



Mathematical principles of road congestion pricing

Wessel J Pienaar*

J Hannelie Nel†

Received: 20 March 2009; Revised: 18 May 2009; Accepted: 22 May 2009

Abstract

This paper briefly considers the objectives of road congestion pricing and identifies prerequisites to the successful application of such a pricing scheme. The paper is divided into two sections. In the first section, a mathematical analysis of the constituents of an optimal road congestion price is offered. The eliminated inefficiency loss achieved by the introduction of a congestion levy is usually evaluated by means of an integral involving marginal trip cost, travel demand and average trip cost in two-dimensional (travel time, traffic flow)-space. In this section we show that this loss may, in fact, be evaluated more easily for a general marginal trip cost function and a linear demand function as the difference between the areas of a rectangle (representing the part of road agency revenue that lies below the original trip cost) and a triangle (representing the loss of consumer surplus of the reduced traffic) in (travel time, traffic flow)-space, eliminating the need to use integration. The next section deals with the application of the illustrated mathematical principles and proofs to a hypothetical case study relating to road congestion pricing in Cape Town.

Key words: Average cost, congestion pricing, demand function, marginal cost, marginal cost pricing, road congestion, traffic flow.

1 Introduction

The argument in favour of road congestion pricing rests on the assertion that the same economic principles that apply to the allocation of other scarce resources should apply to road space as a commodity. The corollary for roads is that it is not sufficient that road users, as a class, on the average, should pay the costs of the road system as a whole; it is necessary that each user should meet the incremental cost resulting from his particular use. This principle is known as *marginal cost pricing*. The implication for road congestion is that if each user has to pay the costs arising from his use of the road, including the congestion costs inflicted by him on other users, the total volume of traffic, as well as its distribution in space and time, would be influenced towards an efficient use of available road space.

The following section of the paper deals with the mathematical principles that underlie the determination of a price that would, if imposed, result in an optimal traffic flow on

*Corresponding author: Department of Logistics, University of Stellenbosch, Private Bag X1, Matieland, 7602, South Africa, email: wpienaar@sun.ac.za

†Department of Logistics, University of Stellenbosch, Private Bag X1, Matieland, 7602, South Africa.

a congested road. In the section thereafter the mathematical principles and proofs are applied to a hypothetical case study relating to road congestion pricing in Cape Town. This case study is intended to serve the purpose of an illustration; it is of limited scope and does not purport to represent the results of an extensive traffic engineering investigation.

2 Analysis of congestion pricing principles

Two causes of urban road traffic congestion and a concomitant inefficient temporal utilisation of available road space are (1) the high demand for travel in private cars while maintaining a low occupancy in these vehicles, and (2) the timewise concentrated nature of the demand for commuter trips, which realises in severe travel peaking. A contributing factor to these phenomena is the absence of a direct pricing mechanism whereby road users are required to bear the costs they impose on others. Among the economic solutions to road congestion, therefore, is pricing aimed at charging road users the marginal cost of road use during peak travel periods [3]. Road congestion pricing is a means of obtaining better value from roads by making efficient use of available road space.

The provision of adequate public transport to divert private vehicle users is a prerequisite to the application of road congestion pricing, otherwise the benefits of a congestion charging scheme would not materialise. A congestion pricing scheme should meet the following requirements [4]:

1. It should discourage trips that are valued at less than the cost they impose, but should not restrain trips that are valued at more than their cost.
2. The charges levied should be related to the distance travelled and time spent on congested roads.
3. It should be equitable in its incidence on individual road users, simple to understand, reliable, reasonably fraud-proof, and publicly acceptable.
4. It should be capable of being implemented, administered and enforced within reasonable limits of effort and expense.

The principles involved are indicated in the following analysis of the benefits resulting from road traffic congestion pricing. In the analysis “trips” refer to private car trips on a specified finite road section.

In Figure 1, q represents traffic flow during the peak hour in passenger car equivalents per lane per hour (pce/l/h); $d(q)$ the demand for trips on the analysed road section; $a(q)$ the average cost per trip incurred by the drivers of private cars; and $m(q)$ the marginal cost per trip. Although these three curves in Figure 1 are similar to the graphical illustrations found in texts dealing with congestion pricing [5, 7, 9], the values quoted in the figure are based on empirical work (see §3).

The average cost per trip $a(q)$ is computed as

$$a(q) = T \cdot t(q) + r \quad (1)$$

where T represents the value of time; t the travel time (as a function of the traffic flow q); and r the vehicle running costs, which are assumed to be perceived by road users as being

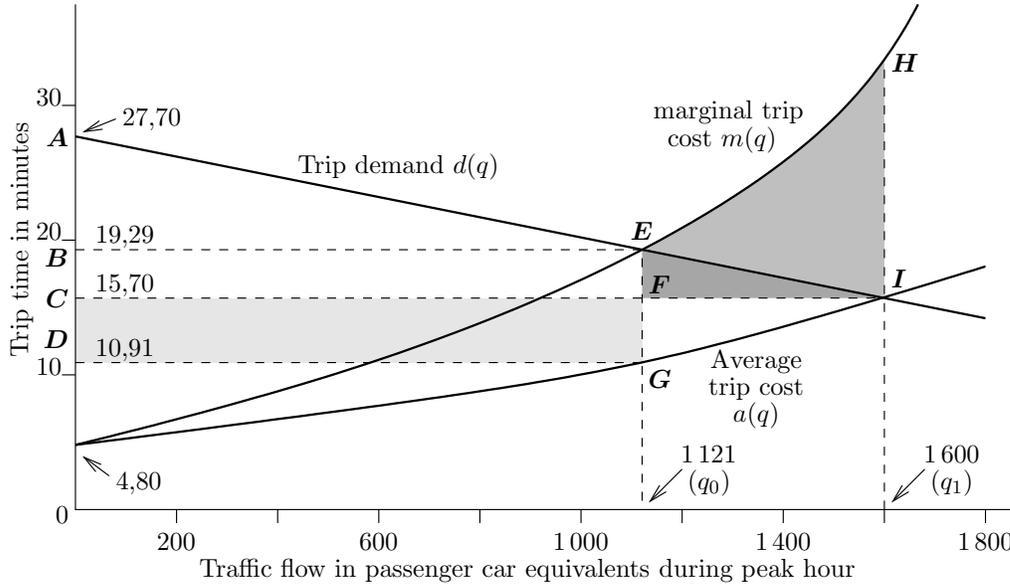


Figure 1: Private car trips to the central business district of Cape Town on the N1-route during the morning peak hour: cost and demand functions.

the same at all traffic levels. Although car running costs rise with increases in travel time in congested urban travel conditions, they are usually regarded by road users as being distance-dependent and not travel-time dependent [6].

In terms of (1), the comparative costs of congestion are taken to be time costs only and the marginal cost per trip is therefore

$$m(q) = \frac{d}{dq}[q \cdot a(q)] = a(q) + q \frac{da}{dq}. \quad (2)$$

The intersection of the demand curve $d(q)$ and the average cost curve $a(q)$ at point I in Figure 1, indicates the equilibrium flow of private vehicles during the morning peak hour, when no congestion charge for access to the central business district (CBD) is levied. In the figure the corresponding traffic flow is $q_1 = 1600$ vehicles per hour. As the cost of congestion has not been taken into account by individual drivers, the traffic flow of 1600 vehicles per hour indicated by I is not efficient. The optimal flow would correspond to the intersection of the demand curve $d(q)$ and the marginal cost curve $m(q)$ at E , and the object of charging a congestion levy would be to achieve the reduction of peak hour flow to that level. This flow is optimal, because at this point no cross-subsidisation takes place: the value of each trip is equal to or exceeds the cost it imposes [2]. In Figure 1 the flow of vehicles would decline from q_1 to q_0 , *i.e.* from 1600 to 1121. The appropriate levy is represented by EG , which is equal to BD , *i.e.* 8,38 minutes.

The levy $\ell(q)$, is equal to the difference between $m(q)$ and $a(q)$. Therefore

$$\ell(q) = m(q) - a(q) = q \frac{da}{dq}, \quad (3)$$

or, from (1),

$$\ell(q) = q \frac{d}{dq} [T \cdot t(q) + r] = qT \frac{dt}{dq}. \quad (4)$$

The imposition of the congestion levy would lead to a revenue of $\ell(q) \cdot q_0$, which is represented by the area $DBEG$ in Figure 1. The decline in traffic from q_1 to q_0 will reduce the user surplus to the extent represented by the area $CBEI$, leaving a net benefit equivalent to the difference of the two areas. This difference is equivalent to the eliminated efficiency loss

$$e = \int_{q_0}^{q_1} [m(q) - d(q)] d(q) \quad (5)$$

achieved through the congestion levy as represented by the area EHI . It can be shown that area EHI is equal to the difference of the two areas $DCFG$ and FEI , or that area $EHI + \text{area } FEI = \text{area } DCFG$. This relationship may be established as follows, by utilising the fundamental theorem of the calculus:

$$\begin{aligned} \text{Area } EHI + \text{area } FEI &= \int_{q_0}^{q_1} (m(q) - a(q_1)) dq \\ &= \int_{q_0}^{q_1} \left(\frac{d}{dq} (qa(q)) - a(q_1) \right) dq \\ &= qa(q) \Big|_{q_0}^{q_1} - a(q_1)q \Big|_{q_0}^{q_1} \\ &= q_1 a(q_1) - q_0 a(q_0) - a(q_1)q_1 + a(q_1)q_0 \\ &= q_0 (a(q_1) - a(q_0)) \\ &= \text{area } DCFG. \end{aligned} \quad (6)$$

3 Congestion pricing example

Although the congestion pricing example illustrated in Figure 1 is hypothetical, the demand curve and cost functions quoted are based on empirical work: In the case forming the basis of this empirical work, the relevant roads authorities intend to jointly introduce a congestion pricing scheme during the morning peak travel period on the CBD bound lanes of the N1-Table Bay Boulevard entrance into Cape Town. The priced section is ten kilometres long and it stretches from the Heerengracht-Table Bay Boulevard Intersection towards the Wingfield Interchange on the N1 Freeway. Motorists who do not want to be subject to the pricing scheme will have to leave the N1 Freeway at the Wingfield Interchange and travel by train or bus to the CBD. All alternative entrances into Cape Town will be priced by the local authority, which will prevent (a) inefficient traffic diversion to other roads, which increases travel distances, and (b) the displacement of the present congestion on the N1 Freeway to other roads. A convenient, safely guarded parking lot will be in operation at the railway station and bus terminal adjacent to the freeway from where those motorists who divert to public transport can commence their onward public transport trip portions.

The trip demand function was approximated from the responses obtained from a sample of 20 car drivers who traverse the section each morning. Their willingness to make use of the

road section at various travel times over the ten kilometres, and related value perceptions can be summarised as follows: In 2000, when this case study was performed, R1,00/minute was a good proxy value for the travel cost of a car commuter on the surveyed road section [6]. Based on the assumed demand function, and with no pricing scheme, the equilibrium q (in pce/l/h) will be 1 600, and the corresponding average travel time will be 15,70 minutes (see Figure 1).

The theoretical capacity of a lane section on a freeway with a design speed of 120 km/h is 2 000 pce/h [8]. However, at this q -value flow is unstable and the capacity flow rate cannot be maintained for a full hour. At a travel time of 15,70 minutes over the 10 km section flow remains fairly stable, which serves as an explanation why the unpriced flow can be in equilibrium at a q -value of 1 600 pce/l/h.

With the imposition of a congestion price equivalent to 8,38 minutes, flow during the morning peak hour will be optimal, with $q = 1\,121$ and $t = 10,91$ (see Figure 1).

Aggregate travel demand of the sample of drivers included in the study remains inelastic (*i.e.* everybody who travels before congestion pricing will travel after the implementation of congestion pricing). One half (10) of the peak hour private car drivers indicated that they would continue to travel during the peak hour in a private car. The other 10 peak hour private car drivers responded as follows: 6 were willing to travel in the hour directly before or directly after the peak hour at approximately 50 per cent of the peak hour congestion price; no respondents were willing to travel more than an hour outside the peak hour, even if no congestion price is charged during those hours; none were willing to travel by train, because Cape Town railway station was deemed too far from their destinations to walk; and 4 were willing to participate in a lift club or a coordinated, convenient and comfortable commuter bus scheme (with bus headways shorter than five minutes, and the CBD bus stops in close proximity of the passengers' destinations). The inelastic aggregate travel demand of the drivers included in the sample can be ascribed to the fact that their commuter trips are made involuntarily (they cannot readily change their places of work and dwelling). Furthermore, 30 per cent of the respondents indicated that they are compelled to commute by car as these are supplied by their employers.

Based on the respondents' stated willingness to undertake trips at various travel times in minutes, the assumed demand function (the number of car trips per lane per hour at different travel times in minutes over the 10 km road section) may be approximated as $27,7 - 0,0075q$.

The average speed-flow relationship on each lane of the entire priced road section is $125 - q^{0,605}$ kilometres per hour where average speed is the average trip speed over the ten kilometre road section in km/h. Therefore, average travel time (in minutes) over the ten kilometre section is $600/(125 - q^{0,605})$ minutes.

This relationship was approximated from the travel times of 15 test runs over the entire road section when different traffic flow rates prevailed. Using this speed-flow relationship, the average travel time in minutes over the entire road section at different traffic flow rates may readily be calculated.

Marginal travel time is derived directly from average travel (or trip) time.

Buses are more efficient in the use of road space than cars. The average car occupancy in the morning peak hour on the analysed road section is 1,25 persons, while the average occupancy of the buses that supplant the cars is 30 passengers. The road space of one bus is 2,4 cars. Therefore, in terms of passenger flow, one bus supplants 10 cars, while vehicle running costs (r) for these types of buses are approximately only three times that of cars [1]. In terms of the practical case illustrated, area $DCFG$ is equivalent to R5 369,59 (or 5 369,59 minutes); area FEI to R859,81; and area EHI to R4 509, 78.

This equivalency of the areas in Figure 1 holds true for any similar average cost, marginal cost and demand function on a road elsewhere, regardless of the shape of the functions, as shown in (6).

4 Conclusions

The optimal traffic flow rate on a roadway is where the marginal road user cost of motorists (MC) and trip demand are equal. When no congestion price is charged, the equilibrium flow of traffic on the roadway will be where average road user cost (AC) and trip demand are equal, which is not optimal. The imposition of a congestion charge equal to the difference between MC and AC at the traffic flow rate where MC equals trip demand, will result in an optimal traffic flow rate. The peak hour revenue of the road authority charging the congestion levy will be equal to the peak hour traffic volume in the pricing situation multiplied by the congestion price (*i.e.* area $DBEG$ in Figure 1). This revenue will be equal to the reduction in the user surplus resulting from the imposition of the congestion charge (*i.e.* area $CBEI$) together with the eliminated efficiency loss (*i.e.* the previous excess of MC over the value of those car trips that have been priced off the road — area EHI). However, it has been shown that area EHI is equal to the difference of areas $DCFG$ and FEI , which represents a simpler way to determine the efficiency loss.

The mathematical principles involved in road congestion pricing, as described in the paper hold true for any capacity constrained facility or service to which congestion (*i.e.* marginal cost) pricing is applied, *e.g.* rail services, airports, seaports, electricity supply and telecommunications.

Road congestion pricing can be instrumental in ensuring that those trips whose cost is more than their value can take place in an alternative way so that, or at a time when, their value exceeds the cost they impose. For example, private car trips whose value lies between the average and marginal cost that they impose can, as a result of congestion pricing, be undertaken outside the peak travel period, or such trip makers may organise lift clubs or divert to public transport. It is conditional to the success of a road congestion pricing scheme that effective public transport is available as an alternative to private car travel.

Acknowledgement

The authors would like to thank Dr Alewyn Burger for his constructive suggestions with respect to elucidation of the mathematical content of the paper.

References

- [1] AFRICON, 2008, *Study on the macro-economic impact of fuel costs in transport*, Contract Report, Department of Transport, Pretoria.
- [2] BUTTON KJ, 1993, *Transport economics*, 2nd Edition, Edward Elgar, Aldershot, Hants.
- [3] FLOOR BC, 1975, *Cost-benefit analysis in the transport sector*, South African Journal of Economics, **43(2)**, pp. 205–220.
- [4] FOURIE SA & FREEMAN PNW, 1980, *Road pricing and supplementary licensing: The current position*, Technical Report RT/9/80, National Institute for Transport and Road Research, CSIR, Pretoria.
- [5] LI MZF, 2002, *The role of speed-flow relationships in congestion pricing implementation with an application to Singapore*, Transport Research Part B, **36**, pp. 731–754.
- [6] PIENAAR WJ, 2000, *The economic evaluation of bus and minibus taxi terminals and transfer facilities*, PhD dissertation, Stellenbosch University, Stellenbosch.
- [7] ROUWENDAL J & VERHOEF ET, 2006, *Basic economic principles of road pricing: From theory to applications*, Transport Policy, **13**, pp. 106–114.
- [8] TRANSPORTION RESEARCH BOARD, 2000, *Highway capacity manual*, National Research Council, Washington (DC).
- [9] YANG H, MENG Q & LEE D-H, 2004, *Trial-and-error implementation of marginal-cost pricing on networks in the absence of demand functions*, Transport Research Part B, **38**, pp. 477–493.

