The Wider Economic Impacts of Transport Infrastructure

Peter Abelson

Department of Economics, University of Sydney

Abstract

The standard economic appraisal of investment in transport infrastructure includes savings in travel times, vehicle operating costs and accident costs and some environmental impacts. There has long been concern that this appraisal framework does not allow adequately for various other (wider) impacts of transport infrastructure on the economy or on the environment. In this paper I examine four categories of potential wider economic impacts: the value of increased output in imperfectly competitive markets, changes in labour supply, agglomeration economies and economic benefits associated with residential development. The paper concludes that these wider economic impacts are generally minor at most. First, although the standard economic appraisal underestimates slightly the unit value of increased output when markets are imperfectly competitive, it also tends to overestimate the quantity of increased output from business travel time savings, so there may be no net underestimation. An exception may occur when new infrastructure opens up a major new region or market. Second, positive welfare effects of changes in labour supply are possible but likely to be minor. Third, the paper finds few benefits for agglomeration economies and, indeed, there may be costs associated with dispersal of activities. Finally mass transit, but not roads, may create some benefits associated with increased residential density.

Key words: Economic appraisal, agglomeration economies

JEL: L91, L92. L98

1 Introduction

The standard economic appraisal of investment in transport infrastructure includes savings in travel times, vehicle operating costs and accident costs and some environmental costs or benefits. There has long been concern, and recently increased concern generated mainly by work in the United Kingdom (UK), that this appraisal framework does not allow adequately for various other (wider) impacts of transport infrastructure on the economy or on the environment. In this paper I examine the potential wider *economic* impacts.[1] Specifically the paper analyses the four main classes of these impacts: the value of output changes due to imperfectly competitive markets, changes in labour supply, agglomeration economies, and possible benefits of scale in residential development.

Section 2 provides a brief history of the debate on wider economic impacts (WEI's) with a focus on recent literature. Section 3 outlines the standard economic appraisal model. This is necessary because a key claim for the WEI's is that they are under-estimated, if not entirely ignored, in the standard appraisal model. Sections 4 to 7 discuss the four main kinds of WEI respectively. Section 8 summarises the main findings.

2 Wider Economic Impacts: History, Claims and Current Situation

There has been a long history of discussions about the economic benefits of transport infrastructure. For example Adler (1987) discusses the role of transport investment in economic development in developing countries. He concluded that when investment in transport infrastructure leads to major new developments or large (non-marginal) increases in output, the net value of this additional output (gross value less all inputs) is the most accurate measure of the value of increased output.

In the late 1980s several studies regressed GDP against levels of infrastructure investment and estimated a high return (up to 60%) to infrastructure investment (Aschauer, 1989, for example). However, most reviews, such as World Bank (1994), discounted these results for three main reasons. First, a common factor may explain output and investment.

[1] In about 2005, various agencies and writers started to use the term "wider economic benefits" to describe the wider economic impacts discussed in this paper (see for example UK Department for Transport, 2005). However, realising that some impacts could be negative the UK Department for Transport (2009) reclassified these wider economic benefits as simply the "wider impacts" of transport. This is rather misleading as many people would expect these impacts to include non-economic as well as economic impacts. Accordingly I use the term "wider economic impacts" to denote the economic impacts of transport investment as distinct from environmental or social impacts.

Second, economic growth may drive investment rather than the reverse. Third, the macroeconomic correlations do not have a clear microeconomic basis.

In the 1990s various reviews of evaluation methods were undertaken in the UK. These examined, inter alia, whether transport improvements could lead to significant regeneration of run-down areas that would not be accounted for in the standard appraisal model. However, this is unlikely where extensive transport networks exist and investment in transport infrastructure would produce only *marginal* reductions in transport costs and this concept was not taken up in the major paper on WEIs by DfT (2005).

In 1999, in a major report to the Standing Advisory Committee on Truck Road Assessment (SACTRA) in the UK, Venables and Gasiorek (1999) argued that transport infrastructure could have other economic benefits that are not captured in regular appraisals. The two main ones were undervaluation of business travel time in imperfectly competitive markets and economies of scale in production when a *substantial* change in transport costs allows firms to reorganize their production methods. They also foreshadowed potential benefits from agglomeration economies. However in their reports to SACTRA neither NERA (1999) nor Ecotec (1999) found any evidence of the impact of transport infrastructure on economic output.

In 2005, the UK Department for Transport (DfT, 2005) detailed how to estimate four wider economic benefits: agglomeration economies, increased competition due to increased transport, increased output in imperfectly competitive markets, and welfare benefits from improved labour supply. It dismissed the second of these impacts as likely to be negligible but found the others likely to be significant, especially agglomeration economies.

This paper drew heavily on research by Graham (2005) that found transport effectively increased urban densities and that this increased output that was not accounted for in standard economic appraisals. Importantly, Graham did not argue that transport changes increase *effective* density, not necessarily actual densities. The **effective employment density** of an area is the employment in that area *plus* the employment in adjacent areas weighted as a function of the generalized cost of transport to the subject area. Thus, the effective density of an area can increase with lower transport costs without any relocation of employment.

The study of the London cross city heavy rail project (Eddington, 2006) provided estimated empirical support for high wider economic benefits which added 57% to the standard measures of benefits (see Table 1), partly because it claimed that the new rail link would reduce effective employment separations across London. This study estimated that, after adjusting for differences in skill levels, wages in central London were 30% higher in inner London than in outer London. The results of this study have been widely cited and various consultants have adopted them as guides to potential wider economic benefits in other, often quite different, environments.

Table 1 Estimated GDP and welfare effects of proposed London Crossrail

Benefits	Welfare £m	GDP £m
Business time savings	4,847	4,847
Commuting time savings	4,152	
Leisure time savings	3,833	
Total time savings (standard method)	12,832	
Increase in labour force participation		872
People working longer		0
Move to more productive jobs		10,772
Agglomeration benefits	3,094	3,094
Increased competition	0	0
Imperfect competition allowance	485	485
Tax benefits of increased GDP	3,580	
Total addition to standard benefits	7,159	
Total benefits (excluding social, environmental)	19,991	20,069

Source: Eddington (2006).

Note the distinction between output and welfare benefits in Table 1. Increased labour force participation increases output, but workers need compensation for loss of leisure. The major welfare benefits are (i) extra output of the existing workforce due to agglomeration effects and (ii) tax benefits due to increases in labour force participation or to workers changing to more productive jobs.

In a follow up paper, DfT (2007a) again emphasized the significant benefits of agglomeration economies, increased output in imperfectly competitive markets and improved labour supply. However, DfT (2007b) summarized proceedings of a seminar on the subject at which various views were expressed. Some participants argued that WEI's should be included in economic appraisals with urgency and that not to do so would risk poor investment decisions. Others argued that WEI's add little, or nothing, to conventional appraisal benefits and that more work was needed to clarify the benefits before they were included. Dft concluded that "the theoretical framework and evidence base would require continued work, building upon emerging academic research and new data sources". DfT (2007c) reached a similar open verdict on WEIs saying that the "evidence base for …wider economic benefits of transport is still relatively new" and called for more research.

In 2009, DFT published a further paper on how to estimate the benefits of agglomeration economies, increased output in imperfectly competitive markets and improved labour supply. However, in describing how to conduct cost-benefit appraisals DFT (2010a) says that these appraisals should include the standard travel time, transport operating costs, accident costs and some environmental impacts such as noise and greenhouse gases. They should "**exclude** any wider economic impacts in the analysis of Monetised Costs and Benefits. However wider economic impacts can be included in the Appraisal Summary Table".

3 Standard Measures of the Benefits of Transport Infrastructure

In the standard appraisal model, the benefits of transport infrastructure include benefits to users of the network and to suppliers of transport services (see DfT, 2010a; Australian Transport Council, 2006). The principal estimated benefits to transport users are savings in travel time and vehicle operating costs and improved safety. Improvements in service frequency, reliability and comfort may also be valued. When there are network effects, transport user benefits are valued across the network. Thus they allow for changes in congestion levels to third parties.

The benefits to transport operators may include revenue gains, savings in infrastructure maintenance and operating costs, and sometimes savings in capital expenses. Environmental benefits and costs, notably noise and carbon emission costs, are also included in many evaluations, but are not relevant to the discussion here where the focus is on economic benefits.

The basic method of estimating user benefits can be described as follows. The perceived price of a trip is the sum of travel time and marginal vehicle operating costs (or fares). This price includes tax on fuel or profit to a transport operator. The generalized *social* cost is the real marginal economic cost of the trip. It includes travel time and marginal vehicle operating costs (whether by private vehicle or public transport) but nets out taxes or charges (such as road tolls) that do not reflect the use of resources.

Figure 1 shows the perceived price and the generalised social cost for a given trip and mode before and after an improvement occurs. There are Q_1 existing trips and Q_2 trips after the improvement. The post-improvement trips include trips diverted from other routes or modes and trips generated by the fall in perceived price.

The benefits to existing trips are the savings in generalised social cost. This equals the shaded are given by $Q_1 \times (GC_1 - GC_2)$.

Trip makers who divert to a new destination, route or mode are assumed to be willing to pay a (perceived) price between P_1 and P_2. If the demand curve is linear, diverted trip makers would be willing to pay an average price of 0.5 $(P_1 + P_2)$. The benefits should therefore be estimated as $0.5(Q_2 - Q_1) \times (P_1 - P_2) + (Q_2 - Q_1)(P_2 - GC_2)$. In practice, these benefits are often estimated as $0.5(Q_2 - Q_1) \times (GC_1 - GC_2)$.

Benefits of new (generated) trips are calculated *in the same way* as benefits of diverted trips. The logic is as before. Some new trips would be generated on the improved route when the cost falls just below P_1 but other trips would be generated only when the cost falls close to P_2.

Significantly, this evaluation model captures the benefits of increased economic output when firms produce with constant returns to scale. This can be illustrated as follows. Suppose that a firm sells 1000 widgets at a price of $100 and has the following cost structure (inclusive of transport costs) per widget:

Labour	$ 50
Capital plant and equipment	$ 10
Materials	$ 20
Transport costs	$ 20
Total cost per widget	$100

Now if transport costs fall to $10 per widget, the firm makes a profit of $10 per widget which is the amount allowed for in the evaluation of existing transport of goods.

I

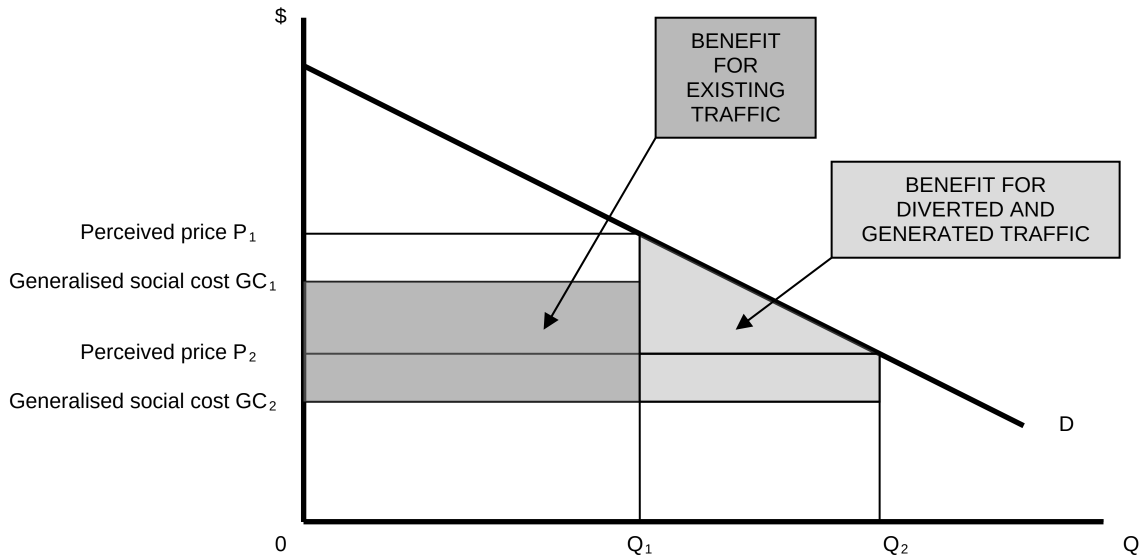

Figure 1 Benefits of existing, diverted and generate traffic

Source: Australian Traffic Council, 2006, Vol. 3, p.67.

In addition, firms that previously could produce and transport widgets at between $100 and $110 per widget can now do so at between $90 and $100 per widget and make an average profit of $5.0 per widget sold. Thus some firms may expand output and others may relocate into this market. Given constant production costs other than transport, the rule of half the savings in transport costs is a correct measure of the benefit of increased output. However, as discussed in Section 4, the benefits are under-estimated if there are economies of scale in production.

Similarly, suppose that a transport improvement reduces the generalized cost of a return trip across a city from $50 to $30 and that someone visits their friend on the other side of the city four times a year instead of twice a year. The standard transport evaluation would ascribe a net value of $10 per generated trip. The person's behaviour indicates that the gross value of each extra trip is worth less than $50 but at least $30. Taking an average gross value of $40 per trip, the net benefit of each extra trip is $10.

In this appraisal model, business and leisure trips (including commuting trips) are treated separately. The value of travel time on business is usually assumed to equal the wage of the trip maker. This may vary with the occupation of the business traveler. DfT (2010b) recommends a mark-up of 24.1% be added for related overhead costs. This valuation assumes a competitive economy in which a person's wage and related overhead reflects the value of their marginal output.

Savings in travel time in leisure hours are valued at the amounts that individuals are willing to pay for them. Based on numerous stated and revealed preference studies, savings in leisure time while travelling are generally valued at about 35% of the wage rate. However, the amounts vary with the disutility of travel, including mode variations and waiting time and uncertainty (Hensher 2001). DfT (2010b) recommends that time waiting for public transport should be valued at two and a half time the normal value of travel time. Unlike business time, an average value for savings in leisure time is usually adopted regardless of the income of the person travelling.

It would follow the value of business time savings should rise with changes in marginal productivity but values of leisure time should not change with increases in incomes. In practice many studies adopt constant values for all time savings. However, DfT (2010b) recommends that business and leisure times should increase with incomes over time.

Of course, not all studies achieve best principles. Significantly, few studies estimate land use changes resulting from improvements in transport infrastructure. Generated trips are a function of changes in transport costs for a fixed set of land uses, but not of changes in land uses. However, this is a weakness in transport/land use modeling rather than in the principles of the standard evaluation model.

4 Output Benefits in Imperfectly Competitive Markets

The standard evaluation model may undervalue output benefits in imperfectly competitive markets in two ways. One arises because prices of goods may exceed marginal costs. The other occurs when production methods change and economies of scale are achieved. We examine both cases.

4.1 Prices in excess of marginal costs

When prices exceed marginal costs, the value of extra goods produced exceeds their marginal cost. Accordingly, wages underestimate the value of output produced by business time savings.

This is illustrated in Figure 2. This shows a demand curve and a marginal revenue curve. Marginal revenue is less than price because firms reduce prices on existing goods in order to increase sales. Firms set output where marginal revenue equals marginal cost and set prices according to what the market will bear. The figure also shows two marginal cost schedules: without the transport improvement (MC_1) and with it (MC_2) respectively.

In this model, lower transport costs reduce the marginal cost of output. The price falls from P_1 to P_2 and quantity increases from Q_1 to Q_2. In the standard evaluation model, the benefits of the transport improvement equal areas $A + E$. Area A reflects the reduction in cost of existing output, measured by wage savings. Area E represents the estimated benefits of increased output as a function of the fall in marginal costs.

However, when price exceeds marginal cost, the benefit of increased output is the difference between what consumers are willing to pay for the goods and the marginal cost of production. In imperfectly competitive markets this benefit equals areas $B + C + D + E$. Accordingly the standard economic evaluation, which values business travel time based on competitive wage rates, underestimates these benefits.

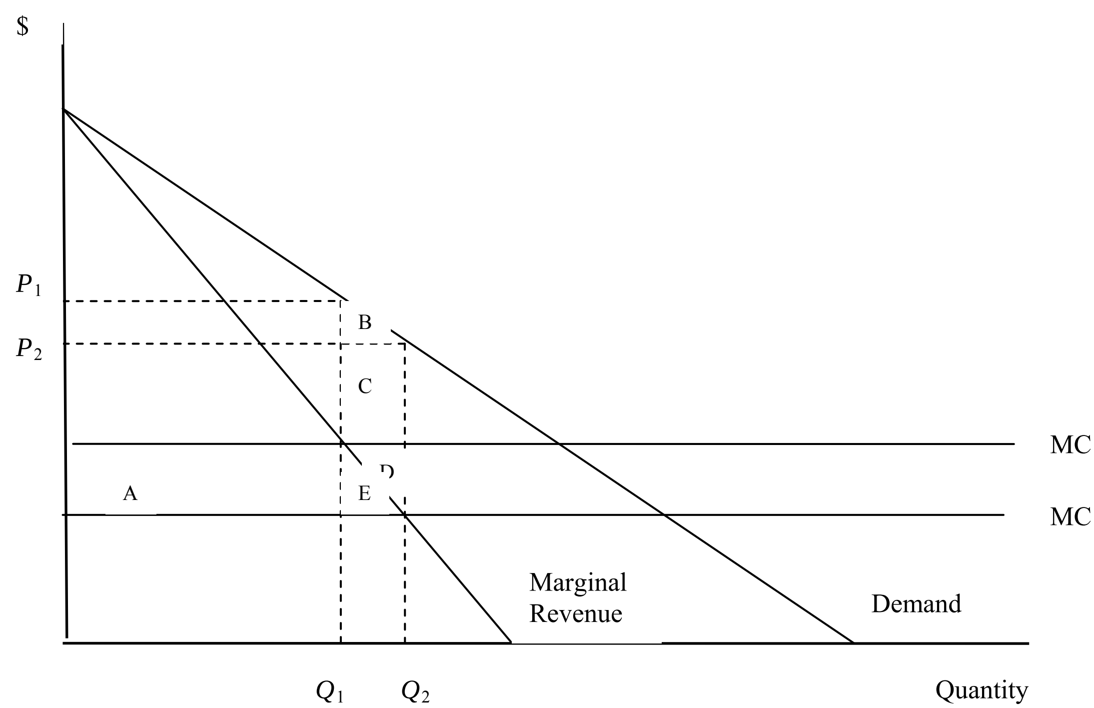

Figure 2 Benefit of increased output in an imperfect market

Turning to estimation issues, following DfT (2005) the benefit of increased output B(ΔQ) can be estimated as:

$$B(\Delta Q) = \Delta Q \times (P_2 - MC_2) \quad (1)$$

where ΔQ is the increase in output. This does not seem quite accurate. Assuming a linear demand curve and constant marginal cost, it would be more accurate to write:

$$B(\Delta Q) = \Delta Q \times [0.5(P_1 + P_2) - MC_2] \quad (2)$$

However, the key point is that the benefit will be a function of the increase in demand for goods and the price mark-up. As DfT (2006) shows, Equation (1) can be transformed to:

$$B(\Delta Q) = BTS \times ED \times (P - MC / P) \quad (3)$$

where

$(P - MC) / P$ = the average price mark-up

BTS = business time savings on existing trips, and

ED = the price elasticity of demand for goods in the imperfect market.

DfT (2005) reviews considerable literature and concludes that the price mark-up in the UK varies from 0% to 35% and averages about 20% and that the aggregate price elasticity of demand for goods (*ED*) is about 0.5. Accordingly, DfT (2005) and more recently DfT (2010) recommend that estimated business times savings for existing trips should be uprated by 10% to allow for the real benefits of business time savings in imperfectly competitive markets. This allows for areas $(C + D + E)$.

Uprating the estimated *unit* value of business time savings by 10% to allow for the differential between marginal revenue product and wages seems a fair order-of-magnitude estimate when markets are imperfectly competitive. However the standard evaluation model overestimates the *amount* of output gained. First, it is clear that some, possibly much, travel time on business is used productively (Hensher, forthcoming). Second, some savings in business time may result in extra leisure rather than extra output. Taking all these factors together, the wage rate plus overhead may be a fair measure of the value of output gained even when markets are not perfectly competitive.

4.2 Economies of scale

Economies of scale may occur as a result of improved transport infrastructure in various ways. Multi-plant firms may reduce the number of plants they operate. Or firms may exploit scale economies via access to a larger market. Increasing returns to scale mean lower average costs, which creates a further source of welfare gain.

The following is an example of economies of scale after a transport improvement and the benefits this creates. Suppose three firms are located at three markets *A*, *B* and *C* and that each firm produces 1000 units at a cost of $100 per unit. Currently freight costs between *A* and *B* and between *B* and *C* are $20 per unit and there are no trips between the three independent markets. Suppose now that after transport improvements freight cost falls to $10 per unit. The firm in *B* can achieve economies of scale by selling to markets *A* and *C*. With the larger market, *B* increases output to 3000 units and reduces costs to $80 + $10 per unit for goods sold in *A* or *C*. The firms in *A* and *C* shut down.

Base case: *A* ------ ($20) --------- *B* ----- ($20) -------- *C*

New case: *A* ------ ($10) --------- *B* ----- ($10) -------- *C*

Using standard analysis, the benefits of generated trips would be: 2000 × (20 – 10) 0.5 = $10,000.

However, the total cost of producing and transporting 3000 units has fallen by more than this. In the base case, the total cost is 3000 × $100 = $300,000 with no transport costs. After the transport improvement, the total cost is (3000 × $80) + (2000 × $10) = $260,000. The net benefit = $40,000.

In summary, when changes in production methods produce economies of scale, the standard evaluation of user benefits underestimates the benefit. The true benefit is the reduction in the sum of production and transport costs. However, because the gains from economies of scale are particular to each situation, these economies cannot be incorporated simply into a standard appraisal. They must be justified and estimated on a case by case basis.

5 Economic Benefits of Improved Labour Supply

The wider impacts literature challenges the assumption in the standard evaluation model that savings in commuting time do not affect labour supply. DfT (2005) identifies three potential sources of increased output from reductions in commuting times:

- Increased participation in the workforce
- Working longer hours in current occupations
- Moving to a more productive higher paid job.

The welfare effect is significantly less than the output effect because extra work means foregoing valued leisure time. In a competitive economy, a person's after-tax wage compensates for loss of leisure. Thus switching from leisure to work provides only a marginal net benefit to those who switch. However the increase in output produces additional tax revenue, which is a social benefit.

5.1 The value of increased participation in the workforce

The percentage change in the value of output in any zone i (ΔO_i) is the product of the percentage change in employment and the average output per new worker:[2]

$$\Delta O_i = (\Delta E/E)_i \times W_i \tag{4}$$

and the change in employment due to increased workforce participation is given by:

$$\Delta E/E = -\Delta T/W \times e_L \tag{5}$$

where

$\Delta E/E$ = the percentage change in employment

T = the average generalised cost of commuting

W = relevant gross wage (see below)

e_L = the elasticity of the labour supply.

The percentage change in employment equals the percentage change in effective wages due to the fall in transport costs ($\Delta T/W$) multiplied by the labour supply elasticity. DfT

[2] This explanation broadly follows DfT (2006) with slightly changed symbols.

(2010) suggests that the appropriate marginal wage could be the median wage × 0.69 to reflect the lower productivity of marginal workers. This equation would be estimated for each set of employment and resident zones.

DfT (2005) recommends a low labour supply elasticity of 0.1 based on the responsiveness of labour supply to various financial incentives, not to commuting costs. To give an example, suppose the average daily wage is $250, transport savings between two areas are $10 per day, and labour supply elasticity is 0.1. The real gross wage after transport costs would rise by 4 per cent. With a supply elasticity of 0.1, the increase in the real wage after transport costs would increase employment by 0.4 per cent.

However as noted, the welfare benefit after allowing for compensating loss of leisure is essentially the taxable component of this extra income. This depends on the tax regime but typically would be about 30% of income generated (see DfT, 2010b). Thus the welfare benefit would be about 0.12% of current employment earnings. For perspective, suppose that base employment is 100,000 workers and the marginal wage is $40,000 per annum, the welfare benefit would be $4.8 million per annum.

5.2 Working longer hours

Turning to labour hours of the existing workforce, DfT (2005) suggests that the impact is likely to be minimal. The evidence, such as it is, indicates that travel time savings leads partly to longer commutes. Households can purchase larger houses at lower prices with little or no increase in travel time. If the longer commuting trips are adequately modeled, the standard appraisal picks up this benefit as consumer surpluses of generated trip km. If faster commutes encourage workers on occasions to increase their working times, the assumption that all commute time is valued as leisure time does undervalue these savings.

5.3 Moving to a more productive higher paid job

The concept here is that savings in commuting time encourage some workers to move to more productive and higher paid employment. To calculate the increase in output, estimates are needed of the number of people who will move and the wage differentials in the two areas. This can be expressed as

$$\Delta O_{\mathrm{i}} = \Delta E_{ij} \times \Delta W_{ij} \quad (6)$$

where

ΔE_{ij} = the change in employment in area i in industry j

ΔW_{ij} = the increase in wage in area i in industry j compared with the initial area.

As before, the main benefit is the taxable component of the extra income. In so far as workers benefit, the standard appraisal estimates the consumer surplus associated with the longer commute.

However DfT (2005) acknowledges that forecasting such employment moves is under-researched and provides little guidance to how this can be done. DfT (2009) advises that the moves to more productive jobs should be calculated only when a land use and transport interaction model is used to forecast employment and residential relocation consequences of the appraised scheme. It advises that "*in the central case the value of extra output should be assumed to be zero*" and that any such estimates should be included only in sensitivity tests.

5.4 Conclusions

The assumption that transport schemes do not affect labour supply is a strong one and may undervalue savings in commuting time. However, the impacts on labour supply seem likely to be marginal and the benefits small. A marginal increase in participation may occur but estimated welfare benefits are small compared with typical capital requirements. There is little evidence that faster commuting times encourage commuters to work longer hours or to move to higher paid jobs.

Moreover, few appraisals can be expected to estimate detailed changes in labour supply as suggested by DfT (2005). A practical option would be to adopt some plausible assumptions. For example, suppose that 80% of commuting time savings should be valued at 35% of the average gross wage and that the balance of 20% of these savings is valued at 50% of this wage to allow for tax benefits and some marginal net benefit to the workforce from additional work. Savings in commuting time would be valued at 38% of the average wage:

$$(0.80 \times \text{O}35W) + (0.20 \times \text{O}50W) = 0.38W \qquad (7)$$

This would represent 8.6% uplift in the value of commuting time savings. While crude, this would be a plausible approximate adjustment for the value of commuting time savings in many circumstances. However, such calculations should be justified by appropriate arguments in each case.

6 Agglomeration Benefits

It has long been observed that incomes are significantly higher in large cities than in smaller ones (O'Sullivan, 2003; Rosenthal and Strange, 2004). This is generally explained by positive externalities from concentration of economic activity. Firms are thought to derive productive advantages from access to suppliers (reducing the price of inputs), access to labour (increasing labour productivity) and access to information (improving technology). Thus, when a firm relocates to an area of agglomeration it may raise the output of *other* firms through one or other of these channels.

In their major survey of agglomeration economies, Rosenthal and Strange (2004) observed that nearly all studies of agglomeration economies are based on *comparisons of metropolitan areas*. They recognize that localization effects (due to industry concentration) may operate for areas within cities but note that this is an under-researched area and cite few studies of this. Nor do they cite any work on effective density, perhaps because such studies were not then available.

The new argument (Graham, 2005, 2006) is that agglomeration economics is a function of "effective density". Effective density allows for interactions with the number of workers in neighbouring areas. Because transport improvements reduce travel cost, they increase effective density even when with no change in location of employment. Graham showed that firm output in 8000 areas (wards) in the UK is a positive function of effective density.

The concept of agglomeration benefits from transport improvements is based on four propositions:

- Output per worker and hence wages are a function of effective density.
- Effective density rises with transport improvements.
- There are positive externalities from transport improvements which increase output for some firms independently of their use of the transport network.
- This increase in output is not included in the standard evaluation of transport.

In the following section I describe how these benefits are estimated. We then discuss various issues with respect to claimed agglomeration benefits.

6.1 Estimating agglomeration benefits

Following DfT (2005) agglomeration benefits (*AB*) for any area can be calculated as:

$$AB_i = \Delta O = O_i \times \Delta ED \times e_P \tag{8}$$

where

O = local economic output

ΔED = percentage change in effective employment density of the area,

e_P = the elasticity of total productivity with respect to effective employment density,

i subscript refers to each area.

For example, if local output equals $10 million, there is a 5% change in effective employment density and the elasticity is 0.05, the gain in output would be $10m × 0.05 × 0.05 = $0.25 million.

The *effective employment density* of an area is employment in that area *plus* employment in adjacent areas weighted as a function of the generalized cost of transport to the subject area. Thus effective density can increase with reductions in transport costs *without any relocation of employment*. Indeed, effective density may increase when actual employment density falls. Note also that the definition of an area is arbitrary.

DfT (2005) recommends that effective employment density should be calculated for each area, however the area is determined, and for each year using the general formula:

$$ED_{t,j} = \Sigma_k \, E_{t,k} T^a_{t,j,k} \tag{9}$$

where

E_k = work-place based employment

$T_{j,k}$ = generalised cost of travel between areas j and k

a = a distant decay parameter

t signifies year.

Estimates of effective density require estimates of generalized transport cost between areas. This may vary by peak and non-peak time of day. Moreover, there is no theoretical basis for the distance decay weighting embodied in the a parameter. Graham (2006) assumed a value of 1.0. Graham et al. (2009) estimated 1.0 for manufacturing, 1.6 for manufacturing and 1.8 for consumer and business service sectors. The higher the value of a, the more rapidly agglomeration effects fall with distance. Presumably the value of *a* is also sensitive to the size of the areas.

Summing agglomeration benefits over all areas and industries,

$$AB = \Sigma_{ij} [O_{i,j} \times E_{i,j} \times (EIP_{i,j} \times \Delta ED_j / ED_j)] \quad (10)$$

where

$O_{i,j}$ = output per worker in industry i and area j,

$E_{i,j}$ = work-placed employment in industry i and area j,

EIP_{ij} = elasticity of productivity with respect to effective density (as measured by ED_{ij}) for industry *i* in area j,

EDj = the effective density of employment in area j, and

ΔED = the change in effective density due to transport project.

Equation (10) assumes that GDP and employment data are available for the relevant areas and that effective densities can be calculated. However, there are few data on the elasticity of productivity with respect to effective density. DfT (2005) recommends a range of values for EIP_{ij} from 0.04 to 0.11 depending on the industry.

Ideally, employment relocations should be modeled but this is rarely done. Also a transport improvement can create a *net reduction in agglomeration* when they service to currently low employment areas. These and other issues are discussed below.

6.2 Transport may disperse activity rather than concentrate it

Most transport improvements disperse both economic activity and resident populations. Over the last 50 years there has been a steady decline in the proportion of employment in central business districts in many cities because of changes in transport infrastructure, especially roads.

It is easy to construct scenarios when transport improvements reduce employment densities. Figure 3 shows three scenarios (*A*), (*B*) and (*C*) before and after transport improvements. In each case the transport improvement encourages firms marked by 'X's to disperse and density of employment falls. To cite another example, most circular roads around a city increase densities on the urban fringe but reduce CBD and intra-city densities.

DfT (2005) recommends that the decentralising effects of transport schemes should be allowed for. Likewise, Graham (2007) notes that transport policies may *reduce* densities and that there may be net agglomeration costs. However, nearly all transport studies are based on a fixed land use – trip matrix. With fixed land uses, effective density always rises with transport improvements. There is no allowance for activity dispersal.

6.3 *Employment concentration and wages may be determined jointly by other factor(s)*

The agglomeration literature assumes that output rises in larger urban areas because of the higher employment. Formally, wages are a function of employment (*N*):

$$W = f(N) \tag{11}$$

However, it is possible (a) that both wages and employment are determined by other factors and (b) that employment is determined by wages because workers move to higher paid areas.

For example, both *W* and *N* may be explained by differences in comparative advantage. Suppose that city (*A*) has a major port or other natural resources and is the seat of a state government and that city *(B)* has minor natural resources. We would expect *A* to have substantially more employment and to pay higher wages than *B*.

Of course, if labour is indifferent between living in the two areas, marginal wage rates would be similar in the two areas. Labour would migrate from *B* to *A* so long as wages were higher in *A*. The fact that wages are higher in *A* means that labour requires compensation to live in *A* because of the costs, be they house prices or transport costs. Differential wages must achieve individual equilibrium. This equilibrium situation is shown in Figure 4. The demand for labour is higher in *A* but so is the labour supply schedule. Accordingly, employment and wages are higher in *A*.

While the level of employment may affect wages via agglomeration factors, the basic drivers of *N* and *W* are the comparative advantages of the two areas which drive the demand for labour and labour supply, which depends inter alia on the environment, house prices and commuting costs. Agglomeration may be only a minor determinant of wage differentials

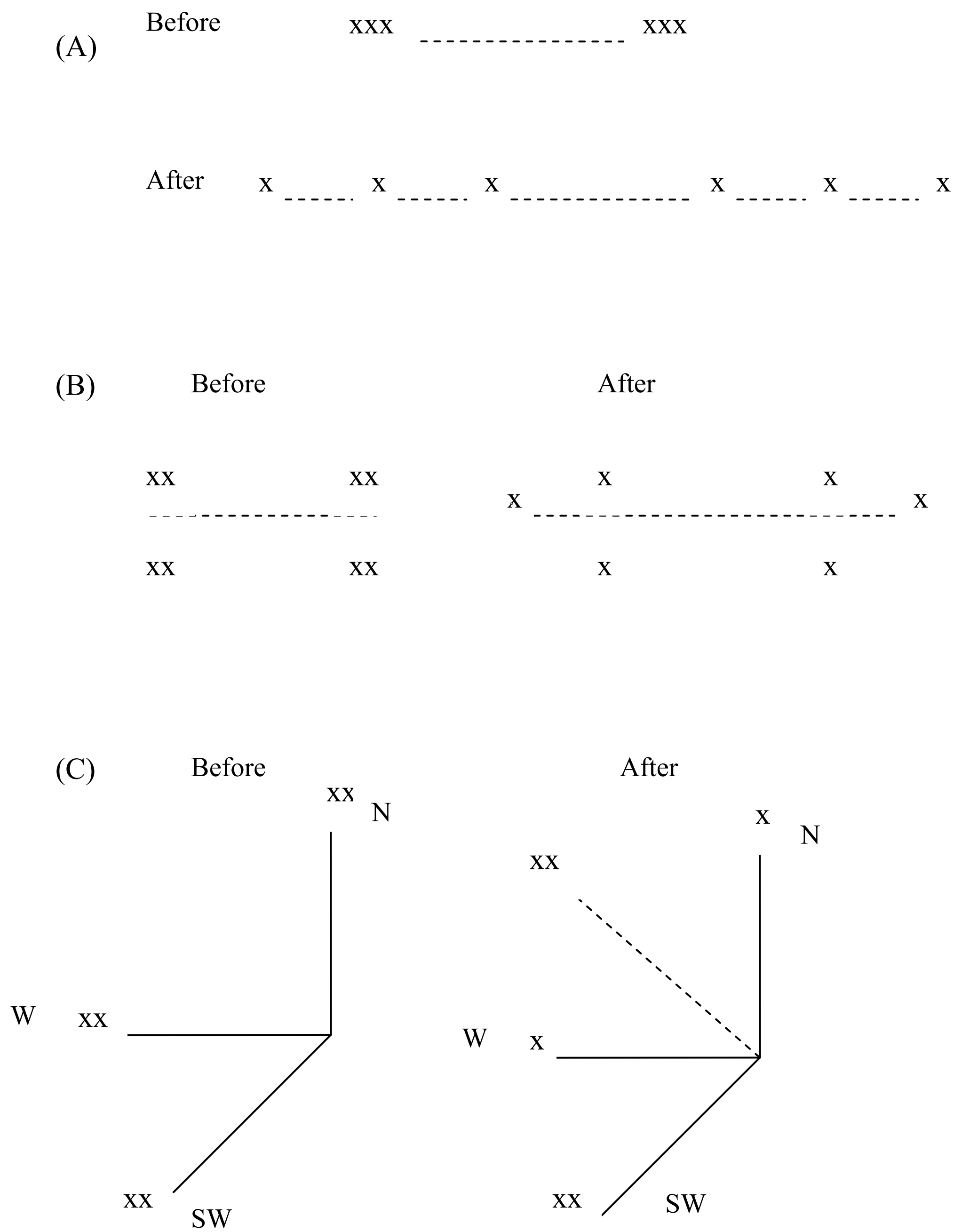

Figure 3 **Three scenarios of firm dispersion (the 'x's represent firms)**

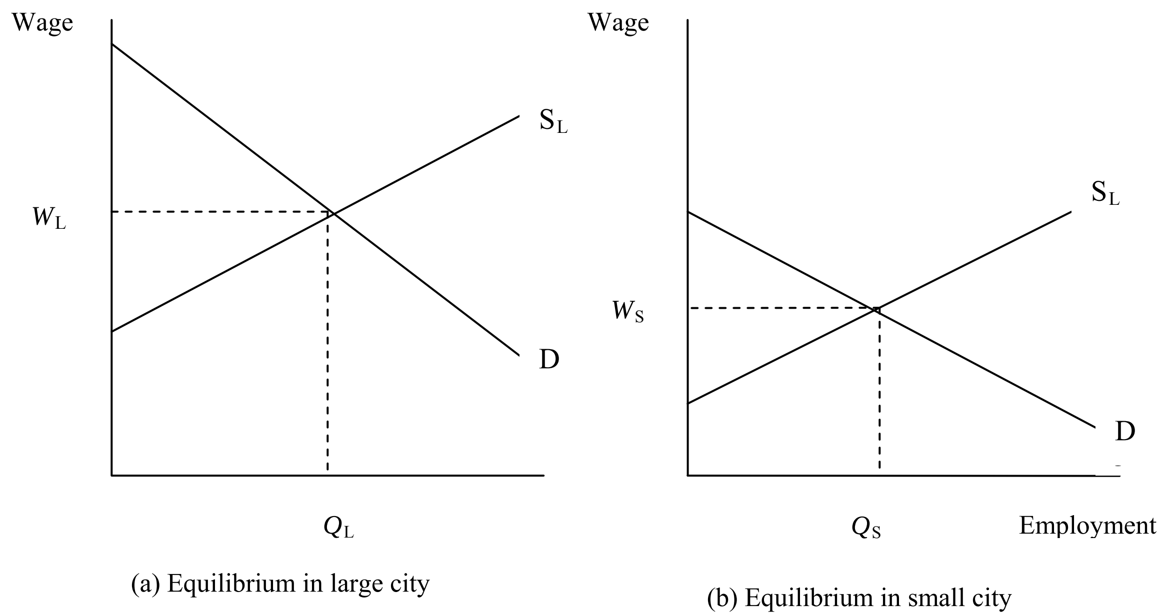

Figure 4 Labour wage and employment equilibrium in large and small cities

6.4 The proposition that higher effective density by itself creates significant external output benefits is not plausible

If employment is constant, it must be shown that a change in effective density raises the average productivity of workers *already in the area*. Proponents of agglomeration economies must argue that the benefits of either existing or new trips are underestimated. It is hard to see why the output of existing workers in an area should increase because of travel time savings on *existing* trips (other than as a direct result of their time savings).

Thus, it must be argued that the standard appraisal underestimates the benefits associated with *new trips for a given workforce*. These trips must generate external output due to one or more of the three drivers of agglomeration economies (improved access to suppliers, labour or information). However, several studies (Goodwin, 1996; Abelson and Hensher, 2001) have found that generated trips are usually small in relation to existing trips. Almost certainly generated business trips are an even smaller proportion of existing business trips. This is especially true for rail. Within cities, few freight or business trips are made by rail. It is implausible that a small number of new trips (*with no employment relocation*) will generate substantial increase in output for the non-trip firms. Given that the direct benefits of generated trips are included in standard transport evaluations, any *additional* output effects seem likely to be small and exceptional.

6.5 The empirical basis for agglomeration economies driven by effective densities is thin

Notwithstanding a priori views about the output benefits of increased effective density, empirical studies could provide evidence for such benefits. However the empirical basis for agglomeration benefits to-date is thin.

Graham (2005, 2006) produced the pioneering work on agglomeration economies related to transport. Further work is emerging slowly. Longworth (2008) found an insignificant relation between income and density in Sydney and sometimes a negative one. Mare and Graham (2009) estimated agglomeration economies for New Zealand. Their results are sensitive to econometric specification but they conclude that the elasticity ranges from 0.032 for agriculture, forestry and fishing to 0.087 for finance and insurance. After allowing for heterogeneous production methods within an industry, Graham et al (2009) found an average agglomeration elasticity of 0.04 in the UK, with elasticities varying from 0.02 for manufacturing and consumer services to 0.08 for business services. These elasticities are significantly lower than those reported in Graham (2006).[3]

Graham (2006) estimated the relationship between output per worker and effective density based on firms within designated industries in 8000 areas (wards) in the UK with an average of 80,000 persons per ward. Multi-plant firms were excluded. In this study firm turnover is a function of labour and capital inputs and effective density estimated over the whole country. The formula included an assumed value of -1 for the decay factor (α).

As Graham (2006) recognized, such work has to deal with various technical issues.

- The concept of an area is arbitrary and has no theoretical basis. Agglomeration economies were based originally on city size. There is little research on the effects of employment densities within cities and no agreement on the size of areas for agglomeration analysis.
- There is no firm basis for the value of the distance decay parameter (α). It might be expected that there would be non-linear (diminishing) returns with distance with a zero

[3] Many consultants have drawn on DfT (2005) to estimate agglomeration benefits in prospective transport schemes, but these are not additional independent evidence for agglomeration economies.

effect beyond a certain point. O'Sullivan (2003) quotes very high decay factors for firms over two kilometers from the basic cluster.

- In many industries, firms are heterogeneous. Thus, density effects may be measuring other factors such as internal economies of scale. Mare and Graham (2009) acknowledge that their "attempts to control of enterprise heterogeneity … were beset by attenuation bias and lack of precision".
- Graham (*ibid*) used turnover as measure of output variable. But turnover reflects prices as well as output. Part of the estimated impact on turnover may reflect higher prices in denser areas.
- Graham (*ibid*) measured capital inputs of firms in dollars rather than in physical units. These monetary estimates are hard to make consistently and accurately and include price effects.

These points suggest that extrapolation of Graham's results should be done with considerable care (as he would doubtless agree).

In conclusion, as DfT (2009) recognizes, transport improvements often disperse employment rather than concentrate it. If the impact on employment distribution is neutral, it needs to be shown that improved travel times are likely to increase the output of firms beyond that associated directly with savings in travel time. The theoretical case for this seems slim. And, for the reasons just given, the empirical evidence also appears slim.

In addition, if agglomeration economies do exist in some cases, estimating them is hard in individual projects. There is a risk of bias against small investment schemes or traffic management schemes where no agglomeration benefits are claimed. But collectively, presumably, such benefits might be claimed.

Finally, if there are agglomeration benefits, it would almost certainly be more efficient to achieve these through planning regulations and incentives promoting employment density. This would be a more direct and cost-effective way to achieve agglomeration benefits.

7 Economies in increased residential densities

An investment in transport infrastructure may increase residential densities without increasing employment densities. This is perhaps most likely with a mass transit scheme. Improved roads are more likely to disperse residences.

DfT (2010c) distinguishes between a residential development that occurs independently of transport investment and development that is 'unlocked' by it. In the former case, the development is a separate issue. In the latter case the residential development should be considered jointly with the transport investment. Suppose a mass transit scheme doubles the transport capacity of an urban corridor and that the population of the corridor doubles from 50,000 to 100,000 households. The 50,000 new dwellings may be supplied in this corridor rather than elsewhere, such as the urban fringe, or some may be additional housing. What benefit would be attributable to the mass transit scheme?

One approach would be to ask potential users of the mass transit scheme the maximum amount that they would be willing to pay for it and to ignore the housing market. This would be appropriate if the market price of the new housing exactly reflected the full public and private costs of creating the housing. In this case there would be no net surplus or deficit from the new housing. However, this assumes no price discontinuities in planning decisions, all planning decisions are efficient (equating marginal social benefit and cost), and all public infrastructure costs associated with alternative housing development are passed on fully to households. These are generally unrealistic assumptions.

If house prices do not reflect real housing costs around the mass transit scheme or elsewhere, the net benefit of housing around the mass transit scheme is the difference between the surplus from developing housing around the transport scheme and the surplus achievable elsewhere. Thus the net social benefit (NSB) of housing attributable to the transport scheme would be:

$$NSB = (P_{\mathrm{H}} - PRIVC - PUBC)_{\mathrm{TC}} - (P_{\mathrm{H}} - PRIVC - PUBC)_{\mathrm{UF}} \qquad (12)$$

where P_{H} denotes the market prices for housing (the willingness-to-pay price), *PRIVC* and *PUBC* denote private and public costs respectively of housing development, subscript *TC* denotes transport corridor and *UF* denotes urban fringe. The two expressions on the right hand side represent the net benefits of housing development in the transport corridor and the urban fringe respectively. If the housing in the corridor is additional housing the second expression on the right hand side falls out. The net benefit could well exceed the amount that rail users would be willing to pay for use of the transit system.

However the price of housing around the mass transit system will be inversely related to expected fares. If expected fares are high, house prices will be lower than otherwise. To account for this relationship and avoid double counting benefits, the expected joint benefit of the mass transit system and the new housing should be the sum of expected mass transit revenues and net housing benefits. Any consumer surplus from use of the rail would be absorbed into the price of housing close to the mass transit system.

DfT (2010c) recognizes such housing benefits but proposes that they be included in qualitative terms in the appraisal summary and not in the analysis of monetised costs and benefits. This appears to reflect difficulties in quantification in estimating the quantity of housing developed due to the transport improvement or in valuing the increase in housing. It is a matter of judgment whether to include or exclude some estimated monetised impact depending on the robustness of the estimate. Where the estimate of housing benefits appears robust, there seems no strong reason to exclude it. However, these benefits would be the exception rather than the rule.

8 Conclusions

Searching for wider economic benefits is something of a holy grail in transport economics. In this paper we find that the standard appraisal model may underestimate some benefits that the net effects are small.

First, although the standard economic appraisal underestimates slightly the unit value of increased output when markets are imperfectly competitive, it also tends to overestimate the quantity of increased output from business travel time savings, so there may be no net underestimation. An exception may occur when new infrastructure opens up a major new region or market.

Second, savings in commuting time may have a small positive effect on labour supply, but evidence is scant. Also, foregone leisure needs to be compensated. The paper estimates that an increased labour supply may raise the value of commuting time savings by 10%, but this should be justified on a case by case basis.

Third, a few transport developments may increase residential densities and provide a net economic benefit when housing prices do not reflect housing costs. However, the link between transport and increased residential densities should be demonstrated.

On the other hand, the paper finds that transport improvements generally do not create agglomeration benefits additional to those estimated in the standard appraisal. The main reasons are that transport infrastructure often disperses employment rather than concentrates it; correlations of wage and employment density overstate the density effect on wages; attribution of significant agglomeration economies to a small number of generated trips is not very plausible; and the empirical basis for agglomeration economies driven by effective densities is thin and subject to unresolved technical issues. This paper concludes that agglomeration benefits should generally not be included in an appraisal.

References

Abelson, P. and D. Hensher (2001): 'Induced travel and user benefits: clarifying definitions and measurement for urban road infrastructure', in K.J. Button and D. A. Hensher (eds), *Handbook of Transport Systems and Traffic Control*, Pergamon, London, ch.9.

Adler, H.A., 1987, *Economic Appraisal of Transport Projects*, published for the World Bank by John Hopkins University Press, Washington.

Aschauer, D. A., 1989, 'Is public expenditure productive?', *Journal of Monetary Economics*, 23, 177-200.

Australian Transport Council (2006): *National Guidelines for Transport System Management in Australia*, Volume 3, *Appraisal of Initiatives,* Australian Transport Council, *Canberra*.

Department for Transport, UK, (2005): *Transport, Wider Economic Benefits and Impacts on GDP*, Department for Transport, London.

Department for Transport, UK, (2007a), *The Additionality of Wider Economic Benefits in Transport Appraisal*, Department for Transport, London.

Department for Transport, UK, (2007b), *Wider Economic Benefits in Transport Appraisal – Summary of the DfT July 2007 Seminar*, Department for Transport, London.

Department for Transport, UK, (2007c), *The NATA Refresh: Reviewing the New Approach to Appraisal*, Department for Transport, London.

Department for Transport, UK, (2009), *The Wider Impacts Sub-Objective,* TAG Unit 3.5.14, Department for Transport, London.

Department for Transport, UK, (2010a), *Cost Benefit Analysis*, TAG Unit 3.5.4, Department for Transport, London.

Department for Transport, UK, (2010b), *Values of Time and Operating Costs*, TAG Unit 3.5.6, Department for Transport, London.

Department for Transport, UK, (2010c), *Appraisal in the Context of Housing Development*, TAG Unit 3.16, Department for Transport, London.

Ecotec Research and Consulting Ltd., 1999, *Review of the Methodology for Assessing the Economic Development Impacts of New Highway Infrastructure*, Report to the (UK) Standing Advisory Committee on Trunk Road Assessment.

Eddington, 2006, *The Eddington Transport Study, London* ,TSO, http://www.dft.gov.uk/about/strategy/eddingtonstudy

Goodwin, P.B. (1996): 'Empirical evidence on induced traffic: a review and synthesis', *Transportation*, 23, 83-101.

Graham, D.J. (2005): *Wider Economic Benefits of Transport Improvement: Link between Agglomeration and Productivity*, Stage 1 Report prepared for the Department for Transport, London.

Graham, D.J. (2006): *Wider Economic Benefits of Transport Improvement: Link between Agglomeration and Productivity*, Stage 2 Report prepared for the Department for Transport, London.

Graham, D.J. (2007): 'Agglomeration, Productivity and Transport Investment', *Journal of Transport Economics and Policy*, 41, 317-43.

Graham, D.J., Gibbons, S. and R. Martin (2009): *Transport investment and the distance decay of agglomeration benefits*, report prepared for the UK Department of Transport.

Hensher, D.A. (2001), 'Measurement of the Valuation of Travel Time Savings', *Journal of Transport Economics and Policy*, 35, 71-98.

Hensher, D.A. (forthcoming): 'Valuation of travel time savings', in A. De Palma, R. Lindsey, E. Quinet and R. Vickerman, (eds.), *Handbook in Transport Economics*, Edward Elgar, Aldershot, UK, ch. 19.

Longworth, T. (2008): Agglomeration externalities – a missing piece of the project appraisal puzzle?, 31st *Australasian Transport Research Forum*, 405-427.

Mare, D.C. and D.J. Graham (2009): *Agglomeration Elasticities in New Zealand*, New Zealand Transport Agency Research Report 376.

NERA (National Economic Research Associates), 1999, *A Framework for Assessing Studies of the Impact of Infrastructure Projects on Economic Activity*, Report to the (UK) Standing Advisory Committee on Trunk Road Assessment.

O'Sullivan, A. (2003): *Urban Economics*, 5th ed., McGraw-Hill, New York.

Rosenthal, S. S. and W.C. Strange (2004): 'Evidence on the Nature and Sources of Agglomeration Economies', in J.V. Henderson and J. F. Thiesse (eds.), *Handbook of Regional and Urban Economics*, Vol. 4, Cities and Geography, Elsevier, Amsterdam, ch. 49.

Venables, A. and M. Gasiorek, 1999, *The Welfare Implications of Transport Improvements in the Presence of Market Failure*, Report to the Standing Advisory Committee in Trunk Road Assessment, Department of Environment Transport and Regions, London.

World Bank, 1994, *World Development Report 1994*, Oxford University Press, Oxford.