ITS

WORKING PAPER
ITS-WP-03-01

Congestion Charging and the Optimal Provision of Public Infrastructure: Theory and Evidence

By

Truong Truong* and David A Hensher

January 2003

ISSN 1440-3501

*ITS Adjunct Fellow, and School of Economics, University of New South Wales.

INSTITUTE OF TRANSPORT STUDIES
The Australian Key Centre in Transport Management

The University of Sydney and Monash University

Established under the Australian Research Council's Key Centre Program.

NUMBER: Working Paper ITS-WP-03-01

TITLE: **Congestion Charging and the Optimal Provision of Public Infrastructure: Theory and Evidence**

ABSTRACT: The paper provides a theoretical framework for analysing the effects of public infrastructure provision on private sector productivity using the example of a transport network. Public infrastructure such as a transport network is assumed to be a (congested) public good. When the provision of this good is at the long run equilibrium level, consumers pay a price which reflects the (individually-determined) marginal productivity of the good and the supplier is also recovering all its opportunity costs. In the traditional literature on transport congestion (Walters, 1961; Mohring and Harwitz, 1962), the concept of infrastructure capacity is often defined in term of the maximum level of traffic flow, which is more of a usage concept rather than a 'capacity' concept. Congestion is then defined in terms of the difference between the marginal social cost of this traffic flow and the marginal private costs. There has been some debate in the literature on the way travel demand in general, and traffic congestion in particular, has been defined in terms of traffic flow because this will tend to give an ambiguous definition of the concept of 'congestion' in some cases. An alternative measure for the concept of traffic demand (and supply), and of 'congestion', is in terms of traffic density or volume rather than in terms of traffic flow. In this paper, we explore this alternative definition of 'capacity utilisation' and of 'congestion' in terms of traffic density. We arrive at an alternative definition for the concept of optimal congestion tax that turns out to be more robust. This is because it can be applied, not only to the situation of 'low congestion' but also to the case of bottleneck or 'hyper-congestion' which is not well analysed in the traditional literature. The paper also illustrates this new concept with some numerical calculations based on empirical observations on an actual road network.

KEY WORDS: Infrastructure provision, congestion pricing, traffic density, short and long run equilibrium.

AUTHORS: Truong Truong and David A Hensher.

CONTACT: Institute of Transport Studies (Sydney & Monash)
The Australian Key Centre in Transport Management, C37
The University of Sydney NSW 2006, Australia

Telephone: +61 9351 0071
Facsimile: +61 9351 0088
Email: itsinfo@its.usyd.edu.au
Internet: http://www.its.usyd.edu.au

DATE: January 2003

1. Introduction

The role that public infrastructure investment can play in increasing private sector productivity is a burgeoning area of research. Although there have been many studies which look at this issue (See Aschauer (1988, 1989a, 1989b), Berndt and Hansson (1981), Nadiri and Mamuneas (1991), Alesina *et al.* (1991), Dixon and McDonald (1991)), these studies often do not take into account the *public* (or semi-public) nature of infrastructure goods such as a transport network. In this paper we look at a transport network as a kind of congested or semi-public good, and derive estimates of a 'congestion index' which can help to determine not only an optimal tax to ration demand to capacity in the short run, but also to provide a signal as to the optimal amount of investment to be put into the transport network.

In the traditional literature on transport congestion (Walters, 1961; Mohring and Harwitz, 1962), the concept of infrastructure capacity is often defined in terms of the maximum level of traffic *flow*, which is more of a usage concept rather than a 'capacity' concept. Congestion is then defined in terms of an increase in the marginal social cost of this traffic *flow* over and above the marginal private costs (measured in terms of the average travel time per trip distance). While this is a useful definition, it does not give adequate recognition to the fact that 'congestion' is truly a phenomenon relating to transport *capacity* rather than transport usage as such, and capacity (or capacity utilisation level) is perhaps more precisely and unambiguously defined in terms of the traffic *density* level rather than traffic flow. In this paper, we explore this alternative definition of transport infrastructure 'capacity' and use this concept to examine the issue of investment in transport infrastructure to relieve 'congestion'. Capacity is defined as a form of *public capital* good, and 'congestion' refers to the situation when the utilisation of this public good becomes partly 'rival'. Congestion results in the actual level of *utilisation* of the system 'capacity' by each user being less than the full maximum potential because of the rivalness in consumption of a semi-public good. Utilisation of system capacity can be measured in terms of the actual speed achievable by each user[1]. Since the change in capacity *utilisation* rate can come about either from a change in the level of demand for travel (trip level) and/or an increase in the system capacity, we distinguish between two different situations: a 'short run' situation where an 'optimal congestion tax' can be used to 'regulate' demand to a given level of system capacity, and a 'long-run' situation, when the issue is an investment in long run capacity to meet with any projected long run demand. We illustrate this with an empirical calculation for an actual road network.

[1] We note that speed is uniquely related to traffic *density* given any particular physical characteristics of the transport system, but since traffic flow = traffic density*speed, there can be two different levels of traffic *flow* corresponding to the same level of traffic speed: one with a high density level (high congestion), and one with a low density level (low congestion) (see Walters (1961), Lindsey and Verhoef (2000)). Flow is thus an ambiguous characterisation of system capacity or capacity utilisation and therefore, an ambiguous measure of 'congestion', as compared to either density or speed.

2. Public Infrastructure as a Congested Public Good

Let G be the stock or *capacity* of a public infrastructure asset available for use and G_i be the 'effective' level of utilisation of this capacity by user i. If G is a pure public good, then by definition:

$$G_i = G, \quad i = 1,...,n \tag{1}$$

where n is the total number of users (that is, every user has unimpeded equal access). On the other hand, if G is a pure private good, then we have instead:

$$\sum_{i=1}^{n} G_i = G. \tag{2}$$

In the general case when G is an impure (partially congested) public good, we have:

$$G_i \leq G, \text{but} \sum_{i=1}^{n} G_i \geq G. \tag{3}$$

As Oakland (1987) pointed out, congested public goods can be treated as though equivalent to a combination of congestion externalities and a pure public good which relates to total system capacity. Thus, if $f^i(.)$ stands for the production function of user i, then we have:

$$f^i = f^i(L_i, K_i, G_1,..., G_n, G) \tag{4}$$

where L_i and K_i stand for the private (labour, capital) inputs, G_j is the effective level of utilisation of public infrastructure by user j, and G is the total system capacity. We can imagine $f^i(.)$ as representing, for example, the utility of a final activity (work, or leisure activity) by a traveller i, who uses a road network to produce a travelling input into this activity. The *effective* travelling input can be defined as $t_i = t_i(G_i, G, k_i, l_i)$, where (k_i, l_i) are some part of the total (K_i, L_i). For example, (k_i) can represent a 'car', (l_i) the driving time by user i and given any level of (k_i), we can define $t_i/l_i = G_i/G$, i.e. the '*effective* travel time input' into the final activity is related, not only to the *actual* travel time l_i, but also to the 'effective level of utilisation of road capacity' (G_i/G). This is because if, for example, there is congestion and it takes 20 minutes to cross a particular segment of the road which would normally take only 10 minutes, then we have: $l_i = 20$, $t_i = 10$, and $t_i/l_i = G_i/G = 0.5$. The effective level of road capacity utilisation in this case is thus only 0.5 or 50%, and this is due the presence of congestion, which increases a 'normal' travel time from 10 minutes (effective level of travel time input) to 20 minutes (actual travel time). The ratio of effective level of road capacity utilisation can be measured by observing the ratio of the *actual* speed (G_i) over the free-flow speed (G). Speed (or its inverse, travel time) can thus be seen as an effective measure of system 'capacity' (or capacity utilisation level).

There are also other important reasons why we define 'capacity' G and capacity utilisation level G_i in terms of speed rather than in terms of 'flow' or in terms of the physical characteristics of the road network, as will be explained in more details in a later section. For the present, however, it is sufficient to say that, as far as each individual traveller i's activity is concerned, the maximum flow or the physical characteristics of the road are 'relevant' to this individual activity only in so far as they have an impact on the ultimate *speed* at which each individual user i can travel at, i.e. on G_i[2]. We have:

$$\partial f^i / \partial G_i > 0; \quad \partial f^i / \partial G_j < 0, \quad \text{for } i \neq j \tag{5}$$

which says that j's utilisation of the system capacity would contribute to congestion and therefore would impact negatively on user i's system capacity utilisation level and hence on her marginal utility (or productivity). An alternative specification for a congested public good is to assume that each effective utilisation of the good is given by the user's own utilisation rate and an overall level of congestion θ:

$$f^i = f^i(L_i, K_i, G_i, \theta) \tag{6}$$

where

$$\theta = \theta\left(\sum_{i=1}^{n} G_i, G\right). \tag{7}$$

We refer to equation (7) as the congestion function. Pareto optimal allocation of resources in the case of a congested public good can now be found by solving the following optimisation problem:

$$\begin{aligned}
&\text{Max } f^i(L_i, K_i, G_i, \theta) \\
&\text{s.t. } f^j(L_j, K_j, G_j, \theta) \geq X^j, \; j \neq i, \; j = 1, \ldots, n. \\
&G_i \leq G, \quad i = 1, \ldots, n. \\
&F\left(\sum_{i=1}^{n} K_i, G\right) \leq 0.
\end{aligned} \tag{8}$$

where X^j is the output or utility level associated with user j, and $F(.)$ is the transformation function between total private capital $(\sum_{i=1}^{n} K_i)$ and total public capital goods. The rate of transformation between these two types of goods will measure the shadow price of public capital for the economy as a whole (or transport network in particular).

[2] In other words, an individual traveller is not so much concerned about how many vehicles there are travelling on the same road, or how many traffic lanes or how wide these traffic lanes are, but rather on how fast she can travel, *given* a particular driving behaviour and safety standard, etc.

The Lagrangian for this optimisation problem is:

$$L = \sum_{i=1}^{n} \lambda^i [X^i - f^i(L_i, K_i, G_i, \theta)] + \sum_{i=1}^{n} \alpha^i (G_i - G) + \mu F(\sum_{i=1}^{n} K_i, G) \tag{9}$$

with $\lambda^i = -1$ and $X^i = 0$ for the reference i. The efficiency conditions with respect to G_i are as follows (assuming that $G_i < G$, and therefore $\alpha^j = 0$ for all i).

$$-\lambda^i f^i_{K_i} + \mu F_P = 0, \tag{10}$$

$$-\lambda^i f^i_{G_i} + \sum_{j=1}^{n} \lambda^j f^j_\theta \theta_1 = 0, \tag{11}$$

$$-\sum_{j=1}^{n} \lambda^j f^j_\theta \theta_2 + \mu F_G = 0, \tag{12}$$

where

$$\begin{aligned} & f^i_x = \partial f^i / \partial x, \quad x = \{K_i, G_i, \theta\}, \\ & \theta_1 = \partial \theta / \partial \left(\sum_{i=1}^{N} G_i \right), \ \ \theta_2 = \partial \theta / \partial (G), \\ & F_P = \partial F / \partial \left(\sum_{i=1}^{N} K_i \right), \ \ F_G = \partial F / \partial G. \end{aligned} \tag{13}$$

Equations (10) and (11) can be combined to give:

$$(f^i_{G_i} / f^i_{K_i}) + \theta_1 \sum_{j=1}^{n} (f^j_\theta / f^j_{K_j}) = 0, \tag{14}$$

and equations (10) and (12) can be combined to give:

$$\theta_2 \sum_{j=1}^{n} (f^j_\theta / f^j_{K_j}) = (F_G / F_P), \tag{15}$$

Equations (14) and (15) can also be combined to give:

$$(f^i_{G_i} / f^i_{K_i}) + (\theta_1 / \theta_2)(F_G / F_P) = 0. \tag{16}$$

If increasing the aggregate utilisation rate and capacity by the same proportion would leave congestion unchanged (i.e. the function $\theta(.)$ is homogeneous of degree zero in its arguments)[3], then we have:

[3] This implies constant returns to scale in the *consumption* (or utilisation) of capacity (taken into account at a particular level of congestion θ), but not necessarily in its *production.*

$$\theta_1 \sum_{i=1}^{n} G_i + \theta_2 G = 0. \tag{17}$$

Multiplying (16) by G_i and summing over all i's using (17), we have:

$$\sum_{i=1}^{n} (f^i_{G_i} / f^i_{K_i}) G_i = G(F_G / F_P) = GP_G, \tag{18}$$

where P_G = (F_G/F_P) is the shadow price of public capital in terms of private capital foregone.

Equation (18) is a special case of the Samuelson condition for the optimal provision of a congested public good[4]. The ratio $(f^i_{G_i} / f^i_{K_i})$ stands for the marginal productivity of a public capital good relative to that of a private capital good for user i. If we adopt the benefit principle of taxation then each individual user should be charged an individualised (or Lindahl) price for the effective utilisation of public infrastructure as follows:

$$\begin{aligned} P^i_G &= (f^i_{G_i} / f^i_{K_i}) \\ T^i_G &= P^i_G G_i \end{aligned} \tag{19}$$

where P^i_G is the per unit price for public infrastructure capacity G and T^i_G is the total contribution from user i towards the total public infrastructure capacity costs[5]. Using (16), we have:

$$\begin{aligned} P^i_G &= -(\theta_1 / \theta_2)(F_G / F_P) = -(\theta_1 / \theta_2) P_G \\ T^i_G &= -(\theta_1 / \theta_2) G_i P_G \end{aligned} \tag{20}$$

3. Application to a Transport Network

To operationalise the model, we need an empirical specification of the congestion function (7). One specification of this function is the form used in many public infrastructure studies such as the one proposed by Shah (1992):

$$G_i = G.(I_i)^\theta \tag{21}$$

Here, I_i <1 stands for an 'index of use' (of the road capacity) by user i, and θ is a 'parameter indicating the degree of publicness of public infrastructure'[6]. In the case of a road network, θ can be used to indicate the 'degree of congestion' on the road[7].

---

4 See Oakland (1987, p. 501).

5 In practice when a price is charged for the use of a public infrastructure, this would consist of both a 'usage' component to cover the short run operating and maintenance costs and a 'capacity' (or capital) component to cover the rental price of infrastructure capital. Here we are concerned only with the capacity component.

To define the index of use, first, we assume that there is a minimum level of traffic density n_0 at which (or below this level), the infrastructure remains a pure public good. When the infrastructure is a pure public good, the index of use by each user should be equal to 1 (i.e. 100 percent utilisation rate): $I_i = 1$ for $n \leq n_0$. Next, when traffic density exceeds this minimum level, $n > n_0$, the ratio (n_0/n) can now be used to indicate the 'relative index of use'. Since the infrastructure is now a semi-public or 'congested' public good, each user's rate of utilisation of the good should be less than 1, and sum up to the same level as when there were only n_0 users, i.e. $I_i = (n_0/n) < 1$ for $n > n_0$.

Capacity utilisation level G_i (or the utilisation index G_i/G) can also be defined as follows. Given any traffic density level n and with an average speed s_i achievable by each user at that density, the average traffic flow will be ns_i. Comparing this with the *potential maximum* traffic flow of ns_{max} achievable if capacity utilisation is 100 percent, i.e. when the speed achievable is a maximum at this density[8] by all users, we can define the 'capacity utilisation ratio' (G_i/G) as being equal simply to this ratio of actual traffic flow over the potential maximum flows. We have $(G_i/G) = ns_i/ns_{max} = s_i/s_{max}$. From this, it can be seen that capacity utilisation ratio can be indicated by ratio of the actual *speed* over the maximum potential speed[9], given any level of traffic density.

Equation (21) can now be interpreted as follows:

- If there is no congestion ($\theta = 0$), then $(G_i/G) = (s_i/s_{max}) = 1$, i.e. every vehicle can travel at the potential maximum speed.

- When there is some congestion on the road ($\theta > 0$), capacity utilisation ratio (and hence speed ratio) will depend on the index of use (relative traffic density n/n_0) as well as the degree of congestion[10]. With $I_i = (n_0/n) < 1$ and $\theta > 1$, $(G_i/G) = (n_0/n)^{\theta} < 1$, which gives: $(n_0/n)G \leq G_i \leq G$ for $(0 \leq \theta \leq 1)$, as is required by equation (3).

- When $\theta = 1$, $G_i = (n_0/n)G$ and $s_i = (n_0/n)s_{max}$. This can be referred to as a situation when 'public' infrastructure good has become a completely rival or pure 'private' good. At this point, the actual traffic flow ns_i reaches the maximum level $n_0s_{\max}$, which is defined in the traditional literature as the maximum 'capacity' of the road. When $\theta = 1$, any percentage increase (or decrease) in traffic density will be matched by an exact and opposite percentage decrease (or increase) in the level of speed for all vehicles[11].

[6] Shah (1992, p. 29).
[7] Shah referred to θ as the degree of 'publicness' of the infrastructure good, but in fact, it can be seen that the greater the value of θ, the *smaller* would be the value of G_i. Hence, it would be more appropriate to refer to θ as the degree of *non*-publicness (or 'rivalness' in consumption) of the public infrastructure good, and in the case of a road network, the degree of 'rivalness in consumption' is in fact the degree of 'congestion' in the network.
[8] Clearly, this can be achieved only if *actual* capacity is increased, and/or actual density is reduced, therefore, we refer to this *flow* as the maximum *potential* flow.
[9] The maximum potential speed is to be determined not only by the physical characteristics of the road, but also by traffic regulation and safety standards.
[10] In fact the degree of congestion is to be defined in accordance with the empirical relationship between speed (capacity utilisation) ratio and relative traffic density level. See the empirical section below.
[11] This implies the elasticity of speed with respect to traffic density (measured by the congestion index θ - see the section below) is now equal exactly to 1.

Road space has become a pure private good, and one vehicle's 'consumption' of this space must be at the full expense of another's. Therefore, the rate at which road space is 'consumed' by *all* users (the traffic flow rate) must remain 'constant' at the maximum 'capacity' level, which is determined by the physical conditions of the road.

- Finally, the case of $\theta > 1$ and $G_i < (n_0/n)G$ (or $s_i < (n_0/n)s_{max}$) can also be described as a situation when 'public' infrastructure good has become, not only a pure private good, but also with *significant negative externality* arising from the use of road space by one user on all others. The negative externality results in the *aggregate* utilisation of road space by all users is now reduced rather than increased as a result of a marginal increase in traffic density. This means the resulting traffic flow will decrease and the travel time-traffic flow curve will become 'backward bending', a situation described in the traditional literature as one of 'hyper-congestion' or bottleneck' congestion.

From equation (21), we can now derive a formula for the congestion index. First, we substitute $I_i = (n_0/n)$ into equation (21), and sum over all *i's*, then taking the logarithm of both sides and re-arranging terms, we have:

$$\theta = 1 - \frac{1}{\ln(n/n_0)}[\ln(\sum_{i=1}^{n} G_i) - \ln(n_0 G)] \tag{22}$$

Alternatively:

$$(n/n_0)^{1-\theta} = (\sum_{i=1}^{n} G_i)/(n_0 G) \tag{23}$$

Using (22)-(23), we can derive:

$$\begin{aligned}(\theta_1/\theta_2) &= [\partial\theta/\partial(\sum_{i=1}^{n} G_i)]/[\partial\theta/\partial(G)] \\ &= -G/(\sum_{i=1}^{n} G_i) \\ &= -(n/n_0)^{\theta-1}/n_0\end{aligned} \tag{24}$$

Substituting this into (20), we obtain:

$$\begin{aligned}P_G^i &= [(n/n_0)^{\theta-1}](P_G/n_0) = (n/n_0)^{\theta}(P_G/n) \\ T_G^i &= G(P_G/n)\end{aligned} \tag{25}$$

Equation (25) shows the (marginal and total) willingness-to-pay (i.e. (P_G^i and $T_G^i = P_G^i G_i$) for effective capacity utilisation G_i by each individual user i at different levels of congestion. We note that equilibrium condition requires that the total willingness-to-pay by *all* users to system capacity (nT_G^i) must be equal to its supply cost (GP_G). From equation (25), we also observe that, firstly, when congestion level is zero (n=n_0 and θ=0) and system capacity is a pure public good, all users can share this capacity without diminishing the level of utilisation of one another, therefore, the individual opportunity cost of capacity utilisation at this zero level of congestion is just P_G^i=(P_G/n)=(P_G/n_0), where P_G is the supply price of system capacity and n=n_0 is the actual number of users. This is the Samuelson condition for the shadow pricing of a public good. Next, when congestion becomes positive (n>n_0, θ>0), system capacity becomes a congested public good, and therefore, each individual user's *effective* utilisation of system capacity is now less than the full maximum level (G_i<G), the willingness-to-pay for capacity utilisation in this case is thus also less than the zero-congestion shadow price, i.e. P_G^i<(P_G/n_0).

4. Congestion Index as an Elasticity of Speed with respect to Traffic Density

*S*ubstituting I_i = (n_0/n) and (G_i/G) = (s_i/s_{max}) into equation (21), we have:

$$s_i = s_{\max}(n_0 / n)^{\theta} \tag{26}$$

From this, we can define the concept of 'elasticity of speed (or *capacity utilisation level*) with respect to traffic *density* level', and which is seen to be equal to just the congestion index θ:

$$\theta = -\frac{n}{s_i}\frac{ds_i}{dn}\bigg|_{s_{\max}, n_0 = constant} \tag{27}$$

Compare this with the traditional concept of elasticity of speed with respect to the level of traffic *flow* (see Walters, 1961, p. 694):

$$\varepsilon = \frac{F}{s_i}\frac{ds_i}{dF}\bigg|_{s_{\max}, n_0 = constant} \tag{28}$$

Here F = ns_i is the traffic flow. We have:

$$\begin{aligned}\varepsilon &= -\frac{F}{s_i}\frac{ds_i}{dF}\bigg|_{s_{\max}, n_0 = constant} = -\frac{ns_i}{s_i}\frac{ds_i}{(nds_i + s_i dn)} \\ &= -\frac{(ds_i / s_i)}{(ds_i / s_i) + (dn / n)} \\ &= \frac{\theta}{1-\theta}\end{aligned} \tag{29}$$

From this, it can be seen that when θ < 1, ε > 0, and when $\theta \to 1$, $\varepsilon \to \infty$. The traditional case of ε = ∞ (when traffic flow reaches the maximum level and the travel time-traffic flow curve starts to become 'backward bending') can thus be described as a situation when the 'public' infrastructure (system capacity) has become a pure private good (θ = 1). The traditional analysis does not proceed beyond this point and the traditional concept of 'optimal congestion tax' (defined in terms of traffic *flow* level rather than in terms of traffic density) also does not apply beyond this point. In our case, however, the analysis can continue beyond this point for $\theta > 1$ and $\varepsilon < 0$. This is because the 'backward bending' traffic flow curve can simply be interpreted as the case when system capacity has become not only a pure 'private' good, but also there is significant negative externality resulting from each individual user's utilisation of system capacity on each other, such that the net result is a reduction in the overall or aggregate utilisation, and hence traffic *flow*[12] level decreases. In this case, it is socially optimal to invest in system capacity expansion rather than in trying to reduce traffic demand. This point will be illustrated further below.

5. Congestion Index as a Measure of Optimal Congestion Tax

In the traditional analysis, optimal congestion tax is defined in terms of traffic flow measure. Given any particular level of traffic flow, the *average* social cost (of traffic flow) is given by the individual (i.e. private marginal) cost of travel, the latter is measured by the individual travel time per unit distance[13]. Given that this travel time is increasing with increasing level of traffic flow (due to congestion), the average social cost is thus rising with increasing level of traffic flow (before the backward bending part sets in). This implies the *marginal* social cost is also rising and is above the average social cost. The percentage difference between the marginal and average social cost is measured by the 'elasticity of average social cost', i.e. the elasticity of travel time, or the negative of the elasticity of travel speed, with respect to the traffic flow level (see Walters, 1961). This is the value of ε as defined by equation (28). The value of ε is then used to define an optimal congestion tax rate to be applied to a particular traffic *flow* to reduce it to a socially optimal level. This is the 'traditional' analysis.

The difficulty of defining optimal congestion tax in terms of ε. (rather than in terms of θ) is that, firstly, it is more difficult to impose a tax on traffic *flow* rather than on traffic density. This is because traffic flow is measured over a *finite interval of time*, and hence the definition of the tax must necessarily be conditional on the definition of this finite interval. Secondly, traffic flow curve can become 'backward bending', i.e. ε becomes negative, and therefore, the definition of an optimal congestion 'tax' which becomes 'negative' can be ambiguous. In contrast, if we define 'optimal congestion tax' in terms of θ rather than in terms of ε, then there is no such difficulty or ambiguity. Traffic volume or density is precisely and unambiguously defined *at any particular point in time*. Secondly, even when the travel time-traffic flow curve becomes backward bending (ε<0) the travel time-traffic density curve remains upward sloping (θ >1) and this value of the congestion index (or speed elasticity) θ continues to bear a particularly meaningful signal for useful analysis. When ε<0 and θ >1, the situation can be described as one where *negative externality* in the utilisation of system capacity has

[12] Traffic flow represents aggregate utilisation of capacity by all users over a fixed interval of time.
[13] Assuming that travel time is the most important component of individual travel costs.

become significant to such an extent that the net marginal social cost of congestion is greater than even the marginal cost of supply of system capacity. Therefore, the value of the optimal congestion tax rate θ is now greater than 1, which implies society is willing to pay to reduce congestion, by investing in an extra unit of capacity rather than merely trying to reduce its utilisation rate. Investing in extra capacity can be referred to as a 'long run' situation. This is in contrast to the 'short run' situation of merely trying to regulate traffic demand (traffic density). Our analysis, is thus applicable to both the long run situation as well as short run analysis, depending on the particular value of congestion index θ derived from empirical analysis.

6. Empirical Application to an Actual Road Network

We want to apply equation (21) to the empirical measurement of the level of congestion for an actual road network. The data we use is a sample of 3,730 road segments (or links) of various types in the Sydney Metropolitan Area for 2001[14]. For each link, we obtained information on the link type (arterial, highway, expressway, freeway, etc.), link length (kms), number of lanes, vehicle density (vehicles per lane per km), travel time and speed, for different time periods of day (AM, Mid-day, PM) and night time (Nite). We have selected freeway conditions only since they relate most appropriately to tollroad settings, the focus of this paper. From this data we first plot the information on vehicle speed versus traffic density for various times of day. This is shown in Figure 1. From this speed-density scatter diagram, we observe that there is a definite (negative) relationship between the (potential)[15] speed s_i achievable at any level of traffic density level n and the actual traffic density, as hypothesised earlier in equation (27). From this scatter diagram, we can also define the 'maximum-capacity' speed $s_{\max}$ (which can be affected, not only by the physical characteristics of the road, but also by traffic regulation) and its corresponding minimum or free flow density n_0. Once these 'reference' values of $s_{\max}$ and n_0 are defined, the value of θ can then be calculated as a function of the speed ratio ($s_i/s_{\max}$) and the density ratio (n/n_0) for each empirical observation as described by equation (27). These empirical values of θ 's are plotted in Figure 2 as a function of the density ratio (n/n_0). Given the congestion index θ, the 'optimal congestion tax rate' is thus also defined by θ. From Figure 2, then, it can be seen that if traffic density on freeway increases to a level which is about 10 times the 'free-flow' density (i.e $n/n_0 = 10$), congestion will then reach a level such that it is optimal to impose a tax rate of θ=0.27 or 27 percent on top of the 'normal' toll to attempt to reduce traffic density to a socially optimal level (which also corresponds to the situation when traffic density will eventually be reduced to the free-flow level (n_0), and congestion is now reduced to zero.

[14] Data was purchased from the Transport Data Centre (within the New South Wales Department of Transport).

[15] The 'potential' speed is defined as the 'maximum achievable' speed at any level of traffic density. This is to eliminate the 'interior' observations (where the actual speed is less than this potential speed due to a whole host of other reasons such as weather condition, traffic accident on the road, individual driver's habit, etc.) and therefore, to retain only the 'boundary' or frontier points which defines the unique relationship of equation (27).

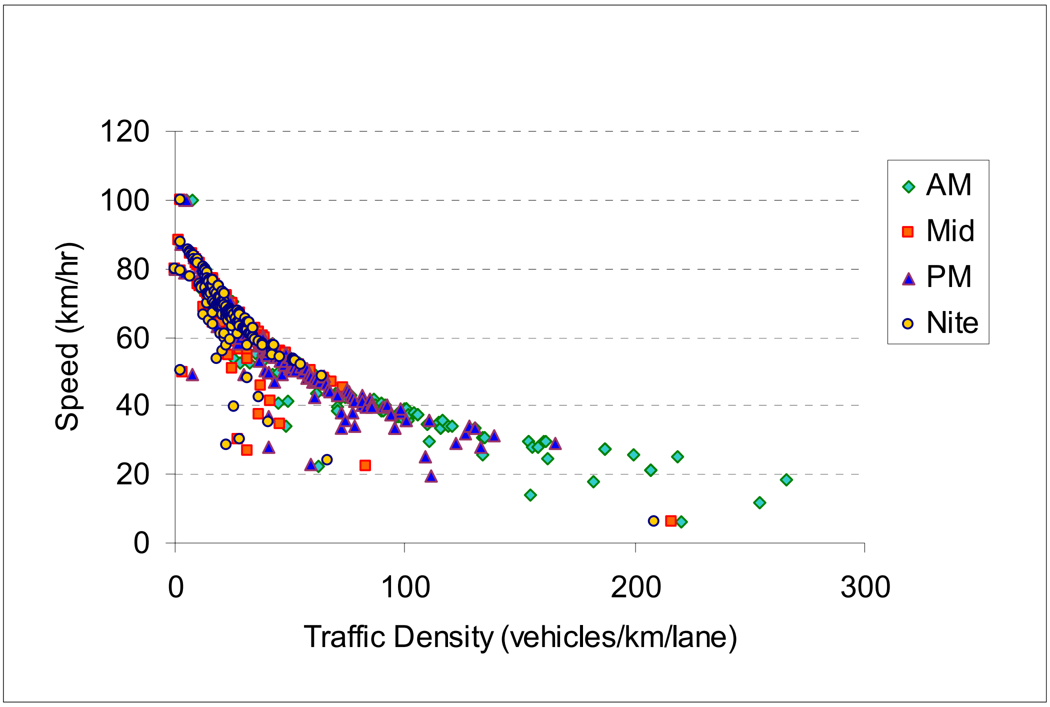

Figure 1. Speed versus Traffic Density on Freeway

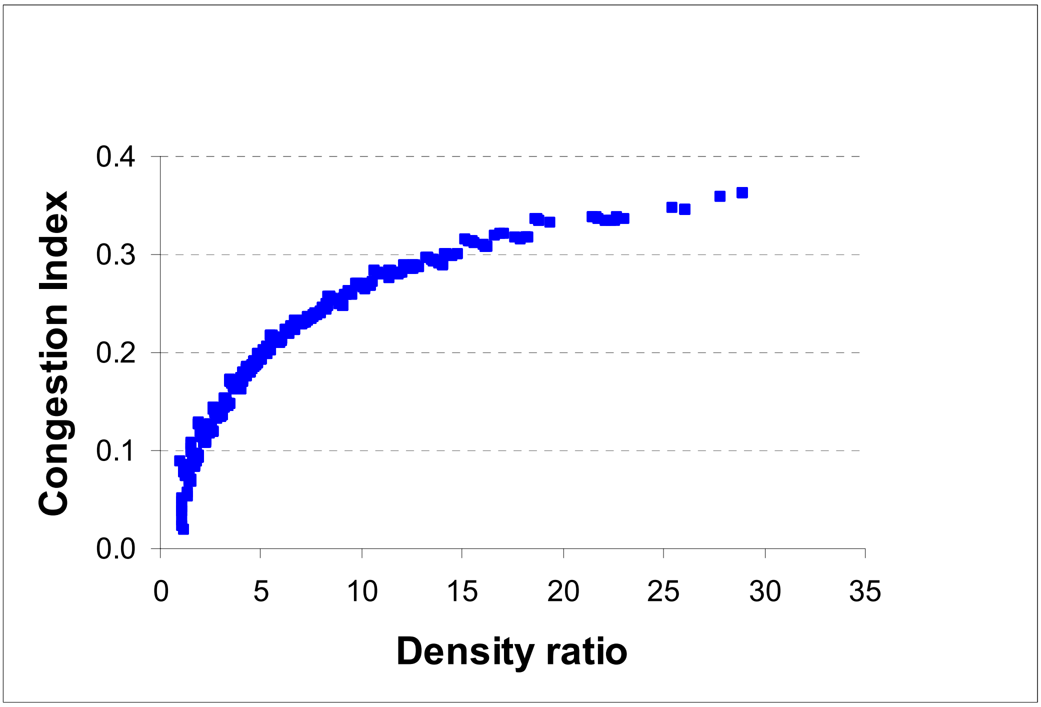

Figure 2. Congestion Index versus Traffic Density on Freeway

7. Conclusions

In this paper, we have set out to explore the role of public infrastructure investment in private sector productivity by constructing a model of private sector activity using public infrastructure as a kind of congested public good input. The paper establishes the conditions under which optimal provision of public infrastructure can be said to have been reached. This is when each individual user pays an 'effective' (but unobserved) price for the use of a public infrastructure 'capital' which is just sufficient to cover its marginal productivity, and the aggregate of all these effective charges is also equal to the supply cost of the infrastructure. When this condition is satisfied, 'congestion' then acts as a kind of implicit tax which regulates each individual user's behaviour to such an extent that the value of this 'tax' gives us an indication of how much individuals are prepared to pay to reduce congestion. We then use the measure of the congestion to define an optimal congestion tax index defined in terms of traffic density level rather than in terms of traffic flow level as was the case in the traditional analysis. When defined this way, the optimal congestion is seen to be more 'robust' and can be applied to both a 'short run' situation where traffic congestion has not reached a 'hyper-congestion' level, In contrast, and a 'long run' situation when congestion has reached a 'bottleneck' level which calls for investment in extra capacity rather than just demand regulation. Although the empirical example we used in this paper illustrates only a 'short run' situation of 'low congestion', the model can also be applied to a 'hyper-congestion' situation if more appropriate data is available in the future.

References

Alesina, A., D.W. Gruen and M. T. Jones (1991) Fiscal Adjustment, the Real Exchange Rate and Australia's External Imbalance, *The Australian Economic Review*, 3rd quarter.

Aschauer, D. A. (1988) The Equilibrium Approach to Fiscal Policy, *Journal of Money, Credit, and Banking* 20, 41-62.

Aschauer, D. A. (1989a) Is Public Expenditure Productive?, *Journal of Monetary Economics* 23, 177-200.

Aschauer, D. A. (1989b) Does Public Capital Crowd Out Private Capital?, *Journal of Monetary Economics* 24, 171-188.

Berndt, E. R. and B. Hansson (1991) Measuring the Contribution of Public Infrastructure Capital in Sweden, *National Bureau of Economic Research (N.B.E.R.) Working Paper 3842*, September.

Dixon, P. B. and D. McDonald (1991) Labour productivity in Australia 1970-71 to 1989-90, in: Economic Planning Advisory Council (EPAC), *Background Paper, No. 8,* AGPS, Canberra, January.

Lindsey, R. and E. Verhoef (2000) Congestion Modelling, in Hensher, D.A. and Button, K. (eds) Transport Modelling, Handbooks in Transport Vol 1, Pergamon Press, Oxford, Ch 21, 353-374.

Nadiri, M. I. and T. P. Mamuneas (1991) The Effects of Public Infrastructure and R&D Capital on the Cost Structure and Performance of U.S. Manufacturing Industries, *N.B.E.R. Working Paper 3887*, September.

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

Oakland, William H. (1987) Theory of Public Goods, Chapter 1 in: A. J. Auerbach and M. Feldstein, eds., *Handbook of Public Economics*, Volume 2 (North Holland, Amsterdam).

Ozbay, K., Bartin, B., and J. Berechman (2001), "Estimation and Evaluation of Full Marginal Costs of Highway Transportation in New Jersey", *Journal of Transportation and Statistics,* 81-103.

Shah, A. (1992) Dynamics of Public Infrastructure, Industrial Productivity and Profitability, *Review of Economics and Statistics* 74, 28-36.

Truong, T.P. and D.A. Hensher (2002) Congestion As The Invisible Hand In The Optimal Provision Of Public Infrastructure Goods, Institute of Transport Studies, The University of Sydney, May.

Walters, A.A. (1961) "The Theory and Measurement of Private and Social Cost of Highway Congestion", *Econometrica*, 29(4), October, 676-99.