25th Australasian Transport Research Forum
Incorporating the BTRE Transport Policy Colloquium
Canberra 2-4 October 2002

Measuring the benefits of new transport services

Mark Harvey
Bureau of Transport and Regional Economics, Canberra

Abstract
Measuring the benefits of a new transport service as part of a cost–benefit analysis entails estimation of the entire consumers' surplus area under the demand curve. This requires knowledge of the demand curve over a substantial part of its length — information that is extremely difficult to obtain. Where the new service has a fairly close substitute already in existence, the difficulty is considerably reduced because the price and other observable characteristics of the substitute are major determinants of the position and shape of the demand curve for the new service.

The paper examines some of the technical issues in inferring consumers' surplus levels for new services from information about existing substitute services. The demand for the new service can be subdivided into traffic transferring from each substitute service and generated traffic. For transferring traffic, it is shown that the total benefits are simply: the saving in the costs for the existing transport alternative, plus the additional costs of the new alternative, plus or minus the value of service quality differences.

The relevant value to use for non-price attributes, whether for transferred or generated traffic, is not the average value for the whole population of freight or passengers, but the average value for traffic that actually uses the new service. If the new service is cheaper and inferior to the existing service, the relevant value of non-price attributes that for traffic at the low end of the distribution of values, and conversely where new service is a dearer, superior one. Inferences about values of non-price attributes can be drawn from the prices of the existing and new services. The relationships derived can be employed to derive values to use in rough cost–benefit analyses and for bounds checking of values from other sources.

The views expressed in this paper are those of the author and do not necessarily represent those of the Bureau of Transport and Regional Economics. The usual caveats apply.

Contact author
Mark Harvey
Bureau of Transport and Regional Economics
GPO Box 501   Canberra, ACT 2601
Phone: 06-6274-6720   Fax: 06-6274-6816
Email: Mark.Harvey@dotars.gov.au

GO BACK TO START SCREEN   GO BACK TO INDEX

Introduction

Measuring the benefits of a new transport service as part of a cost–benefit analysis entails estimation of the entire consumers' surplus area under the demand curve. This requires knowledge of the demand curve over a substantial part of its length — information that is extremely difficult to obtain. Where the new service has a fairly close substitute already in existence, the difficulty is considerably reduced because the price and other observable characteristics of the substitute are major determinants of the position and shape of the demand curve for the new service. An example is BTE (2000), which describes a cost–benefit analysis of a proposed new railway line between Brisbane and Melbourne. The new line would draw freight from faster, more expensive road transport and from the slower, cheaper coastal railway, as well as generate some traffic that would not otherwise exist.

This paper examines some of the technical issues that arise in attempting to measure the benefits of new transport services, with the aim of promoting more robust estimation of these benefits. The paper makes a number of suggestions for analysts carrying out rough cost–benefit analyses where information and resources to gather additional information are limited. A detailed cost–benefit analysis of a major infrastructure project can be a costly exercise, particularly if market surveys are involved. The ability to undertake rough cost–benefit analyses to screen out potential projects that are unlikely to be viable on closer examination is therefore important.

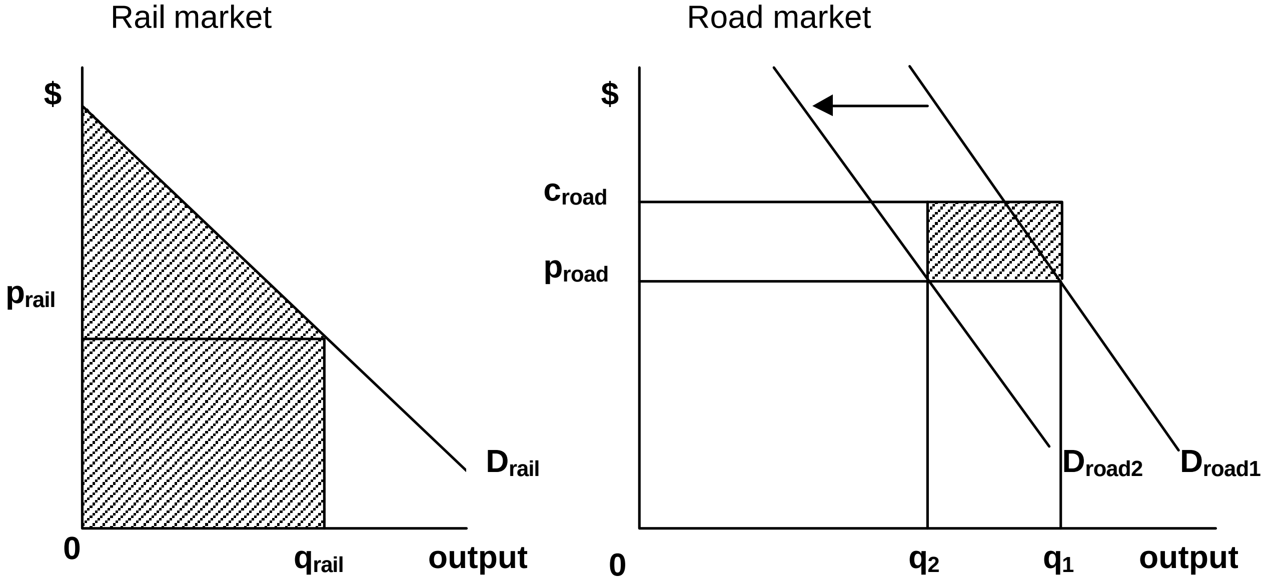
The paper is couched in terms of freight transport, but the discussion is equally applicable to passenger transport.

The problem in a rail context

When undertaking a cost–benefit analysis of a new rail service, the main economic benefits consist of the total willingness-to-pay (WTP) of users of the service, plus an additional amount where road transport is not priced at its marginal social cost. Figure 1 illustrates this, with the benefits being the shaded areas.[1] The two diagrams represent the rail and road markets. The demand curves are in terms of money costs, not generalised costs. Rail must charge a lower price than road in order to compete against road's faster trip times and greater reliability. The quantity for freight carried by rail, q_{rail}, consists of two components, that captured from road (equal to q_2q_1 in the road market diagram) and generated traffic, that is, new traffic that has only come into being as a result of the rail service.

[1] Demand curves throughout the paper have, in most cases, been shown as linear for ease of drawing. There is no suggestion they will be linear in reality.

Figure 1 Main benefits from a new railway line



For each year during which the project operates, the net benefits are found by subtracted from rail's WTP, the railway's operating and externality costs. The capital costs are incurred during the construction phase. These are not the focus of the present paper.

For the road market, it has been assumed that costs are constant. This would be realistic for non-urban roads where there is little congestion and in urban areas where local traffic is so dominant that taking some long-distance trucks off the roads has negligible effect on congestion. The c_{road} line represents marginal social costs, that is, it includes external costs and road damage as well as vehicle operating costs, but excludes all taxes except for those on labour.[2] It is generally accepted that the larger trucks are not meeting the full costs of the damage they do to roads (see BTE 1999). Also, there are externalities of noise, pollution and vibration in urban areas, and accident costs. The price paid for road transport (p_{road}) is that which freight consignors actually pay and would cover all taxes and charges levied on the road haulage industry.

Opening the new railway line causes the demand curve for road transport to shift leftward from D_{road1} to D_{road2}. For each tonne of road freight that moves to rail, society's valuation of the task, as measured by the price, is less than the cost to society, so there is a net gain to society of cost minus price. The shaded rectangle in the road market diagram is the benefit in this market. Note that if price equalled

---

[2] Fuel excise, sales taxes, tariffs and registration charges are *excluded* on the assumption that inputs of fuel, tyres, spare parts, oil, and vehicles, being internationally traded goods, are available in infinitely elastic supply. Labour is assumed to be drawn from other industries. Income and payroll taxes are *included* because industries would employ labour up to the point where value of marginal product equals *pre-tax* wages.

marginal social cost, the road market could be ignored altogether, and if price exceeded cost, the rectangle would be a cost instead of a benefit.[3]

Estimating the WTP area for rail is difficult because it requires knowledge of the demand curve over the whole length of the curve from zero quantity up to q_{rail}. Usually it is difficult enough estimating the elasticity for a short part of the demand curve. Where functions have been fitted to data, estimating total WTP is likely to involve extrapolation well beyond data points. Often these functions are asymptotic to the price axis so an arbitrary decision has to be made about a cut-off point. Having a reliable estimate of consumers' surplus can be particularly important for evaluating projects where there is a minimum scale of investment. In such cases, the situation sometimes arises where financial analysis indicates that the project at the minimum scale is not viable, but that the consumers' surplus benefits are sufficient to make the project economically warranted.[4] The two diagrams in figure 2 show the demand curves for a new railway line. Cost level c_o is the average operating cost, and c_T is the average total cost comprised of operating cost plus the annuitised capital cost per unit of output including a normal rate of return on capital. Both operating and capital costs are assumed to be constant with respect to output up to a rigid capacity constraint. If there were no lumpiness in investment, one could invest in capacity *q* as shown in the left diagram. The project would pass the financial test because revenue (area *B*) would just equal costs (also area *B*), and capital costs, as defined here, includes a normal return on capital. The project would also easily pass the economic test because total willingness-to-pay (areas *A* + *B*) is well above costs (area *B*).

By contrast, say the minimum standard railway track had a capacity of *q** as shown in the right diagram of figure 2. Charging a price that achieves full utilisation of capacity, *p*, causes the enterprise to run at a loss as price is below average total cost. However, as long as the size of area *A* exceeds that of area *E*, the project is warranted on economic grounds.[5]

Such a project would require government assistance to proceed. The result of the cost–benefit analysis, which is likely to hinge on the estimate of total consumers' surplus, would be a critical factor in deciding whether to provide assistance.

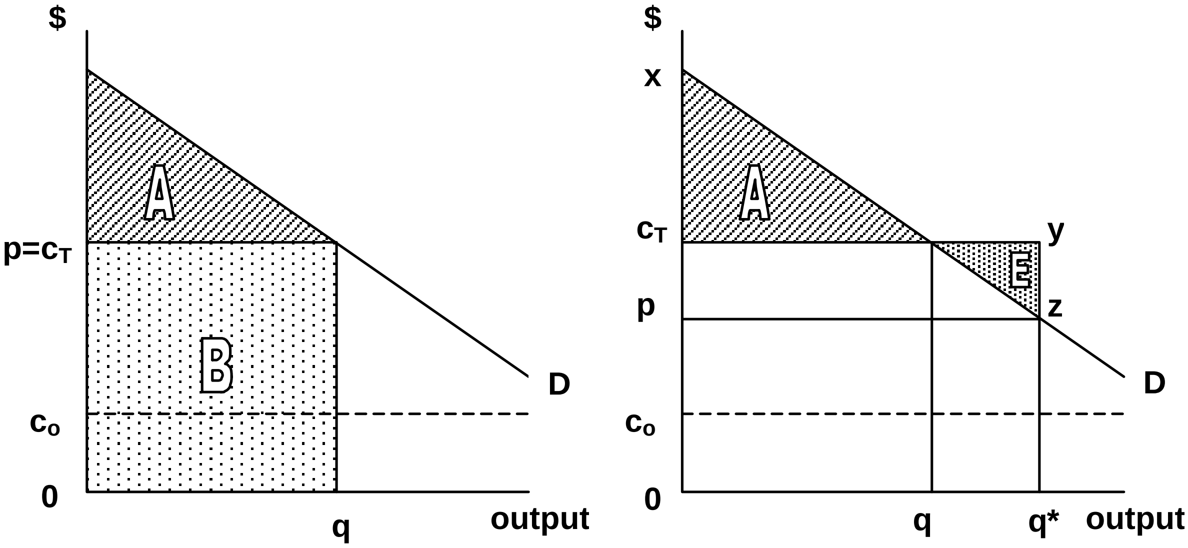
---

3 For a more detailed discussion of the valuation of benefits in related markets see, Boardman et al (1996, chapter 3).

4 The railway operator may be able to capture part of this consumers' surplus through price discrimination, though competition from other modes of transport generally severely limits the ability of railways to do so. Another means is to purchase land close to proposed stations or terminals in order to capture increases in land values that result from the railway project.

5 Annual net benefits from the project equal total willingness-to-pay, area $0q^*zx$, minus total costs, rectangle $0q^*yc_T$, which equals area *A* minus area *E*.

Figure 2 Project benefits where there is a minimum scale of investment



Subdividing the problem

The first step in attempting to estimate total consumers' surplus is to distinguish demands from different sources so they can be considered separately. Separate demand curves can be identified for traffic transferring to the new infrastructure from each substitute service and for generated traffic. 'Generated traffic' is traffic that comes into existence as a result of the project, for example, from new industries appearing or new markets being opened up for existing industries. In the case of BTE 2000, there were four sources of demand: freight that would otherwise travel by road, by the existing coastal railway, and by ship (landbridged containers), as well as generated freight. Since the demand curve for the new service is the horizontal addition of all of these demand curves, consumers' surplus can be estimated separately for each source of demand and summed.

In this paper, estimation of consumers' surplus for transferring traffic is considered first, then for generated traffic. The subscript *N* will be used to indicate the new service for which we wish to estimate benefits, and the subscript *E* for the existing service from which traffic is diverted.

A convenient assumption to make is that the new and existing services are perfect substitutes in all ways except for a limited number of specified, measurable attributes. These attributes might include time taken (determined by both speed and frequency), departure times, reliability, pick up and delivery costs, packaging costs and damage costs. For passengers, the last three would be replaced with other considerations such as comfort and safety. For most of the discussion, for simplicity, it is assumed that time taken is the only non-price attribute, though in places, the conclusions are generalised.

A consequence of this assumption is that the measure of consumers' surplus will be inaccurate to the extent that the new service differs from the existing service in ways that have not been specifically taken into account. In the new rail service case, say only price and time taken were considered in the cost–benefit analysis. If rail charged the same price as road, and despite rail's being slower, some consignors still used rail, this would indicate that there are some attributes of the rail service that some customers value and which would not be measured. It is tempting the say that the assumption leads to a conservative estimate of consumers' surplus, but one cannot be sure that there are no negative attributes that have been ignored.

Transferring traffic

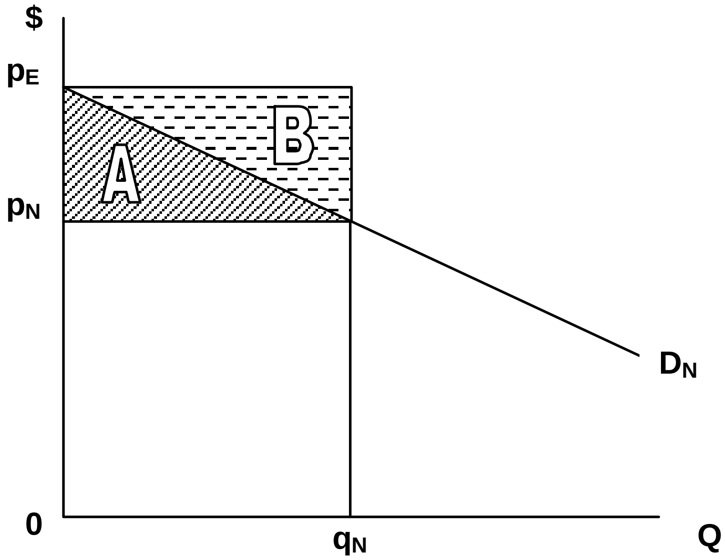
Cheaper, slower alternative case

If the new and existing services were perfect substitutes in all respects except that the new service is slower, then the price (p_E) charged by the existing service would be the maximum price a consignor would be prepared to pay for the new service. This provides the price axis intercept for the demand curve for the new service for traffic that transfers, as shown in figure 3. The consignor of the first tonne of freight that switches to the new service is willing to transfer for a very small price reduction and so must value the time difference at close to zero.[6] The last consignor to transfer would value the time difference at p_E minus p_N, where p_N is the price for the new rail service. For all freight that transfers, therefore, the difference between the height of the demand curve (D_N) and p_E represents the value the consignor places on time lost from using the slower new service $v_i(t_E - t_N) = v_i t \leq 0$ (where, for the *i*th tonne of freight that transfers, v_i is the value of time, t_E the time taken for the existing service, t_N the time taken for the new service, and t the time saving, which is negative in this case. The area of triangle *B* is total cost of the time lost by consignors from switching modes, and the consumers' surplus gain (area *A*) is $(p_E - p_N + \bar{v}t)q_N$ where $\bar{v}$ is the average value of time for freight that transfers, and q_N the total traffic that transfers.

6 It is assumed throughout this paper that there are no negative values of time. There are circumstances in which consignors of freight or passengers would have negative values of time, but they are rare enough not warrant consideration here.

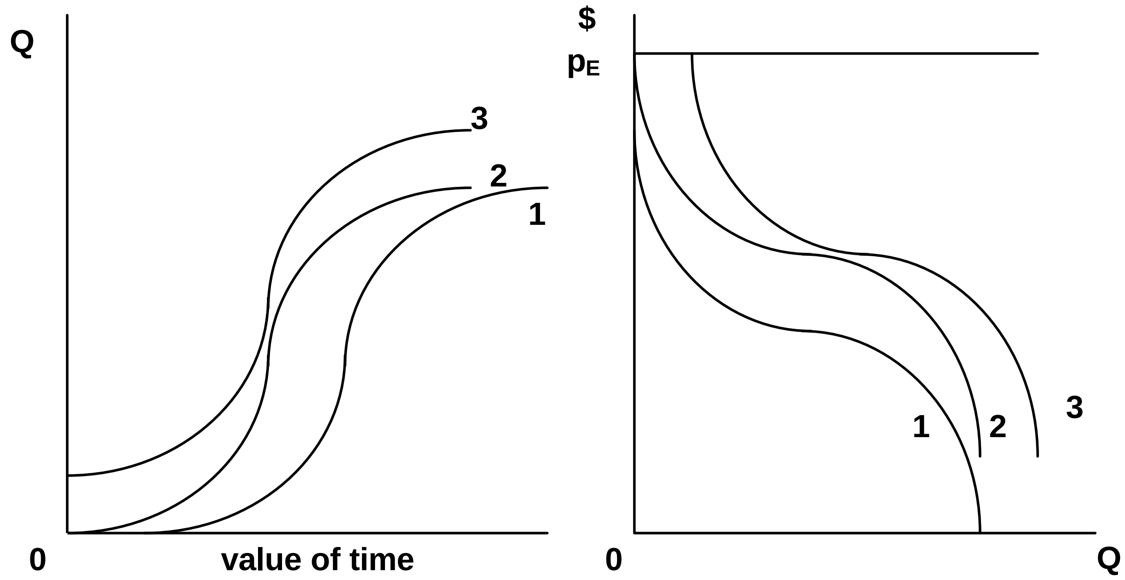
This same result can be derived using the generalised cost approach. The generalised cost to a consignor using road transport is $p_E + \bar{v}t_E$ and that for rail is $p_N + \bar{v}t_N$. The net gain to consignors who transfer is the saving in generalised costs multiplied by the quantity of freight that transfers $[(p_E + \bar{v}t_E) - (p_N + \bar{v}t_N)]q_N$ = $(p_E - p_N + \bar{v}t)q_N$.

Figure 3 Demand by transferring traffic (cheaper, slower alternative)



One important conclusion from the derivation via figure 3 is that the relevant value of time is not the average for whole population of freight that travels by existing service between the origin and destination in question, but only for the freight that transfers. Freight that transfers to a slower alternative would be that having the lowest values of time. So $\bar{v}$ would lie below the average value for the whole population, except in the case where the whole population transferred. To use the population average value of time would lead to understatement of project benefits. The shape of the demand curve for transferring traffic would reflect the lower, left-hand tail of the cumulative frequency distribution for value of time for traffic using the existing mode. Three such hypothetical curves are illustrated in figure 4, with the resultant demand curves. The cumulative frequency distribution curves on the left side represent the number of tonnes of freight having a value of time less than the corresponding value of time on the horizontal axis.

Figure 4 Cumulative frequency distributions of values of time and implications for shapes of demand curves



If the cumulative frequency distribution curve passed through the origin as does curve two, and the difference between p_E and p_N was not too great, an analyst lacking knowledge of the appropriate value of time might be justified in using the 'rule-of-a-half': $\bar{v} = -(p_E - p_N)/2t$ and consumers' surplus $= (p_E - p_N)q_N/2$. This is equivalent to assuming a linear demand curve starting from p_E as shown in figure 3 and implies a uniform distribution of values of time (a cumulative frequency distribution curve that is a ray from the origin). However, there is no guarantee that the curve will pass through the origin. There could be no consignors with values of time at or close to zero (curve 1) or many (curve 3). All we can be sure of is that the relevant average value of time ranges between zero and $-(p_E - p_N)/t$, and by implication, the consumers' surplus area ranges between zero and the sum of areas A and B in figure 3, $(p_E - p_N)q_N$. Even so, this is useful information for carrying out rough cost–benefit analyses as well as for bounds checking in situations where information about values of time is available. For a rough cost–benefit analysis, the rule-of-a-half with sensitivity testing for average values of time of zero and $-(p_E - p_N)/t$ would not be unreasonable. If the analyst does have information about values of time for traffic that transfers, a boundary check can be carried out that the value employed does not exceed $-(p_E - p_N)/t$.

To produce a formal expression for consumers' surplus, say the analyst has information about the frequency distribution of values of time for users of the existing service:

$q(v)$ where $\int_0^{-(p_E - p_N)/t} q(v)dv = q_N$ (1) and $\int_0^{\infty} q(v)dv = q_T$, where q_T is the total number of tonnes travelling by the existing alternative, which could potentially transfer. The consumers' surplus area is:

$$(p_E - p_N)q_N + t\int_0^{-(p_E - p_N)/t} vq(v)dv \,. \quad (2)$$

Since the average value of time for traffic that transfers is $\bar{v} = \int_0^{-(p_E - p_N)/t} vq(v)dv \Big/ q_N$, equation (2) reduces to $(p_E - p_N + \bar{v}t)q_N$, the earlier result.

Total net benefits

Cost–benefit analysis entails adding up all the benefits and costs to all members of society. The total annual net benefit after the project commences operation is shown in table 1.

Table 1 Components of annual benefit summed

Benefit	Formula
Consumer surplus in the rail market	$(p_E - p_N + \bar{v}t)q_N$
plus producer surplus of the rail operator and taxes transferred to the government: where c_N is average annual operating costs after deducting taxes on non-labour inputs	$(p_N - c_N)q_N$
plus benefit in the road market: where c_E is the average annual social cost of road transport (excluding externalities) and x_E is the average annual road externality cost	$(c_E + x_E - p_E)q_N$
minus rail externalities: where x_N is the average annual rail externality cost	$-x_N q_N$
Total net annual benefit	$(c_E + x_E - c_N - x_N + \bar{v}t)q_N$

When the benefits are combined, the price terms cancel out leaving the simple and intuitively obvious result that the annual net benefit from traffic transferring from an existing transport service to a new service is:

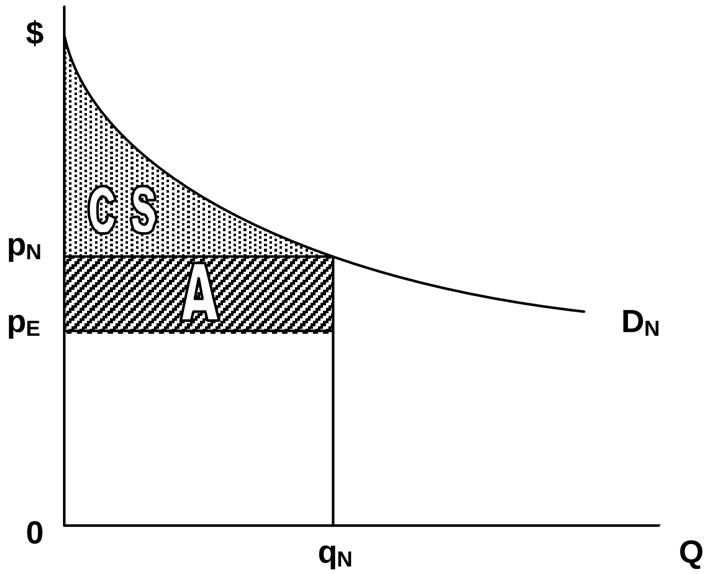
- the saving in costs to society from reduced use of the existing service;
- less the costs to society from use of the new service;
- less the cost of time lost.

It can be shown that this same result applies where traffic transfers to a dearer, superior transport alternative and where the costs of the existing transport alternative are rising. The latter case is examined in the appendix where the new service is a cheaper, slower alternative.

Dearer, faster alternative case

In BTE 2000, traffic that transferred from the existing coastal railway to the new inland railway was transferring to a dearer but faster alternative. Consignors should be willing to pay more for the new, faster service than they do for the existing service. Figure 5 shows the relationship between the demand curve for the new service and the prices of the new and existing services.

Figure 5 Demand by transferring traffic (dearer, superior alternative)



The dearer, faster mode case is more difficult because the characteristics of the existing alternative provide less information about the nature the demand curve above p_N. The vertical distance between the demand curve and p_E still represents the value of the time saving for the consignors transferring. However, in contrast to the cheaper, slower service case, those that transfer come from the upper, right-hand section of the cumulative frequency distribution for value of time, starting from the right. The consignor willing to pay the highest price to switch to the new service would be the one with the highest value of time. As price is lowered, those with progressively lower values would transfer. The last consignor to transfer would have a value of time of $(p_N - p_E)/t$.[7] This provides a lower bound to the value of time to use in the formula $(c_E + x_E - c_N - x_N + \bar{v}t)q_N$, derived in table 1. It should be noted that at this lower bound, consumers' surplus is zero.

7 This is consistent with the cheaper, slower alternative case, discussed previously where the last consignor to transfer has value of time of $-(p_E - p_N)/t$ where *t*<*0*.

It is possible to do better if an estimate is available for the population average value of time and this is greater than $(p_N - p_E)/t$. The average value of time for the whole population of users of the existing alternative must be less than or equal to the average value of time for the traffic that transfers. Hence use of the greater of the population average value of time and $(p_N - p_E)/t$ will lead to a conservative estimate of project benefits.

In the dearer, faster case, the general expression for consumers' surplus, equation 2, undergoes a slight change to account for the fact that the transferring traffic comes from the upper end of the distribution of values of time.

$$(p_E - p_N)q_N + t\int_{(p_N - p_E)/t}^{\infty} vq(v)dv \quad (3) \quad \text{where } \int_{(p_N - p_E)/t}^{\infty} q(v)dv = q_N . \quad (4)$$

The first term in equation (3) is negative, so that the area of rectangle *A* in figure 5 is subtracted from the area between the demand curve and p_E, to leave the consumers' surplus area (marked *CS*), above p_N.

Since the average value of time for traffic that transfers is $\bar{v} = \int_{(p_N - p_E)/t}^{\infty} vq(v)dv \Big/ q_N$,

equation (3) reduces to $(p_E - p_N + \bar{v}t)q_N$.

Some generalisations

The generalised costs of transport can be written as:

$G = p + \sum_{j=1}^{m} v_j a_j$ where there are *m* non-price attributes and v_j and a_j are the values and amounts of the *j*th attribute, respectively. Attributes are defined like price and time, so that less is preferred to more, that is, any increase the level of the attribute raises generalised costs. For example, other attributes might include unreliability (not reliability), expected value of damage costs, and packing costs. A consignor will only transfer to the new alternative if:

$$G_N - G_E = p_N - p_E + \sum_{j=1}^{m} v_j \left(a_{jN} - a_{jE}\right) \leq 0 .$$

For a cheaper, inferior alternative: $p_N - p_E < 0$ and $a_{jN} - a_{jE} > 0$, hence

$p_E - p_N > \sum_{j=1}^{m} v_j \left(a_{jN} - a_{jE}\right)$. (Note that p_E and p_N have been switched around here to keep the price difference positive.)

For a dearer, superior alternative: $p_N - p_E > 0$ and $a_{jN} - a_{jE} < 0$, hence

$$p_N - p_E < \sum_{j=1}^{m} v_j \left(a_{jN} - a_{jE}\right)$$

These relationships are useful for bounds checking, and for rough cost–benefit analyses, they supply some starting values. In the absence of information about the values of non-price attributes, for a cheaper, inferior alternative, $(p_E - p_N)q_N/2$ would not be an unreasonable amount to deduct for the cost to consignors of the reduction in service quality, provided $p_E - p_N$ is not very large. For a dearer, superior alternative, $(p_N - p_E)q_N$ is the minimum amount to add on to account for the value to consignors of the improvement in service quality. However, if population estimates of average values of attributes are available and $\sum_{j=1}^{m} \bar{v}_j (a_{jN} - a_{jE}) > p_N - p_E$ where the $\bar{v}_j$ are the population averages, then these can be used to derive a conservative estimate.

As already discussed in the context of time savings, the values of attributes employed should be those for traffic that actually transfers. In the multiple attribute case, there is an added complication that the order in which units of freight appear in the frequency distributions for each attribute may not be the same. For example, the tonne for which the consignor has the highest value of time may not also have the highest value of unreliability, though it is unlikely.

Where there is a mixture of gains and losses in non-price attributes from switching to the new transport alternative, the value of $\sum_{j=1}^{m} v_j (a_{jN} - a_{jE})$ will show overall, whether the alternative is inferior or superior.

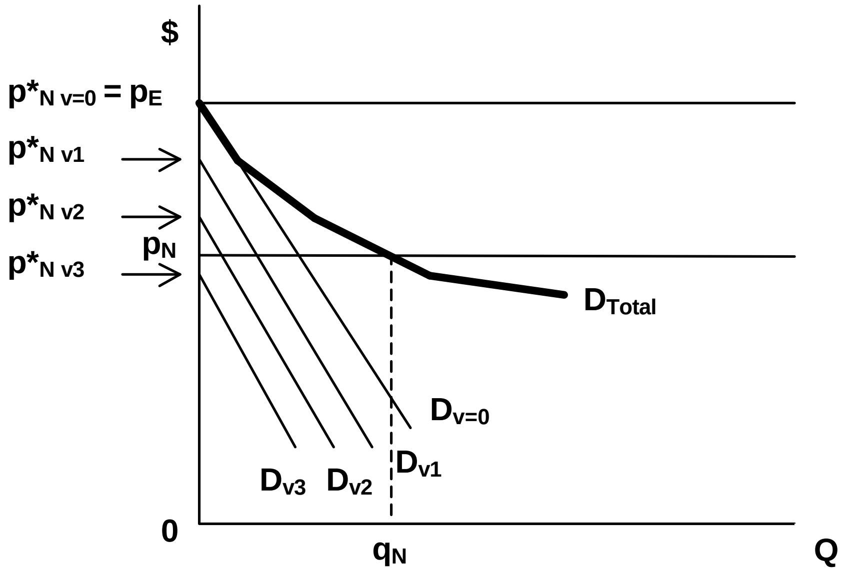
Generated traffic

Cheaper, slower alternative

For generated traffic in the case of a cheaper, slower alternative, the existing alternative is not sufficiently attractive for potential consignors to use. So p_E is the upper limit on consignors' valuations of the new service. A consignor's WTP for the new service will be affected by their value of time. If that value were zero, the first generated consignor to use the new service would be willing to pay almost p_E. Then there would be other consignors with positive values of time whose WTPs would lie below p_E . For groups of consignors with positive values of time, the WTP for the first unit of freight would be less than p_E by an amount equal to the time cost. Segmenting the market into groups of consignors according to their values of time (v_i), the demand curve for each group *i* would have an intercept on the price axis of $p_{Ni}^* = p_E + v_i (t_E - t_N) = p_E + v_i t$, where *t* (<0) is the time saved had they switched from the alternative service. For a group with a given value of time, p_{Ni}^* is the WTP for the first unit of generated freight.

A family of these curves for groups with differing values of time is shown in figure 6 together with the market demand curve, which is the horizontal summation of these curves. For (potential) consignors with a value of time greater than $-\left(p_E - p_N\right)/t$ (the negative sign is to cancel out the negative value of *t*), their demand curves lie below p_N at all points and so they do not ship any freight at all (as in the case of the demand curve for group *v3* in figure 6). If the demand curves are shaped as in figure 6, the market demand curve will be strongly convex. This seems a credible hypothesis. If a new transport alternative with significantly inferior non-price characteristics offers customers only a small discount on the price charged by the existing service, then very little new traffic is likely to be generated. Only as the discount becomes large enough to offset the new alternative's relative disadvantages in the eyes of a large number of potential new customers, will significant new demand be generated.

Figure 6 Demand curves for generated traffic (cheaper, slower alternative)



The likely convexity of the total demand curve for generated traffic implies that using the rule-of-a-half (that is, an assumption of a linear demand curve) is likely to overstate benefits. However, provided the gap between p_E and p_N is not too great, the assumption of linearity may not be unreasonable for the demand curves for groups of potential customers with the same values of time, especially for a rough cost–benefit analysis. On this assumption, the area of the consumers' surplus triangle for the demand curve for value-of-time group *i* is:

$\left(\frac{p_{Ni}^{*} - p_N}{2}\right)q_i = \left(\frac{p_E + v_i t - p_N}{2}\right)q_i = \left(\frac{p_E - p_N}{2}\right)q_i + \left(\frac{v_i t}{2}\right)q_i$, where q_i is the number of units of freight consigned by value-of-time group *i* when price p_N is charged.

Summing over all q_i : $\left(\frac{p_E - p_N}{2}\right)q_N + \left(\frac{t}{2}\right)\sum_i v_i q_i$ where $\sum_i q_i = q_N$. (5)

The values of time in this case range from zero to $-(p_E - p_N)/t$. If a continuous frequency distribution for values of time, $q(v)$, were assumed, then equation (5) would become:

$$\left(\frac{p_E - p_N}{2}\right)q_N + \left(\frac{t}{2}\right)\int_0^{-(p_E-p_N)/t} vq(v)dv \quad (6) \quad \text{where } \int_0^{-(p_E-p_N)/t} q(v)dv = q_N. \quad (7)$$

or $(p_E - p_N + \bar{v}t)q_N/2$ (8) where $\bar{v} = \int_0^{-(p_E-p_N)/t} vq(v)dv \Big/ q_N$ is the average value of time for the generated traffic.

This amounts to the full consumers' surplus triangle as estimated by the rule-of-a-half, minus half the total value of the time lost by the generated traffic, had they switched from the existing alternative.

If a uniform frequency distribution is assumed for units of freight having values of time from zero to $-(p_E - p_N)/t$, then $\bar{v} = -(p_E - p_N)/2t$, and equation (8) for consumers' surplus conveniently reduces to $(p_E - p_N)q_N/4$ — a rule-of-a-quarter. If the frequency distribution rises with value of time, then the market demand curve becomes more convex and the fraction of $(p_E - p_N)q_N$ becomes smaller. For rough cost–benefit analyses, a consumers' surplus of $(p_E - p_N)q_N/4$ might be assumed, with sensitivity tests at zero and $(p_E - p_N)q_N/2$.

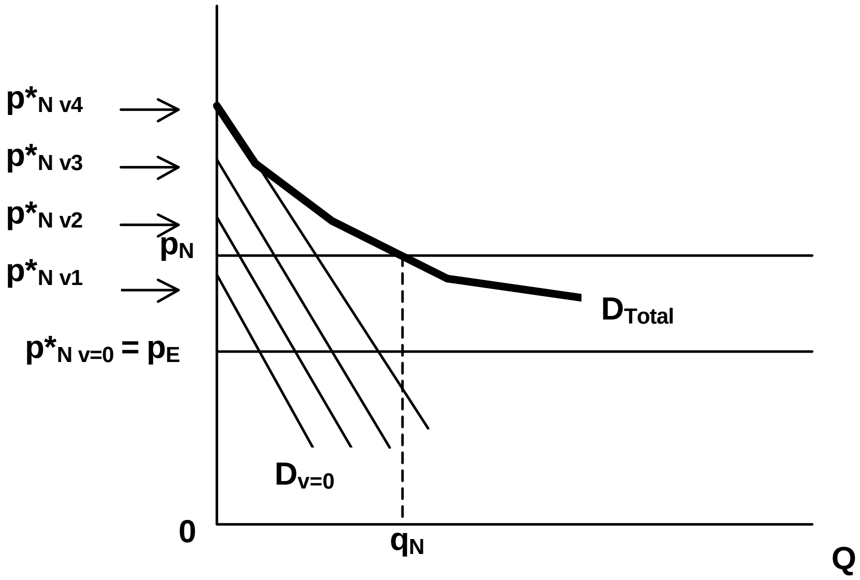
Total net benefits

For estimating the full net benefits to society from generated traffic using a new service, the producers' surplus to the operator of the new service and taxes paid to the government for non-labour inputs need to be added, and the costs of operation and externalities subtracted. This nets out to WTP for the generated demand less the social costs of operation and externalities.

Dearer, faster alternative case

Generated traffic for a dearer, faster alternative can be considered in the same manner by subdividing it according to value of time. Figure 7 shows a family of demand curves for new faster alternative for which p_N exceeds p_E. The values of the intercepts with the price axis are calculated in the same way $p^*_{Ni} = p_E + v_i(t_E - t_N) = p_E + v_i t$. Only traffic with a value of time greater than $(p_N - p_E)/t$ will materialise.

Figure 7 Demand curves for generated traffic (dearer, faster alternative)



On the assumption that demand curves for value-of-time groups are linear, the formula for total consumers' surplus is the same as derived above for the cheaper, slower case, with the important exception that the summation is over values of time from $(p_N - p_E)/t$ to infinity, that is:

$$\left(\frac{p_E - p_N}{2}\right)q_N + \left(\frac{t}{2}\right)\int_{(p_N - p_E)/t}^{\infty} vq(v)dv \quad (9) \quad \text{where } \int_{(p_N - p_E)/t}^{\infty} q(v)dv = q_N \quad (10)$$

or $(p_E - p_N + \bar{v}t)q_N/2$ (11) where $\bar{v} = \int_{(p_N - p_E)/t}^{\infty} vq(v)dv \Big/ q_N$ is the average value of time for the generated traffic.

Since the average value of time for the total population of potential generated traffic must be below that for the traffic that is actually generated, use of the greater of the population estimate and $(p_N - p_E)/t$ will lead to a conservative estimate of project benefits.

Some generalisations

Generalising to the multiple attribute case, equations (6) and (9) can be written as:

$$\left(\frac{p_E - p_N}{2}\right)q_N + \left[\begin{matrix}\text{half the total value of changes in non-price attributes had} \\ \text{the generated traffic switched from the existing alternative}\end{matrix}\right]. \quad (12)$$

The values placed on those attributes should be for the actual traffic using the new service, not population averages. For a cheaper, inferior alternative, the consumers' surplus can be approximated as $(p_E - p_N)q_N/4$ in the multi-attribute case. For the dearer, superior option, the lower bound for total consumers' surplus is zero. If population estimates of attribute values are available and they produce a positive value for consumers' surplus in equation 12, then this can be used as a conservative estimate.

Conclusion

Evaluating the benefits of a new transport service is more difficult than for an improvement to an existing service because the entire consumer surplus triangle has to be estimated. If the new service will compete with an existing service, some useful inferences can be drawn by comparing the characteristics of the two services. These inferences can be used as guides to likely values of consumers' surpluses for rough cost–benefit analyses and for checks of reasonableness where the analyst has more information.

It is hoped that the material presented in this paper exploring the technical issues underlying the estimation of consumers' surpluses of new transport services will contribute to the goal of achieving better economic evaluations of transport infrastructure projects.

Appendix: Increasing short-run costs for the existing alternative

If the new project diverts a significant amount of traffic from a congested road causing a reduction in congestion, the price of road transport will fall. The upper diagram in figure 8, represents the road market with a rising short-run marginal social cost curve and average cost curve below it. In the absence of congestion pricing, road users pay the short-run average cost.[8] It is assumed for simplicity that there are no taxes on fuel and other inputs.

The introduction of the new service causes the demand curve for road transport to shift leftward causing price to fall from p_{E1} to p_{E2} and quantity from q_{E1} to q_{E2}. This quantity transfers to the rail market represented by the lower diagram in figure 8. The welfare gain in the road market is the area traced out by the gap between marginal social cost and the price paid, that is, the sum of areas *A* and *B*.[9]

In the rail market diagram, the road prices p_{E1} and p_{E2} are both shown, along with the price charged by the new rail service, p_N. The path followed by price in the road market as traffic shifts is traced out by the $p_{E1}y$. The price and quantity in the road market follows the average cost curve, so $p_{E1}y$ is the same as the average cost curve between q_{E1} and q_{E2}, but reversed. The area $B+E$ is the same in both diagrams.

With p_{E1} the price in the road market, D_{N1} is the demand curve in the rail market. The price in the road market is the price axis intercept for the rail demand curve. As the price in the road market falls, so does the rail market demand curve, coming to rest at D_{N2} where the road price is p_{E2}. However, at the same time as the rail demand curve has been falling, the market has been moving along the demand curve as traffic switches mode. With p_N charged in the rail market, the quantity of transferring traffic is given by the intersection of p_N with D_{N2}, point *z*. The path traced out by price and quantity in the rail market $p_{E1}z$, is the general equilibrium

8 In the normal congestion pricing diagram, the vertical axis represents generalised cost and the SRAC and SRMC curves include the cost of time that each user incurs himself or herself. In order to make the prices in the upper and lower parts of the figure consistent, the vertical axes in both diagrams show money costs and the demand curves are assumed to be adjusted so they allow for the effects on quantity demanded of journey time changing with the level of congestion. The SRMC curve, however, includes the externality of the cost of longer trip times imposed by each additional user on intra-marginal users.

9 This is the welfare gain area identified by Harberger (1972, pp. 261-3). Since $MC - AC = q\frac{dAC}{dq}$, for each unit that transfers, the welfare gain to society, $MC - AC$, equals the saving to intra-marginal users — the reduction in average cost times the number of users. Hence it can be shown that areas $A + B = (p_{E1} - p_{E2})q_{E2} + E + B$.

adjustment schedule (GEAS).[10] The consumers' surplus gain in the rail market is the area between the GEAS and p_N (marked *CS*).[11]

For each unit of traffic that transfers, the value of time is the vertical distance between the road price, traced out by $p_{E1}y$, and $p_{E1}z$, the GEAS. The final tonne that transfers has a value of time of $-(p_{E2} - p_N)/t$. The total cost of the time lost, –*vtq*, is the triangular area created by the gap between p_{E1}y and the GEAS.

The total welfare gain from the new rail project is shown in table 2.

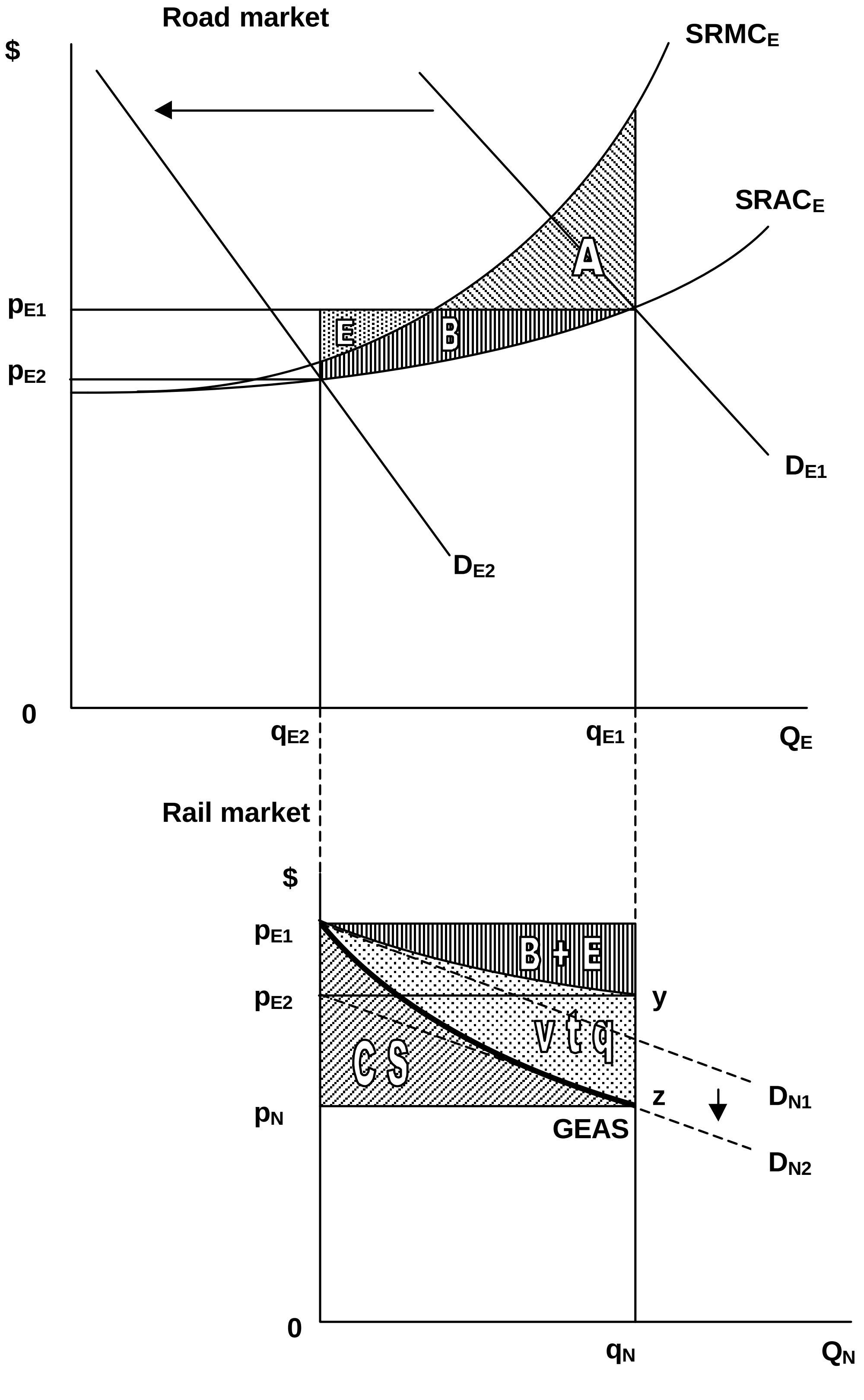
Table 2 Components of annual benefit in figure 8 summed

Benefit	Formula
Consumer surplus in the rail market	$p_{E1}q_N - B - E + \bar{v}tq_N - p_N q_N$
plus producer surplus of the rail operator and taxes transferred to the government: where c_N is average annual operating costs after deducting taxes on non-labour inputs	$(p_N - c_N)q_N$
plus benefit in the road market: where c_E is the average annual social cost of road transport excluding externality costs for the traffic that transfers, and x_E is the average annual road externality cost for the traffic that transfers	*A*+*B* $= c_E q_N + x_E q_N - p_{E1}q_N + E + B$
minus rail externality costs: where x_N is the average annual rail externality cost	$-x_N q_N$
Total net annual benefit	$(c_E + x_E - c_N - x_N + \bar{v}t)q_N$

10 The GEAS term was coined by Dr Robert Albon.

11 If this model were applied in practice, the analyst should ensure that the integrability condition is fulfilled, that is, $\frac{\partial q_1}{\partial p_2} = \frac{\partial q_2}{\partial p_1}$ for all pairs of related markets 1 and 2. This condition will be fulfilled for Hicksian compensated demand curves if the demand curves are derived from a utility function, due to the symetry of the Slutsky matrix. For Marshallian uncompensated demand curves, a further condition is necessary: either income elasticities have to be unitary, or income effects so small as to be negligible. If the integrability the condition was not met and the new service was drawing traffic away from a number of related markets, for example, different roads in a network, the value of the total welfare change would be affected by the order in which the welfare changes in markets was evaluated. This is the result of the line-integral nature of multiple-good consumers' surplus. For a basic discussion see Boardman et al. (1996, chapter 3, especially n. 42) and for an advanced discussion, see Johansson (1987).

Figure 8 Benefits with increasing costs for the existing service (cheaper, slower alternative)



The area under the SRMC curve between q_{E2} to q_{E1} in the road market is the total social cost saved by taking $q_{e1} - q_{E2} = q_N$ of freight off the road. Letting $c_E + x_E$ be the average social cost for this quantity, the total cost saved in the road market is $c_E q_N + x_E q_N$ and area $A = c_E q_N + x_E q_N - p_{E1} q_N + E$, which has been substituted in table 2.

It can be seen that this reduces to the same result as derived previously for the constant cost case.

It is important to note that in the increasing road cost case, the upper limit for the value of time is $-(p_{E2} - p_N)/t$. The relevant road price is that which prevails *after* the new alternative has been introduced, not before.

References

Boardman, A.E., Greenberg, D.H., Vining, A.R. and Weimer, D.L. 1996, *Cost–benefit analysis: concepts and practice*, Prentice Hall, Upper Saddle River, NJ.

Bureau of Transport Economics (BTE) 1999, *Competitive neutrality between road and rail*, Working Paper 40, BTE, Canberra.

Bureau of Transport Economics (BTE) 2000, *Brisbane–Melbourne rail link: economic analysis*, Working Paper 45, BTE, Canberra.

Harberger, A.C. 1972, *Project evaluation: collected papers*, Macmillan, London.

Johansson, P. 1987, *The economic theory and measurement of environmental benefits*, Cambridge University Press, Cambridge.