

TRANSPORT NOTES

TRANSPORT ECONOMICS, POLICY AND POVERTY THEMATIC GROUP



THE WORLD BANK, WASHINGTON, DC

Transport Note No. TRN-11

January 2005

Notes on the Economic Evaluation of Transport Projects

In response to many requests for help in the application of both conventional cost benefit analysis in transport and addressing of the newer topics of interest, we have prepared a series of Economic Evaluation Notes that provide guidance on some of issues that have proven more difficult to deal with.

*The **Economic Evaluation Notes** are arranged in three groups. The first group (TRN-6 to TRN-10) provides **criteria** for selection a particular evaluation technique or approach; the second (TRN-11 to TRN-17) addresses the selection of values of various **inputs** to the evaluation, and the third (TRN-18 to TRN-26) deals with specific **problematic issues** in economic evaluation. The Notes are preceded by a **Framework** (TRN-5), that provides the context within which we use economic evaluation in the transport sector.*

The main text of most of the Notes was prepared for the Transport and Urban Development Department (TUDTR) of the World Bank by Peter Mackie, John Nellthorp and James Laird, at the Institute for Transport Studies (ITS), University of Leeds, UK (The draft text of Note 21 was prepared for ITS by I.T. Transport Ltd). TUDTR staff have made a few changes to the draft Notes as prepared by ITS.

The Notes will be revised periodically and we welcome comments on what changes become necessary. Suggestions for additional Notes or for changes or additions to existing Notes should be sent to rcarruthers@worldbank.org

TREATMENT OF INDUCED TRAFFIC

Induced traffic can be an important part of the economic appraisal particularly when the objective of the investment is to stimulate economic development. Its importance, however, is not restricted to such situations. The omission of induced traffic from the economic appraisal, or its incorrect treatment, may lead to either over or underestimations in the user benefits (consumer surplus) of an investment.

In this note we address this issue by considering: the importance of induced traffic for the economic appraisal (Section [1](#)); what constitutes induced traffic (Section [2](#)); the situations in which induced traffic is likely to be relevant (Section [3](#)) and the manner in which it can be modelled (Section [4](#)) and user benefits calculated when it is present (Section [5](#))

Given the importance of including induced traffic in the evaluation of transport investments, and the many uncertainties related to how it should be evaluated, this Note is longer than many of the others in this series and goes into considerable detail of when the standard method of dealing with induced traffic might break down.

This Note includes three Annexes. The first shows the relative importance of including the benefits of induced traffic in the evaluation of an urban transport project. The second shows where the standard "rule of one half" breaks down in some situations that are often present in World Bank projects, while the third shows a numeric integration technique that can be used as a valid alternative to the rule of one half in many of these situations (and coincidentally, provides a more precise evaluation even where the "rule of one half" gives an acceptable estimation).

THE IMPORTANCE OF INDUCED TRAFFIC

Historically, projects have been justified on the basis of benefits to existing traffic only. The impacts of induced traffic were not incorporated into the economic analysis unless the project could not be justified on the basis of existing traffic alone, the basis for such an approach is that there is uncertainty associated with the level of traffic that will be generated by a project. However, apart from the potential of underestimating the benefits of the project such an approach may also have the following important limitations and consequences:

- ❑ User benefits may in fact be overestimated (not underestimated) if underlying traffic conditions are congested;

- ❑ A sub-optimal design standard for the project maybe promoted, i.e. the project could be under designed.
- ❑ A potentially good project or scheme maybe completely rejected;
- ❑ A potentially bad project could be promoted;
- ❑ The relationship between the infrastructure and the wider economy could be neglected. In the absence of a detailed economic model the amount of induced traffic can be an interesting indicator and a proxy measure of the wider economic impacts (providing demand in the system is being correctly modelled);
- ❑ A significant change in fiscal benefits that may accrue to certain bodies maybe omitted, for example Government tax revenue in the case of “free” roads and additional financial income in the case of commercial facilities such as railways, toll roads and bridges; and
- ❑ An underestimation of environmental impacts could occur

The consequences of omitting induced traffic from the appraisal are therefore significant. The remainder of this Note discusses the situations when induced traffic is relevant to the appraisal and the manner that user benefits should be calculated in the presence of such traffic.

WHAT IS INDUCED TRAFFIC?

When a new transport facility or service becomes available the users of the transport system can alter their behaviour in a number of manners:

- ❑ Change their route
- ❑ Change mode
- ❑ Change destination to one easily reachable using the new system
- ❑ Change their trip making frequency
- ❑ Change the time of travel

Additionally, the transport project may result in an altering of land use patterns. For example, a new road and river crossing may facilitate economic development that would not have otherwise occurred, by say improving accessibility to markets. The impact of changing land use patterns is discussed further in Note 19: *Projects with Significant Restructuring Effects*.

Transport users can be categorised a number of ways (see also Note 12: *Demand Forecasting Errors*).

The basis for the classification in [Table 1](#) is the previous behaviour of the traveller or traffic and the manner in which they alter their behaviour as a consequence of the project. As can be seen from this table Induced Traffic is therefore defined as the additional traffic (in person or vehicle kilometres) that has been induced by the project through mode changes, destination changes, trip re-timing, trip frequency changes or new trips associated with different land uses.

Table 1. Traffic Classification by Behavioural Response

Behavioural Response	Classification from the perspective of	
	Demand associated with the project	Demand within the entire multi-modal transport system
No change in behaviour	Base Load or Traffic	Normal Load or Traffic
Route change (same origin and destination after route change)	Re-assigned or Diverted Traffic	
Mode change	Induced Traffic	
Destination change		
Time of travel change		
Trip frequency increase		Generated Traffic
Generated or new (e.g. from different land use patterns)		

Note: the term *traffic* is used to represent all forms of transport traffic including pedestrian traffic, road traffic, railway traffic and shipping traffic

WHEN IS INDUCED TRAFFIC LIKELY TO BE RELEVANT.

There are two principal situations when induced traffic is likely to be relevant for an economic appraisal (see also [Box 1](#)):

- ❑ Firstly, when the benefits that will accrue to the Induced Traffic are significant compared to the benefits that will accrue to Base and Re-assigned Traffic; and
- ❑ Secondly, when the Induced Traffic imposes significant costs on the Base and Re-assigned Traffic (e.g. a congestion or overcrowding cost).

Such situations are likely to occur within the following scenarios:

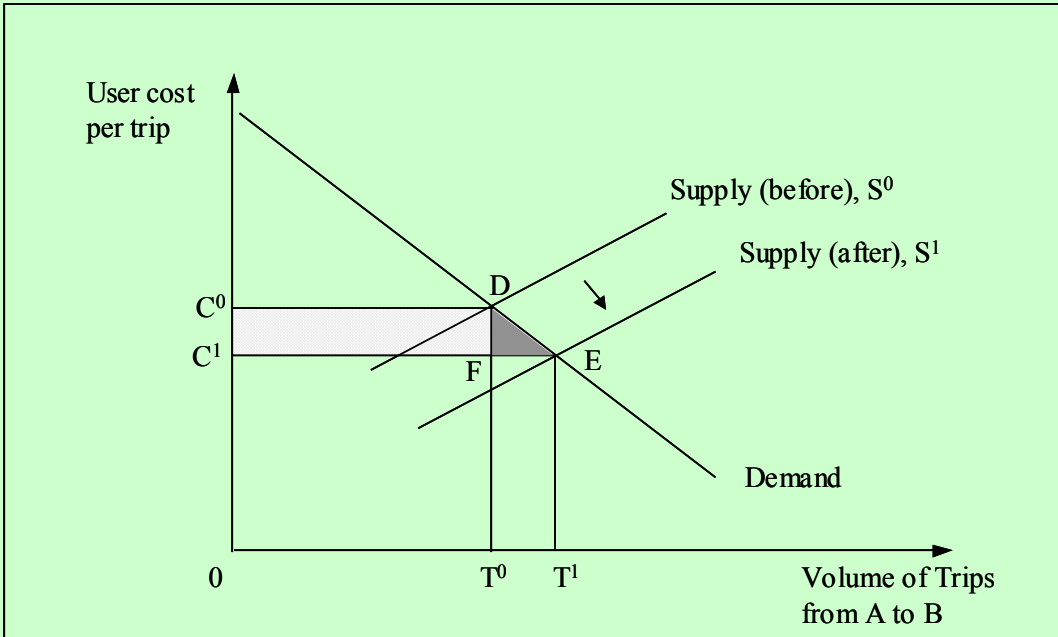
- ❑ Significant time savings for individual origin to destination movements occur (e.g. a new river or estuarial crossings where none previously existed)
- ❑ Significant cost (financial) savings for individual origin to destination movements occur (e.g. situations where transport costs form a large proportion of the total delivered price of the product shipped or the total cost of the trip purpose activity);
- ❑ High elasticity of demand is present (for example within a highly congested urban environment);
- ❑ Heavily congested conditions (steeply upward-sloping supply schedules);
- ❑ Changes in land use patterns will occur (i.e. structural economic shifts);
- ❑ There is little or no existing (Base) traffic (e.g. transport projects that break new ground).

The inclusion of Induced traffic within a project appraisal is also important when one of the objectives of the project is to stimulate economic development, particularly trade (trucks, freight trains and ships).

Box 1. Measurement of Induced Traffic Benefits

Induced traffic contributes to the consumer surplus of a transport investment in the manner illustrated in Figure 1. In this figure the aggregate benefit to travellers between A and B arising as a result of a fall in the cost of a trip (from C^0 to C^1) is C^0DEC^1 . The contribution of induced traffic (the difference between T^1 and T^0) to the user benefit is the area DEF, whilst the contribution of the Base and Re-assigned traffic is the area C^0DFC^1 .

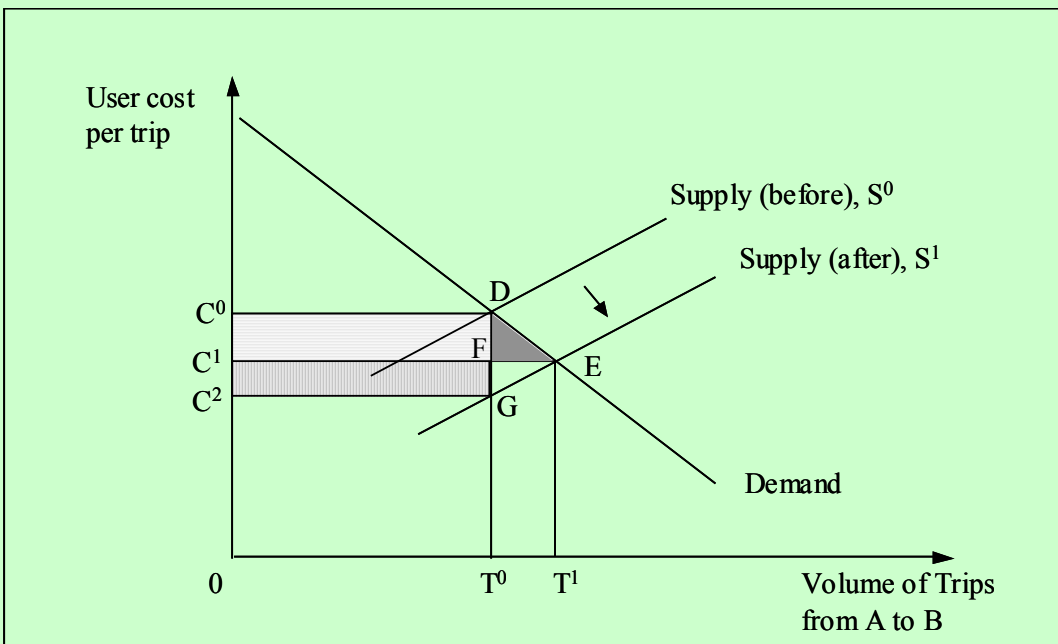
Figure 1. User Benefit Measure



If, however, induced traffic had been excluded from the appraisal, that is a Fixed Trip Matrix (FTM) assumption had been made, the benefit derived by that traffic, area DEF, would be excluded from the analysis. The inclusion of induced traffic is therefore important in situations where its benefit is large.

An additional error can also occur under the FTM assumption in the presence of congestion or overcrowding, as the benefit to the Base and Re-Assigned traffic can be overestimated. This is illustrated in Figure 2. Under the FTM assumption the benefit to Base and Re-Assigned traffic would be estimated as C^0DGC^2 , whilst the correct measure is $CODEC^1$.

Figure 2. Potential Error In User Benefit Measure Associated with Omitting Induced Traffic



In this example, the exclusion of induced traffic from the analysis would lead Total User Benefits being overestimated, as the error associated with the estimate of benefit to Base and Re-Assigned traffic (C^1FGC^2) is greater than the error associated with the Induced traffic (DEF). The relative size of these errors is, however, dependent upon the characteristics of the transport market (depicted by the shapes of the demand and supply curves in Figure 1 and the net effect can go either way).

MODELLING INDUCED TRAFFIC

The complexity of the demand forecasting process and the scale of the data requirements vary when forecasting the different categories of traffic (see [Table 1](#)). Base traffic is the least onerous form of traffic to model, as simple traffic count and origin-destination data in the existing situation provides all the information required. Modelling re-assigned traffic is now also a standard process for which off the shelf software exists (e.g. HDM4 for road based traffic). If only Base Traffic and Re-assigned Traffic is considered within the appraisal the implicit assumption is that all origins, destinations, time of travel choices and mode of travel choices remain fixed. Such an approach is known as the **Fixed Trip Matrix (FTM)** approach, as the matrix of origin-destination demands is the same in the Do Minimum as in the Do Something.

The modelling of Induced Traffic is the most complex of all, as strictly speaking mode choices, destination choices, time of travel choices, changes in trip frequency and importantly land use changes all need to be forecast. A consequence of Induced Traffic is that origin-destination demands will vary between the Do Minimum and Do Something. Methods used to model Induced Traffic are therefore termed **Variable Trip Matrix (VTM)** approaches. Without doubt VTM (or induced traffic) modelling can be complex, but as discussed in the previous sections can also be essential. In practice, however, when modelling induced traffic a range of modelling approaches are available that include both simple methods and more complex methods. These are set out below:

- The simple elasticity approach, where all responses are subsumed into a single elasticity. Different elasticities can be used for different journey purposes (e.g. freight related travel, travel in the course of business and non-working time travel). Such an approach only requires data from the mode affected by the project (e.g. rail travel data for a rail project or road travel data for a

road project) and is most applicable to a situation where an existing network is being enhanced through improved quality of service (e.g. increased frequency of train services, or upgrading of a road from single to dual carriageway);

- ❑ The inclusion of a single behavioural response (such as mode choice). Typically, such an approach is adopted when forecasting the demand for improved public transport services, such as guided bus or Light Rapid Transit (LRT), where some of the demand will be extracted from competing modes (road and other public transport services). Such an approach requires travel data on all modes of travel being considered; and
- ❑ A staged model (also known as a four stage model) that incorporates all behavioural responses. This is the most complex form of transport model and can be expensive in terms of both data and resources in development (calibration and validation) and operation (long model run times). However, models such as these may be needed for large projects in congested urban areas (e.g. new metro systems or new urban motorways).

Recent computing advances have also allowed the operationalisation of Land Use and Transport Interaction (LUTI) models and Computable General Equilibrium (CGE) models which not only model the impact on travel but also the impacts on land use, the labour market and property rents. Such models, however, represent the state of the art and their use is currently relatively limited. They are discussed further in Note 19: *Projects with Significant Restructuring Effects*. Further information regarding the calculation of induced traffic demand forecasts and errors associated with such calculations is also contained in Note 12: *Demand Forecasting Errors*.

A final key issue associated with modelling induced traffic and capturing the benefits of it is that the model area, for both the modelling exercise and the cost benefit analysis, need to be sufficiently large to ensure that all benefits or costs are included within the appraisal. This may seem an obvious statement to make, but in situations where projects have international consequences, such as transit traffic in a land locked country, such issues may be overlooked. The Note 12: *Demand Forecasting Errors* also contains a small discussion on the definition of the model area.

In constructing defensible models, it is always advisable to validate the model against the existing situation (see Note 12: *Demand Forecasting Errors*). Sometimes it is also useful to seek out evidence from comparable situations elsewhere, though caution should be exercised when considering the transferability of travel behaviour particularly between countries. Peer review of the demand forecasts by appropriately experienced individuals will add further to their credibility.

HOW TO CALCULATE USER BENEFIT IN THE PRESENCE OF INDUCED TRAFFIC

In the majority of situations the calculation of the user benefit associated with induced traffic is relatively straight forward and utilises the Rule of the Half (RoH) methodology. This method is presented in the *Framework* but for convenience the formula is also reproduced in [Box 2](#). It should be noted that this method (and all other methods) require reliable demand forecasts of the volume of induced traffic. Such demand forecasts can be quite complex to obtain.

Box 2. Rule of a Half

User Benefit_{ij} = ½(C_{ij}⁰-C_{ij}¹)(T_{ij}⁰+T_{ij}¹)

Where: C_{ij}⁰ is Cost between origin (i) and destination (j) before investment
 C_{ij}¹ is Cost between origin (i) and destination (j) after investment
 T_{ij}⁰ is Demand between origin (i) and destination (j) before investment
 T_{ij}¹ is Demand between origin (i) and destination (j) after investment

Operationalising the Rule of Half (RoH) is a relatively straight forward procedure, though the following properties of the RoH should be borne in mind:

- ❑ User benefits should be calculated on a matrix basis and not a link basis. That is user benefits must be calculated on an origin-destination (i-j) pair basis, implying the cost used in the

calculation is the travel time, vehicle operating costs and money costs required to travel between origin (i) and destination (j) by mode (m). This contrasts to the Fixed Trip Matrix (FTM) situation, i.e. no induced traffic, where user benefits can be calculated for each link in the network and summed, instead of on an i-j pair basis;

- ❑ User benefits can be calculated separately for each mode and time period, even in situations where demand switches between modes and time periods. This is a particularly useful property of the RoH as often the demand forecasting process will use different models to represent different time periods and modes;
- ❑ User benefits associated with the individual components of generalised cost (e.g. time, vehicle operating costs and money costs) can be calculated and summed to give total user benefits:

$$\text{RoH}_{\text{time}} = \frac{1}{2}((H_0 - H_1) \times \text{VoT})(T_0 + T_1)$$

$$\text{RoH}_{\text{VOCs}} = \frac{1}{2}(\text{VOC}_0 - \text{VOC}_1)(T_0 + T_1)$$

$$\text{RoH}_{\text{user charges}} = \frac{1}{2}(U_0 - U_1)(T_0 + T_1)$$

where: H is the travel time per trip in hours,

VoT is the value of travel time in currency units per hour.

VOC is the vehicle operating costs for motorised transport in currency units per trip

U is the user charge in currency units per trip

Subscripts for origin (i), destination (j), mode (m) and for different trip purposes (which would carry different values of time and operating cost) have been omitted for simplicity.

- ❑ User benefits/disbenefits associated with money costs (e.g. road tolls and fares), when calculated under the RoH and variable demand, do not net out with changes in the fare revenue element of the producer surplus calculation (i.e. they are not transfer payments); and
- ❑ Technically, there is no unique attribution of user benefits between modes or indeed between origin-destination pairs, because it is not possible to identify an individual on the do-something network and trace back to find out what mode he/she used in the do-minimum.

Off the shelf software exists to calculate network wide user benefits using a matrix based Rule of a Half approach (for example the United Kingdom program, Transport User Benefit Appraisal (TUBA) [1]).

EXCEPTIONS TO THE RULE OF A HALF

The Rule of a Half breaks down when the assumptions upon which the methodology is based are undermined. The two principal assumptions are:

- ❑ The demand curve is linear; and
- ❑ Demand for travel exists in the before and after situation (by mode and time period).

The circumstances in which these assumptions may break down are set out below. In each situation advice is given regarding the method that should be adopted for the calculation of user benefits.

Large changes in the generalised cost: The bigger the proportionate reduction in generalised cost brought about by a transport infrastructure project, the less reliable the linear approximation to the demand curve becomes. The recommendation here is that, as a rule of thumb, if the project results in a >25% reduction in average generalised cost from origin to destination for trips using the project, this feature should be reported as part of the cost benefit analysis. In such situations user benefits

should be estimated using the method of numerical integration detailed in Nellthorp and Hyman (2001) [2] and attached as [Annex 1](#) to this Note. This approach requires repeated model runs to sketch in the unknown part of the demand curve and is a pragmatic way of estimating the shape of the demand curve.

New modes or existing modes become redundant: Introduction of completely new modes in the do-something scenario - for example, high speed rail, urban light rapid transit, or even a new conventional railway - where none exists in the do-minimum. Alternatively, the removal of a mode (e.g. closure of a railway branch line) in the do-something where it exists in the do-minimum. In such situations either the "before" or the "after" cost does not exist, so the Rule of the Half breaks down. Instead of using the Rule of a Half, user benefits should be calculated through the method of numerical integration as set out in Nellthorp and Hyman (2001) [2] and attached as [Annex 1](#) to this Note. Again numerical integration is a simple pragmatic solution rather than the theoretically best approach - which requires the ability to integrate the demand function - which is often more complicated to do.

There is no existing demand: such projects would include projects that break new ground or open up areas for development such as a new freight railway or low volume rural roads and feeder roads. As with a new mode, in such situations the "before" cost does not exist, so the Rule of the Half breaks down. If the Rule of a Half is applied in such circumstances it will invariably result in an *overestimation error*. However, in some situations (notably producers who are attracted to enter and serve a "world" market for their product which involves a given fixed price) the rule can lead to *underestimation* (see Gannon (1998) [3]). Numerical integration once again offers an option for the calculation of user benefits (see [Annex 1](#)), however, in such situations it maybe more appropriate to estimate some of the wider economic impacts associated with the economic (re-)generation associated with transport projects. This issue is also discussed in Note 21: *Low Volume Rural Roads*.

Significant land use changes and/or structural economic shifts: this special circumstance and its treatment are discussed in detail in Note 21: *Low Volume Rural Roads* and Note 19: *Projects with Significant Expected Restructuring Effects*.

FURTHER READING

[1] Department for Transport, *Transport Users Benefit Appraisal (TUBA)*, Integrated Transport Economics and Appraisal, Department for Transport, UK, ¹Available online at http://www.dft.gov.uk/stellent/groups/dft_control/documents/contentservertemplate/dft_index.hcst?n=7910&l=4

[2] Nellthorp J and Hyman G (2001), 'Alternatives to the RoH in Matrix based Appraisal', *Proceedings of European Transport Conference 10-12 September 2001*. London: Association of European Transport.

[3] Gannon, C. (March 1998), *Evaluation of Generated Traffic Benefits: A Note on Qualifications for the "Rule of One-Half"*. TWU Department, World Bank, Washington DC.

[4] Mackie, P.J.(1996), Induced Traffic and Economic Appraisal. *Transportation* 23: 103-119

[5] SACTRA (The Standing Advisory Committee on Trunk Road Assessment) (1994) *Trunk Roads and the Generation of Traffic*. London: The Stationery Office.

[6] Department for Transport, UK (2000), Guidance on the Methodology for Multi-Modal Studies (GOMMS), Volume 2, Appendix F (Transport Cost Benefit Analysis Background Material). [Available online at http://www.dft.gov.uk/stellent/groups/dft_control/documents/contentservertemplate/dft_index.hcst?n=7955&l=1]¹

ANNEX 1

THE TREATMENT OF INDUCED TRAFFIC IN THE SHIJIAZHUANG URBAN TRANSPORT PROJECT

The Shijiazhuang Urban Transport Project is a collection of measures designed to speed up the flow of buses, bicycles, cars, taxis, and trucks; to reduce operating costs of buses, cars taxis, and trucks; to reduce accidents; and to improve air quality. The measures included upgraded arterial roads both entering the city and within the city; new multi-level interchanges, intersection channelization and bus priority schemes. The economic evaluation included specific treatment of the possible impacts of induced traffic.

The economic benefit calculations are based on three simplifying assumptions.

- First, it is assumed that the generated traffic response will occur in the midday hours, and not in the peak hours of travel. In the peak hour, the great majority of trips are work trips and school trips, and it is unlikely that more of these types of trips will be made in response to a more freely flowing traffic stream. In contrast, during the midday hours, there are more shopping and personal trips being made, including trips to and from work at lunch time. These types of trips are more likely to be increased if travel costs decline.
- Second, it is assumed that marginal increases in traffic during the midday will not result in significant changes in travel times or costs for midday tripmakers. This assumption is based on the observation that there are slightly more trips in the 4.5 peak hours than in the 7.25 midday hours, so that traffic is only 55 percent as heavy during the midday hours. Because traffic is less dense during the midday hours, it is reasonable to assume that marginal increases in trips during those hours will not result in an increased average trip time for all midday tripmakers.
- Third, it is assumed that the midday travel demand elasticity is $-.70$. In combination with the assumed peak hour demand elasticity of $.00$, this results in an aggregate 24-hour travel demand elasticity of $-.33$.

These three assumptions imply that generated trips will have a positive effect on overall project economics. Those who make more trips will be better off, and they will not impose incremental travel costs on other tripmakers during the midday.

ALTERNATIVE MORE PESSIMISTIC CALCULATIONS

A more pessimistic view of generated traffic would anticipate that the incremental generated trips would slow down the traffic flow and impose costs on other tripmakers. In this scenario, the net effect of generated traffic reflects both gains to the generated tripmakers and losses to the other tripmakers.

To illustrate how including costs of generated traffic might impact the economic analysis, Table 1 illustrates the net economic impacts of generated traffic under a range of assumptions of how the existing traffic is effected. Specifically, the second column of Table 1 shows how the economic analysis would change if benefits to the existing traffic stream were eroded by either 2, 4, 6, 8, or 10 percent as a result of the generated traffic that has been projected. The third column of Table 1 then considers the impacts of losing half of the generated traffic as a result of the overall traffic slowdown.

In the extreme situation where generated traffic reduces benefits to the existing traffic by 10 percent, and where generated traffic levels are half of those originally projected (because speed and cost improvements for travelers are less than originally expected), the NPV of the project is still strongly positive, at about 85 percent of the base case level of 3.3 billion Yuan.

Table 1. Alternative Estimates of Project NPV (billions of Yuan) and IRR (%), Driven by Variation in the Level of Generated Traffic, and by the Amount of Reduction in Benefits for the Existing Traffic Stream, Show that Project Economics Remain Strongly Positive in All Cases

Reduction in Cost and Time Savings for non-Generated Traffic	Level of Generated Traffic	
	100% of Generated Traffic Projected in Economic Analysis Report	50% of Generated Traffic Projected in Economic Analysis Report
0 Percent Reduction	3.35 29.7%	3.32 29.5%
2 Percent Reduction	3.26 29.2%	3.23 29.1%
4 Percent Reduction	3.16 28.8%	3.13 28.7%
6 Percent Reduction	3.07 28.3%	3.04 28.2%
8 Percent Reduction	2.97 27.9%	2.94 27.8%
10 Percent Reduction	2.87 27.5%	2.85 27.3%

ANNEX 2

EVALUATION OF GENERATED TRAFFIC BENEFITS

A Note on Qualifications for the "Rule of One-Half"^{1,2}

This Annex is based on a Note written by Colin Gannon in 1998, then an Economic Advisor at the World Bank, in response to questions from regional transport staff regarding the evaluation of induced freight demand)

Background

The "rule of one-half" might well be regarded as a cornerstone in the methodology for estimation of cost-saving benefits of interventions in the transport sector. If an intervention, such as a road improvement project, deregulation policy, or increased private sector participation in the supply of infrastructure services yields lower costs of transport then, plainly, all existing traffic benefits from the full amount of the reduction in transport costs: benefits are transport resource cost savings. But, for generated traffic there are no transport cost savings since no transport costs are incurred pre-improvement. Benefits to generated traffic involve increases in producers' surpluses (excess profits to operators and or shippers) or consumers' surpluses (to private users). Moreover, the level of these benefits will vary from very slightly less than the full amount for new users who were on the "verge" of using the transport prior to its improvement, to very close to zero for those new users who are on the "verge" of not using the transport after its improvement. If the distribution of demand (strictly willingness to pay) by new users is uniform over this range, then, on average, the generated traffic benefit is one-half of the full amount of the reduction. This is the well known "Rule of One-Half".³

Expressed in formal quantitative terms, a reduction in total transport costs from T_1 to T_2 , provides "savings" of $(T_1 - T_2)$ and increases quantity flows from Q_1 to Q_2 , i.e., involves generated traffic of $(Q_2 - Q_1)$.

The conventional method for estimating the associated user benefits is:

For existing traffic (Q_1), benefits (B_E) are given by $B_E = (T_1 - T_2) * Q_1$

For generated traffic ($Q_2 - Q_1$), benefits (B_G) are given by $B_G = 1/2 (T_1 - T_2) * (Q_2 - Q_1)$

This note is concerned with the conditions under which the "rule of one-half" measure for generated benefits, B_G , is a strong and a weak estimator of generated traffic benefits.

A robust estimated specification of the demand schedule for transport services between an origin-destination pair is invariably not available; adequate data do not exist or are costly and time consuming to secure. Typically, existing traffic levels are observed and associated user transport costs calculated. The new, lower, transport costs as a result of an intervention are estimated (e.g., for road sealing via the HDM model) and the magnitude of expected generated traffic estimated -- explicitly or implicitly by application of a "reasonable" price elasticity, taking into account local circumstances (substitutes, share of transport in total costs, etc.)

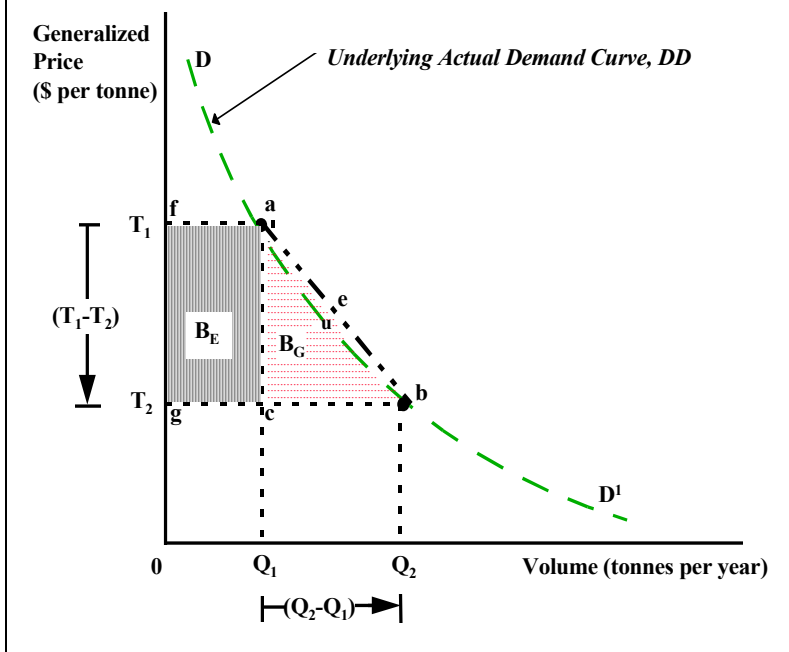
¹ Colin Gannon, TWU Department (February, 1998).

² The origin of this note was a review of the freight demand modeling undertaken as part of the feasibility study for a proposed 40 Km. bridge across the Rio de la Plata between Buenos Aires and Colonia and, in particular, the generation of freight traffic between Argentina and Brazil. The need to take a systematic and analytical look at the rule of one-half was prompted by a concern that this rule applied to total logistics costs could significantly underestimate generated traffic benefits. This note looks into (the transport) part of that concern.

³ This "rule of one-half" applies, of course, to a price (or consumer cost) fall in any market. The "rule" derives from the adoption of a *linear* approximation to the demand schedule for a product or service over the range of the price fall. Note that in the transport context, users of transport services (e.g., a road link) incur "user costs" that, in general, comprise vehicle operating costs, time and other "quality of service" costs (e.g., safety and reliability) and in some cases direct charges (for example, fares or tariffs for "hired services" or additional separate tolls). These total user generalized costs are the "price" paid by users for the transport service.

In order to estimate benefits associated with *generated* traffic, an approximation to the unknown transport demand schedule (or its price elasticity) is required. The conventional approach is to assume a linear relationship ("straight line demand curve"). This leads to the "rule of one-half" as the estimator for generated benefits. This is illustrated in Figure 1. The underlying actual demand curve between a and b is DD^1 ; the linear approximation (over the range of transport cost change) is the broken line aeb. The shaded area aebc represents the estimate of generated benefits (B_G). The actual benefits are represented by the area aubc -- these are "slightly" over (under) estimated, depending on whether the "actual" unknown demand curve is convex or (concave).

Figure 3. Diagrammatic Representation of "Rule of One-Half"

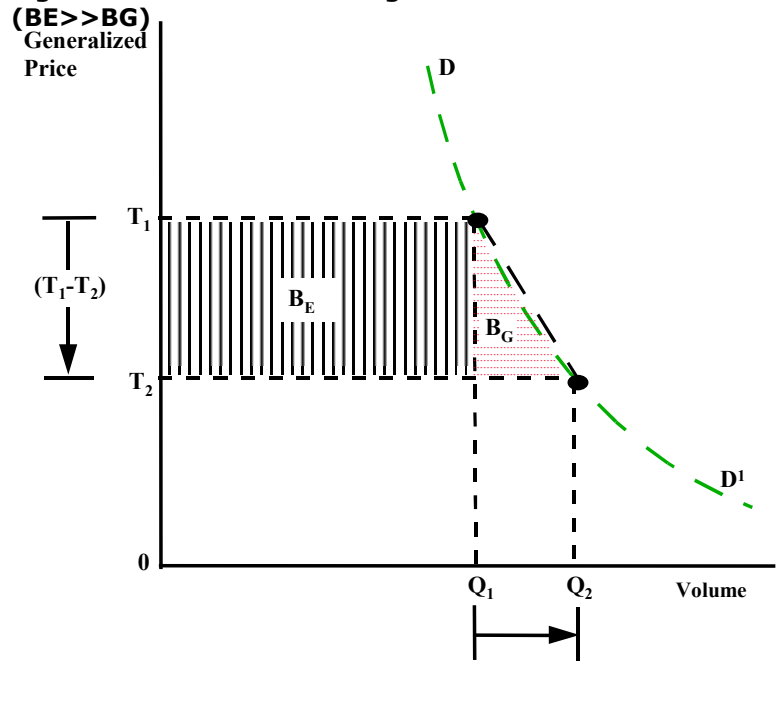


The benefits (B_E) to existing traffic (Q_1) are the resource savings represented by the rectangular area fagc.

Concern over the "estimation error" for B_G is normally low since the lion's share of benefits is usually expected to accrue to existing traffic (B_E). This is the case shown in Figure 2. The rectangular area efca "dominates" the triangular area aebc, ($B_E >> B_G$). However, this is not always the case. In Figure 3, benefits to generated traffic are larger than the benefits to existing traffic.

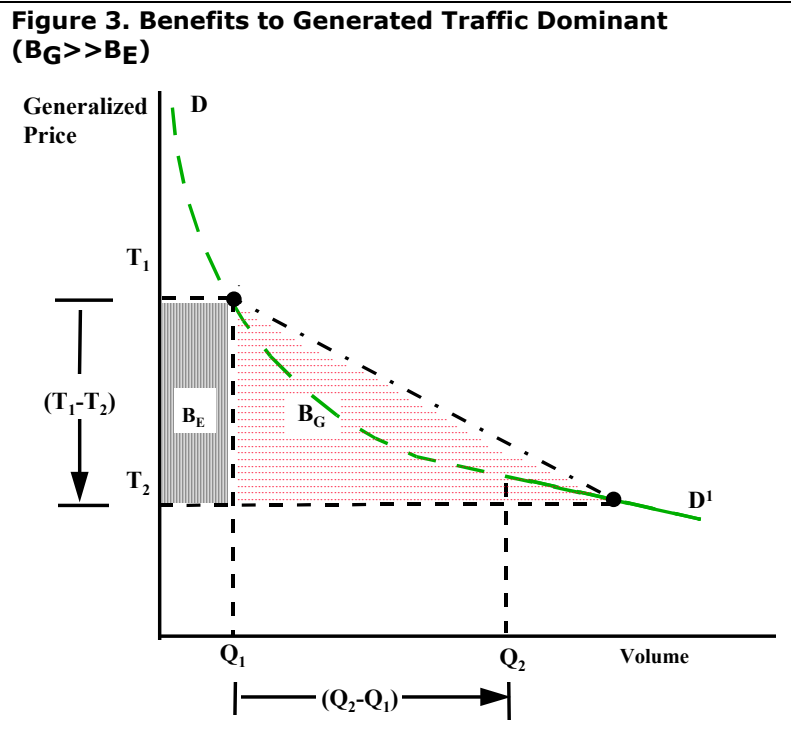
While the relative shares of benefits to existing traffic and benefits to generated traffic are an empirical matter, there are several classes of circumstances under which generated benefits are likely to be substantial, and possibly dominant. For example, in large "well developed" markets, in contexts where

Figure 2. Benefits to Existing Traffic Dominant



the transport service/route has "close" substitutes (and hence its demand relatively price elastic), and in cases where transport costs represent a large proportion of the total delivered (c.i.f.) price of the product shipped or total cost of the trip purpose activity.

Another circumstance, which is the focus of this note, is where there is **no existing traffic**, i.e. *all* traffic associated with a fall in transport costs is predicted generated traffic.⁴ In a freight context this can arise when the delivered price of a good (say, steel) shipped from region A to region B is initially above the local market price in B due to high transport costs--and no shipments of the good are made: existing traffic is zero. However, a transport improvement may lower transport costs sufficiently to attract exports from producers in region A to region B. The question is: what are the benefits of the generated freight traffic? In a somewhat different retailing context, the demand by a consumer(s) for a particular good (tinned meat) will typically involve a *maximum price*, above which the consumer(s) are not willing to buy any quantity of the good--given the local price of "close" substitutes (fish).



If high transport costs result in the delivered price in region B exceeding this maximum consumer purchase price, then there will be no freight movements of this good from A to B. But if a transport improvement (and a competitive market structure in transport services) allows the delivered price to consumers to be pushed below this maximum price, shipments will be made and traffic will be generated. What are the associated benefits for this type of generated traffic? In particular, is the "rule of one-half" a reliable estimator?

SITUATIONS IN WHICH ALL TRAFFIC IS GENERATED

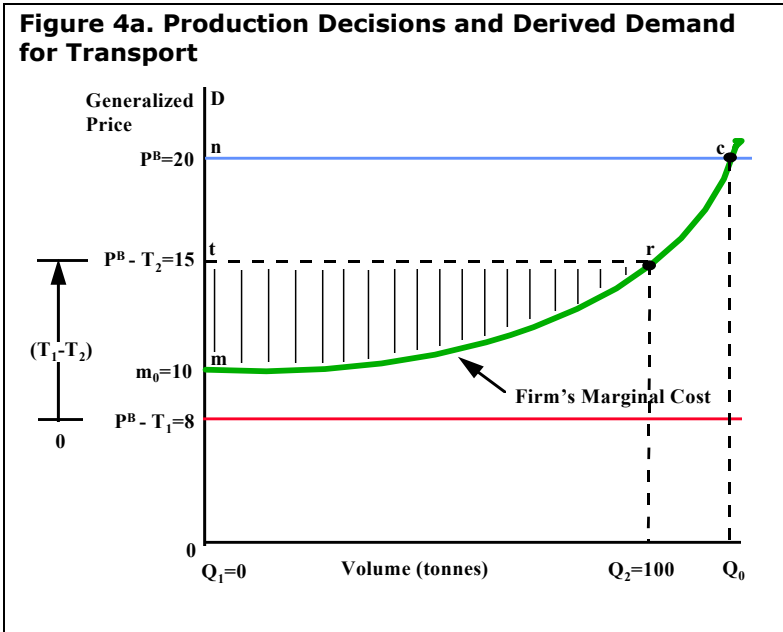
Freight Shipments by Individual Producer

There are two main basic microeconomic models which provide the building blocks for the derived demand for freight transport: one at the level of an individual producer or manufacturing plant, the other at the aggregate market level of interregional trade.⁵ These are examined in turn with respect to the benefits for generated freight traffic.

⁴ In the context of this discussion, the nature or source of the "generated" traffic is important (see Section II) but not exclusive; "generated" traffic herein encompasses "diverted" (from another route, time period, or mode) and "induced" (from a change in production, land use/location, or consumption). Moreover, changes (typically growth) in traffic over time are set aside; the key issue examined here is generated traffic benefits in any one time period.

⁵ These basic models apply to a given spatial pattern of economic activity; i.e., locations of firms/establishments are fixed. Models of industrial location (based, for example, on Weber, Lösch, and Hoover) provide insights into the longer term spatial structure of freight flows, based on firms optimizing overall costs, including transport.

The output and shipment decisions of an individual firm are driven by the objective of increasing profits, which depend on sales revenues less production *and* transport (delivery/logistics) costs. If transport costs to ship to a particular market (say, in region B) are "too high" (relative to the f.o.b. market price of the product), it is unprofitable to serve that market -- and no shipments are made by a firm in A to the market in region B. Alternatively, if transport costs from A to B are "very low", provided the firm's unit production costs are less than the prevailing market price in B (i.e., the firm is "regionally competitive"), profits can be earned and freight movements will take place from A to B. This situation is illustrated in Figure 4a.



The given market price in B is P^B (say, \$20/tonne); demand in B is "perfectly elastic" at this regional (or "world") price, i.e., the firm can sell as much or as little as it wishes at that price, in B. Initially, the transport cost to ship from A to B is T_1 (say, \$12/tonne). Thus, the net price to the firm, after transport costs is $P^B - T_1$ (i.e., $20 - 12 = \$8$ /tonne). The minimum marginal production cost of the firm in A is m_0 (say \$10/tonne) and the firm's marginal production cost schedule is shown by the curve m_0rc . At this level of transport costs, the firm makes no shipments to B; *existing* freight traffic (Q_1) is zero. Now, suppose a transport improvement can be made which would reduce transport costs from T_1 (\$12/tonne) to T_2 (say, \$5/tonne), i.e., a transport cost "reduction" of $T_2 - T_1 = 12 - 5 = \$7$ /tonne. The net price to the firm increases to $P^B - T_2 = 20 - 5 = \$15$ /tonne, and the firm sets its output at Q_2 (say, 100 tonnes) where its marginal production cost has risen to the net price (marginal revenue) of $P^B - T_2$ or \$15/tonne. *Generated* traffic is Q_2 (100 tonnes). The profits to the firm from serving market B are the sum of the profits on each unit shipped; this sum is represented by the vertically shaded area tm_0r in Figure 4a. These profits (producer's surplus) are the benefits associated with the generated traffic.

Application of the "rule of one-half" in this situation, i.e., $1/2 (T_1 - T_2) * (Q_2 - 0)$, or $1/2 (12 - 5) * 100 = \$350$, in general, would be an unreliable poor estimator of the actual net benefits. Moreover, the "rule of one-half" may yield an over-estimate (shown in Figure 4b) or an under-estimate (as shown in Figure 4c). It is not possible to draw any strong conclusions regarding the magnitude of the estimation error. However, the smaller the difference between the post-improvement net price ($P^B - T_2$) and the minimum marginal production cost (m_0), the greater the likelihood of overestimation by the "rule of one-half", as shown in Figure 4b. Contrawise, the larger this difference -- and hence the smaller the initial gap between the minimum marginal production cost (m_0) and the pre-improvement net revenue available ($P^B - T_1$), the more likely is underestimation. This is shown in Figure 4c. Underestimation is reinforced the more elastic the firm's marginal cost of supply schedule. In addition, if reduced transport costs also lower the delivered price of the firm's inputs, then the underestimation is greater since the firm's marginal cost of supply schedule would be lower and output would expand (to Q_3 in Figure 4c).

The influence of transport /logistic costs on the supply decisions by a producer in

serving a market (as discussed above) translate, of course, to the firm's derived demand schedule for transport services. This schedule follows the firm's marginal production cost schedule, but inverted. This is shown in Figure 5 for a case in which adoption of the rule of one-half would lead to significant underestimation of the benefits of a drop in transport costs-- partly due to the elastic supply response.

Figure 4b. "Rule of One-Half" Overestimates Benefits

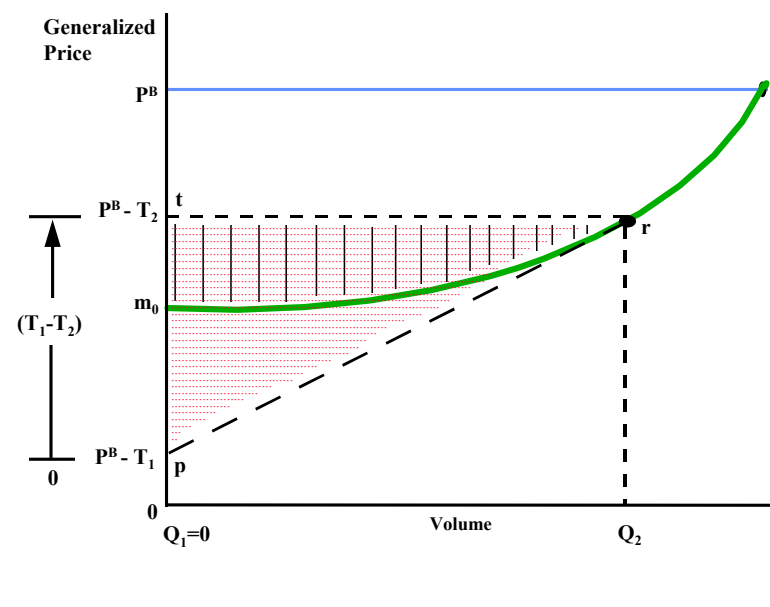
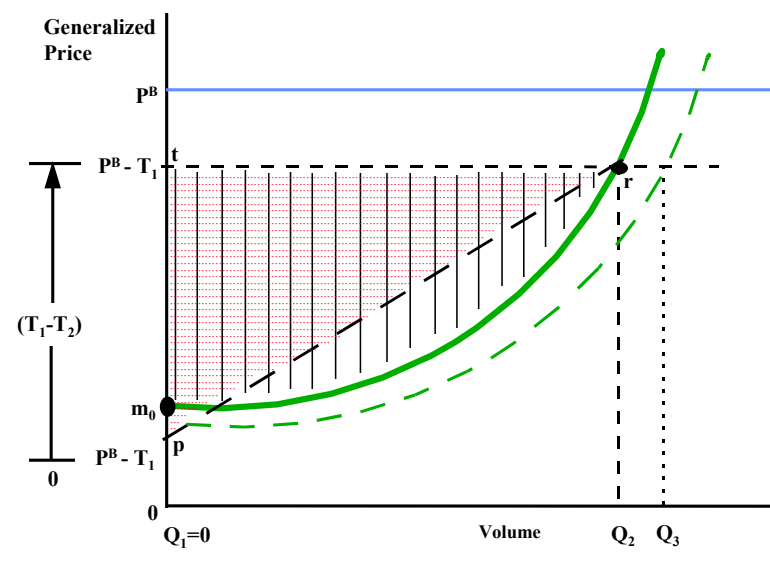
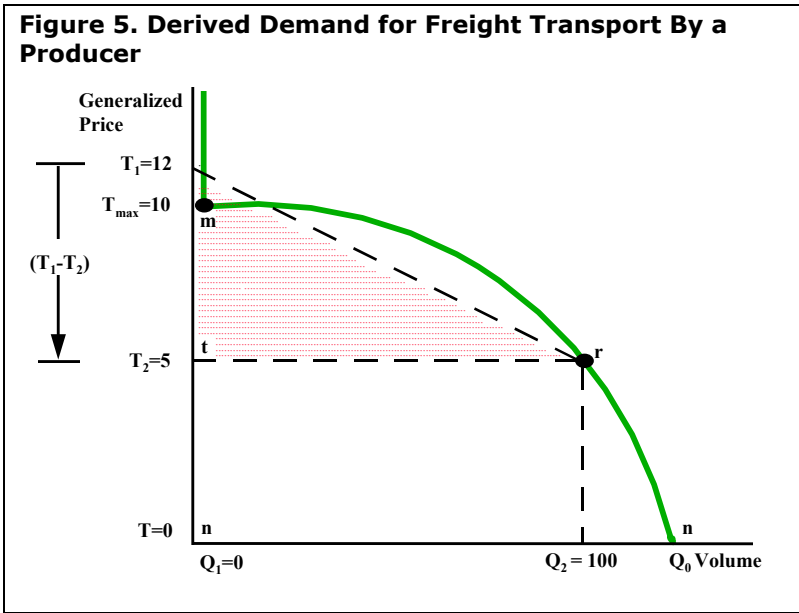


Figure 4c. "Rule of One Half" Underestimates Benefits

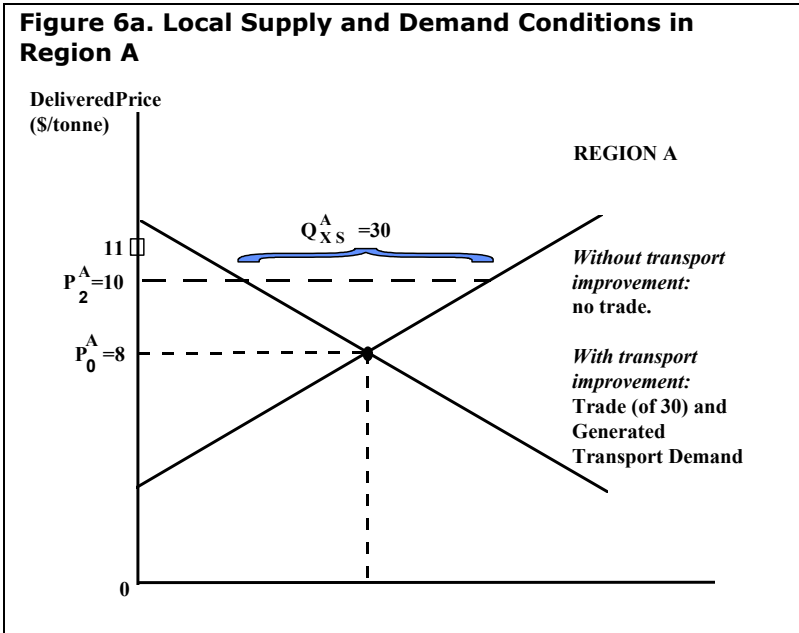


Interregional trade -- and hence the associated derived demand for freight transport services -- takes place provided the no-trade (autarkic) local market equilibrium prices differ by *more* than the cost of shipping the commodity from the low price region to the high price region.⁶ The level of trade flow/quantity of shipments is determined by the local demand and supply elasticities and the level (or supply schedule) of transport costs.⁴ These circumstances are set out diagrammatically in Figures 6a and 6b. The autarkic local prices are P_0^A and P_0^B . If their difference ($P_0^B - P_0^A$) is *less* (greater) than prevailing transport cost from A to B, then there *will not* (will) be shipments



of the commodity from A to B, and vice versa, i.e., if there are no existing freight movements, then transport costs are "too high" -- relative to the difference in local prices. For various levels of transport improvement, i.e., reductions in transport cost *in excess of* ($P_0^B - P_0^A$), there is a corresponding level of freight shipment representing net exports from A becoming net imports by B. This relationship represents the derived demand schedule for transport services.

This is illustrated by the numerical example also set out in Figures 6a and 6b. With no trade, local prices are \$20/tonne (B) and \$8/tonne (A), and the prevailing transport cost, say \$15/tonne, exceeds ($P_0^B - P_0^A$) ($20 - 8$) = \$12/tonne. If a transport improvement yields a substantial fall in transport costs, say ($T_2 - T_1 = 9$ /tonne), a level of import/export shipments equal to $Q_{1/2} = Q_{1/2}$



The basis of this demand for transport services is illustrated by the numerical example also set out in Figures 6a and 6b. Existing local product prices are \$20/tonne in B and \$8/tonne in A, and prevailing transport costs are \$15/tonne -- which exceeds

the difference in local prices ($20 - 8 = \$12$ /tonne). There are no freight shipments of the product from A to B. If a transport improvement can reduce transport costs to T_2 (\$6/tonne), i.e., $T_1 - T_2 = 15 - 6 = 9$ /tonne, then trade will take place and traffic will be "generated": equilibrium levels of local prices with exports balanced by imports will emerge. In the case illustrated, at a new local price of $P_{1/2} = 10$ (\$10/tonne) in A, excess local supply, $Q_{1/2}$ (30 tonnes), represents exports from A which must balance the imports by B that correspond to the excess local demand $Q_{1/2}$ (30 tonnes), in B at local price $P_{1/2} = P_{1/2} + T_2 = 10 + 6 = \16 /tonne. This "interregional equilibrium" trade outcome

⁶ Other factors influencing trade such as tariffs, quotes, non-tariff barriers, export taxes and so on are set aside here.

represents *one* point on the derived demand for freight transport services (for the particular commodity) from A to B. It is represented by the point *e* on the demand for transport schedule shown in Figure 7. Other points on the demand for transport schedule, for other levels of the price of transport services, correspond to other equilibrium values of local prices and quantities of freight movement.

In these circumstances, the flawed basis of "the rule of one-half" (area *nrd* in Figure 7) invariably *overestimates* the generated traffic benefits (area *trd*); the error is directly related to the gap between the pre-improvement transport cost (T_1) and the pre-improvement difference in the local prices in each region ($P_0^B - P_0^A$).

TRANSPORT DEMAND BY INDIVIDUAL BUYER

The demand for a particular commodity by an individual consumer depends primarily on its (delivered) price, the consumer's income, the relative prices of substitutes and, of course, the consumer's preferences.

Transport improvements lower the price of "imported" goods (e.g., tinned beef) relative to the price of local goods (e.g., fish). As noted in paragraph 7, if the delivered price of a commodity is "too high", imports of that commodity will not be made (i.e., existing import freight movements of that commodity are zero)⁷. But, for a "sufficiently large" reduction in transport costs (money, import cost, time, reliability, etc.,) some transport activity may be generated. These circumstances are shown in Figure 8a.

Figure 6b. Local Quantity (tonnes/month) in Region B

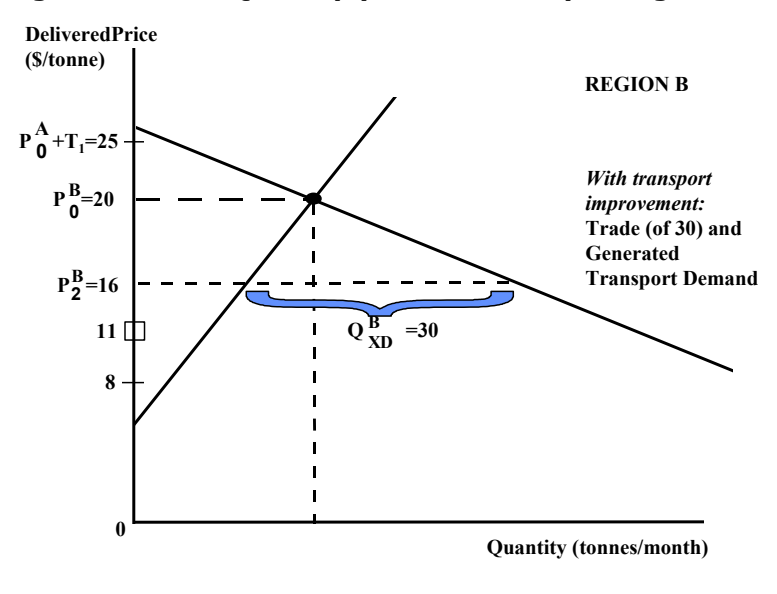
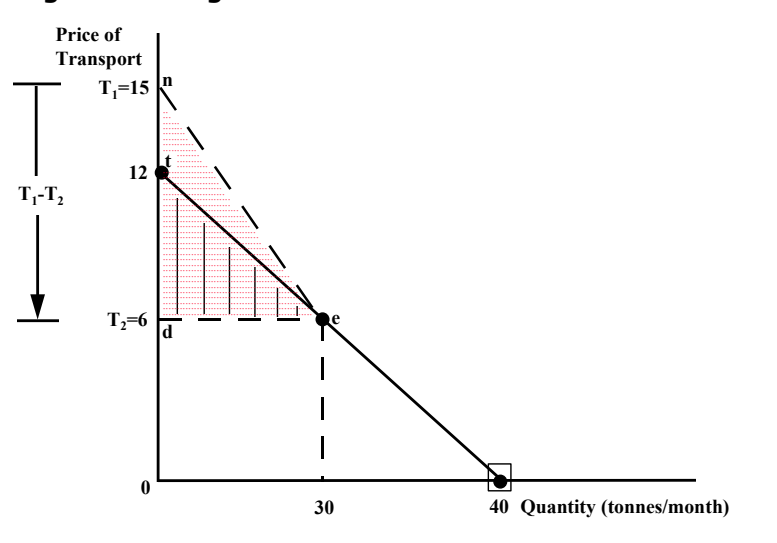


Figure 7. Derived Demand for Freight Transport from Region A to Region B



⁷ Note that these circumstances can be reinterpreted to characterize personal travel from B to A. For example, a work trip from B to A, "imports" income from A; the traveler residing in B chooses between local (net) income and income from work in A, net of commuting transport costs from B to A (ignoring all other differences, e.g., in the disutility of work). Only if commuting cost is "sufficiently low" will a worker commute from B to A, for given wage rates in A and B. Alternatively, consumers in B may travel in person to A to "buy" certain products (e.g., household goods) or services (e.g., health, education, and social/community activities).

The user's demand schedule for a commodity or trip purpose (or for a group of users) is shown graphically by the curve md^1 . If the delivered price exceeds P_{max} (\$10/lb), there is no quantity demanded. Thus, if the f.o.b price in source region A is \$2/lb and prevailing transport costs from a to B are \$9/lb, the delivered price in B (\$11/lb) implies that there is no demand for the commodity in B; existing traffic (for this commodity from A to B) is zero. A transport improvement that reduces the transport cost from A to B to \$1/lb, results in generated traffic of 20 lb per week. The net benefit associated with this generated traffic is the consumer surplus represented by the shaded area mnt , i.e., these benefits are shaped by the new (post-improvement) level of transport costs (\$/lb, for a given f.o.b. price, and the "maximum consumer price", (P_{max}) -- and not be the change/savings in transport costs (T_1-T_2) or $(9-1) = \$8/lb$. As shown in Figure 8b, use of the rule of one-half, here indicated by the shaded area snl , will always *overestimate* the benefits associated with generated traffic. Note that the "better estimator" of these benefits is a rule of one-half applied to the area mnt , i.e., $1/2 (10-3) * 20 = \$70$. The overestimation error can be expressed as $1/2 [P_{1/2} - P_{1/2}] * QG$ or, as a percentage of the "standard rule of one-half" estimator, $[P_{1/2} - P_{1/2}] / [T_1 - T_2] = [P^A + T_1 - P_{1/2}] / [T_1 - T_2]$. Except for the maximum price $P_{1/2}$, all these parameters are "known". Thus, for the illustrative numerical example here, the percentage error would be $[11 - P_{1/2}] / [9 - 1]$; with $P_{1/2}$ estimated to be 10 the error is 12.5%.

Figure 8a. Influence of a Maximum Consumer Price

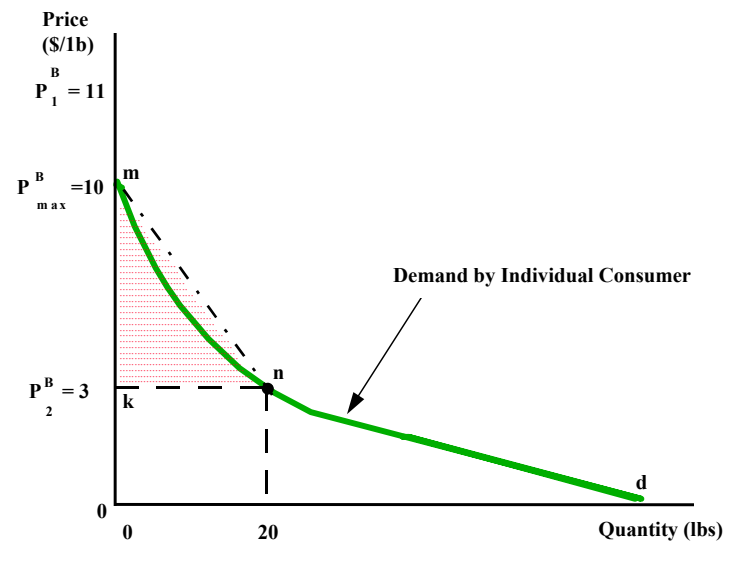
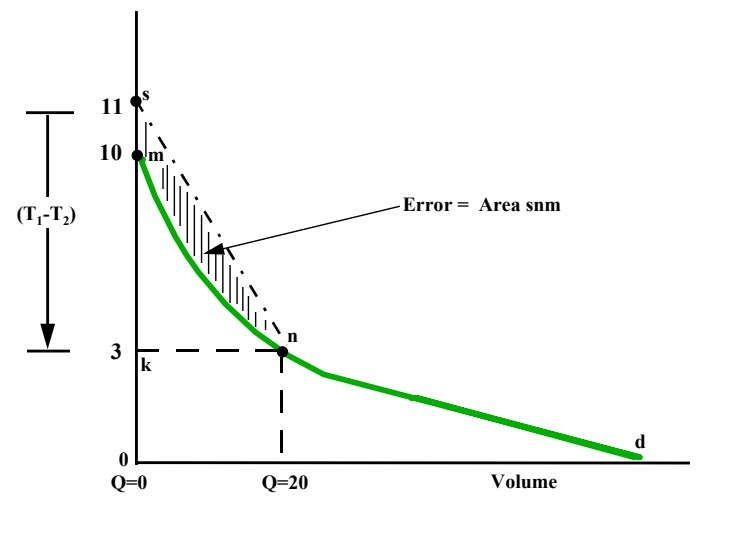


Figure 8b. Maximum Consumer Price and Error of Rule of One-Half



CONCLUSIONS AND PRACTICAL GUIDELINES

- ❑ A transport improvement consists of a reduction in the cost of supplying transport services (money, time, reliability, etc.). Some costs (time) are incurred by users, other costs are incurred by operators (vehicle operating costs); changes in the latter will be passed on to users in full if (i) users themselves provide the transport services (private automobile or ancillary trucking or (ii) if the market structure for transport services is highly competitive, but only in part, if not.
- ❑ The net benefits of transport improvements ultimately accrue as surpluses
 - to users (consumers, passengers, and shippers and ancillary operators)
 - to providers (for-hire operators)
 - suppliers of inputs (government, labor)

- to third parties, (landowners/adjacent residents)
- Use of improved transport facilities typically consists of pre-improvement "base" or "existing" traffic and "new" or "generated" traffic. The (net) benefit of the improvement associated with existing traffic is clear cut - it is the resource cost savings of serving that traffic. If the transport services market is competitive, all the savings will be passed on to users as consumers' surplus, otherwise some of the savings will be retained by operators as producers' surplus/excess profits. (The latter, in turn, may be "captured" as additional factor incomes, if the associated input markets are not competitive). In either case, taxes on transport inputs (e.g., fuel) or outputs (e.g., exports) will be transferred to government with a net loss in total surplus.
- The benefits to generated traffic are not clear cut. These benefits depend on properties (shape/elasticity) of the underlying demand schedule for the transport services, in addition to the levels of total transport user cost, without and with the improvement.
- If pre-improvement existing traffic is substantial, and hence generated traffic is relatively "marginal" to it, then benefits to existing traffic tend to be dominant and the rule of one-half will normally be a "good" estimator of the benefits to generated traffic; the error associated with a linear approximation of the unknown transport demand schedule will be relatively small.
- If without an improvement there is no relevant traffic, i.e., any traffic with an improvement will be generated, then the rule of one-half may be a poor estimator of the benefits associated with generated traffic. The reason is straight forward: the rule of one-half is driven by transport savings, but *for generated traffic there are no savings* since pre-improvement there were no transport costs for this traffic. *The logical basis of the rule of one-half for generated traffic "dissolves" when there is no existing traffic to tie-in the highest marginal value in use of generated traffic to the pre-improvement level of transport costs.*
- Use of the "Rule of One-Half" in situations where essentially all of the traffic is generated should be avoided. If applied, in most situations it will involve an overestimation error. However, in some situations (notably producers who are attracted to enter and serve a "world" market for their product which involves a given fixed price) the rule can lead to underestimation.
- Unfortunately, in situations dominated by generated traffic it is necessary to get a handle on properties of the market for transport services beyond an approximation to the price elasticity of demand. The most important of these properties is some idea of the price at which demand for transport associated with the improvement becomes zero. In some contexts, for example, an improvement which involves a new link that attracts traffic from an existing link which is a "close" substitute, the generalized cost on the existing link will be a good estimator of that maximum price.
- In contexts where transport investments/improvements "induce structural shifts" in the locations of economic activity (as well as "generating" increases in their levels), the economic logic of the "Rule of One-Half", as well as its use as a crude surrogate estimator for generated/induced traffic benefits, collapses. Some form of land use/location/transport/logistics interaction model is required if an attempt appears warranted to estimate the more pervasive ("general equilibrium") sources of such benefits.

Colin Gannon (TWUTD), March 1998

COLOMBIA: APPLICATION TO RAIL UPGRADING⁸

Outline of Proposed Project

- ❑ Colombia is a producer of high quality coal; the mine gate production cost is US\$35 per tonne.
- ❑ Land transport costs to deliver the coal to the port by a poorly maintained railway (or 1300 km by road) are US\$30 per tonne.
- ❑ Existing total transport costs are therefore ($\$30 + \10) = US\$40 per tonne.
- ❑ The "world price" of the coal is \$65 per tonne; world demand is perfectly elastic at this price (i.e., additional supply from Colombia will not lower this price).
- ❑ Delivered price of Colombian coal to "world markets" by existing transport is $35 + (30 + 10) = \text{US\$}40$ per tonne.
- ❑ Since the existing delivered price exceeds the prevailing world price, there are no exports of coal from Colombia.
- ❑ A railway improvement project would reduce land transport costs from US\$30 per tonne to US\$15 per tonne.
- ❑ The existing mine has a production output capacity of 1.0 million tonnes per annum.

Question: What is the economic benefit to Colombia, and the mining industry in particular, of the railway improvement?

Answer:

- ❑ In keeping with the focus on generated traffic benefits in this Note, questions of economic - financial costs, the time profile of benefits, and expanding the productive capacity of the mine are set aside. The analysis here is directed at estimation of rail traffic benefits (in any one year).
- ❑ Under these circumstances, the Colombia case is an example of the situation of "Freight Shipments by an Individual Producer" examined in the Note (paras. 9-13). However, the Colombian coal mine example presents a "special case", in as much of the supply schedule of the mine is:

Output Level tonnes p.a).	Marginal Cost (US\$ per tonne)
0 - 1.0 million	35
>1.0 million	"infinite"; production above 1.0 million tonnes not possible

⁸ This Colombia case was posed to me also by Robin Carruthers.

□ For this “special case”, the marginal cost supply schedule becomes “vertical” at 1.0 million tonnes, as shown in Figure 4a - Colombia, below.

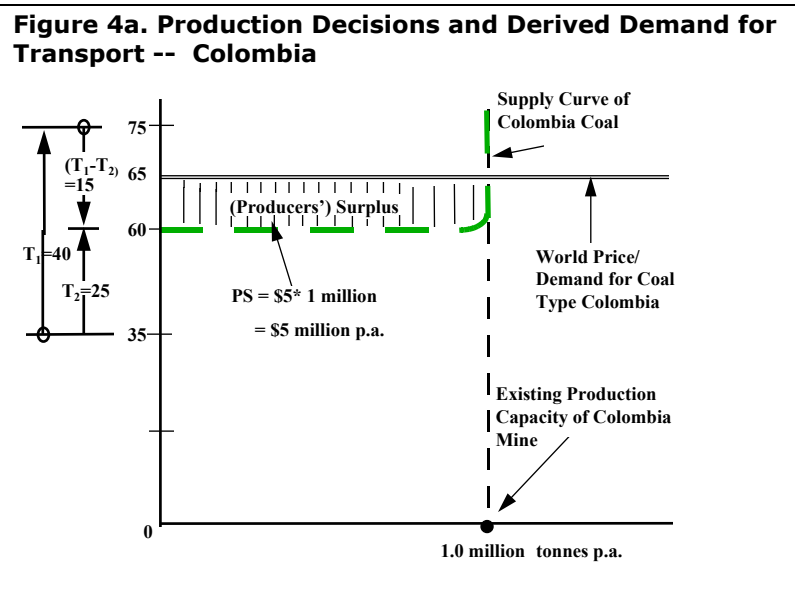
□ Provided transport costs remain at the reduced level of $T_2 = \$25/\text{tonne}$, i.e., transport operators - rail or shipping - cannot/do not try to extract rents/profits that arise, then:

Generated traffic on Colombia rail (“no road”) = 1 million tonnes p.a.

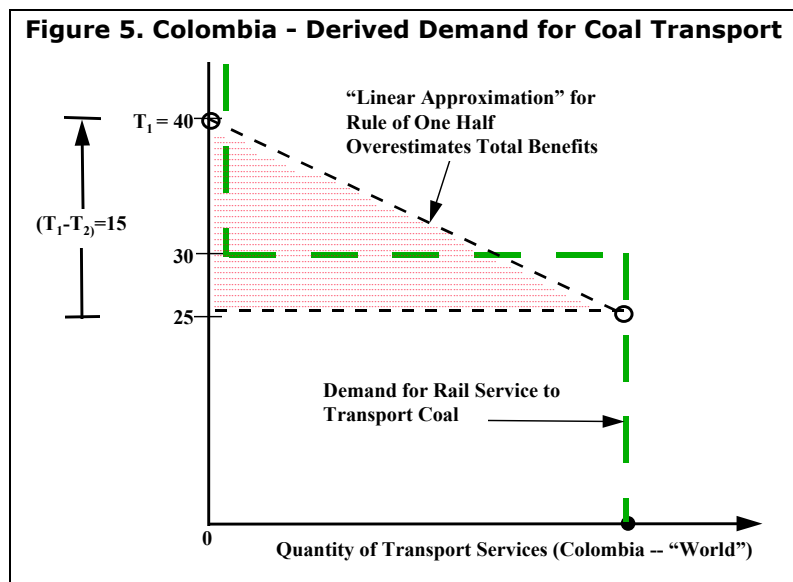
Benefits of generated traffic accrue as producers’ surplus = \$5 million p.a.

This surplus can be :

- (i) retained by mine owners/factory owners (miners);
- (ii) taxed away/lump sum or per unit;
- (iii) does not matter here given inelastic supply;
- (iv) captured by lump-sum royalties via rail access/all or nothing.



□ The “Rule of One-Half ” overestimates total benefits as illustrated in Figure 5 - Colombia.



□ The use of the “Rule of One-Half” is ill-advised here for the reasons presented in the text of the Note. If T_2 happened to be \$20/tonne, and not \$25/tonne, the “Rule of One-Half” would, “by coincidence”, give an exact estimate of the generated traffic benefits ($1/2 * 20 * \$1.0 \text{ million} = \10 million p.a.) since total surplus is then $(65-55) * 1.0 \text{ million} = \10 million, p.a.

ANNEX 3

ALTERNATIVES TO THE RULE OF A HALF IN MATRIX-BASED APPRAISAL

(this Annex is based on a paper by John Nellthorp and Geoff Hyman[†](Institute for Transport Studies, University of Leeds, UK and Department of Transport, Local Government and the Regions, UK) presented to the European Transport Conference, Cambridge, 10-12 September 2001)

Background

The 'rule of a half' is in widespread use as a measure of the user benefits in transport appraisal. In the UK, it is the measure that is recognised in the official methodology for multi-modal studies (DETR, 2000, Appendix F Section 5) and which is built into the TUBA appraisal software (White, Gordon and Gray, 2001 at this conference).

Rule of a half:

$$RoH = 0.5 \sum_{i,j,m} (T_{ijm}^0 + T_{ijm}^1) (GC_{ijm}^0 - GC_{ijm}^1) \tag{1}$$

where GC is the generalised cost of ij travel by mode m;

T is the number of ij trips per period by mode m;

Superscripts 0 and 1 denote the Do-Minimum and Do-Something scenarios respectively; and

CS and GC are at market prices.

However, in three plausible types of situations, the rule of a half is known to break down. These are:

- ❑ large generalised cost changes (for example, as a result of a new estuarial crossing, or a substantial new toll, between i and j on existing mode m);
- ❑ the introduction of new modes (for example, the introduction of LRT between i and j);
- ❑ new generators or attractors of trips (for example, due to a major development at a particular location).

All three are of increasing relevance at the present time, as transport policy moves towards new-mode solutions, road infrastructure charging and integrated transport-land use strategies in a number of countries.

This paper puts forward an alternative benefit measure - called 'numerical integration' - which can be used selectively as an alternative to the rule of a half. It is valid whatever type of demand model is being used, and extends rather than overturns the logic of the rule of a half. Numerical integration is shown to improve the accuracy of benefit estimation in the first two of the problematic cases above, and the analysis offers some insights into how to deal with case (iii) land use change (Sections 2,3 and 4).

Issues which arise when implementing numerical integration are addressed within each Section, and the specific question of implementation in software is considered. A spin-off from the development of numerical integration is a mathematical proof which extends the generality of the rule of a half (see Section 5). Finally, conclusions are drawn for future appraisal practice (Section 6).

The paper reports on work commissioned by DTLR (then DETR) from the Institute for Transport Studies, to examine the problem and recommend practical solutions. It is expected that the official advice on these matters will be prepared shortly. Exposure of this paper to comment and discussion will help, it is hoped, to inform that advice and give professionals an opportunity to input.

LARGE COST CHANGES

Reasons for the Breakdown of the Rule of a Half

In order to explain the breakdown of the rule of a half (RoH) when cost changes are large, we begin by setting out the standard justification for the use of the RoH.

The first and simplest version of the argument for the RoH is based on the assumption of a single generalised cost change (eg. a reduction in travel time from i to j by road). Supply conditions are assumed to change in one market only and no demand curves in any market shift. There is only one product to consider in the appraisal, with no close competition and no complementarity with other products. A good example of this situation in the real world would be a road improvement project between two rural communities, where the best alternative route is so long that no-one would consider using it.

In essence, the rule of a half (RoH) is a linear approximation to the Marshallian consumer surplus measure of benefits¹ (ΔCS):

$$\Delta CS_{ijm} = \int_{GC_{ijm}^1}^{GC_{ijm}^0} D(GC_{ijm}) dGC \tag{2}$$

$$As \ GC_{ijm}^0 - GC_{ijm}^1 \rightarrow 0, \ RoH_{ijm} \rightarrow \Delta CS_{ijm} \tag{3}$$

- where GC is the generalised cost of ij travel by mode m;
- D is the demand for ijm trips per period (a function of GC_{ijm});
- Superscripts ⁰ and ¹ denote the Do-Minimum and Do-Something scenarios respectively, as before; and
- CS and GC are both at market prices.

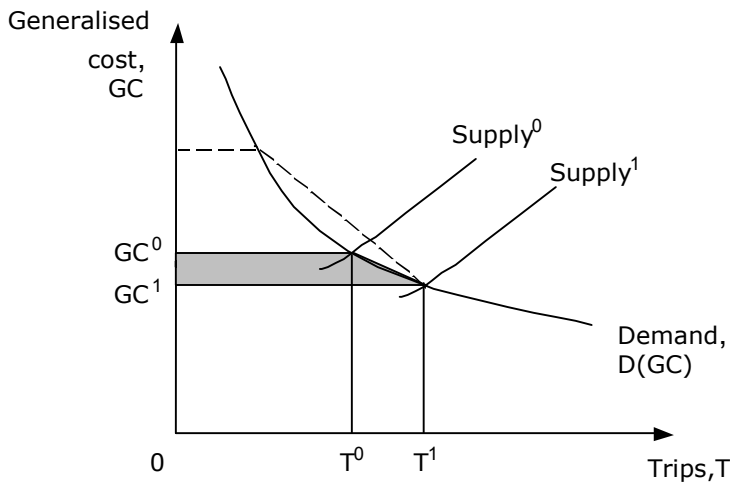


Fig 1. The 'Rule of a Half' to Consumer Surplus Change

This is illustrated in Figure 1 for a strategy which shifts supply conditions.

The RoH is a good approximation when price changes are small. However, when price change is large (eg. shown by the dotted line), the linear approximation becomes inaccurate. How inaccurate becomes clear in Section 2.3 when a numerical example is considered

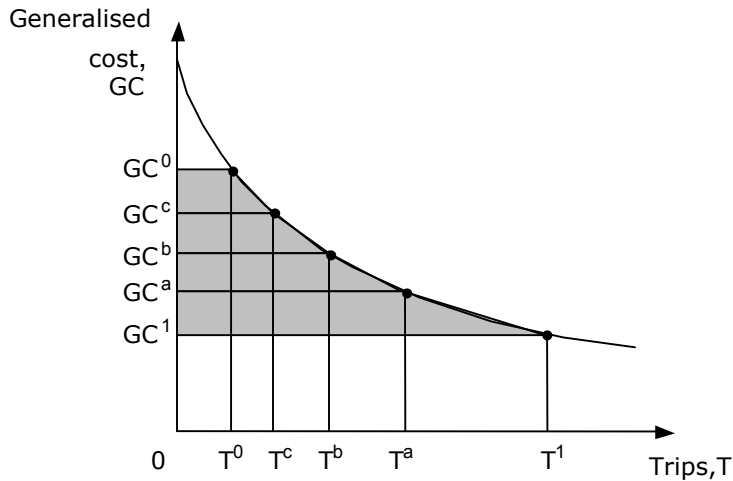
The second version of the justification for the RoH is more complex. In situations where demand curves can shift - for example, decongestion on a road link following the introduction of a parallel LRT - the rule of a half can still be applied to estimate the user benefits, subject to certain conditions holding. The principal conditions are that the change in costs remains small, and that there is symmetry of substitution between the services involved (Jones, 1977; Glaister, 1981). When the generalised cost change is large, this logic also breaks down.

As part of the background for our work, Hyman (2001) has provided a mathematical proof of the validity of the RoH for the joint consumption of related products - see Section 5. However, the conclusion remains that the larger cost change, the less reliable (potentially) the RoH becomes: an alternative benefit measure is needed.

Numerical Integration

'Numerical integration' involves defining a set of trapeziums which together approximate the change in consumer surplus. The Do-Something and Do-Minimum points used for the RoH calculation (T^0, P^0) and (T^1, P^1) are retained, and supplemented by additional points. Figure 2 illustrates how the method works - the shaded areas indicate the (estimated) user benefits.

Each trapezium, and in fact each triangle, can be calculated using the rule of a half, so the process is simple once the additional points have been defined.



**Fig 2 - Numerical Integration:
Existing Mode (large cost change)**

For numerical integration with three additional points (T^a_{ijm}, GC^a_{ijm}) , (T^b_{ijm}, GC^b_{ijm}) and (T^c_{ijm}, GC^c_{ijm}) . Let these GC levels be:

$$\begin{aligned} GC^a_{ijm} &= 0.25*(GC^0_{ijm}-GC^1_{ijm}) + GC^1_{ijm} \\ GC^b_{ijm} &= 0.5*(GC^0_{ijm}-GC^1_{ijm}) + GC^1_{ijm} \\ GC^c_{ijm} &= 0.75*(GC^0_{ijm}-GC^1_{ijm}) + GC^1_{ijm} \end{aligned} \quad (4)$$

Given these three levels of generalised cost between i and j by mode m, the forecasting model must be run to determine the corresponding T^a , T^b and T^c .

The numerical integration function (NI) is then:

$$NI = 0.5 \left[\begin{aligned} &(T^0 + T^c)(GC^0 - GC^c) + (T^c + T^b)(GC^c - GC^b) \\ &+ (T^b + T^a)(GC^b - GC^a) + (T^a + T^1)(GC^a - GC^1) \end{aligned} \right] \quad (5)$$

with the ijm subscripts omitted for clarity.

Three advantages of numerical integration that are immediately apparent are:

- it appears to be general - no specific analytical forms of the demand function need to be assumed, hence NI can be used with logit demand models, negative exponential and any other functional forms;
- related to this, knowledge of the demand function is not required in order to estimate the user benefits - this could be a big advantage for more complex demand models, where there may not even be a single explicit demand function in each market;
- numerical integration is highly consistent with the standard treatment of existing modes, as it requires only explicit trip and cost matrices for the mode of interest.

Worked Example: An Estuary Crossing

In this section, we use a worked example to explore how numerical integration can be applied in matrix-based appraisal. The example uses a very small matrix, but the method can be applied to matrices of any size - we return to the practical issues raised in large matrices at the end.

Suppose that there are two zones in an urban network, separated physically by a river or estuary. Let these be zones 1&2. The shortest route linking 1&2 crosses the river between two neighbouring zones, 3&4. The figure below (Fig 3) illustrates the situation. A project is proposed to link the first two zones directly, across the estuary. To appraise this project, an estimate of the potential user benefits is needed.

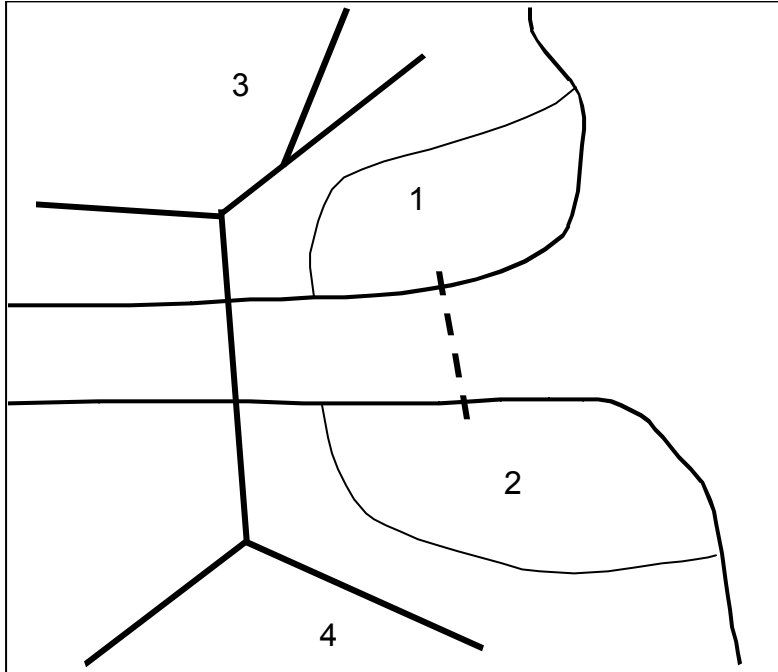


Figure 3. Estuarial Crossing in a Network

User benefits may arise across a range of modes, including private car, bus, rail (if the new crossing carries a railway), cycling and walking. Benefits relating to all these modes are relevant in multi-modal appraisal. For simplicity however, we will confine the analysis to just one of these - private car - and to one trip purpose - non-working- and one time of day - peak.

In order to make the benefit calculations we require some information about travel conditions for this demand segment between zones 1&2, so we assume the following (Table 1). It was also supposed that the crossing is not tolled.

Table 1. Estuary Crossing Example: Short Scenario Descriptions

Do-Minimum:		
No bridge between 1&2		
Time ₁₂	50 mins	=>speed= 24 miles/hour
Distance ₁₂	20 miles	
Trips/hour ₁₂	1000	
Do-Something:		
Bridge between 1&2		
Time ₁₂	10 mins	=>speed= 30 miles/hour
Distance ₁₂	5 miles	
Trips/hour ₁₂		

Values of time and vehicle operating costs were taken from the Transport Economics Note (DETR, 2001), yielding the following generalised costs (GC₁₂) at 1998 prices and values, assuming that all trips used the lowest generalised cost route.

Table 2. Estuary Crossing: Generalised Costs, Pence Per Trip

	Do-Minimum	Do-Something
Time	549.9	110.0
VOC	315.7	72.6
TOTAL	865.7	182.6

To simulate the transport modelling stage, a base number of trips from zone 1 to 2 was assumed: 1000 per period in the Do-Minimum scenario. A demand function was taken from the SATURN modelling procedures for elastic assignment ($T=T^0 \exp(-\beta(GC-GC^0))$) using a demand sensitivity coefficient $\beta=0.0037$. A number of further simplifying assumptions were made:

- ❑ travel between zones 3 and 4 will continue to be quickest and cheapest via the old bridge across the estuary, not through zones 1&2, and will be unaffected by the project to any significant degree;
- ❑ intra-zonal travel will not be affected by the project;
- ❑ travel from zones 1 to 4 and 2 to 3 (and vice versa) will be made easier by the project, but given the overall distance involved generalised costs will not fall by a large amount.

The projected demand response to various large cost reductions including the extremely large cost reduction shown in Table 1, was as shown in Table 3.

Table 3a. Estuary Crossing: Demand Response to Large Cost Reductions (zone 1 to zone 2)

% cost Δ	GC⁰	GC¹	T⁰	T¹
-80	865.7	173.1	1000	6487.3
-50	865.7	432.8	1000	3217.7
-33	865.7	580.0	1000	2162.6

Table 3b. Estuary Crossing: Demand Response Across the Network

ΔTrips_{ij}	Zones i			
Zones j	1	2	3	4
1	-	↑	-	↑ RoH
2	↑	-	↑ RoH	-
3	-	↑ RoH	-	-
4	↑ RoH	-	-	-

Key: ↑ indicates response to large cost change

↑ RoH indicates trip response to cost change is small - RoH may be

Therefore to estimate the user benefits of the project for this demand segment (car, non-working, peak hour), the rule of a half can be used to estimate the benefits for most cells in the matrix. However, for (1,2) and (2,1) we have the opportunity to test alternative benefit measures. These are:

- ❑ the rule of a half - ie a conventional appraisal;
- ❑ the integral consumer surplus - the theoretically correct benefit measure, not usually calculated because of the practical difficulties; and

Table 4: Estuary Crossing: Alternatives Measures of User Benefits

%GC change	RoH	ΔCS	NI		%difference vs ΔCS:		
			1 point	3 points	RoH	NI 1 point	NI 3 points
-80	4836236	3234250	3665008	3344102	49.5%	13.3%	3.4%
-50	1289942	1070385	1126977	1084646	20.5%	5.3%	1.3%
-33	553871.3	507481.1	519238.2	510430.6	9.1%	2.3%	0.6%

numerical integration- with one or three additional points.

Table 4 gives the benefit estimation results and the error due to the use of firstly the rule of a half and secondly the new benefit measure, numerical integration. These results suggest that the RoH is seriously inaccurate (>10%) for cost changes of >33%. This finding is supported by wider experimentation, in which not only the size of the cost change but also the elasticity coefficient on the demand function were allowed to vary.

It should be noted also that we found before that large changes in trips are no more or less problematic than large changes in costs, in terms of the inaccuracy they cause in the RoH. Therefore the suggested rule for use of NI relate to large trip changes as well as large cost changes. These rules and other practical implementation issues are discussed in the following Section.

Implementation Issues

The key issues identified during our work were:

- When is it advisable to use numerical integration in place of the RoH?
- How many additional (T,GC) points are needed?
- What is involved - in practical terms - in obtaining Trip estimates for these points?

On the first and second issues, the following rules for intervention are proposed. These are based on experimentation, and except in the case of extremely curved demand functions (for example, an exact reverse L shaped curve), they should usually be adequate to report the benefits accurately to within +/-10% (Nellthorp and Mackie, 2001).

Table 5. Suggested Rules for Implementation

Magnitude of cost and trip changes	User benefit estimator
If % change in $P_{ijm} < 33\%$ AND % change in $T_{ijm} < 33\%$ then	RoH
If % change in $P_{ijm} \geq 33\%$ OR % change in $T_{ijm} \geq 33\%$ then	Numerical integration with 1-3 additional points

In choosing the number of additional (T,GC) points, we have found that NI with 1 additional point brings acceptable accuracy (+/-10%) for cost changes up to roughly 70%. For larger cost changes, NI with 3 additional points brings improved accuracy, at a cost in terms of calculations necessary. Table 4 illustrates this using the numerical example.

The third issue warrants some discussion too. Additional data to implement NI would be generated by re-running the forecasting model for different levels of perceived (generalised) cost for the ijm option concerned. In doing this, care should be taken to ensure that perceived costs are measured in the standard unit of account for TUBA inputs, which is factor cost for work travel and market prices for non-work travel.

When specifying the additional model runs, it is suggested that the definition of the project should be used to help identify the intermediate levels of generalised cost. Thus in the estuary crossing example, the GC level halfway between GC^0 and GC^1 was obtained by setting ij journey time and distance approximately halfway between their original two levels. It is not necessary for the intermediate points to be exactly evenly spaced - indeed it is recognised that given the need for model equilibrium it may be impossible to fix absolutely exactly on the target value. Numerical integration can give improvements in accuracy without such rigid precision in these GC levels.

These rules could be refined in future in the light of practice using this method with real multi-modal study data. There is obvious scope for adjustment of the number of additional data points, and/or the threshold values of '% change in trips' and '% change in cost', and/or the use of a single threshold.

Disaggregation into Components

This is fairly straightforward, although as with the conventional RoH calculations, taxes and unperceived costs complicate matters slightly. The NI_{ijm} calculations for each component are:

$$\text{user charges: } 0.5 \left[\begin{aligned} & (T^0 + T^c)(M^0 - M^c) + (T^c + T^b)(M^c - M^b) \\ & + (T^b + T^a)(M^b - M^a) + (T^a + T^1)(M^a - M^1) \end{aligned} \right] \quad (6)$$

$$\text{fuel VOCs: } 0.5 \left[\begin{aligned} & (T^0 + T^c)(F^0 - F^c) + (T^c + T^b)(F^c - F^b) \\ & + (T^b + T^a)(F^b - F^a) + (T^a + T^1)(F^a - F^1) \end{aligned} \right]$$

$$\text{non-fuel VOCs: } 0.5 \left[\begin{aligned} & (T^0 + T^c)(N^0 - N^c) + (T^c + T^b)(N^c - N^b) \\ & + (T^b + T^a)(N^b - N^a) + (T^a + T^1)(N^a - N^1) \end{aligned} \right]$$

$$\text{travel time: } 0.5 \left[\begin{aligned} & (T^0 + T^c)(V^0 - V^c) + (T^c + T^b)(V^c - V^b) \\ & + (T^b + T^a)(V^b - V^a) + (T^a + T^1)(V^a - V^1) \end{aligned} \right]$$

The components of user benefit to enter into the TEE table are then:

for work trips:

$$\text{user charges: } (1 + t)NI_{ijm_user\ charges} \quad (7)$$

$$\text{fuel VOCs: } (1 + t)NI_{ijm_fuel\ VOCs}$$

$$\text{non-fuel VOCs: } (1 + t)NI_{ijm_non-fuel\ VOCs}$$

$$\text{travel time: } (1 + t)NI_{ijm_travel\ time}$$

for non-work trips:

$$\text{user charges: } NI_{ijm_user\ charges}$$

$$\text{fuel VOCs: } NI_{ijm_fuel\ VOCs}$$

$$\text{non-fuel VOCs: } NI_{ijm_non-fuel\ VOCs} - \sum_{ijm} T_{ijm}^1 N_{ijm}^1 - T_{ijm}^0 N_{ijm}^0$$

$$\text{travel time: } NI_{ijm_travel\ time}$$

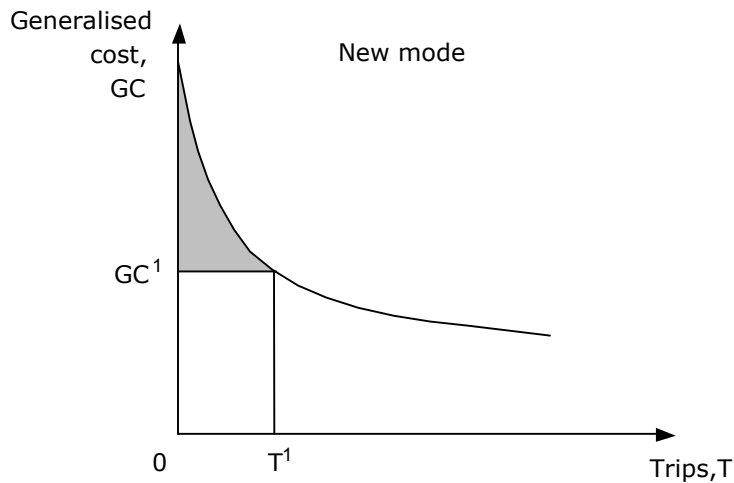
NEW MODES

Reasons for the Breakdown of the RoH

For a new mode between i and j, the trips and costs in the do-something scenario are known from the transport model outputs: (T_{ijm}^1, GC_{ijm}^1) . In the do-minimum scenario, trips on the new mode $T_{ijm}^0=0$. However, the do-minimum generalised cost for this mode between i and j is undefined (or infinite), because the mode does not exist for that ij pair in that scenario. This implies that the RoH is also undefined (or infinite).

Consumer surplus for a new mode is given by the integral above the Do-Something generalised cost - see Figure 4. This can be a definite integral if the demand curve intersects the GC axis or an improper integral if the demand curve is asymptotic to it. In general, user benefit of a new option is defined as:

$$\Delta CS_{ijM} = \int_{GC_{ijM}^1}^{\infty} D_{ijM}(GC_{ijM}) dGC_{ijM} \quad (8)$$



**Fig 4 - Consumer Surplus:
new mode**

The question is: how best to estimate this consumer surplus?

Solutions Proposed Previously

The first and in some ways most appealing answer is: calculate the integral directly. There are some significant obstacles to this, however:

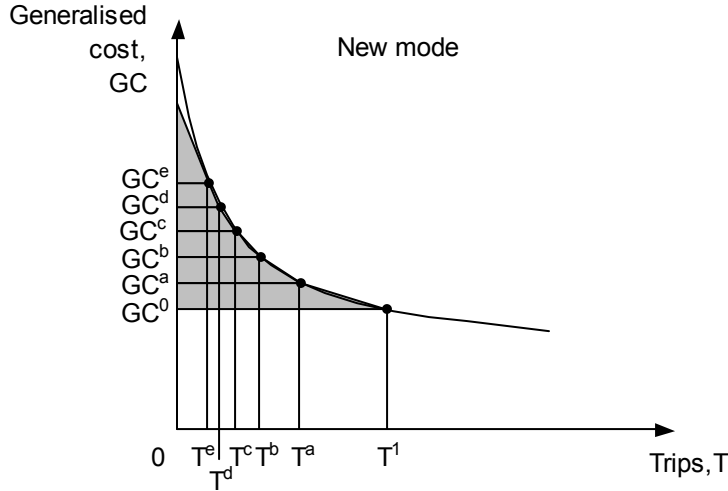
- ❑ the demand function in the transport model may not be expressed in such a way that it can readily be integrated: hierarchical choices and discontinuous functions are key sources of difficulty;
- ❑ the values and coefficients in the transport model must be identical with those used in appraisal, to ensure consistency across modes and cells in the appraisal;
- ❑ the functions of modelling and evaluation are often separated physically and in time, which places a great deal of emphasis on the data transfer between them: to pass non-standard information such as the specification of potentially complex demand functions between the two could be an invitation for errors and misinterpretations.

In view of these difficulties with the integral consumer surplus, the Common Appraisal Framework (MVA, OFTPA and ITS, 1994, Appendix D) proposed a number of pragmatic alternatives. These included:

- ❑ the rough estimation of demand curves for the new mode by fixing a cost intercept on the GC axis and then connecting this to the known point (T_{1ijM}, GC_{1ijM}) by a straight line;
- ❑ assuming that the demand function could be represented by a binary logit function and using professional judgement to estimate the scaling factor λ ;
- ❑ using the rule of a half further up the choice hierarchy, to avoid having to calculate CS for new modes at all.

Each of these has some more and some less apparent weaknesses, many of which were recognised in the Common Appraisal Framework itself. A full critical discussion is given in our report (Nellthorp and Mackie, 2001). DTLR was still looking, therefore, for a more robust, standardised approach to estimating the consumer surplus for new modes. Ideally, the approach should be implementable in software - for example as an extension of TUBA.

Numerical Integration for New Modes



**Fig 5 - Numerical integration:
new mode**

Numerical integration provides a way of obtaining most of the accuracy of the integral CS without requiring knowledge of the demand function.

To obtain an estimate of user benefits by numerical integration in the case of a new mode M, it is necessary to determine GC_{ijM} , and the corresponding trip matrices T_{ijM} , for a number of hypothetical levels of perceived generalised cost. Let us call these hypothetical scenarios a,b,c and so on. The additional input data required is:

$$\begin{aligned}
 &(T^a_{ijM}, GC^a_{ijM}) \\
 &(T^b_{ijM}, GC^b_{ijM}) \\
 &(T^c_{ijM}, GC^c_{ijM}) \\
 &(T^d_{ijM}, GC^d_{ijM}) \\
 &(T^e_{ijM}, GC^e_{ijM})
 \end{aligned}
 \tag{9}$$

for all those i-j pairs between which the new mode M is introduced. Equal spacing of data points in terms of cost is recommended between GC^e and GC^1 . Hence $GC^d_{ijM} = 0.8*(GC^e_{ijM} - GC^1_{ijM}) + GC^1_{ijM}$ and so on.

This data would be generated by re-running the forecasting model for different levels of perceived generalised cost for the new mode. In doing this, care should be taken to ensure that perceived costs are measured in the standard unit of account for TUBA inputs, which is factor cost for work travel and market prices for non-work travel. TUBA itself makes the correction to market prices for work travel (Mott MacDonald, 2001).

The numerical integration function is:

$$NI \approx 0.5 \left[T^d \left(\frac{GC^e T^d - GC^d T^d}{T^d - T^e} \right) + (T^d + T^c)(GC^d - GC^c) + (T^c + T^b)(GC^c - GC^b) \right. \\ \left. + (T^b + T^a)(GC^b - GC^a) + (T^a + T^1)(GC^a - GC^1) \right] \quad (10)$$

Incorporating this function directly in matrix-based appraisal would involve programming the software to accept multiple scenarios a to e simultaneously as input data. If this were found to be too difficult, for example if the TUBA data structure is now rigidly fixed in terms of the number of scenarios that can be manipulated at once, an alternative would be to implement the method as a series of TUBA runs with pairs of scenarios input as do-minimum and do-something. One of these runs would serve to calculate the upper triangle, the remainder would calculate the trapeziums. The results, calculated using the RoH, would be summed to obtain the same result as equation (10).

Worked Example: A New LRT

Numerical integration has been tested using a hypothetical new LRT route from a suburb i to the city centre zone j. For simplicity, we consider only individuals for whom car is available and we assume that the existing bus option is so infrequent and of such poor quality that it is not a realistic alternative: car and LRT are therefore the only choices.

In general, generalised cost is the sum of: Money cost (fares and VOCs); In vehicle time; Walk time (access and egress); Wait time; Modal constant. Demand was modelled very crudely using a binary logit model taken from the Manchester Metrolink Monitoring Study (Oscar Faber, 1996, Volume 2 Tables C5/6), whose scaling parameter is equivalent to -0.042 utils per minute.

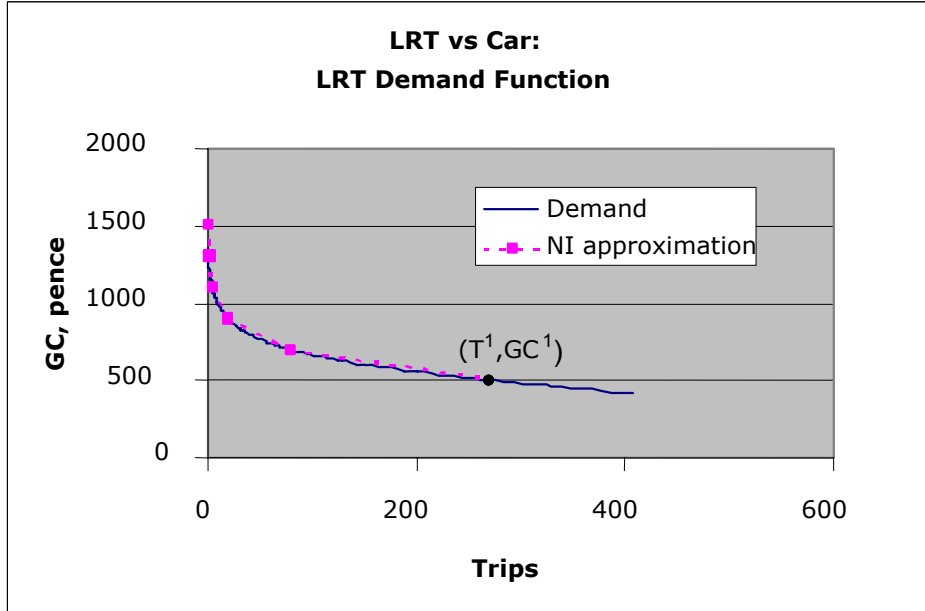
Assumptions were made about the characteristics of the alternatives as follows (Table 6).

Table 6. LRT Example - Characteristics of Car and LRT

GC _{ijm} components	Mode		Units
	Car	LRT	
Money cost (fares & VOCs, parking charges)	200	80	pence per one-way trip
In-vehicle time	15	20	mins per one-way trip
Walk time (access and egress)	5	10	mins per one-way trip
Wait time	0	7.5	mins per one-way trip
Modal constant	-	-13.9	pence per one-way trip

This model and assumptions implied the demand relationship for LRT shown in Figure 6. The prediction in the do-something scenario is for LRT to take a 26.8% mode share.

Figure 6. LRT Example: Logit Demand Function and NI Approximation



It can be shown that the integral consumer surplus calculation for LRT is:

$$\int_{GC_{ijLRT}^1}^{\infty} CS = \frac{T_{ijLRT}^0}{\lambda} \ln(1 - S_{ijLRT}^1) = 43077.5 \text{ pence} \quad (11)$$

where λ is the scaling factor in utils per pence = -0.0072;
 S_{LRT}^1 is the market share taken by LRT in the do-something.

The points for numerical integration (NI) are as shown in Table 7. Applying the NI formula above gives NI = 47874.6 pence.

Table 7. LRT Example - Points for Numerical Integration

	T	GC
e	0.25	1509
d	1	1308
c	5	1107
b	19	905
a	78	704
1	268	503

Thus NI overstates the integral CS by 11.1%, when it is applied with 5 additional points. An alternative approach might be to apply NI with just 3 points (e, c and a) instead of 5. This overstates the true CS by 22.1%.

It is worth noting that in response to comments on the study report, the leftmost point, (T^e, GC^e) has been defined in terms of GC rather than Trips. This should ease the job of implementation, because it is easier (more direct) to input a particular level of GC_{LRT} into the model than to aim for a particular number of trips, the latter possibly requiring a process trial and error. A tentative assumptions broadly consistent with the analysis do far is that:

$$GC^e = 3 * GC^1 \quad (12)$$

and $T^e = D(GC^e)$.

Another suggestion is that this number may depend on the number of additional points.

For comparison, if a rough estimate is made using a straight line through the known point and just one other point on the demand curve - ie. a simple RoH calculation - the scope for inaccuracy is much greater. The % errors are +91.1% if point c is used and +151.6% if point d is used.

Implementation Issues

The principal implementation issues are:

- How to define the leftmost point?
- How many additional (T,GC) points are needed?

We have already noted that the leftmost point could be found by entering supply conditions such that $GC_M = 3 * GC_M^1$ into the demand model. If there is a demand/supply response which causes equilibrium GC to diverge significantly from the input value (eg. congestion on the new mode), then some adjustment may be needed, although on the whole new modes are probably less likely to suffer from congestion than others.

The number of additional points required has been explored using several numerical examples and typical results suggest that at least 3 points would be needed to reduce the error from 100% to 20%. The addition of further points does not appear to improve the accuracy that much further (not below 10%). In absolute terms, most of the error arises in the lower portion of the demand function: this is, we believe, a consequence of the way the points are specified at equal GC intervals. It may be desirable to refine the advice on this - although whether this is necessary depends partly on the acceptability of a 10-20% error (bearing in mind that the error when new modes are evaluated using other approximate methods is likely to be much greater).

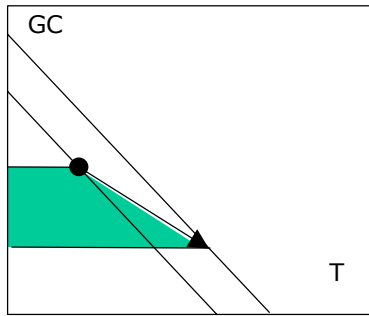
NEW GENERATORS AND ATTRACTORS OF TRIPS

The main focus of interest in transport appraisal is usually the benefit due to the Transport strategy (ie. the change in welfare between the Transport Do-Something scenario and the Transport Do-Minimum). However, other policy issues may arise.

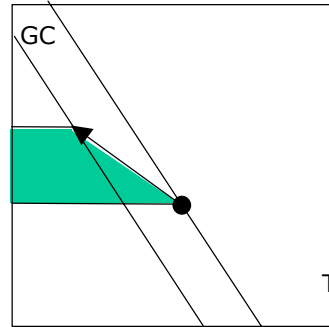
There may also be a Land Use strategy. In England, under the new 'integrated' regional structure, the Regional Planning Guidance (RPG) incorporates the Regional Transport Strategy (RTS) as well as a regional development and land-use planning policy. It is quite possible, therefore, that policy-makers will wish to see an evaluation of the transport strategy and land use strategy together, as well as the more traditional evaluation of transport strategy alone. Such an evaluation would need to address the (dis)benefits associated with new generators and attractors of trips: physically, these could take the form of new business parks or new towns (such as the proposed 'Alconbury' new town in the Cambridge-Huntingdon corridor). In modelling terms, they represent ijm cells in the trip matrix where no trips - or few trips - were made in the Do-Minimum. There is an analogy here to new modes and large trip changes, which we can make use of in thinking about new generators and attractors.

Hyman (2001) supports the conclusions of Jones (1977) and Glaister (1981) that when there is a shift in the partial demand curve for a product - for example ij travel by rail - the RoH can be used to estimate the change in user benefits, provided the income effects are zero. Let us consider what this means in four different appraisal situations.

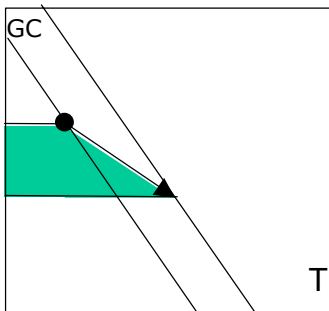
Figure 7. Evaluation Using Partial Demand Functions - Market Interactions



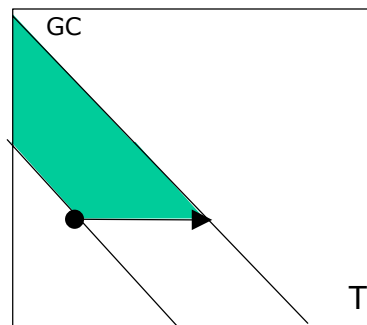
A) Price cut in product:
e.g. Variable Matrix Appraisal



B) Price rise in substitute:
e.g. Multimodal Appraisal



C) Price cut in complement:
e.g. Wider Economic Benefits



D) Rise in income:
e.g. Reference Case

In cases (A), (B) and (C), prices of products change but incomes are assumed to be fixed and the RoH gives the correct change in CS (shaded). However, in case (D) where incomes rise but prices remain fixed, the change in CS is not given by the RoH. In diagram A the shaded area shows the change in consumer surplus for product 1 when there is a reduction in its price and the same time a change in price of a related product. In diagram B the related product 2 is assumed to be a substitute, whose price has risen and the shaded area shows the loss in consumer surplus associated with product 2. In diagram C the related product 2 is assumed to be a complement, whose price has fallen and the shaded area shows the increase in consumer surplus associated with product 2. In diagram D there are no related products and the shaded area shows the increase in consumer surplus associated with product 1 when there is a rise in income.

Now, suppose housing and transport are complementary goods: then by the above reasoning it is appropriate to evaluate price changes for both using the RoH (or NI where necessary). There is no need to hold land use constant when estimating the transport benefits in a Land-Use/Transport Interaction model: they can both be allowed to vary and the RoH will serve to attribute the benefits by 'source' between transport and housing, just as it will between modes of transport. If correct, this suggests an amendment may be needed to the advice in GOMMMS (DETR 2000, Volume 2, pB11, Paragraph 2.44).

Going somewhat further, we hypothesise that:

$$\text{Total benefit} = \text{RoH(Transport)} + \text{RoH(Land \& property)} + \text{RoH(Labour)} \quad (13)$$

RoH(Transport) would be provided by TUBA; the other two would be the subject of separate economic analysis and would be brought together to inform the 'Wider economic impacts' line in the New Approach to Appraisal (NATA). This makes an interesting comparison with Martinez and Araya (2000, Fig 1), who conclude that transport sector benefits and land use benefits may not be strictly additive

in the presence of transport externalities such as noise nuisance. The full report (Nellthorp and Mackie, 2001) gives some further discussion on this topic.

GENERALISATION OF THE RULE OF A HALF

Hyman (2001) gives a mathematical proof of the validity of the RoH in the case of joint consumption of related products, such as the multi-modal appraisal context. There are two related strands of logic.

The first assumes that the willingness to pay for any basket of products can be expressed as a C^2 (twice continuously differentiable) function of the quantities of each product and proceeds to establish sufficient conditions for the exact application of the rule of a half in the case of multiple related products.

The second uses the multivariate Taylor series expansion of $f()$, a C^2 scalar function:

$$f(\mathbf{Y}) = f(\mathbf{X}) + \sum_k (Y_k - X_k) \frac{\partial f}{\partial X_k} + \frac{1}{2} \sum_{kl} (Y_k - X_k)(Y_l - X_l) \frac{\partial^2 f}{\partial X_k \partial X_l} + O(3) \tag{14}$$

to demonstrate that:

$$f(\mathbf{Y}) - f(\mathbf{X}) = \frac{1}{2} \sum_k (Y_k - X_k) \left(\frac{\partial f}{\partial X_k} + \frac{\partial f}{\partial Y_k} \right) + O(3) \tag{15}$$

Note that the 1st order terms are identical to the rule of a half, the second order terms are zero and the third order terms are a possible source of error

Hence in the demand shift/multiple price change case, the RoH gives an *approximation* that is accurate to second order. The rule can be applied either indirectly to general twice differentiable multi-product WTP functions of quantities or directly to twice differentiable multi-product consumer surplus functions of prices.

When dealing with large price changes or new products the accuracy of the RoH may be insufficient. Provided that we are still dealing with twice differentiable functions, numerical integration may be used to provide the required accuracy.

These results are significant because they verify that the benefit measures embedded in TUBA are appropriate to estimate benefits on inter-related modes, subject to extension to include NI. In particular, they demonstrate that the position taken in GOMMS (DETR, 2000, Appendix F Para 4.18) is a rather cautious one: it is important to ask why is there believed to be a problem with non-uniqueness of attribution of benefits to mode? The results also provide the basis for the novel discussion of the land-use change issue in Section 4.

CONCLUSIONS

Alternatives to the rule of a half are needed if we are to have truly general appraisal methods. However, this paper has argued that substantial progress can be made by extending the logic of the rule of half and using the technique of 'numerical integration' in selected cases where the RoH breaks down.

Numerical integration appears to be applicable to the estimation of consumer surplus for both existing and new modes, and represents a natural generalisation of the RoH. By applying the RoH in a series of incremental stages a simple, unified and accurate treatment is both possible and practicable.

Numerical integration appears to offer distinct advantages over the alternatives considered. It appears to be general, so that no specific analytical forms of demand function need to be assumed. It is highly consistent with the standard treatment of existing modes, as it requires only explicit trip and cost

matrices for the mode of interest, and the accuracy of numerical integration is controllable by varying the number of points.

For large cost changes (>33%) in matrix-based appraisal, we recommend that numerical integration be used with 1-3 additional points. It appears that with 1 point the NI error is about 1/4 that of the RoH, whilst with 3 points it is 1/16 that of the RoH.

For new modes, we recommend that numerical integration be used with 3-5 additional points. This should be sufficient to obtain results accurate to +/-10-20%, a substantial improvement over the other techniques tested.

We recommend that with regard to new generators and attractors, the RoH and NI (as above) be applied as usual, with the land use strategy in place in the do-something if it forms part of an integrated regional strategy. Whatever the land use benefits are, it appears that the transport benefits can be estimated separately using TUBA/NI methods. Similarly, even when the price of the own mode affects the demand on another mode, it is legitimate to add the benefits of one mode to another.

Some refinement of the rules for application may be desirable in the light of experience. Nevertheless, application of these methods should help to ensure that we do not overestimate the benefits of new modes, including LRTs, and conversely that we do not overestimate the disbenefits from large cost increases - eg. due to road user charging - in the future.

NOTE

1. The Marshallian consumer surplus measure is itself a second choice after the 'ideal' measure of benefits, which is the compensating variation (CV) (see Jones 1977, Chapter 9; or Glaister 1981, Chapters 2&4). However, provided that the income effects of small changes in the transport system are zero, the consumer surplus can be used to represent CV without introducing any inaccuracy.

REFERENCES

Department of the Environment, Transport and the Regions (DETR) (2000), *Guidance on the Methodology for Multi-Modal Studies*, DETR, London.

Department of the Environment, Transport and the Regions (DETR) (2001), *Transport Economics Note*, DETR, London.

Hotelling H (1938), The general welfare in relation to problems of taxation and to railway utility rates, *Econometrica* 6(3), p242-269.

Hyman (2001), The Applicability of the Rule of a Half to the Consumer Surplus of a Set of Related Products, ITEA Division, DTLR.

Jones IS (1977), *Urban Transport Appraisal*, Macmillan, London.

Martinez FJ and Araya CA (2000), Transport and land-use benefits under location externalities, *Environment and Planning A* 32(9), p1611-1624.

Mott MacDonald (2001), *TUBA User Manual (Version 1.2)*, Mott MacDonald, Winchester.

MVA, Oscar Faber TPA and ITS Leeds (1994), *Common Appraisal Framework for Urban Transport Projects*. Final Report. Department of Transport and Birmingham City Council.

Nellthorp J and Mackie,PJ (2001), *Alternatives to the Rule of a Half in Matrix Based Appraisal*. Final Report prepared for the Department of the Environment, Transport and the Regions, with contributions by G Hyman and JJ Bates. Institute for Transport Studies, University of Leeds.

Oscar Faber (1996), *Metrolink Monitoring Study*, Volume Two: Demand Modelling Pre and Post Implementation, Independent Report for the Department of Transport and Greater Manchester PTE, GMPTe, Manchester.

Reaume (1974), Cost-benefit techniques and consumer surplus: a clarificatory analysis, *Public Finance* 29(2).

Sugden, R. and Williams, A. (1978), *The Principles of Practical Cost-Benefit Analysis*, Oxford University Press.

White C, Gordon A and Gray P (2001), Economic Appraisal of Multi-Modal Transport Investments: the Development of TUBA. *Proceedings of the European Transport Conference*.

AUTHORS' ACKNOWLEDGEMENTS

The authors of this Annex have benefitted from comments and suggestions by John Bates. They would also like to thank Peter Mackie for his contributions and Robert Sugden and Dave Milne for helpful discussions. John Nellthorp also acknowledges DTLR's funding of this work. The conclusions are, however, the authors' own.