road at every moment of the day. But this presupposes that road users are able and willing to take account of such a highly complicated system. In practice, of course, they are not. If the price system is complicated, road users will probably find simple 'rule of thumb' methods to tell them approximately what the average prices are and roughly what the prices of particular journeys are likely to be, and they will act accordingly. If this is so, the complicated system may be no more efficient than a simpler system."

Small and Verhoef:

"The more important source of constraints is not technology, but rather the kind of complexity that users and political representatives will tolerate."

Jonas Eliasson:

"While it is an important consideration that the system must be sufficiently simple for the presumptive users to understand, policy makers often seem to underestimate people's cognitive ability. The Singapore system and the U.S. "value pricing" of roads, for example, appear complex at first glance. The charge is differentiated by time and location, and on top of that may change fairly often. Despite this, it turns out that users are able to grasp and adapt to the system."

Peter Bonsall and others:

"It is concluded that people have a strong preference for simple tariffs, but they are able to respond to quite complex tariffs provided they have a clear and logical structure."

Source:Smeed, Walters, Roth, others (1964), p. 42; Small, Verhoef (2007), p. 153; Eliasson (2010), 0p. 6-7; Bonsall, others (2007), p. 672..

The application of dynamic pricing (pricing that may vary every few quickly with immediate traffic conditions) in United States "value pricing" schemes involving HOT lanes and tolled express lanes is of particular interest. The evidence available indicates that, while consumers prefer simple to complex pricing arrangements, drivers have coped well with these arrangements and responded rationally to them. Indeed, there is no evidence available that would indicate drivers have experienced more difficulty adjusting to dynamic pricing than to pricing that involves time- and location-based pricing schedules that are adjusted at pre-determined intervals, typically three-monthly. In turn, there is no available evidence that drivers have experienced more adjustment problems with schedules of prices varying by time of day than with flat tolls (Bonsall, others, 2007, pp. 672, 675-676)

3.6 Welfare Gains, Compensation of Losers, and Administration Costs

3.6.1 Winners and Losers

The static model of congestion based on speed-flow relationships that is widely used to illustrate the concept of congestion pricing indicates that government revenue from congestion pricing (a transfer, not a social cost) would be larger than the losses of the losers, and both would be much larger than the nett benefits (welfare gains) from application of this policy instrument. This suggests that government would collect adequate revenue from congestion pricing to finance measures that could provide compensating value to losers.

To formulate compensatory arrangements, it would be necessary to identify the winners and losers before any revenue recycling. In the context of the static speed-flow model, they would be as follows.

Winners from congestion pricing, before taking into account revenue use/recycling, would include:

- road-users whose time and fuel savings exceed charges
- bus passengers before congestion pricing who get shorter and more reliable trip times, without crowding (assuming more services)
- residential property owners who gain from property value appreciation, because of locations in or adjacent to major activity centres, or near (but too near) access points to higher quality public transport services
- the government collecting congestion revenue – the big winner.

Nett losers from congestion pricing before revenue recycling would include:

- car users who are priced-off to another time, route or mode because their willingness to pay is less than the charge

- the component of the priced group (those staying on the priced roads), who pay more to meet congestion charges than they save in time value and fuel
- unpriced travellers, who previously used other routes, times, or modes, but encounter crowding when joined by some of the priced-off group
- owners of residential property in areas remote from major activity centres and not well-served by public transport connections or uncrowded road links to such centres.[7]

However, there are three issues that require investigation. First, the basic, static, speed-flow model of congestion has pertinent deficiencies. Second, the basic model does not take into account costs of running and complying with a congestion pricing system. Third, it is important to use the revenue to compensate users in ways that are consistent with, not detrimental to an efficient allocation of resources. The implications of these issues have been discussed below.

Care should be taken to avoid overreliance on the basic static, speed-flow-based model of the economics of congestion and congestion pricing. Its deficiencies led to formulation of the bottleneck model by William Vickrey (1969), its development by Richard Arnott, André de Palma and Robin Lindsey (1990, 1993, 1994, 1998), and formulation of other dynamic models combining elements of speed-flow and bottleneck models, such as those developed by Xuehao Chu (1995), Se-il Mun (1999) and Erik Verfhoef (2003).

The basic static model does not take into account trip re-scheduling within peak and shoulder periods, just rescheduling to off-peak periods or shifts from cars to other transport modes. Simulations by Arnott, de Palma and Lindsey (1993) indicated that more than half of the welfare gains from congestion pricing could come from re-scheduling of trips within peak and shoulder periods. These gains would be triggered by variable pricing over peak and shoulder periods.

In addition, the basic static model does not allow for congestion (or hypercongestion) resulting from bottlenecks, or hypercongestion caused by very high traffic volumes, when both traffic flows as well as speed decline as additional users access crowded roads. In bottleneck situations, it may be optimal to eliminate queues completely rather than only partly (Parry, 2009, p. 467; Fosgerau, Van Dender, 2010, p. 6; Arnott, de Palma and Lindsey, 1994, p. 158). This deficiency of the basic static model also results in understatement of welfare gains from congestion pricing.

In bottleneck situations, which can be common occurrences in various locations at peak times, and allowing for re-scheduling within peak and shoulder periods, congestion charges would replace queuing time (time and schedule delay) losses. This would leave road-users' private costs inclusive of charges unchanged on average from private costs including queuing costs, before use of revenue from charges. Therefore, welfare gains from congestion pricing could be approximately the same as revenue from charges (Arnott, de Palma, Lindsey, 1994, p. 144; Small, Verhoef, 2007, pp. 130, 132). This is very

---

[7] Land values may also decline for residential properties in the central business district for reasons set out in Ergas (2011). However, as also explained there, it is important not to double count the welfare effects of changes in transport times by counting them both directly and through changes in land values.

different to the indication from the basic static model that welfare gains would typically be much smaller than revenue transfers.

In the more realistic case of hetereogenous road-users in bottleneck situations, the optimal time-variant pricing arrangement would still eliminate (substitute for) queuing time losses. However, additional efficiency gains would result from extra price-induced changes to travel schedules (including reliability) in the context of differing values of time delays, arrival times, being late, and being early. These extra changes involve variations in trade-offs between costs of changing travel times and travel schedules (Small, Verhoef, 2007, pp. 134, 137).

Another issue in the case of heterogenous road-users is that there are winners and losers from efficient, time-variant congestion pricing. The distributional outcome depends on differences between road-users in respect of values placed on time and schedules and the relative value of time and schedules. Elimination of queuing through pricing would benefit those with high values of time and disadvantage those with low time values (This also occurs in the speed-flow model of congestion). Alteration to schedules induced by pricing would favour those with high schedule change costs. They would not necessariliy be high income groups, because low income groups might have much less scheduling flexibility because of rigidities in work starting and finishing times charges (Arnott, de Palma, Lindsey, 1994, p. 157-158; Small, Verhoef, 2007, pp. 133-134).

On balance, Kenneth Small and Erik Verhoef (2007, p. 130) and Ian Parry (2009) argued that the insights of the bottleneck model should moderate concerns regarding distributional and acceptability issues. Parry (2009, p. 467) explained:

> "….. welfare gains from bottleneck pricing can be roughly the same as the toll revenue collected, unlike in the traditional (speed-flow) model, where welfare gains are typically much smaller than revenue transfers. Because it is optimal to eliminate bottleneck queues completely rather than partly, welfare gains are first order unlike in the static (speed-flow) model where they are second order. In principle, this should reduce public opposition to tolls designed to alleviate specific bottlenecks in the road network."

A caveat is that work on some alternative dynamic models (see Chu, 1995; Verhoef, 2003) has indicated that the bottleneck model may exaggerate the extent to which travel delays can be eliminated without increasing scheduling change costs. This could lead to overestimation of benefits from optimal time-variant pricing and underestimation of the increase in the cost of travel (inclusive of congestion pricing). On the other hand, theese other dynamic models may be inherently biased the other way (Small, Verhoef, 2007, p. 136).

Before revenue recycling, winners are more likely to emerge when time-variant pricing is applied to target bottleneck-related congestion or hypercongestion (traffic flows and speeds decline with car numbers) resulting from bottlenecks or other influences, so that charges on average replace costs of travel delays and schedule disruption. For all forms of congestion, the most likely winners from congestion pricing without recycling are those with high time values, who save more in time and fuel than they pay in congestion charges. High time values typically reflect high incomes, but those without high incomes may have high time values in specific circumstances, such as when schedules are relatively important.

Having regard to the insights of the speed-flow and bottleneck models of congestion, it can be deduced that a congestion pricing regime that is sufficiently well-designed to yield nett social benefit would provide gains to winners that are more than large enough to compensate those who would be nett losers before use of government revenue from the regime. The implementing government is the big winner. Its revenue from congestion charges would be adequate to compensate the losers, if it chooses to do so. This has been clearly articulated by eminent transport economist Kenneth Button, as shown in Box 6.

Box 6 **The Big Winner, Government Collects Enough to Compensate Losers**

"What congestion charges do in theoretical terms is to produce a Hicks-Kaldor-Scitovsky (potential) improvement in welfare. This means there would be an aggregate social gain from such charges. What they do not do is produce a Pareto improvement that would require that no one is made worse off by the charges. This is, to reiterate, the situation because the main immediate direct beneficiaries of road pricing are the pricing authorities that take the windfall net revenues of pricing. It is a Hicks-Kaldor-Scitovsky improvement because these authorities could compensate those road users who lose out, pay for the administration of the system, and still have surplus revenue. Whether this compensation is actually given is a crucial matter when looking at the implications of the distribution (of) gains from road pricing. It also determines the degree of opposition to it."

Source: Button (2006a), p. 226. Words in brackets added.

3.6.2 System Costs

The absence of administration and compliance costs from the static, speed-flow model of congestion is an important issue for policy formulation. If revenues are not adequate, administration costs are not tightly controlled, potential revenues are greatly sacrificed through exemptions and discounts, or nett revenues are wasted on misconceived outlays, welfare gains from congestion pricing could be completely dissipated (Lindsey, 2006, pp. 306-307; Small, Verhoef, 2007, p. 147).

The discussion above regarding the magnitude of potential revenue relative to losses of losers should ameliorate concerns regarding adequacy of revenue to provide compensatory value to losers.

While the London congestion charging scheme, and to a lesser extent the Stockholm scheme have been criticised because of high administration costs and the extent of exemptions and discounts provided, it is noteworthy that benefit-cost analyses of both schemes have indicated the schemes have yielded nett social benefits. However, these congestion pricing systems were not compared with alternative pricing regimes with different revenue recycling arrangements. There is considerable evidence available that more efficient pricing regimes could have been put in place, involving superior price differentiation and fewer exemptions.

It is not appropriate to include congestion pricing revenue collected by government in the estimation of the welfare or efficiency gain if that revenue is to be used for purposes that result in economic waste (costs in excess of benefits). Indeed, it is just as important to recycle revenue from congestion pricing efficiently, as it is to design, implement and administer congestion pricing efficiently (Small, Verhoef, p. 147). Efficient recycling of revenue has been discussed in some depth in sections 4 and 6.

4 Important Complements to Congestion Pricing

Congestion pricing is a necessary condition for alleviation of congestion in a way that is both effective and consistent with an efficient allocation/use of resources (efficient). However, it is not sufficient for efficient alleviation of congestion, at least in the medium- to long-term.

Congestion pricing in its ideal form, in combination public transport pricing on a social marginal cost basis, would (in the absence of countervailing market and policy failures) reduce traffic congestion and public transport passenger congestion to the efficient or optimal level in the short-term in the context of existing infrastructure. Reducing congestion to the efficient or optimal level in the medium- to long-term would require ideal pricing of urban roads and public transport, plus efficient provision of urban transport infrastructure (roads and public transport). Underpricing or underinvestment would result in too much congestion, while overpricing or overinvestment mean would mean too little congestion from a social perspective. In each case, resources would be allocated inefficiently.

To ensure that ideal congestion pricing and public transport pricing achieve are efficient and effective in alleviating congestion, it may also appropriate to accompany introduction of pricing with information programs to address market failures in provision of information regarding the full range of urban travel options. The case for this measure could be expected to diminish over time, as urban travellers learn more about alternative transport routes, modes, and times.

To alleviate congestion efficiently in the long-term, adjustments may be required to land-use (land-zoning) regulations to eliminate impediments to increases in commercial and/or residential density around access points to major public transport corridors. Otherwise, the congestion alleviation effects of congestion pricing and accompanying public transport upgrades could be partly frustrated.

Complements to congestion pricing are important for acceptability and equity reasons, as well as to improve the efficiency of resource allocation. The availability of good alternatives to using priced roads and good knowledge of them by the community can ease adverse distributional effects of congestion pricing and improve community acceptance of it.

A package of complementary policy instruments would be required to alleviate congestion in a way that is effective, improves the efficiency of resource allocation in the medium- to long-term, and is equitable and politically acceptable. The importance of applying a package of policy instruments to the task of congestion alleviation has been emphasised by, Richard Arnott, a major contributor to the economics of congestion pricing, as indicated in Box 7.

Box 7 An Anti-Congestion Policy Cocktail

"……. city (congestion) tolls are only one element of an effective policy cocktail for dealing with urban traffic congestion. Urban transport economists should broaden their horizons beyond congestion pricing to give due attention to the myriad other congestion-relief policies whose effectiveness can only be improved by the application of sound economics."

Source: Arnott (2005), p. 11.

4.1 Congestion Pricing and Road Capacity

4.1.1 Road Capacity, Triple convergence and Induced Demand

There is a widespread view among urban transport planners and urban planners that adding capacity to an unpriced road network in congested areas would attract more traffic, wiping out potential congestion-relieving effects of costly capacity increases. As a result, proponents of this view have portrayed investment in new urban arterial road capacity as futile, self-defeating, and wasteful. This view has been labelled the "induced demand" or "induced traffic theory". It has spawned clichés such as, "You can't build your way out of congestion" and "Build it and they will come".

In some cases, adherents to the "induced demand theory" have stretched it further to suggest that adding arterial reducing road capacity in the context of static demand would increase congestion. Moreover, they have suggested that not increasing arterial capacity in the context of growing demand for urban transport would actually reduce congestion.

The "induced demand theory" is based on a misinterpretation or misapplication of Anthony Downs' (1962) "triple convergence" analysis.

Downs pointed out that expansion of an urban expressway or new expressway in the context of pre-existing congestion on the network would offer a faster trip than previously on the pre-expansion road, other routes, and other transport modes. In addition, the expanded road would offer a faster trip at the peak of the peak than previously. At the peak of the peak and shoulders of the peak, traffic would "converge" on the expanded road from other times, other roads and other transport modes (hence the term, "triple convergence"). Downs explained that switching would occur until a new traffic equilibrium had been established, in which it was equally unattractive to travel to a work, education or event destination via alternative routes and modes and at alternative times. Downs explained that in the new equilibrium, if public transport pricing and services remained unchanged:

- the new or expanded expressway would be crowded during peak periods, almost certainly operating beyond its optimal capacity
- the new equilibrium would probably be established before congestion on the network reached the level applying before provision of new arterial road capacity
- the average commuting time for car and bus users could be expected to be less than before provision of the new capacity
- the peak period could narrow
- the proportion of commuters riding buses and trains would fall
- the proportion of off-peak travellers using buses and trains would fall even more than the proportion of peak period commuters using buses and trains, because the relative advantage of driving improves more off-peak when there is no congestion, with the result that public transport usage becomes more concentrated in peak times
- the per capita cost of providing public transport services would rise, with higher subsidies required to maintain fares and services.

Downs (1962) observed that if public transport fares were raised to cover per capita cost increases or if a combination of service cuts and lower fare rises were implemented, more patronage would be lost by public transport to car-use. In those circumstances, it is conceivable that congestion and car and bus commuting times could worsen following the expansion of arterial road capacity. However, he explained that this could occur only in cities in which segregated rail transport facilities carried a high proportion of commuters before the new or expanded expressway is opened.

"Triple convergence" involves short-run diversion or redistribution of urban travel movements that had been occurring during and around peak periods prior to expansion of arterial road capacity. Anthony Downs (2004, pp. 84, 104) emphasised that "triple convergence" effects should not be categorised as "induced demand", a term he reserved for long-run increases in usage of the road network resulting from the increase in road capacity. These increases are associated with improved potential mobility on the network and attraction of new people and businesses.

The "induced demand theory" ignores the distinction between short-term "triple convergence" and long-term "genuine induced demand". Downs (2004, pp. 84, 104-107) explained that only the latter could conceivably lead to congestion on the expanded network as bad or worse than before the addition to road network capacity. However, he did not consider this was likely. If it did occur, he argued that the worsening of congestion could be expected to take the form of a longer period of very intense congestion (broadening of the peak), rather than a greater level of intensity at the peak of the peak period.

Some analysts have referred to redistributed and new demand triggered by an increase in road capacity as “latent demand”. This has been explained by eminent transport economists Kenneth Small and Erik Verhoef (2007) as shown in Box 8.

Box 8 **Latent Demand**

“….. induced traffic …. represents the release of potential traffic that is deterred by congestion itself. Such potential traffic consists of people who, because of congestion, now choose an alternative route, mode, time of day or home or workplace location, or do not travel at all. It is sometimes called *latent demand* because in the initial situation it is unobserved; yet, in a sense, it lies just beneath the surface. Unfortunately, latent demand is prevalent in just those areas where capacity expansion seems most needed – namely, in high-density urban areas during times when congestion is severe. Under such conditions, the release of latent demand is likely to undo much of the congestion relief that capacity expansion or demand reduction might otherwise bring about.”

Source: Small and Verhoef, 2007, p. 173.

The available empirical evidence published has not resolved the issue of whether or not there is substance in the claim of the "induced demand theory" that adding road capacity in congested areas would attract more traffic, wiping out potential congestion-relieving effects of costly capacity increases. Important contributing issues appear to be conceptual, methodological and data problems. The last of these has exacerbated the others.

An important methodological issue has been causality. It is not clear to what extent increased traffic is caused by more road capacity, rather than increased capacity being a response to previous and anticipated increases in traffic. If governments increase road capacity because of past and anticipated traffic growth, it would be invalid to argue that traffic growth following additions to capacity resulted from the capacity increase.

Another methodological issue relates to the attraction of an increase in capacity being related to the benefits it provides, rather than the number and length of additional lanes. A capacity increase could provide benefits that differ greatly, depending on location and other circumstances.

A conceptual and methodological issue has been imperfect distinction between "triple convergence" and "genuine induced demand", as discussed above.

Efforts have been made to address some of these issues in studies undertaken over the past decade.

Robert Cervero (2003) applied an elaborate model designed to address the causality and benefit issues. He analysed effects of 24 freeway expansion projects in California between 1980 and 1994. He found that 6-8 years after motorway expansion, 20 per cent of the added capacity had been "preserved", 40 per cent had been absorbed by traffic growth arising from population and income growth, 31 per cent had been taken-up as a result of behavioural shifts, and 9 per cent because of land-use shifts. However, behavioural shifts could involve redistribution or diversion of demand, as well as some genuine "induced demand". Land-use shifts would appear to be linked to "genuine induced demand".

Kent Hymel, Kenneth Small and Kurt Van Dender (2010) used cross-sectional data at the state level for the United States for the period 1966 to 2004 in an econometric model designed to estimate "induced demand" effects associated with capacity increases and "rebound effects" associated with improved fuel economy. They attempted to decompose "induced demand" effects into those arising from increases in urban lane-kilometres per adult and road kilometres per hectare. Hymel, Small and Van Dender (2010) estimated "induced demand" elasticities of around 0.032 in the short-term and 0.16 in the long-term This means a 10 per cent increase in urban arterial road capacity would result in an increase in vehicle kilometres travelled in the region of 3.2 per cent in the short-term and 16 per cent in the long-term. For the end of the period, the estimates were 3.7 per cent and 18.6 per cent, respectively.

The estimates provided by Hymel, Small and Van Dender (2010) were at the lower end of the range of previous estimates of "induced demand". They suggested this could be explained mainly by their use of a measure of vehicle kilometres travelled based on state-wide aggregates, with some contribution from use of more control variables and data relating to a long time period. They acknowledged that percentage changes in vehicle kilometres travelled in a state would understate the effect in a major urban area.

Gilles Duranton and Matthew Turner (2011) applied various econometric techniques to estimate "induced demand" effects of increases in lane kilometres of urban arterial roads in metropolitan statistical areas across the continental United States, using data for 1983, 1993 and 2003. The estimated elasticities of vehicle kilometres travelled on interstate highways in the urbanised parts of metropolitan statistical areas relative to lane kilometres was around 1. For a broad class of major roads within the urban part of metropolitan statistical areas, the estimated elasticities ranged from 0.67 in 2003 to 0.89 in 1983. Duranton and Turner (2011) claimed that truck traffic accounted for about 29 per cent of the estimated elasticities, migration accounted for about 6 per cent, and changes in household behaviour (not described) contributed between 9 and 39 per cent. They claimed diversion from other roads was relatively minor, accounting for 0-10 per cent of the estimated elasticities. They did not discuss diversion from other travel times and other modes of transport. Presumably, diversion from other modes of transport would be included changes in household behaviour.

The extreme version of the "induced demand theory" that restricting road capacity would alleviate congestion is not credible. The apparent rationale is that making driving less attractive by allowing congestion to worsen in the context of reduced road capacity or capacity growing slower than demand will frustrate commuters into switching to other travel modes and times. There is no doubt that as congestion worsens, some commuters would change the time they drive, broadening the peak period, or would switch from driving to another mode of transport. Also, in the longer-term, some would re-locate residences or jobs. However, these adjustments would not be anywhere near enough to prevent more intense congestion and/or a broader peak period. It defies logic to suggest that letting congestion worsen in the short-term will reduce congestion in the medium-term.

Proponents of the "induced demand theory" neglect benefits from additional capacity. More use of an expanded road means benefits have accrued to those accommodated by the extra capacity, including drivers shifting from less convenient travel options and others representing additions to population. Benefits have also accrued to travellers on other routes and modes or at other times that have become less crowded. Therefore, to the extent that congestion levels are restored on an expanded arterial road or even the road network, the outcome should not be interpreted as failure (Downs, 2004, pp. 106-106; Taylor, 2002, p. 13; Martin, Kain, 2008, pp. 27, 34).

4.1.2 Congestion Pricing, Triple Convergence and Induced Demand

“Triple convergence” and “induced demand” analyses have been undertaken in the context (or assumption) of zero prices for access to roads. Congestion pricing would moderate the tendency of “triple convergence” and “genuine induced demand” to frustrate congestion alleviation (Downs, 2004, pp. 86-87).

Congestion pricing is not only a mechanism for potentially achieving efficient use of existing road facilities, but also a device for potentially attaining efficient use an expanded road network. Because congestion pricing, if properly designed, requires road-users to internalise or take into account the congestion costs they impose on others, as well as the costs they experience individually, new road capacity in a differentially priced network would divert or redistribute traffic on the network (triple convergence) or attract new traffic to the network (induced demand) only to the extent that social marginal benefits exceed social marginal costs, including costs internalised by charges. Congestion pricing would alleviate concerns about the futility of expansion of urban arterial road capacity, as well as alleviating traffic congestion.

4.1.3 Synergies from Combining Congestion Pricing and Arterial Road Improvements

Congestion pricing and selected improvements to the urban arterial road network are complementary elements of an efficient anti-congestion package. By-pass and ring-road capacity and debottlenecking of radial road capacity are likely to be particularly important urban road network improvements. Congestion pricing and selected road improvements would each enhance the efficiency and effectiveness of the other.

Because congestion has short-term and medium- to long-term dimensions, a single instrument is not enough to reduce congestion to optimal levels in the short-term and in the medium- to long term. Moreover, the suitability of instruments for pursuit of specific targets and the contribution of each instrument to pursuit of other targets should be considered. These important principles were pointed out in a general economic policy context by Jan Tinbergen (1952) and Bent Hansen (1955) around 60 years ago.

Congestion pricing could be set in the short-term to reduce congestion to the optimal level in the context of existing urban transport facilities. By moderating “triple convergence” and “genuine induced demand” effects, congestion pricing would release greater value from road investments. This was recognised in 1964 by the Smeed review of road pricing in the United Kingdom. Richard Arnott (2005) elaborated on this point. In December 2009, the Henry Tax Review (Henry, others, 2009, Part 2, p. 381) expressed a similar view. These observations appear in Box 9.

Box 9 **Congestion Pricing Adds Value to Urban Arterial Roads**

Smeed, Walters, Roth, others:

“In considering road pricing as a means of regulating traffic congestion, the panel have made the point that pricing by itself cannot produce a 'cure' for congestion. The proposed charge for use of congested roads should not be regarded as an alternative to new and better roads; it is rather a means of better value from the roads that already exist and from those that are yet to be built."

Richard Arnott:

“The approach of traffic engineers in the 1960s was to build our way out of the problem, by building more and better-engineered freeways, highways and roads. The standard cost-benefit rule employed was to expand capacity when the discounted value of travel time savings from doing so exceeds construction costs. Application of this rule results in efficient choices when prices are right. But urban auto travel was substantially under-priced. In this situation, application of the standard rule results in too much capacity. To see this consider the effects of an incremental road expansion. If traffic flow is fixed at the pre-expansion level, the expansion reduces congestion and trip price, inducing additional drivers to use the road. The increase in flow that accompanies an increase in capacity is known as latent demand. With under-priced congestion, trip price is below marginal social cost. The marginal social benefit from these additional drivers using the road is simply the trip price. Thus, the additional drivers who use the road add to the deadweight loss due to under-priced congestion, which dissipates the benefit from the road expansion. Latent demand therefore weakens the effectiveness of road building in reducing traffic congestion when congestion is under-priced, but not when it is properly priced. Traffic engineers should therefore support congestion pricing since it renders road construction more effective.”

Ken Henry and others:

“Introducing congestion pricing does not negate the need for expanded supply of roads in many cases, or other non-price measures. However, pricing is needed to leverage the value of urban road space, to ensure that investment in road capacity is put to its highest value use.”

Source: Smeed, Walters, Roth, others (1964), p. ii; Arnott (2005), p. 9.; Henry, others (2009), Part 2, p. 381.

Over time, optimal charges would need to rise on existing roads already priced and be applied to some other roads that were previously unpriced. Traffic would be further re-distributed and deterred as prices rose. Beyond the short-term, this would be aided by residential and employment re-location. Subject to these adjustments, government revenue from congestion pricing would rise.

High and rising revenues from existing roads would act as a signal that an expansion of capacity may be warranted. However, there is not a competitive road market, just a network controlled by a state government and a small number of local governments with considerable monopoly power. David Newbery and Georgina Santos explained that a government might exercise its monopoly power in pricing the metropolitan road network to exploit road-users. The result would be inefficient overpricing of, and underinvestment in metropolitan arterial roads. Therefore, Newbery and Santos argued that pricing of roads should be subject to economic regulation like network utilities to prevent the exercise of that monopoly power (Newbery, 2005; Newbery and Santos, 1999). The Henry Tax Review also made this point in the Australian context (Henry, others, 2009, Part 2, p. 406).

Then, decisions to invest in extra capacity of various types in various locations should be based on detailed social benefit-cost analysis. The analysis would include consideration of the appropriate pricing of the extra capacity in the context of differential network pricing and the appropriate pricing of the rest of the urban network in the context of the extra capacity. In some cases, such as provision of an outer by-pass or ring-road, low or zero pricing may be appropriate.

Appropriate road investments would increase the value from congestion pricing in relieving congestion and improving the allocation of resources. They would do so by addressing inefficiencies associated with monopoly pricing, and by providing substitutes for driving on existing crowded roads at peak times.

Provision of more capacity would reduce congestion, meaning congestion charges and revenue should fall initially. Because road investment is typically “lumpy”, excess capacity may exist in some locations initially. In those locations, charges would be zero. However, charges would become positive when traffic grows sufficiently for congestion to rise above optimal levels.

Congestion pricing could be applied to target reduction of congestion to the optimal level in the short-term, with existing urban transport facilities. Investment in urban road infrastructure and other urban transport infrastructure could be undertaken to target congestion alleviation in the medium- to long-term. These policy instruments would interact. At any point of time, congestion pricing would affect the appropriate investment for the medium- to long-term, while available capacity, reflecting past investment, would affect optimal pricing levels. Congestion pricing and road investments are complementary anti-congestion measures.

4.2 Congestion Pricing and Public Transport

Subsidised public transport has been widely advocated by transport and urban planners as a substitute for provision of road capacity and application of congestion pricing. The bases for this advocacy appear to be:

- the perceived futility of arterial road capacity improvements because of “induced demand” effects
- public transport’s superior people moving potential

- subsidised public transport changes effective relative prices (including value of time) of public transport and car-use, reducing demand for road space for cars
- subsidising public transport is politically more attractive than pricing an urban road network.

There are good reasons for challenging these views. The "induced demand" issue has to be faced in the case of public transport that is under-priced to alleviate congestion, as well as for under-priced roads. Public transport supply is much more likely to be an efficient and effective congestion-alleviating tool as a complement to (part of a package of policy instruments with) congestion pricing and arterial road improvements, than as a largely stand-alone policy instrument. Subsidised under-pricing of public transport is a poor substitute for congestion pricing.

4.2.1 Subsidised Public Transport and Induced Demand

To the extent that capital and operating subsidies for public transport attract passengers from cars, congestion is temporarily eased on radial roads. This lures some passengers back from public transport to driving on these roads, and attracts car-users from alternative routes, travel times and modes. Consistent application of induced demand theory reasoning indicates that congestion would tend to be restored despite more subsidies to public transport. Therefore, "triple convergence" and "genuine induced demand" or "latent demand" issues are as relevant to public transport policy as to urban arterial road provision policy in congested areas (Taylor, 2002, p. 13; Downs, 2004, pp. 85, 121; Small and Verhoef, 2007, pp. 173, 174; Duranton, Turner, 2011, p. 2634).

In a recent quantitative investigation of the "induced demand" theory, which was reviewed above, Gilles Duranton and Matthew Turner (2011, pp. 2644, 2645) found that "triple convergence" and "genuine induced demand" are triggered as strongly by subsidised public transport expansions as by increases in capacity of urban arterial roads. A related finding was that no evidence could be found that increases in subsidised public transport facilities and services affected vehicle kilometres travelled on urban arterial roads.

It is noteworthy that "triple convergence" and "genuine induced demand" or "latent demand" are also applicable to most "demand management" devices (see Box 10). These devices include travel option information programs, to induce people to drive less, parking levies and supply restraints in major activity centres, and staggering of working hours (Downs, 2004, pp. 85-87, 180; Smeed, Walters, Roth, others, 1964, p. 11; Small, Verhoef, 2007, pp. 173, 174). Congestion pricing is the notable exception. It would counteract, not trigger, "triple convergence" and "genuine induced demand" (Downs, 2004, pp. 86-87).

Box 10 **Demand Management Policies Trigger Induced Demand**

"……. the problem of releasing latent demand occurs not only with policies that shift the cost (*supply*) curve to the right (eg, road capacity expansion), but also with those that shift the demand curve to the left. Thus demand management policies (*excluding congestion pricing*) and improved transit, when implemented as measures to relieve congestion, are also vulnerable to induced traffic. Empirical evidence supports this conclusion. Sherret (1975) analyses trip rates on the Bay Area Rapid Transit (BART line) between Oakland and San Francisco during the first few months after it opened, and on the parallel San Francisco-Oakland Bay Bridge; he finds that the diversion of 8,750 trips to BART was soon followed by the generation of 7,000 new automobile trips, so that traffic levels at the busiest hours showed only small reductions"

Source: Small and Verhoef (2007), p. 174. Bracketed words in italics added by ACIL Tasman.

4.2.2 Potential Rationales for Public Transport Subsidies

Provision of heavy subsidies to public transport is common around the world. In many medium sized and large cities in Europe and the United States, public transport fares cover less than 40 per cent of operating costs.[8] In two cities in the United States, fares cover as little as 4 per cent of operating costs. Moreover, fares make no contribution to capital costs associated with public transport (Gomez-Ibanez, 1999, p. 111; Downs, 2004, p. 144; Parry, Small, pp. 700-701; 2009; O'Toole, pp. 221-222). In Australian metropolitan cities, subsidies cover 100 per cent of capital costs of public transport and up to about 77 per cent of operating costs.

The prevalence of public transport subsidies should not be taken as indication of their justification in economic terms. Pertinent arguments follow.

Subsidies and Economies of Scale and Density

Economies of scale in public transport have been cited as a justification for subsidies. Efficient allocation of resources requires pricing equal to social marginal costs in all markets, but economies of scale mean that marginal cost is lower than average cost, with the result that losses are incurred if marginal cost pricing is adopted. Therefore, subsidies are required to permit marginal cost pricing.

While economies of scale associated with fleet operation are exhausted with just a few vehicles, they could derive from large fixed costs and spare capacity associated with lumpy investments in exclusive facilities like railway lines, bus-ways and bus lanes (Mohring, 1999, p. 188; Gomez-Ibanez, 1999, pp. 100, 112). However, after analysing various studies of bus and passenger-rail systems, Jose Gomez-Ibanez (1999, pp. 112-113) concluded that such economies of scale are insufficient to justify large subsidies to public transport.

Herbert Mohring (1999, pp. 188-189; 1972, pp. 591-604) identified another source of economies of scale in public transport, which he called "economies of density". He explained that an increase in demand for service that leads to full capacity induces provision of additional services, which reduces waiting times between services. This reduces the effective marginal cost of public transport use for all passengers. The declining marginal social cost of public transport use associated with these "density economies" justifies subsidies, which according to Herbert Mohring would be "substantial". However, the magnitude of subsidies based on economies of density is subject to debate (Gomez-Ibanez, pp. 113-114). Ian Parry and Kenneth Small (2009, p. 700) pointed out that the magnitude of subsidies justified by economies of scale and density would depend on location, time of day, form of public transport and how public transport operators respond to increases in passenger demand (more routes, more services or higher loadings).

---

[8] New York City is the extreme outlier in the United States, with fares covering 70 per cent of operating costs. Las Vegas and Washington DC rank second and third in the United States with fares covering 41 per cent and 40 per cent of operating costs, respectively.

On balance, economies of scale/density may justify subsidies for public transport, but it seems unlikely that large subsidies would be warranted. The application of congestion pricing would reinforce that view, because it would trigger a cycle of higher demand for, and better economic performance of public transport allowing lower subsidies, as explained below.

Subsidies as a Second-Best Anti-Congestion Measure

An argument derived from the economic theory of the "second-best" has been suggested as a justification for additional subsidies for public transport. It starts with the observation that car-driving in congested conditions is priced below social marginal cost. This occurs because fuel and other motoring taxes are not closely linked to use of busy roads at peak times and associated congestion costs. The argument continues that if it is not possible to implement "first-best" pricing equal to social marginal cost, including congestion costs, the efficiency of resource allocation might still be improved, but to a lesser extent, by subsidising public transport, which is a substitute for driving. The idea is to lower the effective price of public transport patronage relative to car-use and thereby induce less road-use at congested locations and times.

The size of the "second-best" subsidy would depend on a multiplicity of factors (Gomez-Ibanez, 1999, p. 114; Parry, Small, 2009, p. 7000:

- the relative sensitivity of car and public transport use to the effective price of public transport services
- the extent of under-pricing of car use compared to social marginal cost
- the relative magnitude of car and public transport usage
- location
- time of day
- form of public transport
- public transport operators' responses to increases in passenger demand
- external costs of public transport operation, such as traffic congestion, passenger crowding and longer boarding and alighting times.

One reason why subsidising public transport is only "second-best" is that it increases the attraction of public transport relative to all alternatives. Some increase in public transport patronage will be at the expense of walking, cycling, driving off-peak, and driving on less congested routes. Another reason is that subsidies tend to increase the overall demand for travel, including peak period trips. These effects reduce the efficiency of resource-use.

A second reason for public transport subsidies being a "second best" strategy is that the "induced demand theory" is just as applicable to public transport subsidies as to increases in road capacity, as explained above. It seems the oft-repeated cliché used against road-building, "You can't build your way out of congestion", applies with equal force to provision of public transport facilities. However, its companion anti-road cliché, "Build it and they will come" does not apply to public transport infrastructure. According to eminent transport economist Kenneth Small, extensions of heavily subsidised public transport facilities have not attracted sufficient numbers from cars to make a significant dent in congestion anywhere, as stated in Box 11.

Box 11 Costly Failure of Public Transport Subsidies to Cure Congestion

"Observers of city life have long looked to mass transit to create urban vitality. Transit is supposed to promote a healthy high-density street life within economically vital business and retail districts, and to concentrate new developments into attractive patterns. Above all, it's supposed to limit road congestion without resorting to ugly high-volume roads everywhere.

These goals have been frustrated by the limited ability of mass transit to attract travellers out of automobiles and by the enormous expense of building and operating mass transit. While many recently built transit systems have achieved some desirable effects, none have seriously lessened traffic congestion. Furthermore, few cities have been able to afford a system extensive enough to make more than a small change in urban form; and the share of trips by mass transit continues to fall virtually everywhere."

Source: Small (2005), p. 10.

Similarly, Anthony Downs (2004, pp. 120-122, 138-141, 345-346) observed that enormous public transport subsidies in various countries had not been able to obtain significant cuts in the proportion of peak period trips by car. To achieve this, governments might need to cover all of the operating and capital costs of public transport and pay people an appreciable amount to use the service (Mohring, 1999, p. 192)

An economic study of 25 rail transit systems in United States' cities found that in all but one case, social costs of those systems exceeded the social benefits they provided. This conclusion held under various restructures of networks and prices (Winston, Maheshri, 2007; Winston, 2006, p. 70).

José Gómez-Ibañez (1999, pp. 114-117) pointed out that the available evidence suggested that in a medium-sized metropolitan area, where the sensitivity of car-use to the effective price of public transport services appears to be very low, and car-use is relatively high, significant subsidies to public transport to compensate for under-pricing of car use are not justified by "second-best" considerations.

In contrast, using data for three large metropolitan areas, Washington DC, Los Angeles and London, Ian Parry and Kenneth Small (2009) found that extending fare subsidies beyond 50 per cent of operating costs (and 100 per cent of capital costs), and often well beyond, is welfare improving at the margin across public transport modes, periods and cities. They explained that the finding was robust to plausible alternative assumptions in respect of parameters and agency behaviour. However, Parry and Small recognised that these results did not allow for two frequent, important criticisms of public transport subsidies. These criticisms were also highlighted by José Gómez-Ibañez (1999, p. 119).

First, they noted that some of the subsidy may be lost to management inefficiencies or captured by labour. They said that this was supported by empirical evidence. (A study by IPART showed that Sydney's rail service is less technically efficient than others in Australai, and recent NSW Government statements have focussed attention on the desirability of improve Railcorp's efficiency.). Therefore, their result was most applicable to public transport agencies provided with strong incentives to minimise costs.

Second, the marginal cost of public funds should be taken into account. Bigger public transport subsidies mean higher taxes, and cuts in government programs. Higher taxes mean greater economic damage (inefficiency or deadweight losses) from taxation. This varies with the type of tax, but is typically substantial. Cuts in highly valued government programs also mean less efficient use of resources.

In any event, the "second-best" argument for public transport subsidies depends critically on the premise that it is not possible to implement the superior option of congestion pricing. Parry and Small (2009, p. 722) concluded:

> "Finally, we reemphasise that transit subsidies are only a second-best response to automobile externalities. Our results should not divert attention from the much larger welfare gains to be had from pricing those externalities more directly. …….. Such gains are on an entirely different scale from those achievable by reforming transit prices."

This is consistent with earlier modelling undertaken by Ian Parry (2002), a prominent urban transport and environmental economist at Resources for the Future. Parry's modelling revealed that subsidising public transport provides only a tenth to a quarter of the gains to the community from a properly designed system of congestion charges.

Reasonably sophisticated congestion pricing is now feasible. Highly sophisticated congestion pricing is likely to become feasible in the foreseeable future. Therefore, pursuit of a "second-best" solution through public transport subsidies does not make economic sense.

Another important issue is the distributional consequences of public transport subsidies. They favour those who live in close proximity to the better public transport services and work in the central business district. Subsidised improvements to facilities and services deliver windfall gains to residents near train stations and boarding points for express bus services with lanes. Beneficiaries tend to be middle to high income groups. Others have to pay through higher taxes and fewer government services.

José Gómez-Ibañez (1999, pp. 117-118) commented:

> "Many of the benefits of low fares go to people who are not poor, and the poor who benefit might prefer alternatives. For example, in the largest U.S. metropolitan areas, the average household income of urban public transport users is similar to the average household income of all metropolitan residents because transit patronage is dominated by commuters to the central business district, many of whom are highly paid. In the smaller and lower-density metropolitan areas, by contrast, most public transit users are poor. The low densities usually make it expensive to provide convenient transit service; many riders might prefer to receive the subsidy in cash or in the form of a transportation voucher that could be used to buy a car, taxi rides, or other services besides conventional transit.
>
> In theory, one could target transit subsidies more efficiently by giving discount fares or transportation vouchers only to the poor. But such policies are seldom adopted because focusing benefits more narrowly also reduces political support."

Parry and Small (2009, p. 722) commented that for distributional reasons it might be appropriate to provide lower subsidies for public transport services patronised by higher income groups and higher subsidies for services patronised by lower income groups.

Similar distributional issues have arisen in Australia, where public transport subsidies mainly benefit higher income people. Australian Bureau of Statistics surveys of household expenditures have found that households with incomes in the top 20 per cent make public transport expenditures that are about three to four times those made by households with incomes in the bottom 20 per cent. It appears that this has occurred because high income suburbs located short to moderate distances to central business districts have better public transport services to those locations where higher income jobs tend to be concentrated.

4.2.3 Congestion Pricing Improves Viability of Public Transport

Kenneth Small (2004, pp. 134, 151; 2005, p. 12) argued that time- and location-variant congestion charges trigger a "virtuous cycle" of higher demand for, better services, and improved economic performance of public transport. There are short term and longer term aspects of this cycle.

The initial shift to public transport induced by congestion pricing facilitates better spatial and time coverage of public transport services, and greater service frequencies. Less congestion means faster on-road public transport services and lower operating costs. The improved services further increase demand for public transport. These effects allow increased capture of economies of scale/density, allowing lower fares, which attract further patronage. Improving viability arising from greater patronage and lower unit costs means public transport services can be expanded and subsidies reduced.

Meanwhile, the tendency of freed-up road space to attract public transport passengers back to cars ("triple convergence" and "genuine induced demand") is managed by congestion pricing of the urban road network. So, the cycle of improving demand for, and performance of public transport persists.

Changes in the full costs of driving in busy locations at peak times resulting from congestion charges and congestion reductions change the relative attraction of various residential locations. People with high values of time might be prepared to move away from major activity centres, particularly the central business district, because their travel costs fall. People with low values of time might decide to move closer to major activity centres, because their travel costs rise. Residential and business location in areas that are well-served by public transport would become more attractive. Land values would increase in areas that have become more attractive and decline in areas where attraction has diminished. Urban density would tend to increase, except in the city centre and urban fringe areas. Demand for public transport would rise, triggering a further cycle of higher demand for public transport, better services, and improved financial performance, allowing further cuts in subsidies.

A simplified form of congestion pricing applying in Singapore triggered such a "virtuous cycle" of higher demand for, better services, and improved economic performance of public transport. Public transport in Singapore is now profitable and the government continues to make unsubsidised improvements to infrastructure and services (Santos, Li, Koh, 2004, p. 231).

Of course, Singapore implemented congestion charging when public transport use was high relative to Sydney and other Australian cities. It also was assisted by a much higher degree of concentration of employment than is the case in Sydney. This contained the magnitude of investments in public transport and roads required to support congestion charging.

Kenneth Small (2005, p. 10), commented:

> 'Rather than mass transit (public transport) being the solution to congestion, perhaps congestion pricing – a measure often viewed as an alternative to transit – could be transit's saviour."

With time- and location-variant congestion pricing, there is no justification for subsidising public transport to alleviate congestion, but subsidies may still be justifiable to allow social marginal cost pricing in the context of economies of scale/density. Pricing all transport at the relevant social marginal cost will yield an efficient modal split (Proost, Van Dender, 2004, pp. 171-172, 176 fn 15; Small, Verhoef, 2007, p. 156). Stef Proost and Kurt Van Dender (2004, p. 167) added:

> "Policy makers often do not appreciate that the introduction of road (congestion) pricing may be an opportunity to correct (subsidised) public transport prices as well."

4.2.4 More Public Transport Services and Less Subsidies Complement Congestion Pricing

Public transport subsidies perform poorly in respect of economic efficiency and equity. However, public transport infrastructure and services comprise an important component of a package of complementary measures to tackle congestion. Just as selected road improvements, such as by-pass and ring-roads, provide through-traffic with alternatives to use of radial roads passing through or near major activity centres, public transport services provide an alternative to driving to those centres, particularly the central business district.

A network-wide congestion-pricing regime would trigger a "virtuous cycle" of higher demand for and greater viability of public transport, as occurred in Singapore. This would allow better services, lower fares and lower subsidies for public transport consistent with efficient pricing.

Congestion pricing would also effectively manage "triple convergence" on, and "genuine induced demand" for road space arising from provision of additional road and public transport capacity. The influence of the "induced demand theory" and fear of congestion charges have greatly helped governments to maintain their zeal for provision of public transport facilities and services. Rising subsidies have been a by-product of provision of more facilities and services and acquiescence to technical inefficiency in their operation. However, both the induced demand theory, when applied consistently, and congestion charges undermine the case for such subsidies.

4.3 Congestion Pricing and Information on Transport Options

4.3.1 Rationales for Urban Transport Information

Deficiencies in available information regarding benefits, costs and availability of travel options have been nominated as one cause of socially excessive car-use in major urban areas. Various education/information programs have been proposed to correct the perceived inadequacies. They range from dissemination of better, generally available information to provision of personalised advice.

Particular attention has been focused on schemes involving personalised travel plans, which involve provision of individualised analysis and advice directed towards encouraging shifts from single-occupancy vehicles to car-pooling, other modes of travel (particularly public transport) and working from home on occasions. These programs may be community-based or centred on work-places. Such schemes tend to be relatively labour intensive and therefore expensive, particularly programs involving the general public.

These programs have been promoted by governments and transport and urban planners as measures to alleviate traffic congestion, reduce vehicle emissions, increase fitness, and improve the viability of public transport. They have been described as (car-use) demand management measures, and regarded as complements to public transport subsidies. Together, information on urban transport options and subsidies for public transport have been depicted as a substitute for congestion pricing.

An economic rationale for provision of information on urban transport options could be correction of a possible information market failure regarding travel choices. This might arise from commuters' lack of knowledge of availability and benefits of alternatives to alternatives to car-use, search inertia, and search costs for individuals.

Of course, eliminating informational deficiencies would not remove the primary causes of socially excessive congestion and emissions. These are the absence of prices to internalise costs imposed on others through overuse of scarce road space and the atmosphere. These matters should be addressed whether or not an information market failure exists.

Moreover, it has not been demonstrated that travel behaviour has been significantly distorted by an information market failure. Existing travel behaviour could simply reflect consumer preferences in the context of underpricing of road space and the atmosphere.

Travel options education schemes are poor substitutes for pricing. It has been shown below that benefit/cost ratios of these schemes could be expected to be below one, even in the context of current substantial subsidies for public transport. In addition, without appropriate congestion pricing measures, travel information schemes would be undermined by "induced travel" in peak times and locations. To the extent that better information encourages travel mode and time changes, unpriced road space is freed-up at busy times and locations. This attracts drivers from other travel times, modes and routes (Downs 2004).

Travel choice information schemes are likely to be more effective as complements to congestion and emissions pricing. Together, they could address any information market failure interacting with congestion market failure, and "induced travel" effects would not be an issue. These points have been discussed further below.

4.3.2 Examples of Existing Information Schemes

Australian Schemes

TravelSmart is a personalised travel plan program developed by the Western Australian Department of Transport in 1997. The program has since spread to other Australian capital cities.

The Australian, Queensland, Victorian, Western Australian, South Australian, and Australian Capital Territory Governments have established the National Travel Behaviour Change (NTBC) project, which is based on *TravelSmart* principles. The total cost of the NTBC is approximately $18.3 million over the period 2008-2012, including funding of $6.4 million from the Greenhouse Gas Abatement Program. The Commonwealth has supported the NTBC project since 2003.

TravelSmart brings together many community and government based travel behaviour change programs encouraging people to switch from use of private cars. *TravelSmart's* most widely applied household travel behaviour change tools are based on either *Travel Blending*® or *IndiMark*®.

Travel Blending® involves an analysis of people's travel behaviour through the use of a 7-day travel diary. This is followed by detailed and customised suggestions on how their travel behaviour can be modified through a process of monitoring and feedback.

IndiMark®, which is similar to *Travel Blending*®, is a direct marketing method targeted at changing travel behaviour. It is based on the premise that a majority of commuters perceive public transport to be much worse than it actually is. The program involves four main steps and includes rewards or incentives for participants.

U.K. Schemes

In the United Kingdom, community-based, personalised information programs and work-place based schemes have been implemented.

National planning guidelines state that all planning applications with significant transport implications should be covered by a travel plan. So, firms expanding or relocating are likely to be required by local government to have a travel plan (United Kingdom, Department for Transport, 2008).

U.S. Schemes

Workplace travel plans were mandated in some parts of the United States from the late-1980s. However, businesses objected strongly, because of the costs the requirement imposed on them. Mandated work-place programs have been abandoned nearly everywhere in the United States (Victorian Competition and Efficiency Commission, 2006).

4.3.3 Effects of Existing Urban Transport Information Schemes

Some transport departments in Australia have reported 10-20 per cent reductions in urban vehicle trips or kilometres travelled by car, and increases of 20-30 per cent in public transport trips or kilometres travelled in metropolitan areas among TravelSmart participants who have persisted with the program. Reported increases in cycling trips among persevering participants have ranged from 6-90 per cent.

In the UK, following travel awareness campaigns, a program based on a targeted, personalised, customised market approach to reduce car-use was developed and piloted in several cities to reduce car-use. Reductions of car trips by about 11 per cent and vehicle kilometres by about 12 per cent were reported. Mandated work-place based schemes were reported to have reduced car trips to and from work by up to 15 per cent (United Kingdom, Department for Transport, 2008). A meta-analysis of reviews of work travel plans in the United Kingdom indicated an 11 percentage point reduction in car-use to travel to and from work from 64 per cent to 53 per cent (Bamberg and Möser 2007).

Taylor and Ampt (2003) reported that benefit/cost studies of the Dulwich, Adelaide and South Perth trials had produced benefit cost ratios of 5.7 and 11-15, respectively. This indicated that the schemes improve economic efficiency.

However, Taylor and Ampt (2003) emphasised that these results should not be considered representative of the whole community. They explained that the programs tapped into the desires of a significant minority wanting to make or be seen to make contributions to improve the environment and reduce resource consumption.

The results of the trials and programs were clearly influenced by self-selection bias. This was exacerbated by relatively small numbers of program participants. Moreover, it appears that results were included only for those participants who passed the initial screening stage and subsequently continued to participate in the program until the final stage.

Morton and Mees (2005) strongly criticised the conduct of *TravelSmart* trials in South Perth and Alamein, Victoria, arguing that there were potentially significant sources of systematic bias making these schemes seem more attractive than they really were. They said researchers and policy makers should be sceptical of claims made in respect of *TravelSmart* schemes, because there was no reliable evidence that they produced real changes in behaviour.

Moreover, none of the assessments of urban transport information programs or the critiques of those assessments took into account “induced travel” effects generated by travel options education programs in the absence of congestion and emissions pricing. “Induced travel” effects could ensure that benefits are negligible.

In addition, per participant costs of personalised urban transport information programs are high. They are estimated to be $43 to $82 per participating household (Taylor and Ampt, 2003; United Kingdom, Department for Transport, 2008). Some economies could be made in the case of work-place based schemes, but relatively high costs could still be incurred. Allowance would also have to made for the substantial marginal cost of public funds because information programs are funded by taxpayers.

Therefore, benefit/cost ratios could be below one. Also, expensive personalised schemes targeting particular areas may not be regarded as equitable.

4.3.1 Urban Transport Information Programs with Congestion Pricing

The benefits of travel choice information programs could be expected to increase in the context of congestion pricing (Downs 2004). Information and congestion pricing would be mutually supporting. Together, they would address interacting congestion and information market failures.

Better information would facilitate and enhance changes in travel behaviour triggered by congestion pricing. At the same time, congestion pricing would avoid benefits of information programs being undermined by “induced travel”. Pricing would also strengthen incentives for individuals to search for information to facilitate changes in travel behaviour. Then, less labour intensive, and therefore, less costly information programs on aggregate and per participant bases may be required to correct any information market failure and facilitate switching. Also, these programs would be required only on a short-term basis while urban travel patterns are being adjusted.

Broadly based travel options information programs would probably be regarded as equitable because information would be available to all, unlike current narrowly focussed, personalised programs.

4.4 Land-Use Regulations and Congestion Pricing

Transport and urban planners have typically advocated deployment of land-use regulations to increase land-use densities to allow increased frequency and density of public transport services and lower fares. They have also typically advocated increases in public transport subsidies to provide more frequent public transport services and a denser network of services to complement regulated increase in land-use density. They have envisaged that the increases in public transport services and lower fares supported by higher densities would reduce traffic congestion. This packaging of subsidised public transport services and regulated increases in land-use densities has been depicted as an alternative to congestion pricing as an anti-congestion policy instrument, with and without road capacity increases.

These views are based on confused thinking. A stinging critique of widespread deployment in the United States of a policy package comprising heavily subsidised public transport and regulated increases in land-use densities has been provided by Randall O'Toole (2007, 2009). He documented persistent widespread failure of these measures to alleviate congestion and substantial economic waste associated with their deployment.

As explained above, subsidising public transport is a poor substitute for congestion pricing, but better public transport services would complement congestion pricing well. As explained below, regulated increases in land-use density (with or without additional public transport subsidies) would be a poor substitute for congestion pricing and improvements in arterial road capacity and public transport services. However, a land-use regulation regime not impeding increases in density, rather than one inhibiting or mandating higher density, could complement congestion pricing with scaled down public transport subsidies.

Various studies have confirmed that clustering high-density housing near access points to good public transport services (transit oriented development) and raising the commercial density of central business districts and other major activity centres (regional activity centres) are likely to increase public transport usage. These studies have also indicated that increases in commercial density seem to be more important than higher residential densities. In addition, they have revealed that improving public transport serving residential areas and major activity centres encourages increases in density in both types of location (Downs, 2004, pp. 212-213, 227).

Increased demand for public transport arising from higher densities would allow operators to capture economies of scale/density with the result that fares could be lowered or subsidies reduced. Better public transport services would improve the viability of transit oriented developments and regional activity centres.

However, there are important provisos regarding the support and economies that transit oriented developments and public transport facilities can provide to each other. These problems help explain why few effective transit oriented developments have been built elsewhere (Downs, 2004, pp. 203, 211-212, 226).

First, political problems may arise because of local resistance to higher density development in existing residential areas, particularly because density has to be very high to provide a substantial boost to public transport use. This resistence derives from local preferences for lower density.

Second, transit-oriented developments would have to include substantial car parking for non-residents to provide a viable market catchment for public transport and other on-site facilities. To avoid detracting from the residential appeal of the development and its attraction to those wishing to walk to its facilities, car parking has to be provided above or below street-level. However, this would substantially raise the cost of establishing transit-oriented developments.

Third, the cost of upgrading or replacing infrastructure in established areas could be higher than the cost of providing new infrastructure in "greenfield" areas (O'Toole, 2001, p. 25; Productivity Commission, 2004, p. 136; Wood, 2004; Bruegmann, 2005, p. 139). Pertinent observations by Bernard Salt, a prominent Australian demographer have been re-produced in Box 12.

Box 12 **Infrastructure Costs: Re-development Vs Greenfields Areas**

"Metropolitan and even coastal planning since the turn of the century has steered policy in favour of urban consolidation……But neither the planning nor any other community has actually completed a study that measures the real or notional savings achieved by the implementation of urban consolidation over the continuation of urban sprawl……The consolidationists, of course, claim that no such study is needed. They don't need facts to prove what they already know and believe."

Source: Salt (2006), p. 26.

Fourth, the issue of who bears the cost of upgrading infrastructure in existing residential areas to cope with substantial increases in density at transit oriented development sites would have to be addressed.

Similar problems might be encountered when increasing densities of major activity centres in established areas. However, such problems have not been given adequate attention by proponents of regulated increases in land-use density.

Even more problematic is the effect of higher densities on traffic congestion. Key issues are the timing of effects and the amount and location of car-use.

Anthony Downs (2004, pp. 201-203) and Brian Taylor (2002. p. 14) explained that significant metropolitan-wide effects on transport patterns from regulated changes in land-use density would be achievable only in the long-term, because population growth and new residential development and redevelopment could occur only incrementally. This simply reflects the long lifetime of the existing housing stock and the inefficiency of accelerating its replacement merely to increase density. Downs (2004, p. 203) summarised:

> "In short, it is extremely difficult to increase substantially the average density of an entire metropolitan area – including existing settlements – through marginal growth or new in-fill development"

If governments wanted significant medium-term effects they would have to force the pace of re-development to establish transit-oriented developments and expand or establish regional activity centres. Significant <u>short-term</u> effects would <u>not</u> be possible and even if they were, it is likely that they would be extremely inefficient.

Anthony Downs (2004, pp. 399-401) explained that high densities of transit oriented developments and regional activity centres, and substantial car-use by travellers to and from these locations could mean intensification of <u>local</u> traffic congestion. Downs' analysis and his survey of other United States studies indicated that the proportion of residents of transit oriented development sites choosing to use public transport is likely to be less than 25 per cent and may be less than 20 per cent. If so, congestion in the vicinity of these sites would certainly intensify.

Moderate-density, metropolitan-fringe residential developments supported by public transport subsidies could be expected to suffer higher leakages from public transport to cars than would occur with transit-oriented developments. It is inevitable that such developments will need to be supported by good road links to major activity centres and by-pass road capacity.

Superficial analysis might lead to an expectation that an increase in the proportion of public transport usage associated with high density transit oriented developments and regional activity centres would reduce <u>regional</u> congestion. However, Brian Taylor (2002, p. 14), Anthony Downs (2004, p. 401), and Edward Glaeser and Matthew Kahn (2003, p. 35) have expressed strong doubts as to whether this occurs in practice. Their analyses of density and congestion in United States cities indicated that higher residential and commercial density typically means greater congestion, whether in older, central city areas, or in newer outlying areas. Brian Taylor (2002, pp. 14-15) explained that increases in urban density added to <u>traffic density</u> and therefore, congestion. This makes walking, cycling and public transport more attractive, but the extent of switching to other modes typically is not enough to offset the increase in traffic density. A nett increase in congestion is the inevitable result.

It is clear that policies to regulate higher residential and commercial densities in conjunction with provision of additional public transport policies are not effective ways of alleviating congestion. Mandating higher densities could change overall densities only very slowly. It could worsen congestion, rather than alleviate it. Meanwhile, a policy of increasing public transport subsidies in an attempt to alleviate congestion is likely to be futile and/or fiscally unsustainable.

Ultimately, even putting aside the fact that it is likely to be ineffective in reducing congestion, the basic problem with mandating increases in land-use density is that it is more likely to aggravate overall inefficiency than alleviate it. Absent other distortions, land-use density would be determined by relative prices of land and of capital, the costs of travel, and consumer preferences for higher or lower density. With relatively abundant land, and strong consumer preference for living space, it is quite possible the outcome would be characterised by low density settlement patterns, as has characterised Australian cities since the 19th century.

Thus, as Ergas (2012) explains, land is a normal good – its demand rises with income. So long as households value extending the urban fringe at more than its costs, the mere fact that those extensions involve increases in social overhead capital and in travel times should be neither here nor there. Similar considerations apply to housing density: so long as local residents, in opposing densification, face both the costs and the benefits to which their decision will give rise, there is nothing inherently undesirable about that.

However, there are factors that in practice, can induce an inefficiently high level of urban dispersion. Three such factors are especially important in Australia. The first and likely largest is tax preferences to owner-occupied housing. These amount to a sizeable subsidy to land use, as land is a large part of the purchase of housing.

A second set of inducements to excessive dispersion comes from the under-pricing of goods and services that are complements to extensive land use. These include local public goods, where homeowners do not face the incremental costs arising from their settlement decisions, and the absence of congestion charges on roads, which create a wedge between the social cost of an added resident at the city fringe and the private cost that added resident bears.

Third, misalignments between the distribution of costs and that of benefits distort decision-making in a way that could induce excess dispersion. In particular, the costs of densification, in terms of reduced amenity, typically fall on local communities; but given constraints on council taxing and spending powers, local homeowners are unlikely to capture much of the benefits. This can create incentives to oppose rezoning decisions even if those decisions' overall benefits exceed their overall costs.

Given these factors, it is not necessarily inefficient for governments to 'lean against' the pressures to population dispersion. What is questionable, however, is the efficiency of the ways in which denser settlement has been pursued. As a general matter, urban growth boundaries are not an efficient response to distortions such as the absence of congestion charging. Using a growth boundary to force development in a central city may avoid inefficient commuting, but it risks doing so at the expense of inefficient use of capital. In other words, densification policies operate by increasing the price of land relative to that of capital. This induces substitution of capital for land, most obviously in the form of high rise developments. It risks unnecessarily increasing costs, compared to the option of directly tackling congestion through congestion charging.

Moreover, urban growth restrictions will distort the supply of housing, not only reducing its aggregate amount, but also shifting its composition away from low-cost, low-quality homes. The costs of these policies therefore fall disproportionately on low-income households. While congestion charging also raises equity issues, the discussion above has highlighted ways in which those issues can be addressed without compromising the efficiency gains congestion charging can bring.

Overall, as Ergas (2012) puts it:

> "Land is a resource like any other; the goal of policy should be to ensure it is used efficiently. Policy ought to aim at maximizing land's unimproved value, which implies ensuring its allocation to most highly valued uses."

Rather than distorting that allocation, congestion charges are a more direct and far better way of requiring people to take into account the congestion and settlement pattern effects of their road-use decisions and the congestion consequences of their land-use decisions. Moreover, they do not over-ride local preferences and would eliminate misdirection of consumer preferences as a result of under-pricing of crowded roads. Also, congestion prices do not distort the price of land relative to the price of capital.

Clearly, regulating higher density is poor policy. In contrast, not impeding higher densities induced by congestion pricing would help alleviate congestion, improve public transport's services and viability, and reduce the burden of funding high public transport subsidies. Congestion pricing would increase the value of land close to regional activity centres, transport oriented development sites, train stations and bus stations/stops. This would encourage development or redevelopment of such land at higher densities. These changes in land-use density would increase density of demand for public transport, facilitating more frequent services and a denser services network, and improving financial performance, which would allow cuts in public transport subsidies. Such effects represent part of a "virtuous cycle" of higher demand for, and better economic performance of public transport triggered by congestion pricing (Small, 2004, p. 134), as discussed above.

5 Cost Recovery from Congestion Pricing

The purpose of congestion pricing (and complementary policy instruments) is to alleviate congestion to improve the efficiency of allocation of resources. The ideal structure would involve prices that vary across the urban road network and over time. Revenue to recover cost of provision of roads is a by-product of congestion pricing, but not the primary purpose. In contrast, the primary purpose of tolls on new roads in Sydney and elsewhere in Australia, has been recovery of costs, including risk-related returns to capital. Typically tolls do not vary with congestion or any determinant of it, such as time of day.

The better (or poorer) are the alternatives to congested roads at the time congestion pricing is implemented, the greater (or smaller) would be the reduction in congestion, the larger (or lower) the nett social gain, and the smaller (or greater) would be the revenue yield. Of course, high revenues arising from poor alternatives would signal the potential need for more capacity and provide resources to provide it. Therefore, the expected revenue yield of congestion pricing is a matter of considerable theoretical and practical interest, even though the primary purpose of congestion pricing is to alleviate congestion.

Herbert Mohring and Mitchell Harwitz (1962) demonstrated that under a range of assumptions, congestion charges equal to marginal external social costs (optimal congestion pricing) would yield revenue just sufficient to cover costs of providing and maintaining the urban arterial road system in the context of constant-economies (neither nett economies nor nett diseconomies of scale). In the case of diseconomies of scale, more than enough revenue would be raised, while there would be a revenue shortfall with economies of scale.

An important extension of the Mohring-Harwitz analysis was the inclusion road damage costs. David Newbery (1988, 1989), Kenneth Small and Clifford Winston (1988), and Kenneth Small, Clifford Winston and Carol Evans (1989) demonstrated that with plausible assumptions, broadly speaking, the sum of congestion and road damage charges would just cover the cost of providing and maintaining an urban arterial road network subjected to damage by heavy vehicles and congestion by all vehicles.

Subsequent substantial theoretical work has found the Mohring-Harwitz result to be surprisingly robust as assumptions were relaxed. The major exception related to the assumption of continuous small increments in capacity. Lumpiness or indivisibilities in capacity would cause the Mohring-Harwitz cost recovery theorem to break down (de Palma, Lindsey, 2007; Verhoef, 2007). However, this problem this may not be a problem in aggregate. Erik Verhoef (2007, p. 87) concluded in a comprehensive review of the relevant literature:

> "But even with lumpy capacity, pooling of surpluses and deficits across roads and over time would make aggregate self-financing less unlikely than for every individual road segment at every instant. The empirical evidence, insofar as available, suggests that the extent to which the constant-economies condition is fulfilled, necessary for exact self-financing, may vary over applications, but on the whole it appears to be a reasonable approximation. "

Complementary public transport improvements have to be funded too. We have already explained that using congestion charging revenue to subsidise public transport would waste resources. However, with congestion charging in place and public transport pricing simultaneously reformed, the viability of public transport would be significantly improved. Therefore, it would be better placed to support capital raisings on commercial terms.

6 Distributional and Acceptability Issues

6.1 Distributional/Acceptability Obstacle to Congestion Pricing

Inadequate public support for congestion pricing has been the most important obstacle to implementation of this anti-congestion measure. Lack of support has stemmed primarily from fears among a majority of voters that that they would be made worse-off by congestion pricing, and additional concerns regarding disadvantage to lower income groups.

Simplified forms of congestion pricing that had been proposed for Edinburgh and Manchester were rejected in referenda in 2005 and 2008, respectively. In 2007, following an online petition against road pricing, the United Kingdom Government shelved the concept of a nation-wide regime. In 2008, the New York state legislature blocked a plan by the Mayor of New York City to implement a simplified congestion pricing system in Manhattan.

Other governments, including some state governments in Australia, have shelved congestion pricing proposals and investigations because of concerns about community opposition that could be inflamed by scare campaigns by special interest groups.

In contrast, the Mayor of London implemented an area pricing version of congestion pricing despite strong opposition. The scheme subsequently gained majority support. This has been attributed to effective prior consultation and communication with the community, provision of substantial additional bus capacity in time for implementation of pricing, various exemptions and price concessions, commitment to earmarking of most of the nett revenue for public transport, and a significant reduction in congestion.

In the case of Stockholm, there was majority community opposition prior to a congestion pricing trial, which was inflamed by scare campaigns by the automobile association, chamber of commerce and a morning daily newspaper. A referendum following the pricing trial provided sufficient support for congestion pricing for the government to implement congestion pricing permanently . This change in public opinion has been attributed to effective communication, provision of extra public transport capacity before the trial, effective prior communication with the public, trouble-free implementation of the cordon pricing scheme, and a significant reduction in congestion. Community support for the Stockholm scheme increased after the referendum. This may have been influenced by the government's decision, in response to public feedback, to switch revenue earmarking to roads from public transport, which had already been upgraded substantially.

Governments might be concerned about distributional issues for at least two reasons. First, distributional changes may cause conflict with the ruling party's moral judgements regarding an equitable or fair distribution of income and wealth. Second, distributional changes may adversely affect a large block of voters, and others may be concerned about effects on road-users with low incomes, raising concerns about community acceptance of the changes.

In sub-section 3.5.3, it was explained that a congestion pricing regime that is sufficiently well-designed to yield nett social benefits would provide gains to winners that are more than large enough to compensate those who would be nett losers before use of government revenue from the regime. Of particular importance are the points that the implementing government would be the big winner, and its congestion pricing revenue would be adequate to compensate the losers.Therefore, an important key to dealing with distributional concerns and winning acceptance of congestion charging is astute use of revenue. This is widely acknowledged. An example of such an observation by economists based at Resources for the Future appears in Box 13.

Box 13 **Concerns about Low Income Car-Users**

"Since everyone pays the same charge regardless of income, there are concerns that low-income motorists will suffer disproportionately. Compensation of potential losers, by appropriate spending or other recycling of (congestion) toll revenues, may critically determine the overall political feasibility of congestion pricing. This has been confirmed by various surveys, where support for congestion pricing increased with explicit proposals for using the revenues in, for example, other tax reductions or investment in roads or public transport……a sufficiently progressive recycling of toll (congestion charge) revenues could ensure that all income groups benefit overall."

Source: Safirova, Parry, others (2004), p. 180.

6.2 Concerns about Effects on Retail Activity

The lobbying of business associations can have an important influence on political acceptability of congestion pricing. Congestion pricing proposals have typically attracted complaints from retailers that their businesses would be undermined. The available evidence from the Singapore, London and Stockholm provides little or no support for such complaints.

There is no evidence of adverse effects on the retail sector in Singapore. Indeed, adjustments to Singapore's electronic road pricing scheme in 2005 seemed to be designed to reduce congestion in the shopping precinct that was impeding retail activity and to redistribute traffic in a way that increased retail activity (Chin, 2010, pp. 9-10).

Quddus, Carmel and Bell (2007) analysed the effects of the London congestion charging on retail activity in the central London area, which includes but is larger than the current charging area. Analysis of sales data for the period 2002 to 2007 found a statistically significant negative effect for the John Lewis Oxford Street store, but no significant effect on overall retail sales in the central London area.

Effects on retail activity in central Stockholm were similarly neutral. In the case of Stockholm, Daunfelt, Rudholm and Rämne (2009) found that the congestion pricing trial did not have any measureable effect on retail activity.

6.3 Winning-over Potential Losers without Sacrificing Efficiency

Distributional issues could be resolved and acceptance bought in ways that either enhance or reduce the efficiency of allocation/use of resources. It is important to resolve distributional and acceptance issues in ways that do not detract from the economic efficiency gains from congestion pricing and preferably further improve the efficiency of resource allocation. This requires astute recycling or use of the revenue from congestion pricing and public transport subsidy savings yielded by congestion pricing (Schade, Schlag, 2003; Button, 2006a; de Palma, Lindsey, Proost, 2007a; Parry, 2009).

Congestion charging concessions and exemptions would not be astute. They would reduce the extent of alleviation of congestion and undermine revenue yield and economic efficiency. This occurred in London and Stockholm, particularly London. Moreover, Button (2006a, p. 237) doubted that heavy discounts for local car-using residents could be justifiable on equity grounds in central London, where incomes are high relative to the rest of the United Kingdom. Equity concerns are better dealt with by measures that enhance or do not detract from efficiency, rather than by exemptions and concessions.

While congestion pricing would have to be accompanied by improvements to public transport services, involving increases in capacity, service frequency, and density, it would not be astute to increase the already very large subsidies to public transport. As explained in sub-section 4.2, while public transport improvements and congestion pricing are complements, public transport subsidies and congestion charges are substitutes. Public transport subsidies (other than those related to economies of scale and density) are a form of "second-best" pricing, a substantially inferior substitute for ("first-best") congestion pricing. If congestion pricing was to be implemented, subsidies to public transport to alleviate congestion would be redundant. Persistence with such subsidies in the context of "first-best" congestion pricing would result in less congestion, more public transport use than optimal levels, and persistence of x-inefficiency involving disincentives to provide quality services at low cost. Resources would be wasted.

Persistence with public transport subsidies after implementation of congestion pricing may have adverse effects on equity, as well as adverse effects on the efficiency of resource allocation. As pointed out in sub-section 4.2.2 above, higher income groups benefit more from public transport subsidies than lower income groups.

Ideally, pricing of road-use to reflect marginal social costs should be accompanied by pricing of public transport services to reflect marginal social costs of provision. Cars, buses and trucks should bear charges consistent with congestion, road damage and other costs imposed on others, and public transport passengers should pay fares that reflect marginal social cost of provision of public transport services (Small, Verhoef, 2007).

Prior to implementation of simplified congestion pricing in London and Stockholm, promises were made to earmark nett revenues from congestion charges for public transport. While increases in public transport capacity were clearly appropriate, increases in public transport subsidies were not, for reasons explained above. However, these revenue recycling measures enhanced community acceptance of congestion pricing in London and Stockholm.

Experience with simplified congestion pricing in London and Stockholm has shown that community support can be won if pricing is packaged with appropriate alternatives to use of priced roads, supported by competent communication and implementation. In both cities, community support increased significantly following implementation, indicating the importance to the public of experiencing the operation of the system and complements to it. In the case of Stockholm, community support apparently increased further after the referendum because of the government's decision, in response to public feedback, to switch nett revenue earmarking to roads from public transport, which had already been upgraded substantially.

Because Singapore does not have the democratic tradition of the United Kingdom, Sweden and Australia, it has been argued by some observers that community acceptability issues would not have been as large an obstacle in Singapore as in these other countries. However, whether for acceptability or economic efficiency reasons, Singapore's government has complemented pricing with provision of high quality public transport facilities (not subsidised) and road infrastructure, following positive benefit-cost analysis results.

There is no doubt that congestion pricing should be accompanied by improvements to complementary infrastructure, such as by-pass roads, tackling bottlenecks and improvements to public transport to ensure that congestion is reduced to the optimal level in the short-term, medium-term, and long-term. Revenue from congestion pricing and savings from lower public transport subsidies could be recycled (not hypothecated) to fund investments in transport infrastructure that would complement congestion pricing and could survive screening by comparative social benefit/cost analyses. Alternatively, these revenues and savings could be applied to funding of cuts to taxes that otherwise particularly interfere with efficient allocation of resources.

If revenue gains were applied in these ways, opinion surveys in various locations have indicated strong majority support would be forthcoming for congestion pricing (Schade, Schlag, 2003; Sariforova, Parry, others, 2004; Santos, Li, Koh, 2004; Button, 2006a; Layton, 2002; Market & Communications Research, 2005).

In the Australian context, another astute move would be cuts in Commonwealth fuel or income tax to "make room" for congestion pricing. Road-users outside major urban areas would benefit from a lower tax take, without any offsetting congestion charges. Low value of time residents of major urban areas would pay more overall, if they incur congestion charges, but the amount of revenue available to state and local governments from these charges to upgrade transport infrastructure and perhaps reduce inefficient state road taxes would exceed the nett increase in congestion price and tax combined. The most inefficient of these state taxes are stamp duty on vehicle transfers and vehicle and third party insurance policies.

Not only would this tax and pricing reform alleviate distributional and acceptability concerns, but also it would provide economic efficiency gains. In addition to efficiency gains from alleviation of congestion, efficiency improvements would result from reduction of overall tax system inefficiencies and other tax and expenditure distortions caused by serious vertical intergovernmental fiscal imbalance (vifi) in the Australian federal system of government. Vifi would decline because there would be an effective transfer of revenue raising capacity from the Commonwealth to the states, and the additional state revenue would be raised in a relatively efficient way. Such astute action by the Commonwealth Government would simultaneously advance microeconomic, tax, and fiscal federalism reform objectives (Willett, 2006).

The Henry Tax Review appeared to support the idea of recycling congestion pricing revenue accruing to state and local governments through provision of complementary infrastructure satisfying comparative benefit-cost analyses, state cuts in stamp duty, plus Commonwealth fuel tax cuts. It was recommended that loss of Commonwealth fuel tax revenue be made up through relatively efficient broad-based taxes (Henry, others, 2009, Part 2, pp. 392, 398,-400, 406).

The Commonwealth Government has already set a precedent for tax cuts to make room for pricing of external costs of motoring by deciding to cut fuel tax to make room for carbon pricing. Sir Alan Walters (2002), a seminal contributor to the congestion pricing literature, proposed such an arrangement to accompany congestion pricing in the United Kingdom . Walters (2002, p. 40) argued:

> "This is a massive 'rebalancing' operation. At present anyone who travels on an uncongested road is paying far too much in petrol tax. Their vehicles do not get in the way and delay other road users. The real cost of using uncongested roads is very low and the high petrol duties are, in a pejorative sense, discriminatory against people who live and work in rural areas.
>
> In order to get political support for such proposals, there must be substantial reduction in the overall level of fuel taxes."

Sir Alan Walters arguments apply in Australia too, even though fuel tax is much lower here. A fuel tax cut to make room for congestion pricing could also have positive equity effects in metropolitan areas. It is most likely to favour those who make inter-suburban trips to work (limiting use of roads heavily congested before congestion pricing), and who have poor access to public transport, long drives to work, and older vehicles with poorer fuel economy. This would tend to benefit lower income groups more than high income groups.

6.4 Timing of Provision of Complementary Infrastructure

Another important issue in gaining acceptance of congestion pricing is the timing of provision of complementary infrastructure. Acceptability would be undermined if provision was delayed until substantial revenue and subsidy savings had been accumulated. It would be important to have adequate complementary infrastructure in place no later than commencement of congestion charging. However, that would present a funding problem.

A solution would be to borrow funds to bring forward construction of complementary transport infrastructure. The potential for congestion charging revenue and subsidy savings to service the debt was discussed in section 5.

6.5 Monopoly Pricing Inconsistent with Astute Revenue Recycling

David Newbery and Georgina Santos warned against a policy very different to astute recycling of revenue to gain acceptance for congestion pricing. Such a policy could involve exercising the government's monopoly power in respect of the metropolitan road network to exploit road-users in applying congestion pricing to raise revenue, rather than to reduce congestion to the optimal level in the medium- to long-term. The result would be inefficient overpricing of, and underinvestment in metropolitan arterial roads. Therefore, they argued priced roads should be subject to economic regulation like network utilities (Newbery, 2005; Newbery and Santos, 1999).

The Henry Tax Review made a similar point in the Australian context (Henry, others, 2009, Part 2, Detailed Analysis, p. 406):

> "Institutions are not currently set up to support efficient road pricing, nor efficient investment and operation of roads. Regulation will be necessary to ensure that road agencies do not exploit their market power in setting prices, and to ensure that prices reflect social marginal costs. Institutions will need to be vigilant to adjust prices in response to changing conditions."

6.6 Attitudinal Changes

Alex Anas and Robin Lindsey (2011, p. 81) reviewed various studies of attitudes to road pricing. They explained that it had been found that people became more favourably inclined toward road pricing after experiencing its benefits. This had been observed in the case of the London and Stockholm congestion pricing regimes, HOT lane and tolled express lane schemes in the United States, and Norwegian (cost recovery, not congestion alleviation) toll rings.

Anas and Lindsey (2011. p. 81) said that evidence suggested it was best to hold a referendum after a congestion pricing trial, as in Stockholm, rather than before proposed implementation of congestion pricing, as in Edinburgh and Manchester. A similar comment was made by Schuitema, Steg and Forward (2010, p. 108). However, this view seems to overlook the point that the Swedish Government could have chosen to reject congestion pricing on the basis of equivocal results of the referendum. While a small minority of voters in Stockholm city supported congestion pricing, a solid majority in the remainder of the Stockholm metropolitan area voted against it (Schuitema, Steg and Forward, 2010, p. 102).

An alternative approach would be to follow the lead of London City Council, which proceeded with simplified congestion pricing without a trial and referendum. This would require the political courage to ignore the public opinion polls and the noise of vested interests, in anticipation of a turn-around in attitudes following successful implementation of congestion pricing.

7 Tailoring an Anti-Congestion Package including Congestion Pricing to Sydney's Circumstances

Sydney has a serious and worsening traffic congestion problem. Packages of policy instruments deployed previously have not been effective and efficient means of addressing this problem. Some of these instruments have been deployed as substitutes for congestion pricing. They have been poor substitutes. However, revised versions of some of these instruments would be good complements to congestion pricing.

Congestion pricing is a necessary condition for alleviation of congestion in a way that is both effective and efficient (consistent with an efficient allocation/use of resources), but it is not sufficient for efficient alleviation of congestion, at least in the medium- to long-term. A package of complementary policy instruments would be required to alleviate congestion in a way that is effective and efficient in the medium- to long-term. Complementary measures are also essential for an equitable and politically acceptable outcome.

Transport for NSW (2012) presented a smorgasbord of anti-congestion policy options in its *NSW Long Term Transport Master Plan Discussion Paper,* although the paper focussed particularly on infrastructure provision. The paper did not discuss prioritisation of these measures or how they might complement or substitute for each other. The list included:

- road upgrades, including provision of new links, additional capacity on some existing road segments, and debottlenecking
- clearways in peak periods
- public transport upgrades, including provision of new facilities and routes and debottlenecking parts of the network
- re-allocation of road space to light rail, buses and other vehicles carrying multiple passengers
- carpooling
- public transport information programs
- facilities for cyclists
- flexible working hours
- various electronic aids to facilitate traffic flows
- HOT lanes
- cordon and area pricing schemes
- distance-based road pricing
- location-based road pricing
- time-variable road pricing
- parking space levies
- integrated planning of land-use, roads and public transport.

There would be several important issues to be addressed in formulating a congestion pricing regime and complementary anti-congestion policy measures for Sydney. Some issues are interrelated. Considerable thoughtful analysis would be required to resolve the various issues.

Introductory discussions of some of these issues follow. Recommendations on the best course of action for establishment of congestion pricing in Sydney have not been provided. It would be beyond the scope of this discussion paper.

7.1 Efficiency Vs Simplicity

7.1.1 General Principles

An ideal congestion pricing regime from an economic efficiency perspective would operate on a network-wide basis and involve variable charges. The price variations would reflect changes in marginal external congestion costs as the degree of congestion varies over time, across locations, and with the occurrence of events, such as severe weather conditions, crashes, and major sporting activities. In some cases, prices might also vary between lanes on the same road to accommodate heterogeneity of road-users.

Departures from the theoretical ideal involve sacrifice of some potential efficiency benefits. This sacrifice should be compared with any cost savings (including risk reduction) from simplification that would accrue to government and road users. Pursuit of simplicity or pragmatism is worthwhile only to the extent that it provides administrative and compliance savings that exceed losses from less precise targeting of deadweight losses (inefficiencies) from traffic congestion. Pursuit of simplicity for simplicity's sake would not be sensible.

Analyses of highly "simplified" (much less sophisticated than the ideal) or "pragmatic" congestion pricing regimes, such as those applying in Singapore, London, and Stockholm (discussed in section 3) have shown that important economic efficiency gains can be made by applying regimes that are not close to the ideal. While simulation studies have indicated that important economic efficiency gains would be achievable from application of more sophisticated charging systems in major urban areas, practical issues typically have not been taken into account. Congestion pricing regimes approaching the theoretical ideal have not yet been implemented anywhere.

It seems to have been taken for granted (not estimated) by policy makers that, from a social perspective, the benefits of the chosen degrees of "simplification" (departures from sophistication) would exceed the costs (efficiency losses) of moving away from network-wide, time- and location-variable pricing. The adoption of this assumption may have been influenced by two considerations.

First, while there are technologies that could be adapted and deployed to support network-wide variable congestion pricing, involving variations in prices according to the degree of congestion, time, location, distance, and vehicle type, this has not yet been done anywhere. There would be challenges and risks in adapting and deploying available technologies to support such a scheme. In contrast, "simplified" congestion pricing supported by less sophisticated technologies has already been implemented in Singapore, London and Stockholm. Therefore, such "simplified" congestion pricing involves fewer design challenges and risks for government officials.

Second, it has been suggested that a "simplified" scheme is necessary because road-users lack the cognitive ability to respond sensibly to prices that are finely differentiated by time and location, particularly if prices vary from day to day as well as within each day. However, others have cast doubts on this view. Research by Peter Bonsall and others (2007, pp. 672, 681) indicated that people preferred simple pricing structures, but they are able to respond sensibly to quite complex price structures provided they are clear and logical, and have been properly communicated to road-users. Jonas Eliasson (2010, pp. 6-7) argued that policymakers often seemed to underestimate cognitive abilities of people. He cited examples of road-users being able to grasp and adapt to apparently complex pricing arrangements on tolled express lanes subject to continuously variable pricing ("dynamic pricing"), and under Singapore's multi-zone cordon pricing regime involving pricing that varies by time and location in schedules that are adjusted every three months and for school holidays.

Jonas Eliasson (2010, p.7) added:

> "Forcing the system to be too simple too early in the design process is likely to cause design restrictions that are difficult to solve. The reluctance of many politicians and planners to consider 'too complicated systems' can lead to the point where the system becomes so simplified that it will not deliver the promised congestion reduction. This will not only be a waste of resources – it will also lead to low acceptability of the charges."

Pursuit of simplicity for simplicity's sake can be counterproductive. A quest for "simplicity" in congestion pricing may add to complexity, if not carefully thought-through by policy makers. Departures from network-wide variable pricing require difficult (certainly not simple) decisions regarding roads and parts of a metropolitan area that should be included in or excluded from the pricing regime. Considerable care is required to ensure that these decisions do not inadvertently shift congestion (not just traffic) from some locations and times to others. A related issue is that "simplification" decisions may perversely add to complexity by necessitating various initial adjustments and ongoing evolutionary modifications to deal with overlooked traffic diversion effects. This is an aspect of the sacrifice of potential economic efficiency benefits through "simplification".

7.1.2 Relevance to Sydney

The pitfalls of jumping to the conclusion that, on balance, "simplification" *per se* is advantageous apply in Sydney, just as they do elsewhere. A decision to opt for a "simplified" congestion pricing scheme in Sydney should be based on thorough comparative analysis of implications of different forms of "simplification" in the context of Sydney's particular circumstances, not on assumptions. Appropriate benchmarks for the comparative analysis should be business as usual policy and sophisticated network-wide variable congestion pricing with complementary measures.

A benchmark involving network-wide variable congestion pricing is a genuine option, not simply a theoretical ideal. In 2003, the *Ministerial Inquiry into Sustainable Transport in New South Wales* headed by Tom Parry (2003, pp. xxi, 79) proposed application of a network-wide, variable congestion pricing regime (potentially in combination with road damage pricing) in "Greater Sydney"[9] within 5-10 years, to be accompanied by rationalisation of existing taxation of motorists. The Ministerial Inquiry argued that cordon and area pricing schemes, the most commonly discussed "simplified" versions of congestion pricing, were not suited to Sydney's circumstances. The argument was extended to multiple, geographically dispersed, non-concentric priced areas. It was based on employment locations and journeys to and from work being widely dispersed, and the radial nature of the public transport network in Sydney (Parry, 2003, pp. 75-81).

One form of "simplification" or policy pragmatism that should be included in the comparative analysis is to put in place complements to congestion pricing while working on or waiting for others to undertake the task of adapting global positioning system technologies to support sophisticated network-wide variable congestion pricing. The benefits of this form of "simplification" would include the option values of waiting and taking incremental steps in the context of uncertainty, a substantial degree of investment irreversibility, and flexibility in timing of actions that characterise anti-congestion policy and associated investments. Avinish Dixit and Robert Pindyk (1994) and other contributors to the "real options" literature have shown that option values arising in such circumstances are important and should be taken into account in assessing and planning investments, policy measures, and the timing of these initiatives.

7.2 Spatial Configuration and Coverage

7.2.1 General Principles

Network-wide variable congestion pricing is the ideal form of congestion pricing from economic efficiency and effectiveness perspectives. It also has a substantial practical advantage over "simplified" schemes. It does not require decisions to be made regarding spatial configuration and coverage – where boundaries should be drawn, which roads should be included, whether to define a single cordon concentric cordons, adjacent cordons, or spatially dispersed cordons, etc. However, network-wide variable congestion pricing would involve practical issues relating to risk and cost of adapting and deploying suitable existing technologies.

Dedicated short range communications (dsrc) technology and automated number plate recognition (anpr) technology have been successfully deployed for Singapore's multi-zone cordon pricing regime (dsrc), London's area pricing system (anpr), Stockholm's cordon pricing regime (anpr), and "dynamic pricing" (prices changing every few minutes according to traffic conditions beforehand) of HOT and tolled express lanes in the United States (dsrc). These technologies have also been used to support distance-linked tolling. However, it may be very costly to deploy these technologies for a network-wide variable congestion pricing system, because of the large number of detection facilities required.

It is widely accepted that global positioning system technology has great potential as a support for network-wide variable congestion pricing, as well as for other purposes such as distance-linked insurance and navigation. However, the technology has not yet been adapted and deployed to support

---

[9] "Greater Sydney" was defined by the Ministerial Inquiry to include Sydney, Wollongong, the Blue Mountains, the Central Coast, Newcastle, and parts of the Hunter, Southern Highlands and Shoalhaven regions.

congestion pricing in practice (A brief outline and discussion of this and other technologies suitable for congestion pricing have been provided in Appendix A.).

Therefore, some entity would have to take on the risk of adapting, refining and deploying the technology to support network-wide variable congestion pricing. The alternative is to wait for another entity to do it, before deciding to proceed. The incentive to wait could be strong. This would delay application of the technology, and cause uncertainty regarding timing of availability of suitable proven systems.

A decision to eschew a network-wide variable pricing arrangement for "simplified" congestion pricing would mean decisions have to be made regarding spatial configuration and coverage of the scheme. The "simplified" options include area pricing (pricing presence within a cordon) as in London, pricing for crossing a single cordon as in Stockholm, a multiple zone cordon pricing approach as in Singapore (includes pricing of some road links to the three cordoned zones), pricing of crossings of two or more concentric cordons, a combination of area and cordon pricing, which was the initial congestion proposed for New York by its mayor (subsequently changed to a single cordon, then abandoned), and pricing of selected urban arterial road links.

Bernhard Oehry (2010, pp. 15-16) stressed that the success of a cordon or area pricing scheme depends critically on selection of a boundary or cordon that captures a large fraction of commuter traffic, without disrupting traffic that is less relevant to congestion. "Logical" boundaries include natural barriers, such as the waterways surrounding Stockholm, or built structures, such as the ring road around central London. These "logical" boundaries were actually adopted for the Stockholm and London schemes, respectively.

The relationship between such features and other aspects of natural and built topography is also important. Stockholm's natural boundary and river crossings meant only 18 boundary check points were required for its cordon pricing system. Singapore's topography and road network meant that only 15 boundary check points have so far been established for that city's multi-zone cordon pricing system. In London, 174 boundary checkpoints have been established because of different topographical and road network features.

Robin Lindsey (2008) investigated how "simplified" congestion pricing regimes using technologies already established for congestion pricing (dsrc and anpr), might be designed for three Canadian cities experiencing serious and worsening congestion, on the assumption that government would not be prepared to adapt global positioning system technology to support network-wide variable congestion pricing. He explained that topographical differences between these cities meant that a different "simplified" scheme would have to be applied in each case.

Lindsey (2008, pp. 246-247) argued that a cordon or area pricing arrangement would be well-suited to Île de Montréal, an island of 495 square kilometres upon which the city of Montréal is located, and which is linked to the mainland by 17 bridges and a tunnel. Lindsey based this view on the existence of a natural cordon provided by the surrounding waterways, limited access points provided by the river crossings, and congestion across the entire road network.

In contrast, Lindsey (2008, pp. 258-262) argued that a cordon or area pricing arrangement would not be suitable for Vancouver, because there is no natural cordon (in contrast to Île de Montréal and Stockholm), congestion is not concentrated around the central business district, commuting patterns are widely dispersed because of relatively decentralised employment patterns, and these patterns were strengthening. However, he argued the case for road pricing in Vancouver was as strong as for Montréal. Lindsey could only suggest variable tolling of all of greater Vancouver's major bridges (22) and tunnels (20, which he said "would intercept a large fraction of total traffic."

In Toronto, like Vancouver, congestion is not concentrated around the central business district, commuting patterns being widely dispersed. Also, like Vancouver, Toronto does not have a natural boundary for a cordon or area pricing scheme. However, unlike Vancouver, Toronto does not have a set of bridges and tunnels involved in most trips, other than those made over short distances. In this case, Lindsey (2008, pp. 262-267) suggested conversion of Toronto's 450-kilometre network of HOV lanes to HOT lanes with dynamic pricing (variable pricing that adjusts quickly to traffic conditions experienced just beforehand).

7.2.2 Relevance to Sydney

Determination of the suitability of "simplified" congestion pricing systems in Sydney would require a detailed analysis of metropolitan Sydney's road and public transport networks, passenger and freight movements, passenger transport the relative merits of mode shares for various routes origination points and destinations, demographic trends, and trends in residential and commercial location decisions. It also would require access to a sophisticated traffic model that would have to be modified to accommodate the influence of price signals. These activities are beyond the scope of this paper.

However, some indication of the suitability of "simplified" congestion pricing systems in Sydney, relative to the benchmark of network-wide variable congestion pricing, can be gleaned from high level assessments of transport issues in Sydney provided by:

- Transport for NSW (2012) in NSW Long Term Transport Master Plan Discussion Paper
- GHD (2011) in Preliminary Transport Baseline Report for NSW
- Tom Parry and others (2003) in Ministerial Inquiry into Sustainable Transport in New South Wales.

Box 14 contains excerpts from the discussion paper released by Transport for NSW (2012).

A review of the very concise information in the *Preliminary Transport Baseline Report for NSW* prepared by GHD (2011) indicates that by 2016, heavy congestion (vehicle/capacity ratios of 0.8 or above) is expected on:

- the M4 motorway (city West Link)
- the M5 (Liverpool-airport area)
- part of the M2 (Blacktown-Lane Cove Road)
- Pennant Hills Road

- Princes Highway south of Hurstville
- the Eastern Distributor from the airport area to the city
- parts of the Pacific Highway around Chatswood
- Military Road/Spit Bridge
- two north-south roads, Pymble-Ryde—Strathfield-Hurstville, and Epping-Bankstown
- road segments within and around the inner city and North Sydney.

There are also many road segments with little or no congestion expected by 2016. A notable example is most of the M7 motorway, which provides the western segment of the Sydney orbital network.

Box 14 Transport for NSW Assessment of Transport Issues in Sydney

Strategic Transport Corridors

"Transport for NSW has developed a system of 46 strategic transport corridors. Each corridor may include a number of different modes of transport, such as road transport, trains, or light rail. The corridors are the primary links between Sydney's major centres. They include a combination of existing routes and also transport corridors that will be important in the future.

Transport for NSW has reviewed the performance of each of these corridors, based on morning peak load factors for the rail network, travel speed and reliability for buses, and the volume and capacity on the road system. At present, five corridors are highly constrained and 11 corridors have medium constraints. ……… The areas of concern are the city connections to the Northern Beaches, Inner West, and Sydney Airport, as well as Macquarie Park to the North West Growth Centre, and the airport links to the west. These corridors highlight the extent of the challenge to make the transport network perform better in the future and clearly indicate where future investment is needed."

Network of Motorways and Major Roads

"The Sydney road network consists of motorways and freeways, arterial roads and local roads. Arterial roads provide the essential links between local precincts and the motorway network. The established network of major roads includes the 110-kilometre Sydney orbital network. This comprises the M7, M2, the Lane Cove Tunnel, the Sydney Harbour Tunnel, the Eastern Distributor, the M5 East and M5. There are motorway connections to the regional road network formed by the F3 to the north, the M4 to the west, the F5 to the south-west and the F6 to the south. This network connects the major commercial centres running from Port Botany through the Sydney CBD to Macquarie Park with north-west and south-west Sydney. The motorway network is heavily used throughout the day and night by freight vehicles, commercial traffic and commuters.

Sydney's road system and the motorway network are vital to the economic development of Sydney and the State. They provide access to jobs and links to Port Botany and Sydney Airport. They support the major freight task required to service the needs of Sydney and beyond. To do this effectively, the motorway network needs to be well connected to the major traffic generating precincts and it needs to have the capacity to meet the demand for those trips that cannot be made by public transport.

Apart from the Sydney orbital, the current motorway network has a number of vital missing connections. There is currently no motorway connecting the growing employment lands and population along the M4 with the Sydney CBD, Port Botany and Sydney Airport. There is no motorway connection between the Sydney orbital and the F3 in the north and no motorway connection to the F6 in the south.

There is heavy congestion in peak hours on significant sections of the orbital. This indicates that current capacity falls well short of the needs of a growing economy. Widening the existing motorway network and building new connections can provide for faster travel across the city. An extension of the M4 would provide a motorway standard link between the Western Sydney Employment Area, Parramatta, Sydney CBD, Sydney Airport and Port Botany. Duplicating the M5 East would alleviate congestion near Sydney Airport and Port Botany and allow quicker journeys to warehousing and freight distribution locations in South West Sydney. An F3 to M2 link would provide motorway standard travel across Sydney connecting to the Central Coast and Newcastle.

The motorway network can potentially be managed more effectively to get the best possible performance out of the existing infrastructure, by managing the spread and minimising the growth in congestion."

Sydney City Centre

"More than 300,000 people work in the city centre compared to 47,000 in North Sydney, 43,000 in Parramatta and 39,000 in Macquarie Park. …… On a typical weekday, the Sydney city centre grows from about 60,000 to half a million people representing the highest concentration of people and trips in the country. …. More than 500,000 trips are made to the city centre each weekday, including over 200,000 people arriving in the two hour morning peak. Approximately three-quarters of these people use public transport.

The need to move so many people to a single location in a narrow space of time places strain not only on the transport network in the city centre, but on the whole metropolitan system. There is a substantial amount of internal travel with over 1.3 million trips (1.2 million walking) made each day within the city centre. Bus and car congestion in the Sydney CBD is already a problem and CBD access presents a challenge for train services as passenger numbers grow."

Sydney Airport and Port Botany

"Sydney Airport accounts for around 45 per cent of Australia's international passenger movements and airfreight tonneage. Port Botany is Australia's second largest container port, handling about one-third of all containerized cargo shipped into and out of Australia. …….. The increase in traffic movements from both the airport and the port will place additional pressure on the ground transport networks that feed these gateways. Accommodating movements between Sydney Airport, Port Botany and Western Sydney will be particularly difficult on existing networks, but is essential to support growth in the State's economy."

Other Major Activity Centres

"Sydney's largest regional cities are Parramatta, Liverpool and Penrith. Around 40 per cent of employment in Sydney is based in these and other major centres, particularly those located in the eastern part of the metropolitan area. Employment in regional cities and other major centres is forecast to increase by 25 per cent between 2011 and 2031. In some other parts of Sydney, low density residential development and a relatively low proportion of centre based employment combine to create car dependency and associated problems such as social isolation."

Growth Areas

." Connecting residential areas, particularly growth areas, to centres of employment, health and education will be important in increasing community wellbeing and quality of life across Metropolitan Sydney."

"Over the next 20 to 30 years, the South West Growth Centre will be home to more than 300,000 people with a further 200,000 living in the North-West growth centre. The NSW Government is already constructing the South West Rail Link and is committed to building the North West Rail Link, which is in the detailed planning stage. Ensuring these train connections deliver accessible, well-connected and integrated transport are significant challenges in for these areas. The NSW Government is also expanding the M2 and M5 motorways to help improve access to these areas."

Source: Transport for NSW (2012), pp. 37-39, 45-46.

These high level reviews indicated that congestion is not focussed in one geographical area, and it does not derive from travel to access one geographical area. Congested road links and locations are widely dispersed. This appears to be linked to wide dispersion of employment locations.

The available information supports the advice of the *Ministerial Inquiry into Sustainable Transport in New South Wales*, which was released in December 2003. The authors argued that a cordon or area pricing arrangement for inner Sydney's central business district would not be effective in reducing congestion, as only a small proportion of traffic on major roads would be charged, because traffic on congested arterials is going to multiple destinations. Moreover, the authors pointed out that a single cordon could have perverse outcomes, such as inducing relocation of offices to locations well away from the central business district that are not well served by the radial public transport network. This would encourage growth of car-use, leading to greater congestion (Parry, 2003, pp. 73-74).

This might suggest multiple cordoned areas, based on major activity centres, for cordon or area pricing in the Sydney metropolitan area. However, the report of the Ministerial Inquiry identified problems with this concept too (Parry, 2003, pp. 76-77):

> "…the Sydney CBD, Inner Sydney and North Sydney areas are well serviced by public transport (reflected in the share of journeys to work accounted for by public transport). However, these areas only account for 25 per cent of employment positions in the greater metropolitan region. …….. the radial nature of Sydney's public transport network results in greater reliance on private transport for getting to and from work the further out from the inner city. Unless commuters are on the transport corridor, a viable alternative to private transport is unlikely to be available today. ……. If implemented in major employment areas, what will cordon (or area) pricing achieve? Given the absence of circumferential and viable alternative public transport, cordon pricing is unlikely at least in the short- to medium-term, to see a marked reduction in private transport."

Another issue with pricing based on a single or multiple cordons is the absence of natural features or built structures in metropolitan Sydney (unlike Stockholm, Singapore and Montréal) that would facilitate establishment of a cordon or cordons with relatively few access points to be controlled. This is an important requirement for a cordon-based pricing system without high administration (establishment and operating) costs. Lyn Martin and Peter Kain (2008, p. 120) of the Bureau of Infrastructure Transport and Regional Economics noted that this was a potential problem for a Sydney cordon-based scheme. The importance of natural features and built structures for cordon-based schemes in general has been emphasised by Bernhard Oehry (2010) and Robin Lindsey (2008) as explained in the preceding sub-section. It was not discussed by the *Ministerial Inquiry into Sustainable Transport in New South Wales*.

An additional issue not discussed by the Ministerial Inquiry is that multiple dispersed cordons would not be consistent with the pursuit of "simplicity" underpinning adoption of cordon and area pricing systems. Multiple dispersed cordons would involve greater risk of shifting congestion. Also, consequent initial and ongoing evolutionary modifications to deal with this problem could perversely add to complexity, relative to the benchmark of network-wide variable congestion pricing.

The Ministerial Inquiry's report described cordon and area pricing schemes as "second best" measures. It proposed application of network-wide, variable congestion pricing (potentially in combination with road damage pricing) in "Greater Sydney"[10] within 5-10 years, to be accompanied by rationalisation of existing taxation of motorists. It pointed out that its preferred congestion pricing regime would allow more efficient finer differentiation of pricing than cordon-based systems, and could be extended to include pricing of other external costs, as well as avoiding the deficiencies of cordon-based systems discussed above. The Ministerial Inquiry proposed re-negotiation of existing tolling arrangements, and phasing-in arrangements (Parry, 2003, pp. xxi, 77-81).

---

[10] "Greater Sydney" was defined by the Ministerial Inquiry to include Sydney, Wollongong, the Blue Mountains, the Central Coast, Newcastle, and parts of the Hunter, Southern Highlands and Shoalhaven regions.

The Ministerial Inquiry envisaged that network-wide variable congestion pricing would be supported by dsrc or anpr technology. However, it acknowledged that a "vast" number of detection points would be required to avoid circumnavigation of these points. This would limit the arterial roads that could be included in the pricing regime "until this problem can be overcome" (Parry, 2003, p. 77). The Ministerial Inquiry did not suggest how the problem might be resolved.

Two potential solutions would be to work on, or wait for others to undertake the task of adapting global position system technologies to support network-wide variable congestion pricing. In the meantime, the time consuming and expensive task of upgrading or otherwise providing complements to congestion pricing could be undertaken.

A variant of that approach would be to phase-in congestion pricing via variable cordon-based pricing, involving geographically dispersed cordons, plus variable pricing of selected links with cordoned areas. This would be supported by established anpr or dsrc road pricing technologies. The phasing-in would follow provision of infrastructure complements to pricing. Later, coverage would be expanded through application of global position system technology to support network-wide variable congestion pricing.

A precedent for an evolutionary approach to congestion pricing has been provided by Singapore. It started with a paper-based area pricing regime, moved to variable pricing of crossings of a single cordon and selected arterial roads, based on dsrc technology, and then to differential pricing of three adjacent zones within the original cordon and refinement of pricing coverage of arterial roads. In the meantime, roads and public transport were progressively improved, with major upgrades occurring before significant changes to the spatial configuration of the scheme.

Of course, it would be more difficult to establish a reasonably efficient and effective evolutionary path including cordon-based pricing in Sydney than Singapore, for two reasons. First, the natural and built topography in Sydney is not well-suited to cordon-based pricing. Second, multiple dispersed cordons would be required, because of wide dispersion of employment.

7.3 Determination of Variable Prices

Ideally, congestion charges should reflect marginal external costs of congestion so that road-users are required to internalise them – consider costs imposed on others as if they incurred those costs themselves. This would induce road-users to adjust their travel behaviour in a socially optimal way.

Pricing on the basis of marginal external costs of congestion means charges will vary over time, across locations, and with events that affect demand for, and supply of road space, because the degree of congestion and marginal external costs of congestion vary with these circumstances.

Marginal external costs and efficient prices to internalise them cannot be determined just by reference to conditions at particular times and locations. The determination would have to take into account the effect of that traffic on road-users at subsequent times and other locations until the end of the congested period (Vickrey, 1992).

Precise determination of varying charges reflecting marginal external costs of congestion would not be possible. This is not just a problem for network-wide variable congestion pricing ("first-best" pricing).

The price determination task is even more difficult when prices have to be set to take into account effects on unpriced parts of the network when pricing applies to only part of the urban arterial road network(“second-best” pricing). This difficulty is applicable to cordon and area pricing schemes, and to pricing of selected road links.

However, imprecisely determined charges would be much better than none, and more price differentiation is better than less, at least until road-users become so overloaded with information that further variability has little effect on travel behaviour. Certainly, the “simplified” congestion pricing regimes in Singapore, Stockholm, and London have improved the efficiency of resource allocation, even in the case of London where a flat charge has been applied on workdays between 7 am and 6 pm, so that the only variation is between workday daylight hours (£10) and nights and weekends (zero).

Because of the difficulty of determining “first-best” and “second-best” prices, Andre de Palma, Moez Kilani and Robin Lindsey (2005) investigated a form of “third-best” pricing, which they termed “no-queue” pricing. This involved application of time-varying tolls on specific parts of the road network to eliminate queuing on and around those network locations and segments.

They explained that the practicality of this approach was evidenced by experience with dynamically and other variably tolled HOT and express lanes in the United States for which charges were set to achieve pre-determined levels of service. They also referred to experience with Singapore’s multi-zone cordon pricing regime (including variable pricing for cordon crossings and selected roads).

Under the Singapore scheme, charges are set to achieve traffic speed targets. Singapore has progressively increased the degree of price differentiation by time and by location through increases in the number charging zones within the original cordon, additions of detection points, and adjustments to charging times and amounts. Although Singapore adopted a “simplified” congestion pricing regime, the regime has become increasingly sophisticated as periodic adjustments have been made. The evolutionary refinement of price differentiation has allowed charges to be lowered from time to time, as a result of less congestion as indicated by speed improvements.

Several advantages of “no-queue” pricing over “second-best” pricing were nominated by de Palma, Kilani and Lindsey (2005, p. 90).

- “No-queue” pricing is much less computationally demanding, because the only information required is on traffic delays, traffic flows or both on road links affected by queuing, not on other links or on demand for road space.
- Introduction and refinement can be evolutionary.
- If “no-queue” pricing is extended network-wide it would be approximately “first-best” optimal.
- Appreciable efficiency gains can be achieved even on individual links.
- The target of queue elimination is easily defined, visible, verifiable, easily explained, and likely to be readily accepted. With regard to acceptability, standards or targets are usually easier to sell to policy makers and the community than complicated or abstract targets, such as maximisation of social nett benefits or improving the efficiency of resource allocation.

Simulations undertaken by de Palma, Kilani and Lindsey (2005) indicated that “no-queue” pricing could perform reasonably well compared to “second best” variable (discrete step) pricing, even before taking into account the computational and acceptability advantages of "no-queue" pricing. They concluded that their analysis suggested that (p. 96):

> "….. no-queue tolling (time-varying pricing) or similar policy could fruitfully be implemented on selected roads now, rather than waiting a decade or more for satellite-based comprehensive road pricing to become feasible."

While successful adaptation of global positioning system technology would facilitate very fine differentiation of prices according to the degree of congestion or according to time of day, location, or occurrence of congestion-influencing events, charges under network-wide variable congestion pricing would still have to be determined. Aids to price determination, which would also be relevant to "no-queue" pricing could include (Eliasson, 2010, p. 6):

- access to a good transport model or models of the metropolitan area
- modification of the model(s) to facilitate simulation of price effects (Existing models typically have been constructed for purposes other than assessing effects of differentiated pricing.)
- particular attention to reduction of traffic impeded by bottlenecks in the network.

The difficulty of determining prices is not a valid reason to eschew price variability over time and across locations. Variability can still be used to provide incentives to change travel times, routes and modes to alleviate congestion and improve the efficiency of resource allocation.

William Vickrey (1992) emphasised the importance of encouraging shifts in trip times within peak and shoulder periods, not just from these periods to other times of the day or night. Vickrey (1992, pp. 1-2) explained that to achieve this:

> "Charges should vary smoothly over time. Only in this way can everyone be given an incentive to shift the time of travel, if only by small amounts away from the peak. If charges vary discontinuously, excessive incentives are given to rush to get ahead of a jump in the charge, or to lag waiting for a drop in the charge. There is a likelihood of creating mini-peaks just before a scheduled increase or just after a scheduled decrease. Few of those moving at the height of the peak will be willing to make the substantial shift in travel of an hour or more necessary to obtain a lower charge. It will often be easier and cause less disruption to get 12 drivers to shift the time of their trip by 10 minutes each than to get one person to shift his trip by two hours.
>
> If staggering of work hours were to succeed initially in elimination peak congestion, in the absence of road pricing it would then fail, as firms seeing that congestion has been abated then drift back to their preferred times and re-create the congestion. Reasonably smooth variation of the prices over time is needed to prevent this from happening."

7.4 Exemptions and Discounts

If a congestion pricing regime is established in Sydney, it will be important to strictly limit exemptions and discounts.

These concessions would undermine the basic purpose of congestion pricing, which is to alleviate congestion efficiently and effectively. Provision of some exemptions and discounts would encourage political lobbying by others for similar treatment. Therefore, credibility and equity issues would arise.

Exemptions and discounts would also reduce revenue from congestion pricing. There would be indirect revenue losses, as well as direct losses from a smaller charging base. One source of indirect loss would be reduced viability of the public transport system, which would be affected by slower bus trips, reduced loadings, less frequent services, and a less dense network, leading to higher subsidies than in the absence of concessions. Another source of indirect losses would be extra administration costs associated with allowance of exemptions and discounts.

Extraordinary exemptions and discounts have been an important reason for a high cost to revenue ratio for the London congestion charging system. They have also raised the cost to revenue ratios in Stockholm, and to a lesser extent Singapore.

The major reason for exemptions and discounts, particularly in the London case, has been to gain acceptance. However, there are much more efficient ways of doing this that would enhance the efficiency of anti-congestion policy, rather than undermine it. These means include provision of complementary road facilities and public transport services, and cuts in relatively inefficient taxes, as discussed in section 4.

7.5 Existing Tolled Facilities

Implementation of a congestion pricing regime in metropolitan Sydney would be complicated by existing arrangements with private operators of tolled roads. These cover important parts of the road network, particularly the majority of the road segments that Transport for NSW (2012, p. 45) described as the "110-kilometre Sydney orbital network". Tolling arrangements on these roads are not consistent across tolled facilities. Moreover, the tolls were not designed to alleviate congestion.

Transport for NSW (2012, p. 91) pointed out:

> "Currently, around two-third of the motorway network is tolled, but there is inconsistency in pricing. …… Although tolling has been used in Sydney to fund motorway construction, its role in addressing and managing congestion has not been widely discussed. There are nine toll roads in NSW. They are the M2, M5, and M7, the Eastern Distributor, the Sydney Harbour Tunnel and Bridge, the Cross City Tunnel, the Lane Cove Tunnel, and the Falcon street Gateway."

Because of conflict between private operators' profit maximising objective and government's desire to alleviate congestion, provision of privately owned toll-roads to address gaps in the road network and efficient congestion-alleviation are incompatible. If toll-roads are government owned, congestion-alleviation could take precedence over returns on capital on specific new road segments (Downs, 2004, pp. 168-169). In contrast, as explained by Erik Verhoef (2007, p. 87) and discussed in section 5, congestion pricing of the urban arterial network may be compatible with recovery of costs of the network, including a reasonable return on capital.

Over fifty five years ago, Nobel Laureate in economics William Vickrey argued that applying tolls to roads built to alleviate congestion, particularly by-pass roads, while allowing free access to congested roads, was an "outstanding absurdity" of public policy. Vickrey explained that applying congestion charges on busy roads at peak times and using the money to provide toll-free by-pass roads was the appropriate policy. Vickrey's colourful comments have been re-produced in Box 15.

Box 15 An Outstanding Absurdity of Public Policy

"Perhaps some indication of the outstanding absurdities that occur in present utility rate (public facility pricing) structures may be worthwhile in conclusion. For example, in New York a vehicular tunnel was opened a few years ago from the Battery to Brooklyn. Since it is a new facility and undoubtedly much more easy and pleasant to use than the old East River Bridges, it must, forsooth, be made to pay for itself by the imposition of tolls starting at 35 cents, the practical consequence of which is to encourage continued heavy use of the Manhattan Bridge for all trips for which the route is shorter than the tunnel, with the result that the streets near the Manhattan end of the bridge are the scene of some of the worst traffic in the city. Marginal cost considerations would call for the collection of a substantial toll (congestion charge) on the old East River bridges, at least during hours of heavy congestion, and a smaller toll or none at all for the tunnel, even though this might mean that the users of the bridges might be 'paying for' the tunnel."

Source: Vickrey (1955), p. 619.

Implementation of congestion pricing in Sydney would necessitate dealing with existing tolling arrangements with private operators. An appropriate approach would be to determine the congestion pricing regime as if the tolling arrangements did not exist, and then re-negotiate the tolling agreements before implementation of congestion pricing. Profit-neutral "shadow tolls" could be paid to the private operators by government at rates that took into account the effects of variable congestion prices (including zero prices) borne by users of those road segments.

7.6 Occasional User Problem

The design of any congestion pricing system for Sydney would have to address the challenge of occasional users. Bernhard Oehry (2010, pp. 7, 14) has argued that this is an obstacle to a network-wide variable congestion pricing regime because:

> "…… it is difficult to equip occasional users with sophisticated equipment. Occasional user solutions typically are simple booking systems that cannot support a complex, highly differentiated and variable charge."

The argument was terse and requires further investigation. In any event, the problem is relevant only if network-wide variable congestion pricing is under consideration for immediate adoption. It could be expected to diminish in significance if an evolutionary approach to congestion pricing is adopted, so that global positioning technology is adopted in the medium- to long-term, when use of such technology has become more commonplace in vehicles and interested entities have borne the risks and other costs of adapting the technology to support network-wide variable congestion pricing and perhaps other applications such as distance-linked insurance.

7.7 System Costs and Risks

The experiences of London and Stockholm have highlighted the importance of controlling system costs and risks borne by government. Formulating controls would be an important issue for establishment of a congestion pricing system in metropolitan Sydney.

Although the Stockholm's performance on this matter was superior to that in London, Jonas Eliasson (2010) and Carl Hamilton (2010) have documented how costs of the Stockholm system were raised substantially by lack of early clarification of the legal basis for the charge (price versus tax), poor risk allocation between government and the system supplier, excessive technical and administrative back-up arrangements, and unduly high service level specifications.

7.8 Privacy Issues

In a congestion pricing system, it is important to address privacy concerns. The privacy issues are similar to those applying in telecommunications, banking and retail sectors. Technological solutions have been applied in existing "simplified" congestion pricing regimes.

7.9 Revenue Recycling and Complementary Measures

Detailed discussions of the importance of revenue recycling and complementary policy measures have been provided in sections 6 and 4. There, it was emphasised that these are essential for efficient, equitable and acceptable outcomes from an anti-congestion policy based on congestion pricing. It was also emphasised that these complements should be in place prior to the implementation of congestion pricing or prior to elements of its roll-out.

It would be beyond the scope of this discussion paper to nominate specific investments and other measures. The generic measures discussed in sections 4 and 6 provide a starting point for further investigations.

A Technological Options for Congestion Pricing

At present, there are several technologies available to support traffic congestion pricing schemes. They can be grouped into the following categories:

- video-based license-plate recognition systems, using roadside cameras to capture images of vehicle license plates, which are then processed through optical character recognition software to identify a vehicle by its number plate
- dedicated short-range communication systems, for two-way communication between a roadside or gantry beacon and in-vehicle tags or transponders
- wide-area communications-based systems, which utilise a technology in the form of a location monitoring system, complemented with a communications system to manage and enforce charging.

Video-based license plate recognition systems

Video-based license plate recognition systems make use of cameras and sensors mounted on gantries or beacons to locate cars as they enter a zone. Cameras used in this system have an integrated automatic number plate recognition (anpr) system, which locates the number plate in an image and converts it into its alphabetic/numeric characters. The gantries require a connection to a central processing authority which imposes charges.

The London area pricing scheme relies heavily on anpr technology. The detection network consists of 174 video camera sites at access points to the charging zone and 24 cameras within the charging zone. The cameras capture images of vehicles entering the zone and vehicles on public roads within the zone. Anpr technology identifies vehicle number plates with an accuracy rate of 90 per cent. Number plates are checked electronically against a data base of vehicles for which payment has been made, automatic payment registration applies, and exemption or discount entitlements exist. If a match between an identified vehicle and a paid, automatic payment registration, or exempt vehicle is made, the image and vehicle details are discarded. If there is no match, the number is referred for manual checking, and if warranted, issue of a penalty notice.

The Stockholm cordon pricing system is based on anpr technology. There are 18 video camera sites, which cover all entrances/exits to the charging area.

During a trial of the system from January to August 2006, overlapping anpr and dsrc technologies were used. When the system was re-introduced on a permanent basis in August 2007, the dsrc technology was phased out because the trial demonstrated that the anpr technology was adequate for the task and tweaking of the technology subsequent to the trial further improved its effectiveness.

In the case of Stockholm, the standard anpr technology was enhanced by the use of algorithms to make a second attempt at identification of number plates that could not be identified initially because of light and weather conditions, dirty number plates, or poor camera angles. By 2008, the identification accuracy rate of the anpr system, with a small amount of manual support, was regularly between 95 and 99 per cent.

Dedicated short-range communication systems

Dedicated short range communications (dsrc) systems are widely used on toll roads. These systems use road-side equipment or above-road (gantry) equipment to communicate with electronic tags via infra-red or microwave sensors. The tag in a vehicle is queried as it moves past a gantry. These systems typically use an anpr system to identify road-users not equipped with an electronic tag or acting fraudulently.

There is a range of technologies that can be utilised by dsrc. Some tags are passive and are powered by the gantry transmission. Others are active and are capable of two-way communication.

Active tags allow identification and charging of road-users and updating of an account balance on the "smart card" tag. Details are cleared from the system after payment is made and the smart card balance is updated. Two-way technology is used to support Singapore's electronic road pricing regime.

While dsrc systems tend to be more accurate than anpr systems, they do not provide better location data than other gantry/beacon systems.

Wide-area communication-based systems

Wide-area communication-based systems, also known as mobile positioning systems, are newer innovations than anpr and dsrc systems. Wide-area communication-based systems use technologies adapted from other applications. One is global positioning system (gps) technology. The other is mobile telephone technology.

Gps technology uses satellites to identify the location of equipped vehicles and a two-way communications link based on either global system for mobile communications (gsm) or dsrc technologies, which inform a central system of current charge accounts.

An in-vehicle unit contains a gps receiver and computing memory. The gps receiver calculates the location and/or distance travelled. The information is then matched against a record of the locations of all charging points, which are either pre-stored or downloaded via the unit's communication link. Finally, charges are calculated and accrued by the in-vehicle unit, and periodically transmitted to control/enforcement centres, either via gsm or dsrc technology.

Since 2005, Germany's Toll Collect has operated using this technological system to support a distance-based toll for all trucks over 12 tonnes using German roads. The in-vehicle units collect data on accrued tolls based on the toll route network, toll rates per kilometre, emissions class and the number of axles on a truck.

Wide-area communication-based systems offer considerable advantages over charging systems based on dsrc and anpr technologies. They can provide extensive coverage with much smaller road-side infrastructure requirements. In addition, they offer economies of scale.

Current costs of the in-vehicle units have been a major reason why these systems have so far been used only for heavy vehicle charging. However, the cost of these systems has been declining rapidly.

Another key problem with gps-based systems for road pricing is that the presence of tall buildings impedes the "visibility" of satellites and therefore location-determination (the "urban canyon effect"). However, this problem has diminished with the launching of additional satellites.

A third issue has been the accuracy of location of the position of vehicles. This has diminished as technology has quickly improved. Locational errors are now no more than a few metres on average, and the 99th percentile error no more than 25 metres.

Another emerging technology involves the use of cellular networks to track the location of mobile phones, which act as the in-vehicle unit for road-use charging purposes. Extensive trials using mobile phones in 2004 on the 3G network in the U.K. indicated that the location accuracy was in the region of several hundred metres at best. However, the tests did show that as mobile phone technologies progress, this technology has the potential to provide a viable alternative to gps-based systems.

B Bibliography

Anas, A., Lindsey, R. (2011), "Reducing Urban Road Transportation Externalities: Road Pricing in Theory and in Practice", *Review of Environmental Economics and Policy*, 5, 1, Winter, pp. 66-88.

Armelius, H. and Hulkrantz, L. (2006), "The Politico-Economic Link between Public Transport and Road Pricing: An Ex-Ante Study of the Stockholm Road-Pricing Trial", *Transport Policy*, 13, 2, March, pp. 162-172.

Arnott, R. (2005), "City Tolls – One Element of An Effective Policy Cocktail", *CESifo DICE Report*, 3, pp. 5-11.

Arnott, R. (2006), "Spatial Competition between Parking Garages and Downtown Parking Policy", *Transport Policy*, 13, 6, November, pp. 458-469.

Arnott, R. (2007), "Congestion Tolling with Agglomeration Externalities", *Journal of Urban Economics*, 62, 2, September, pp. 187-203.

Arnott, R., de Palma, A., Lindsey, R. (1990), "Economics of a Bottleneck", *Journal of Urban Economics*, 27, 1, January, pp. 111-130.

Arnott, R., de Palma, A., Lindsey, R. (1993), "A Structural Model of Peak-Period Congestion: a Traffic Bottleneck with Elastic Demand", *American Economic Review*, 83, 1, March, pp. 161-179.

Arnott, R., de Palma, A., Lindsey, R. (1994), "The Welfare Effects of Congestion Tolls with Heterogeneous Commuters", *Journal of Transport Economics and Policy*, 28, 2, May, pp. 139-161.

Arnott, R., de Palma, A., Lindsey, R. (1998), "Recent Developments in the Bottleneck Model" in Button, K., Verhoef, E. (eds), *Road Pricing, Traffic Congestion and the Environment*, Cheltenham: Edward Elgar, 1998, pp. 79-110.

Arnott, R., Kraus, M. (1998), "When Are Anonymous Congestion Charges Consistent with Marginal Cost Pricing", *Journal of Public Economics*, 67, 1, January 1998, pp. 45-64.

Arnott, R., Rave, T. and Schöb, R. (2005), *Alleviating Urban Traffic Congestion*, Cambridge MA: MIT Press.

Becker, G. (1998), "Goodbye Tollbooths and Traffic James", *Business Week*, Industrial/Technology Edition, 18 May, p. 26.

Bonsall, P., Shires, J., Maule, J., Matthews, B., Beale, J. (2007), "Responses to Complex Pricing Signals: Theory, Evidence and Implications for Road Pricing", *Transportation Research A*, 41, 7, August, pp. 672-683.

Bruegmann, R. (2005), *Sprawl: A Compact History*, Chicago: University of Chicago Press.

Buchanan, J. (1952), "The Pricing of Highway Services", *National Tax Journal*, 5, 2, June, pp. 97-106.

Button, K. (1998), “Road Pricing and the Alternatives for Controlling Road Traffic Congestion” in Button, K. and Verhoef, E. (eds), *Road Pricing, Traffic Congestion and the Environment*, Cheltenham: Edward Elgar, 1998, pp. 113-135.

Button, K. (2004), “The Rationale for Road Pricing: Standard Theory and Latest Advances” in Santos, G. (ed.), *Road Pricing: Theory and Evidence*, Oxford: Elsevier, 2004, pp. 3-25.

Button, K. (2006a), “How Should the Revenues from Congestion Pricing Be Spent?” in Roth, G. (ed.), *Street Smart: Competition, Entrepreneurship and the Future of Roads*, London: Transaction, 2006, pp. 225-242.

Button, K. (2006b), “The Political Economy of Parking Charges in ‘First’ and ‘Second Best’ Worlds”, *Transport Policy*, 13, 6, November, pp. 470-478.

Button, K. (2008), “London Congestion Charging: Comment”, *Brookings-Wharton Papers on Urban Affairs*, pp. 208-214.

Button, K. and Verhoef, E. (eds) (1998), *Road Pricing, Traffic Congestion and the Environment*, Cheltenham: Edward Elgar.

Cervero, Robert, “Are Induced-Travel Studies Inducing Bad Investments?”, *Access*, 22, Spring 2003, pp. 22-27.

Chin, K-K. (2005), “Road Pricing: Singapore’s 30 Years of Experience”, *CESifo DICE Report*, 3/2005, pp.12-16.

Chin, K-K. (2010), "The Singapore Experience: the Evolution of Technologies, Costs and Benefits, and Lessons Learnt", *Joint Transport Research Centre Round Table*, Paris, 4-5 February 2010, JTRC Discussion Paper 1, 2010, International Transport Forum/OECD.

Chu, X. (1995), “Endogenous Trip Scheduling: the Henderson Approach Re-formulated and Compared with the Vickrey Approach”, *Journal of Urban Economics*, 37, 3, May, pp. 324-343.

Clarke, H., Prentice, D. (2009), *A Conceptual Framework for the Reform of Taxes Related to Roads and Transport*, paper prepared for Henry Tax Review, School of Economics, La Trobe University, June.

Cosgrove, D. (2011), “Long Term Patterns of Public Transport Use”, *Australian Transport Research Forum 2011 Proceedings*, Adelaide, September.

Cosgrove, D., Gargett, D., and others (2002), *Greenhouse Gas Emissions from Transport: Australian Trends to 2020*, Report 107, Bureau of Transport and Regional Economics, Commonwealth of Australia, Canberra.

Cosgrove, D. and Gargett, D. (2007), *Estimating Urban Traffic and Congestion Cost Trends for Australian Cities*, Working Paper 71, Bureau of Transport and Regional Economics, Commonwealth of Australia, Canberra.

Council of Australian Governments, Competition and Regulation Working Group (2006), *Review of Urban Congestion Trends, Impacts and Solutions*, Canberra, December.

Cox, J. (2006), 'Galloping technology drives public transport demand off the rails", *The Age*, 24 January, p. 8.

Daunfelt, S., Rudholm, N., Rämne, U. (2009), "Road Pricing and Retail Revenues: Results from the Stockholm Road Pricing Trial", *Transportation Research A*, 43, 3, March, pp. 306-309.

de Palma, A., Kilani, M., Lindsey, R. (2005), "Comparison of Second-Best and Third-Best Tolling Schemes on a Road Network", *Transportation Research Record*, 1932, pp. 89-96.

de Palma, A. and Lindsey, R. (2007), "Transport User Charges and Cost Recovery", in de Palma, A., Lindsey, R. and Proost, S. (eds), *Investment and the Use of Tax and Toll Revenues in the Transport Sector*, Oxford: Elsevier, 2007 pp. 29-57.

de Palma, A., Lindsey, R. and Proost, S. (2007a), "Synthesis of Case Study Results and Future Prospects" in de Palma, A., Lindsey, R. and Proost, S. (eds), *Investment and the Use of Tax and Toll Revenues in the Transport Sector*, Oxford: Elsevier, 2007, pp. 269-297..

de Palma, A., Lindsey, R. and Proost, S. (eds) (2007b), *Investment and the Use of Tax and Toll Revenues in the Transport Sector*, Oxford: Elsevier.

Dixit, A., Pindyck, R. (1994), *Investment Under Uncertainty*, Princeton: Princeton University Press.

Downs, A. (1962), "The Law of Peak-Hour Expressway Congestion", *Traffic Quarterly*, 16, 3, July, pp. 393-409.

Downs, A. (2004), *Still Stuck in Traffic: Coping with Peak Hour Traffic Congestion*, Washington: Brookings.

Duranton, G., Turner, M. (2011), "The Fundamental Law of Road Congestion: Evidence from US Cities", *American Economic Review*, 101, 6, October, pp. 2616-2652.

Eddington, R. and others (2006a*), The Eddington Transport Study, The Case for Action: Sir Rod Eddington's Advice to Government,* Report to Chancellor of the Exchequer and Secretary of State for Transport, London, December.

Eddington, R. and others (2006b*), The Eddington Transport Study, Main Report: Transport's Role in Sustaining the UK's Productivity and Competitiveness*, Report to Chancellor of the Exchequer and Secretary of State for Transport, London, December.

Eliasson, J. (2008), "Lessons from the Stockholm Congestion Charging Trial", *Transport Policy*, 15, 6, November, pp. 395-404.

Eliasson, J. (2009), "A Cost-Benefit Analysis of the Stockholm Congestion Pricing System", *Transportation Research A*, 43, 4, May, pp. 468-480.

Eliasson, J. (2010), "So You're considering Introducing Congestion Charging? Here's What You Need to Know", *Joint Transport Research Centre Round Table*, Paris, 4-5 February 2010, JTRC Discussion Paper 4, 2010, International Transport Forum/OECD.

Eliasson, J., Hulkrantz, H., Nerhagen, L., Rosqvist, L. (2009), "The Stockholm Congestion-Charging Trial 2006: Overview of Effects", *Transportation Research A*, 43, 3, March, pp. 240-250.

Eliasson, J., Brundell-Freij, K. and Hugosson, M. (2008), “The Stockholm Congestion Charging System: A Summary of the Effects” in Richardson, H. and Bae, C. (eds), *Road Congestion Pricing in Europe: Implications for the United States*, Cheltenham: Edward Elgar, 2008, pp. 293-309.

Ergas, H. (2011), "Dealing with Congestion Efficiently by Steve Meyrick: Discussant Comments " in Productivity Commission, *Proceedings of Roundtable on a Sustainable Population? – Key Policy Issues*, Canberra, 21-22 March 2011, pp. 125-132.

Ergas, H. (2012), “Urban Infrastructure and Land-Use” in Committee for Economic Development of Australia, *A Greater Australia: Population, Policies and Governance*, Research Report, March 2012, pp. 186-201.

Fosgerau, M., Van Dender, K. (2010), “Road Pricing with Complications”, *Joint Transport Research Centre Round Table*, Paris, 4-5 February 2010, JTRC Discussion Paper 2, 2010, International Transport Forum/OECD.

Friedman, M., Boorstin, D. (1952), "How to Plan and Pay for the Safe and Adequate Highways We Need", Epilogue in Roth, G., *Roads in a Market Economy*, Aldershot: Avebury Technical, 1996, pp. 223-245.

Gargett, D. (2012), *Traffic Growth in Australia*, Research Report 127, Bureau of Infrastructure, Transport and Regional Economics, Commonwealth of Australia, Canberra, February.

GHD (2011), *NSW Infrastructure Baseline: Preliminary Transport Baseline Report (Draft)*, Sydney, September.

Glaeser, E. (2011), *Triumph of the City*, London:Macmillan.

Glaeser, E., Kahn, M. (2003), *Sprawl and Urban Growth*, Discussion Paper 2004, Harvard Institute of Economic Research, May.

Gómez-Ibañez, J., “Pricing”, in Gómez-Ibañez, J., Tye, W., Winston, C., *Essays in Transportation Economics and Policy: A Handbook in Honour of John Meyer*, Washington, DC: Brookings Institution, 1999, pp. 99-136.

Gómez-Ibañez, J., Tye, W., and Winston, C. (eds) (1999), *Essays in Transportation Economics and Policy: A Handbook in Honour of John R. Meyer*, Washington DC: Brookings.

Hamilton, C. (2010), "Revisiting the Cost of the Stockholm Congestion Charging System", *Joint Transport Research Centre Round Table*, Paris, 4-5 February 2010, JTRC Discussion Paper 5, 2010, International Transport Forum/OECD.

Hansen, B. (1955), *The Economic Theory of Fiscal Policy*, London: George Allen and Unwin, English edition, 1958.

Harvey, M. (2005), *The Economic Efficiency of Lane Differentiation Policies*, Bureau of Transport and Regional Economics staff paper to Australasian Transport Research Forum, Sydney, September.

Hau, T. (1998), “Congestion Pricing and Road Investment” in Button, K. and Verhoef, E. (eds), *Road Pricing, Traffic Congestion and the Environment*, Cheltenham: Edward Elgar, pp. 39-78.

Henry, K., Harmer, J., Piggott, J., Ridout, H., Smith, G. (2009), *Australia’s Future Tax System: Report to the Treasurer*, Canberra: Commonwealth of Australia, December.

Hubbard, P. (2009), “Urban Congestion – Why ‘Free’ Roads Are Costly”, *Economic Roundup*, 2, 2009, pp. 1-19.

Hymel, K., Small, K., Van Dender, K. (2010), “Induced Demand and Rebound Effects in Road Transport”, *Transportation Research B*, 44, 10, December, pp. 1220-1241.

Jansson, J. (2008), "Public Transport Policy for Central-City Travel in the Light of Recent Experiences of Congestion Charging", *Research in Transportation Economics*, 22, 1, pp. 179-187.

Joint Transport Research Centre (2007), *Transport Infrastructure Charges and Capacity Choice: Self-Financing Road Maintenance and Construction*, Paris: European Conference of Ministers of Transport and OECD.

Kahn Ribeiro, S., Kobayashi, S., others, “Transport and its Infrastructure” in Metz, B., Davidson, O., Bosch, P. Dave, R., Meyer, L. (eds) (2007), *Climate Change 2007 – Mitigation of Climate Change,* Contribution of Working Group III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge: Cambridge University Press.

Karlström, A., Franklin, J. (2010), “Behavioural Adjustments and Equity Effects of Congestion Pricing: Analysis of Morning Commutes during the Stockholm Trial”, *Transportation Research A*, 43, 3, March, pp. 283-296.

Kay, J. (2007), “London’s Congestion Charge Has Taken a Wrong Turn”, *Financial Times*, 20 February.

Kleven, H., Kreiner, C. (2003), “The Marginal Cost of Public Funds in OECD Countries: Hours of Work Versus Labour Force Participation”, *CESifo Working Paper*, 935, April.

Knight, K. (1924), “Some Fallacies in the Interpretation of Social Cost”, *Quarterly Journal of Economics*, 38, 4, August, pp. 582-606.

KPMG Econtech (2010), *CGE Analysis of the Current Australian Tax System*, Final Report to Commonwealth Treasury, Canberra, March

Langer, A., Winston, C. (2008), “Toward a Comprehensive Assessment of Road Pricing Accounting for Land Use”, *Brookings-Wharton Papers on Urban Affairs*, pp. 127-167.

Layton A. (2002), *Public Inquiry on Mechanisms to Fund Queensland’s Roads and Transport Infrastructure: Final Report*, Brisbane: Local Government Association of Queensland, May.

Leape, J. (2006), “The London Congestion Charge”, *Journal of Economic Perspectives*, 20, 4, Fall, pp. 157-176.

Lindsey, R. (2006), “Do Economists Reach a Conclusion on Road Pricing? An Intellectual History of An Idea” *Econ Journal Watch*, 3, 2, May, pp. 292-379.

Lindsey, R. (2008), "Prospects for Urban Road Pricing in Canada", *Brookings-Wharton Papers on Urban Affairs*, 2008, pp. 235-270.

Lindsey, R. (2009), "Cost Recovery from Congestion Tolls with Random Capacity and Demand", *Journal of Urban Economics*, 66, 1, July, pp. 16-24.

Mackie, P. (2005), "The London Congestion Charge: A Tentative Economic Appraisal. A Comment on the Paper by Prud'homme and Bocarejo", *Transport Policy*, 12, 3, May, pp. 288-290.

Market and Communications Research Pty Ltd (2005), *Key Areas of Advocacy Member Feedback*, prepared for RACQ, Brisbane, November.

Martin, L., Kain, P. (2008), *Moving Urban Australia: Can Congestion Charging Unclog Our Roads?*, Working Paper 74, Bureau of Infrastructure, Transport and Regional Economics, Commonwealth of Australia, Canberra, September.

Menon, G. (2006), "Congestion Pricing: the Singapore Experience", in Roth, G. (ed.), *Street Smart: Competition, Entrepreneurship and the Future of Roads*, London: Independent Institute/Transaction, 2006, pp. 117-140.

Meyrick, S. (2011), "Dealing with Congestion Efficiently" in Productivity Commission, *Proceedings of Roundtable on a Sustainable Population? – Key Policy Issues*, Canberra, 21-22 March 2011, pp. 109-123.

Mohring, H. (1965), "Urban Highway Investments" in Dorfman, R. (ed.), *Measuring Benefits of Government Investments*, Washington: Brookings, 1965, pp. 231-291.

Mohring, H. (1972), "Optimisation and Scale Economies in Urban Bus Transportation", *American Economic Review*, 62, 4, September, pp. 591-604.

Mohring, H. (1999), "Congestion" in Gomez-Ibanez, J., Tye, W., Winston, C. (eds), *Essays in Transportation Economics and Policy: A Handbook in Honour of John R. Meyer*, Washington DC: Brookings 1999, pp. 181-221.

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

Morton, A., Mees, P. (2005), *Too Good to be True? An Assessment of the Melbourne Travel Behaviour Modification Pilot*, Proceedings of the 28th Australasian Transport Research Forum, Planning and Transport Research Centre, Perth.

Mun, S. (1999), "Peak-Load Pricing of a Bottleneck with Traffic Jam", *Journal of Urban Economics*, 46, 3, November, pp. 323-349.

Newbery, D. (1988), "Road Damage Externalities and Road User Charges", *Econometrica*, 56, 2, March, pp. 295-316.

Newbery, D. (1989), "Cost Recovery from Optimally Designed Roads", *Economica*, 56, May, pp. 165-185.

Newbery, D. (1990), “Pricing and Congestion: Economic Principles Relevant to Pricing Roads”, *Oxford Review of Economic Policy*, 6, 2, Summer, pp. 22-38.

Newbery, D. (2005), “Why Tax Energy? Towards a More Rational Policy”, *The Energy Journal*, 26, 3, pp. 1-39.

Newbery, D. (2005), “Road User and Congestion Charges” in Cnossen, S. (ed.), *Theory and Practice of Excise Taxation*, Oxford: Oxford University Press, 2005, pp. 197-233.

Newbery, D., Santos, G. (2001), “Road Taxes, Road User Charges and Earmarking”, *Fiscal Studies*, 20, 2, pp, 103-132.

Noll, R. (2008), “London Congestion Charging: Comment”, *Brookings-Wharton Papers on Urban Affairs*, pp. 214-228.

Oehry, B. (2010), “Critical Success Factors for Implementing Road Charging Systems”, *Joint Transport Research Centre Round Table*, Paris, 4-5 February 2010, JTRC Discussion Paper 3, 2010, International Transport Forum/OECD.Olzewski, P., Xie, L. (2005), “Modelling the Effects of Road Pricing on Traffic in Singapore”, *Transportation Research Part A, 39*, pp. 755-772.

O’Toole, R. (2001), “The Folly of ‘Smart Growth’”, *Regulation*, Fall.

O’Toole, R. (2007), *The Best Laid Plans: How Government Planning Harms Your Quality of Life, Your Pocketbook, and Your Future*, Washington DC: Cato.

O'Toole, R. (2009), *Gridlock: Why We're Stuck in Traffic and What to Do About It*, Washington DC: Cato.

Pigou, A. (1920), *The Economics of Welfare*, London: Macmillan.

Parry, I. (2002), “Comparing the Efficiency of Alternative Policies for Reducing Traffic Congestion”, *Journal of Public Economics*, 85, 3, September, pp. 333-362.

Parry, I. (2009), "Pricing Urban Congestion", *Annual Review of Resource Economics*, 1, pp. 461-484.

Parry, I., Bento, A. (2002), "Estimating the Welfare Effect of Congestion Taxes: the Critical Importance of Other Distortions within the Transport System", Journal of Urban Economics, 51, 2, pp. 339-365.

Parry, I., Small, K. (2009), “Should Urban Transit Subsidies Be Reduced?”, *American Economic Review*, 99, 3, June, pp. 700-724.

Parry, I., Walls, M., Harrington, W. (2007), “Automobile Externalities and Policies”, *Journal of Economic Literature*, 45, June, pp. 373-399.

Parry, T., others (2003), *Ministerial Inquiry into Sustainable Transport in New South Wales: A Framework for the Future, Final Report*, State of New South Wales, Sydney, December.

Poole, R. (2007), *A Swedish Take on Congestion Pricing*, Reason Foundation, July.

Productivity Commission (2004), *First Home Ownership*, Inquiry Report 28, Canberra: Commonwealth of Australia, 31 March, p. 136;

Prud'homme, R., Bocarejo, J. (2005), "The London Congestion Charge: A Tentative Economic Appraisal", *Transport Policy*, 12, 3, May, pp. 279-287.

Prud'homme, R., Kopp, P. (2006), *The Stockholm Toll: An Economic Evaluation*, 2nd draft, Swedish Institute for Transport & Communications Analysis, Stockholm, September.

Proost, S., Van Dender, K. (2004), "Marginal Social Cost Pricing for All Transport Modes and the Effects of Modal Budget Constraints" in Santos, G. (ed.), *Road Pricing: Theory and Evidence*, Oxford: Elsevier, 2004, pp. 159-177.

Quddus, M., Carmel, A., Bell, G. (2007), "The Impact of the Congestion Charge on Retail: the London Experience", *Journal of Transport Economics and Policy*, 41, 1, January, pp. 113-133.

Raux, C. (2005), "Comments on 'The London Congestion Charge: A Tentative Economic Appraisal' (Prud'homme and Bocarejo, 2005)", *Transport Policy*, 12, 4, July, pp. 368-371.

Richardson, H., Bae, C. (eds) (2008), *Road Congestion Pricing in Europe: Implications for the United States*, Cheltenham: Edward Elgar.

Roth, G. (1967), *Paying for Roads: The Economics of Traffic Congestion*, Harmondsworth: Penguin.

Roth, G. (ed.) (2006), *Street Smart: Competition, Entrepreneurship and the Future of Roads*, London: Transaction for Independent Institute.

Safirova, E., Parry, I., others (2004), "Welfare and Distributional Effects of Road Pricing Schemes for Metropolitan Washington, DC" in Santos, G. (ed.), *Road Pricing: Theory and Evidence*, Oxford: Elsevier, 2004, pp. 179-206.

Salt, Bernard (2006), "Plot to destroy a sacred site: the backyard", *The Australian*, 15 June, p. 26.

Santos, G. (2004a), "Urban Road Pricing in the U.K.", in Santos, G. (ed.), *Road Pricing: Theory and Evidence*, Oxford: Elsevier, 2004, pp. 251-282.

Santos, G. (ed.) (2004b), *Road Pricing: Theory and Evidence*, Oxford: Elsevier.

Santos, G. (2008a), "The London Experience" in Verhoef, E., Bliemer, M., Steg, L. and van Wee, B. (eds), *Pricing in Road Transport: A Multi-Disciplinary Perspective*, Cheltenham: Edward Elgar, 2008, pp. 273-292.

Santos, G. (2008b), "The London Congestion Charging Scheme, 2003-2006", in Richardson, H., Bae, C. (eds), *Road Congestion Pricing in Europe: Implications for the United States*, Cheltenham: Edward Elgar, 2008, pp. 159-175.

Santos, G. (2008c), "London Congestion Charging", *Brookings-Wharton Papers on Urban Affairs*, pp. 177-207.

Santos, G. (2010), "The Experience of Area Licences and Cordon Tolling to Date", *Resources for the Future Policy Commentary*.

Santos, G., Li, W.-W. and Koh, W. (2004), "Transport Policies in Singapore" in G. Santos (ed.), *Road Pricing: Theory and Evidence*, Oxford: Elsevier, 2004, pp. 209-235.

Schade, J., Schlag, B. (eds) (2003), *Acceptability of Transport Pricing Strategies*, Oxford: Elsevier.

Schuitema, G., Steg, L., Forward, S. (2010), "Explaining Differences in Acceptability before and Acceptance after the Implementation of a Congestion Charge in Stockholm", *Transportation Research A*, 44, 2, February, pp. 99-1-9.

Shoup, D. (2004), *The High Cost of Free Parking*, Chicago: Planners Press.

Shoup, D. (2006), "Cruising for Parking", *Transport Policy*, 13, 6, November, pp. 479-486.

Small, K. (2004), "Road Pricing and Public Transport" in Santos, G. (ed.), *Road Pricing: Theory and Evidence*, Oxford: Elsevier, 2004, pp. 133-158.

Small, K. (2005), "Road Pricing and Public Transit: Unnoticed Lessons from London", *Access*, 26, Spring, pp. 10-15.

Small, K., Verhoef, E. (2007), *The Economics of Urban Transportation*, London: Routledge.

Small, K., Winston, C. (1988), "Optimal Highway Durability", *American Economic Review*, 78, 3, June, pp. 560-569.

Small, K., Winston, C., Evans, C. (1989), *Road Work*, Washington DC: Brookings.

Small, K., Winston, C., Yan, J. (2006), "Differentiated Road Pricing, Express Lanes, and Carpools: Exploiting Heterogeneous Preferences in Policy Design", *Brookings-Wharton Papers on Urban Affairs*, 7, pp. 53-86.

Smeed, J., Walters, A., Roth, G., and others (1964), *Road Pricing: the Economic and Technical Possibilities*, Report of a Panel set up by U.K. Ministry of Transport, London: HMSO.

Soderholm, G. (2007), *The Stockholm Trial: Congestion Charging and Improved Public Transport Aimed at Reducing Traffic Jams and Creating a Better Environment*, Congestion Pricing Workshop, Hubert Humphrey Institute of Public Affairs, University of Minnesota, January.

Sullivan, E. (2006), "HOT Lanes in Southern California", in Roth, G., *Street Smart: Competition, Entrepreneurship and the Future of Roads*, London: Transaction for Independent Institute, 2006, pp. 189-223.

Taylor, B. (2002), "Rethinking Traffic Congestion", *Access*, 21, Fall.

Taylor, M., Ampt, E. (2003), "Travelling Smarter Down Under: Policies for Voluntary Behaviour Change in Australia", *Transport Policy*, 10, 3, July, pp. 165-177.

Tinbergen, J. (1952), *On the Theory of Economic Policy*, Amsterdam: North-Holland.

Transport for London (2007), *Central London Congestion Charging: Impacts Monitoring, Fifth Annual Report*, London, July.

Transport for New South Wales (2012), *NSW Long Term Transport Master Plan: Discussion Paper*, Sydney: New South Wales Government, February.

United Kingdom, Department for Transport (2008), *Travel Plans*, multiple documents at www.dft.gov.uk/pgr/sustainable/travelplans, accessed July.

Verhoef, E. (2003), *Inside the Queue: Hypercongestion and Road Pricing in a Continuous Time – Continuous Place Model of Traffic Congestion*, Tinbergen Institute Discussion Paper, Amsterdam, May.

Verhoef, E. (2007), "Self Financing of Infrastructure Services" in Joint Transport Research Centre, *Transport Infrastructure Charges and Capacity Choice: Self-Financing Road Maintenance and Construction*, Paris: European Conference of Ministers of Transport and OECD, 2007, pp. 63-91.

Verhoef, E. (2008), "Road Transport Pricing: Motivation, Objectives and Design from an Economic Perspective" in Verhoef, E., Bliemer, M., Steg, L. and van Wee, B. (eds), *Pricing in Road Transport: A Multi-Disciplinary Perspective*, Cheltenham: Edward Elgar, 2008, pp. 6-26.

Verhoef, E., Bliemer, M., Steg, L. and van Wee, B. (eds) (2008), *Pricing in Road Transport: A Multi-Disciplinary Perspective*, Cheltenham: Edward Elgar.

Verhoef, E., Small, K. (2004), "Product Differentiation on Roads: Constrained Congestion Pricing with Heterogeneous Users", *Journal of Transport Economics and Policy*, 38, 1, January, pp. 127-156.

Vickrey, W. (1955), "Some Implications of Marginal Cost Pricing for Public Utilities", *The American Economic Review*, 45, 2, May, pp. 605-620.

Vickrey, W. (1959), "Statement on the Pricing of Urban Car Use to the Joint Committee on Washington, DC, Metropolitan Problems", reprinted in edited form in *Journal of Urban Economics*, Vol. 36, Issue 1, July 1994, pp. 42-65.

Vickrey, W. (1963), "Pricing in Urban and Suburban Transport", *The American Economic Review*, 53, 2, May, pp. 452-465.

Vickrey, W. (1969), "Congestion Theory and Transport Investment", *The American Economic Review*, 59, 2, May, pp. 251-260.

Vickrey, W. (1992), *Principles of Efficient Congestion Pricing*, Colombia University, June, at http://www.vtpi.org/vickrey.htm.

Victorian Competition and Efficiency Commission (2006), *Making the Right Choices: Options for Managing Transport Congestion*, State of Victoria, Melbourne, September.

Walters, A. (1954), "Track Costs and Motor Taxation", *The Journal of Industrial Economics*, 2, 2, April, pp. 135-146.

Walters, A. (1961), "The Theory and Measurement of Private and Social Cost of Highway Congestion", *Econometrica*, 29, 4, October, pp. 676-699.

Walters, A. (1968), *The Economics of Road User Charges*, World Bank Occasional Paper 5, Baltimore: Johns Hopkins Press.

Walters, A. (2002), "Petrol Taxes and Congestion Charges", *Economic Affairs*, 22, 1, March, p. 40.

Willett, K. (2003), *Choked with Congestion, But Prescribed the Wrong Medicine*, RACQ staff paper to Travel Demand Management Conference, International Association of Public Transport, Brisbane, September.

Willett, K. (2006), *Stuck in Traffic and Stuck for Solutions: Brisbane's Congestion Crisis*, RACQ staff paper to Australian Institute of Transport Planning and Management (Qld) Seminar, Brisbane, November.

Willett, K. (2009), "Congestion Pricing to Untangle the Traffic Snarl", *Australian Financial Review*, 14 August.

Winston, C. (2006), *Government Failure Versus Market Failure*, Washington DC: AEI-Brookings.

Winston, C. (2010), *Last Exit: Privatisation and Deregulation of the U.S. Transportation System*, Washington: Brookings.

Winston, C., Maheshri, V. (2007), "On the Social Desirability of Urban Rail Transit Systems", *Journal of Urban Economics*, 62, 2, September, pp. 362-382.

Wood, A. (2004), "We're urban sprawlers, so don't cramp our style", *The Australian*, 27 April.

Ziegler, I. (2007), *The Stockholm Road Pricing Model: Outcomes and Lessons Learnt*, AB Storstockholms Lokaltrafik/Stockholm PTA staff paper to Conference on Public Transport: Moving People, Moving Cities, Hensinki.