

# A critical review of the wider economic benefits of transport infrastructure

Peter Abelson\*

*Department of Applied Economics, Macquarie University, Sydney, Australia*

## Abstract

The standard evaluation of transport infrastructure focuses on transport user benefits. Wider economic benefits (WEBs) are claims, often large claims, for additional economic benefits over and above transport user benefits. This paper reviews the four main forms of WEB: agglomeration economies, the value of additional output in imperfectly competitive markets, labour supply effects, and various possible impacts of transport infrastructure on economic development, sometimes described as place making. The paper finds that wider economic benefits in the first three cases are generally likely to be small or non-existent. The impact of transport infrastructure on economic development is highly context development and stronger where there are substantial changes in transport infrastructure. Where a claim for any wider economic benefits is made, it needs to be supported by a transparent explanation of the driver of the benefit and its prospective magnitude.

**Note:** The Institute of Transport and Logistics Studies, Sydney University, published an earlier version titled "The Wider Economic Benefits of Transport: A Review" as Working Paper ITLS-WP-19-10. The current paper is re-structured and contains significant new material.

**Acknowledgements:** I thank Professor Peter Mackie, Dr. Neil Douglas and Dr. Mark Harvey for detailed comments on a previous draft and Professors David Hensher and James Laird for general advice and references. I am fully responsible for the paper as presented.

**Keywords:** Wider Economic Benefits (WEBs) • Transport Infrastructure • Economic Development

## Introduction

Worldwide, investment in transport infrastructure is huge. Ferrari et al (2019, p.241) report that some world regions account for over US \$500bn annual expenditure on transport infrastructure. The OECD reports that 9 countries spend well over 10 billion euros a year on transport infrastructure. China, Japan and the US spend many multiples more. With such vast expenditures, it is especially important that full, informed and transparent evaluations are undertaken before investment decisions are made [1].

There is much international evidence that major transport projects are prone to large cost overruns (Flvbjerg, et al., 2003; Terrill, 2016). This paper reviews the other side: claims of "wider economic benefits" (WEBs). Whereas the standard cost-benefit analysis (CBA) of transport infrastructure focuses on transport user benefits, WEBs are claimed, often substantial, additional economic benefits.

In a standard CBA, following Mackie et al. (2005) and Cedex (2010), the net social benefit (NSB) or overall welfare outcome of investment in transport infrastructure is given by:

To clarify, user benefits are estimated across the whole relevant transport network (somehow determined), not just on a new mode or route and include, in principle at least, estimated benefits for existing and new trips. These benefits are principally savings in travel time and vehicle operating costs. They often include reliability benefits and sometimes travel amenity values. Externalities are principally safety and environmental impacts such as air quality and noise. Some externalities may be negative. Further, in cost-benefit studies, labour is valued at its opportunity cost. Importantly, this means that, if some labour employed in construction or operation of the infrastructure would otherwise be unemployed or under-employed, the evaluation implicitly picks up the economic benefits of additional employment [2].

- Over time, four principal WEBs have been identified and estimated.

\*Address to correspondence: Peter Abelson, Department of Applied Economics, Macquarie University, Sydney, Australia; E-mail: pabelson@appliedeconomics.com.au

- Agglomeration economies: the impacts of increased employment density on productivity
- The value of business travel time savings in imperfectly competitive markets
- Labour market benefits: impacts on labour supply and productivity including tax revenue

Economic development benefits: a range of commercial and/ or residential development benefits not accounted for in transport user benefits or other WEBs. The first three WEBs were identified in the influential seminal paper by the UK Department for Transport and are the most conventional WEBs. Of these three, the largest claims are usually made for agglomeration economies. As described below, the impact of transport infrastructure on economic development has a longer history. Since the 1990s, it has been particularly associated with the New Economic Geography, which seeks to integrate transport infrastructure and spatial activity. A common feature of all WEBs is that they are associated with some form of market failure. But market failures, e.g. costs of externalities, are not necessarily WEBs. Estimated WEBs are common in the UK and Australasia, but apparently less so in other countries. In a recent study of the UK High Speed Rail Phase Two (HS2), the UK Department of Transport, UK estimated that WEBs 1 to 3 would add 24% to the estimated user benefits, with agglomeration benefits making up nearly three-quarters of these benefits. In the London Cross Rail study, the same WEBs added over 50% to the estimated standard set of transport benefits. In some NZ and Australian projects, WEBs including development benefits have added 30% to 50% to transport benefits. However, describe the empirical verifications of WEBs as "debatable". And the UK DfT observed circumspectly that "modelling and valuing wider economic benefits is complex and subject to a high degree of uncertainty".

## Methodology

This paper also advises circumspection. The paper layout is straightforward: a section for each wider economic benefit with longer sections for the most prominent WEBs: possible agglomeration economies and impacts of transport infrastructure on economic development. The final section provides conclusions.

### Agglomeration economies

Agglomeration economies may be dynamic or static. Dynamic economies occur when productivity (output per worker) rises with greater employment within an area. Static economies occur if productivity rises with effective employment density. Effective density is a weighted sum of the employment in a designated area and neighbouring areas, with the latter employment discounted for either distance or trip costs between the areas. Thus, effective density can rise with no increase in actual employment in the relevant area(s). This section comments on some general issues in agglomeration economics, makes brief points on transport infrastructure and dynamic agglomeration economies, and then focuses on the concept of effective density. As Graham and Gibbons note, nearly all applications of agglomeration economies in transport studies are based on effective density. Many studies have found that productivity is related positively with total city employment. In their major survey of agglomeration economies, Rosenthal and Strange found that

doubling city employment is associated with increases in productivity by between 3 and 8 per cent. Using meta-analysis found a productivity elasticity at the lower end of 0.03. The reasoning is based on scale effects: firms derive productive advantages from greater access to suppliers (reducing the price of inputs), labour (increasing labour productivity) and information (improving technology). Thus, when total employment in an area increases, the output of firms may rise through one or other of these channels. It should be noted that most of these findings are based on comparisons of output over metropolitan areas of different sizes. Rosenthal and Strange (ibid) found little research on localisation economies: where productivity varies with employment density in parts of a city. The general approach to estimating agglomeration economies is to estimate output per firm (represented by revenue) within an industry as a function of inputs to the firm (labour, capital and other purchased inputs) and area employment. This may be represented generally.

$$\ln R_{ijn} = \beta_0 + \beta_1 \ln L_i + \beta_2 \ln K_i + \beta_3 \ln O_{pli} + \beta_4 \ln E_{jn} \dots (1)$$

where  $R_{ijn}$  = revenue per firm  $i$  in industry  $j$  in area  $n$ ,  $L_i$  and  $K_i$  are labour and capital inputs employed by firm  $i$ ,  $O_{pli}$  is other purchased inputs, and  $E_{jn}$  is employment in industry  $j$  in area  $n$ . Capital and other inputs are generally presented in dollar units.  $\beta_4$  represents the elasticity impact of increased employment. Sometimes studies adopt wages per worker as the dependent variable, for example. There are several issues regarding such estimates of agglomeration economies. First, the geographical units generally reflect political determinations. A large area may have high total employment but low employment density whereas a small area may have low total employment but high employment density. Economic theory does not tell us which is more important: employment over a large area or density in a small area. Second, valuing output, capital and other inputs in dollar terms (not in physical units) creates a problem where prices of outputs vary. Following the principle of spatial equilibrium, incomes are generally higher in large cities where workers are compensated for the higher costs of commuting and congestion. And following the "urban wage gradient", wages for the same work fall with distance to the CBD. Unless revenues are adjusted for these differentials, estimated productivity differentials are biased. As a micro example, in Sydney, petrol prices are typically 15% lower in the outer suburbs than in inner suburbs. But this surely does not mean that petrol station workers are less productive in less dense employment areas further from the CBD!

Estimating capital inputs is also complicated in absence of ready data sources other than corporate depreciation, which is based on historic costs. And there are often sparse data on intermediate inputs – utility, materials costs etc. Third, productivity has many drivers. As observe, there are "sorting effects": cause and effect must be sorted. Agglomeration economies may reflect natural competitive advantages. Cities often grew up around ports, government centres and high amenity areas. These centres attract population and have a high demand for labour. As Glaeser observed: "Productivity certainly attracts population the basic problem with estimating agglomeration effects on productivity is that population density is not exogenous. People move to places that are more productive." Thus, density and productivity may be simultaneously determined by some third factor. More productive workers may sort into denser (high priced) areas. It cannot be assumed that labour is equally skilled in all centres or that

all jobs even within an industry sector, (such as banking or legal services) are the same across the urban area. It estimated the relationship between mean wages and total employment across urban areas in France with a special attempt to deal with the endogenous quantity and quality of labour. Their main results were as follows:

- The raw elasticity of mean wages to employment was slightly below 0.05.
- Controlling only for the endogenous quantity of labour bias lowers the estimate to about 0.04.
- Controlling only for the endogenous quality of labour bias yields a lower estimate of 0.033.
- Controlling for both sources of bias produces a coefficient of 0.027

Allowing for agglomeration economies to spill over the spatial units of boundaries, their "preferred estimate for the elasticity of wages to local density stands at 0.02". This careful analysis produced productivity results for actual employment density at the lower end of the scale cited above by Rosenthal and Strange (2004) and even lower.

### Dynamic agglomeration, employment density and transport infrastructure

Given a lack of any generalised findings, UK DfT requires that any suggestion of dynamic agglomeration economies must be supported by a clear narrative on employment density and explicit modelling of land use and employment changes as a function of new infrastructure. This seems eminently sensible.

### The concept of effective density

UK DfT introduced the concept of "effective density". Using cross-section analysis and a distance decay factor between employment areas, found that, in the UK, productivity rises with effective density. Proponents of static agglomeration contend that effective density also increases with reductions in generalised trip costs (GTC) which increase interactions between areas and hence productivity. Measured in this way, effective density rises with lower trip costs between areas without any change in actual employment densities.

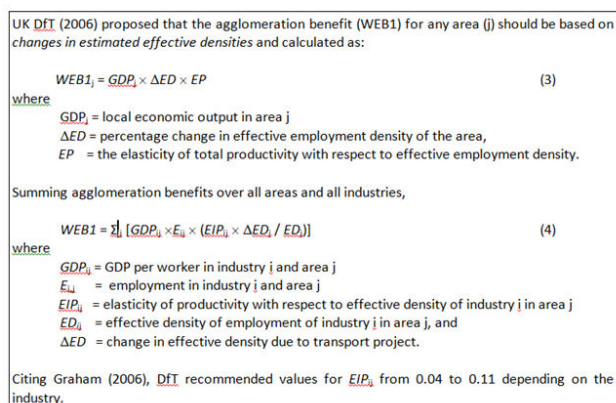
Following UK DfT, effective employment in area  $j$  equals employment in area  $j$  plus employment in adjacent areas ( $k$ ) as a weighted function of generalised trip costs (GTC) between area  $j$  and the other areas.

$$ED_j = E_j + \sum E_k T_{jk} \dots\dots\dots (2)$$

Where,

- $E_j$  = employment in area  $j$
- $E_k$  = employment in neighbouring areas  $k$
- $T_{jk}$  = generalised cost of trips between area  $j$  and areas  $k$
- $a$  = a decay parameter that reflects the lower importance of employment further away.

This shows UK DfT (2006) advice on how to model the benefits of an uplift in effective density (Figure 1).



**Figure1:** Estimating agglomeration benefits based on effective densities.

However, there are many definitions of effective density. Variations include. Employment may be total, or industry, employment in an area. Density figures may be obtained by dividing employment by size of area. Importantly, in most studies of effective density effects (see below), effective density is modelled as a function of distance between areas, which is more readily observable than GTC. There is no theoretical basis for definition of an area or distance decay weighting ( $a$ ). The default parameter value for  $a$  is -1.0, but this may be varied. The higher the value of  $a$ , the more rapidly agglomeration effects fall with distance. Assumed a value of 1.0. found that 1.0 was appropriate for some sectors, but that higher decay factors of 1.6 was more appropriate for manufacturing and 1.8 for consumer and business service sectors. On the other hand, in evaluations of transport projects, estimated changes in effective density are almost always based on changes in generalised trip costs (GTC) between areas. As Graham and Gibbons note "distance is an inappropriate metric for the change in impedance since transport improvements rarely primarily entail reductions in distance". This may be a weighted metric based on the GTC of various travel modes [3]. Where GTC is intended to reflect costs of business-to-business travel, this should reflect the main business modes and times of travel, principally off-peak times. This is rarely done. Critically, is productivity likely to rise with increases in effective density with little or no changes in employment locations? This question sets up three more questions Do lower GTC significantly increase business trips between neighbouring employment areas? Generated business trips are a necessary input to static agglomeration economies, How should these new trips be valued? (iii) Would these extra trips create external agglomeration economies? Most short-distance business trips are made by walking or taxi, usually in off-peak hours; few by bus or train. A large survey of bus and train users in Auckland, Christchurch and Wellington in New Zealand found that company business trips were only 1 per cent of total trips. New transport infrastructure usually reduces door-to-door GTC for such trips only marginally. Studies have found that generated trips are small in relation to existing trips. Thus, lower GTC generally has little impact on short-distance business trips. Second, generated business trips are usually of marginal business importance. Where lower GTC generate new business trips, these are valued in the standard evaluation process by the Rule of a Half as shown in the Annex. Thirdly, a small number of marginal business trips is unlikely to generate significant external agglomeration economies. We conclude

that agglomeration economies associated with generated business trips are likely to be exceptional and small. Despite an extensive general literature on agglomeration economics, there are few empirical studies of these effects. We comment below briefly on four studies of effective density effects: In the UK, Mare and Graham (2009) in New Zealand, the Australian) for the United States [4].

#### Four studies of effective density and productivity

Graham (2006, 2007) estimated the relationship between firm revenue and capital and labour inputs and effective density in designated industries in 8000 areas in the UK. Multi-plant firms and firms with more than 100 employees were excluded. As notes above, effective density was based on an assumed value of -1 for the decay .2 noitauqE ni ) ( rotcaf

Graham reported high agglomeration elasticities, ranging from 0.07 for manufacturing to 0.197 for services, with an average urban elasticity of 0.129. However, after allowing for the heterogeneity of products or services within each industry, estimated revised, lower, agglomeration elasticities ranging from 0.02 for manufacturing and consumer services to 0.08 for business services and averaging 0.04. Also recognised the following issues. The concept of an area is arbitrary and has no theoretical basis. There is no agreement about the size of areas for agglomeration analysis. Little research has been done into the effects of employment densities within cities. There is no firm basis for the distance decay parameter, the value of ( ). In many industries, firms are heterogeneous. Thus, density effects may measure other factors such as internal economies of scale. As noted above, revenue reflects prices. Thus, some of the estimated impact may reflect higher prices in denser employment areas. And there was no allowance for intermediate inputs. Mare and estimated agglomeration elasticities across urban areas in New Zealand drawing on longitudinal microdata on enterprises. The authors regressed gross revenue of firms against labour, capital and intermediates and effective density measures again based on distance. At an aggregate urban level, the study found a high degree of agglomeration with an overall elasticity of 0.17. However, 70% of this was due to observable differences in regional industry composition. And when differences in industry composition were fully controlled for by including fixed effects, the overall elasticity fell by 90% to 0.0015. The authors found that "denser areas are more productive, but this may reflect other factors that are positively associated with both density and productivity. It is more difficult to establish that an increase in density would necessarily lead to an increase in productivity". Critically, they stated that our "attempts to control for enterprise heterogeneity using the 'within enterprise' specification were beset by problems of attenuation bias and lack of precision." Using cross-section analysis, the Australian Bureau of Statistics estimated firm revenue as a function of labour, capital employed (represented by depreciation), intermediate inputs and effective density by industry in the eight Australian capital cities [5]. All values, other than effective density, were estimated in dollar terms. Issues include: the arbitrary size of the areas, the arbitrary distance decay curve, the treatment of firms located in more than one area, the assumption of homogeneous firms within an industry, the high number of zeros in the data base, and scaling the coefficients in the production function to equal 1.0. In effect, profits above a normal return on capital were assumed due to differences in effective density, and not to other factors

such as market power. The agglomeration results were weak with many insignificant and negative agglomeration effects. Of 152 estimates of agglomeration in 8 capital cities over 19 industry sectors, ABS found 42% were positive and significant at the 90% confidence level, 38% were positive but not significant and 29% were negative. The ABS also used fixed effects panel data models with 2006 and 2011 data, but this did not produce significant results estimated the productivity gains from agglomeration for the 50 largest metropolitan areas in the United States based on employment density and accessibility. Productivity represented by wages is modelled as a function of educational attainment, industry specialisation, the cost of living and year specific effects as well as by employment density and accessibility, defined as the number of jobs a representative traveller can reach within 20 or 60 minutes by road travel. The major finding was that productivity gains from urban density were consistent with elasticities between 0.07% and 0.10%. The paper concluded productivity (wages) rises with the number of jobs accessible within 20 minutes of road travel time compared with the number accessible within 60 or more minutes and concluded that this "highlights the importance of investing in efficient transport networks". However, the paper also found that incorporating road network speeds has little effect on the results "which seem to be mainly driven by density effects". Assuming that all productivity drivers are allowed for in the analysis, there does seem to be a density related productivity effect. However, the analysis does not account for public transport and the impact of road travel speeds on productivity is not clear. Conclusions on static agglomeration economies. Most studies of effective density are based on distance factors. There is little evidence that lower GTC without changes in actual employment density have a significant impact on productivity. There appears to be little evidence that improved transport infrastructure generates significant new business-to-business trips. And most such trips would presumably be of marginal business importance and unlikely to generate significant agglomeration economies. As Laird and Venables note "there is a lack of robust ex-post data" to support agglomeration economies. More forcefully, Douglas and O'Keefe observed, static agglomeration "is invisible and largely unprovable". The Australian Transport and Infrastructure Council concluded that "it is bad practice to apply a broad percentage up-lift to the results of the traditional appraisal".

#### Transport infrastructure and business travel time savings

In the standard cost-savings approach (CSA) to valuing business travel time savings, output gained from business travel time savings is valued at the marginal value of the travel time gained. In a competitive market, this is the relevant wage rate plus direct overheads. In imperfectly competitive, premiums are needed to allow for prices set above marginal cost. These price premiums are a WEB. While this basic point about the marginal value of business time saved is accepted, nevertheless, several questions arise. How extensive are imperfect markets and how great are the mark-ups in these markets? But, most critically, do business travellers work while travelling and, if so, what is the net gain in output due to savings in business travel time? And, if not working, does the traveller lose utility by travel time savings that increase his or her work time? UK DfT (2005) found that the price mark-up in the UK varied from 0% in competitive industries to 35% in uncompetitive industries and

averaged about 20%. On the other hand, extra output may reduce market prices. Allowing for this, UK DfT (ibid.) recommended that business time savings should be uprated by an average of 10%. In Australia, KPMG (2017) reached similar conclusions. However, there are two other significant issues in estimating this WEB. The first and most critical issue is the assumption that all time spent in business travel is unproductive. As Chris Foster, a doyen of UK transport economists, observed in relation to the proposed HS2 rail project: "The benefits of faster journey times have been greatly exaggerated by the false assumption that time spent by business people on trains is wasted" (Castles and Parish, 2011, p.iii). This issue is discussed at length. These papers provide substantial evidence that much business travel time is spent productively, especially given the expanding use of digital work instruments. Second, some business travellers (especially self-employed workers) may convert some of the travel time saved into leisure time, or simply enjoy some relaxed time in travel away from an employment environment. It describes how these observations have led to two alternative valuation methods: the employer willingness to pay (WTP) method and the Hensher Equation (HE) method, which seeks to integrate employee preferences with employer values. They conclude (in the Abstract) "that the CSA does not provide a suitable basis for valuing business travel time savings. The HE and WTP approach do not tell the same story in terms of absolute values or variations in values according to key factors, but it is not clear which is the more appropriate. Further detailed exploratory research is needed into how employers value the benefits of employees travel time savings." Allowing for the productive use of business travel time and some gain in leisure values, the average social value of business travel time savings appears likely to be below the relevant wage of business travellers rather than above it as is commonly assumed (e.g. Transport for NSW, 2018). Pending more research, an assumption of no WEB seems more realistic [3].

## Transport infrastructure and labour supply and productivity

The three potential labour supply effects identified in UK DfT (2005) and elsewhere due to lower trip costs are:

- Working longer hours in existing occupations
- Increased participation in the workforce
- Labour moving to a more productive, higher paid, jobs. Working

longer hours represents a marginal behavioural preference, but no change in trip behaviour. On the other hand, increased workforce participation or moves to more productive jobs usually involve new trips. As shown below, in each case the private benefits are included in the standard evaluation approach. But tax revenue needs to be added.

### Working longer hours

When a worker saves commuting time, the standard assumption is that she has a constant working week time and will enjoy a preferred form of leisure to time spent in commuting. Such personal travel time savings are usually estimated in the range of 25% to 50% of prevailing wage rates.

Alternatively, someone may choose to use the commuting time savings by working longer hours, especially part-time workers. Given

a choice between extra leisure and work, the worker is assumed to be indifferent at the margin between leisure and work. If she works longer hours, she gains after-tax wage income but forgoes leisure time. It follows that the value of the travel time saved for the worker is independent of whether she experiences improved leisure or takes on extra work.

However, the increase in output produces additional tax revenue, which is a social benefit, which is a WEB that is not counted in standard transport evaluations. Increased workforce participation may also involve taking on full or part-time work instead of leisure. In standard economic appraisals, the value of additional work is derived from the "rule of a half" the change in trip user costs. Suppose that the GTC falls from \$30 to \$15 per return trip, or over a year from say \$60,000 to \$30,000. Following the rule of a half principle, the average (welfare) benefit per additional work trip would be \$7.50 per day or \$15,000 per annum. As UK DfT observes: "the value to the individual can be no greater than the value of the commuter travel time reduction – otherwise they would not have needed the time saving brought about by the travel time savings to enter the labour market". The annex below provides more explanation of the "rule of a half" valuation principle. Again, a tax benefit accrues to Government due to the additional work, which is a WEB. It is of course also necessary to forecast the extent of increased participation and not simply to make black-box forecasts. In most developed world cities, most workers have several workplace options. Thus, the number of workers entering the workforce because of lower transport costs between their home and their (new) employment location is likely to be small. Observes that there is little evidence of these labour supply effects occurring in the United States except in some rural areas with high unemployment. UK DfT (2018c, p.9) suggests a low labour supply elasticity of 0.1. Thus, suppose the average daily wage after tax is \$250, GTC falls by \$10 per day, and the labour supply elasticity is 0.1. The net wage after transport costs would rise by 0.4 per cent and employment would rise by only 0.04 per cent [2].

## Literature

### Moves to more productive jobs

In principle, lower transport costs may encourage some workers to travel further to higher paying, more productive, jobs. Possibly, also the supply of productive jobs may increase. Of these, the former scenario is more likely. In this case, the valuation principles for worker moves from low paid to higher paid jobs are the same as for entry to work. Unless there are major barriers on access to jobs, the private benefit cannot exceed the savings in GTC and the private benefit is approximated by the "rule of a half". The public benefit is the extra tax revenue. In this case, there is no simple labour supply elasticity to apply to forecast labour supply shifts. The UK DfT (2005) observation that forecasting such employment moves is under-researched appears to be valid today. UK DfT notes that a land use and transport interaction model should be used to forecast employment and residential relocation impacts of the appraised scheme, but it provides no clear model for doing this. However, at least, the evaluation should provide an explicit explanatory narrative and reasoning – not some arbitrary black box assumption [1].

New transport infrastructure may affect labour supply without changes in the locations of businesses which are picked up in WEB4 below. However, the impacts of changes in GTC on labour supply hours require a context and an economic narrative and are likely to be marginal. When labour supply changes, the rule of a half provides a general proxy measure of the private benefits. There would be some additional tax revenue benefit. The latter is an additional economic benefit over standard assumptions for transport evaluation but is generally of minor magnitude.

### Transport infrastructure and economic developments

Many experts have suggested that investment in transport infrastructure may produce development benefits that are not picked up by transport user benefits. The claims include macro and microeconomic arguments. The macroeconomic claim is that investment in transport infrastructure increases aggregate demand and output (GDP) and thus incomes. The microeconomic claims are that commercial and/or residential development may be transport dependent, i.e. they would not occur without the new transport infrastructure). The various possible causes include production due to large non-marginal changes in accessibility and costs, urban regeneration due to increased competition between markets, economies of scale and the exercise of comparative advantages. In this section we examine these macro and microeconomics claims.

### Macro-economic impacts

Aschauer (1989) reported several studies that regressed GDP against levels of infrastructure investment and estimated a return of up to 60% to infrastructure investment. However, a common factor may explain any positive association of output and investment. Alternatively, economic growth may drive investment rather than the reverse. In its major review, the UK Standing Advisory Committee on Trunk Road Appraisal did not support this macroeconomic claim for transport infrastructure. It concluded: "Our studies underline the conclusion that generalisations about the effects of transport on the economy are subject to strong dependence on specific local circumstances and conditions." In a report for the World Bank, Straub (2008) concluded strongly similarly. Re-opened the issue of the macro relationships between transport investment, employment and GDP. However, they found that "the literature does not supply robust answers to many of the key questions". The authors also reviewed how CGE modelling of transport investment tends to produce GDP outcomes greater than net CBA benefits although the latter includes non-market benefits. However, they remarked that it was not clear what assumptions the CGE models made about displacement of other investment or expenditure. In the US, describes how state-based evaluations often include regional development and freight logistics and supply chain connectivity. He also noted that transport improvements can entice more workers into the labour market. However, he found little evidence of these effects, except in rural areas with high unemployment and where labour force participation may rise when employment growth occurs. More substantively, Holmgren and Merkel (2017) report a meta-analysis of 776 estimates of elasticity of production with respect to infrastructure. The estimated elasticity varies from -0.06 to 0.52. They concluded that, the higher the reliability of the estimate, the closer it was to zero. In a recent UK review of transport investment impacts, Melia (2018) found that "none

of the studies reviewed has empirically demonstrated that transport investment boosts national GDP or employment growth" and concludes that "claims made about the national economic benefits of transport investment are not robustly supported by the underlying evidence." In a further survey of this topic, cite several recent studies suggesting an elasticity of 0.1 to 0.2, i.e. a 10 per cent increase in public transport investment would raise GDP by between 1 and 2 per cent. However, they conclude that transport investment has no greater return on GDP than other forms of public investment. Thus, if an impact exists, it may reflect that most investment expenditure has more impact on GDP than does consumption expenditure. In considering possible macroeconomic benefits in addition to benefits estimated in standard evaluation studies, this paper would note two important points. First, the standard transport evaluation approach values labour at its opportunity cost and therefore picks up the resource cost savings of building transport infrastructure, which can be represented as gains in output. Second, it is generally inappropriate to include second-round benefits of investment via a multiplier effect of first round income gains. An alternative investment of equivalent size will create similar (though not necessarily exactly equal) multiplier effects. Consistent with most literature on the subject, this review concludes that there are no special macro-economic benefits of transport investment.

### Dependent commercial development

We now turn to the possible causes of dependent development noted above: non-marginal changes in accessibility, urban regeneration, economies of scale and the exercise of comparative advantage. In all such cases, there needs to be some producer surplus in addition to transport user savings. Sometimes firms relocate to take advantage of transport savings but without changes in production methods or costs. In such cases, any relocated development reflects transport cost savings. As UK DfT notes, in perfectly competitive markets, "user benefits will capture the entire welfare effects of a transport investment". This includes the benefits of induced business developments which are captured by application of the rule of a half to savings in GTC (see Annex below). Thus, we need to examine what special cases may occur.

### Large non-marginal changes in transport costs

It has long been claimed, with good reason, that major transport infrastructure may be necessary to enable resource development in regional areas. It observed that railroads in the nineteenth century and roads in the twentieth century had major impacts on the growth of different areas. Weisbrod (2016) points out that the Erie Canal was fundamental to the development of broad Ohio River Valley. Cite several recent studies indicating substantial commercial / industrial developments and relocations around major transport infrastructure in Europe, China and India and drawing out cause and effect. Where there is no effective access or, in other cases, high cost, unreliable, access, transport infrastructure may be an integral part of the development process. Transport supply may be lumpy and the removal of access barriers may create non-marginal changes in transport costs. In such cases, there may indeed be development benefits (in the form of producer and/or consumer surpluses) over and above transport user savings. The writer spent a year in Thailand in 1971-72 (on a World Bank consultancy to the Thai Department of Highways) trying to identify just such rural development projects.

However, as Ferrari et al. note (ibid, pp 229-240), the main impact of transport investment on economic development occurs in less developed economies. And, where impacts on regional output occur, they may reflect displacement from other regions. Thus, while transport infrastructure may promote development in some regions in developed economies, and create economic surpluses, this needs to be established case by case and not as a general expectation.

### Urban regeneration

In its seminal report on WEBS, UK DfT suggested that improvements in transport infrastructure could stimulate market competition and hence urban regeneration, but the report found few examples of this occurring in the UK. The report concluded that such benefits would rarely be significant and that generally no value should be attached to such possible benefits. Subsequently, few claims have been made for this version of economic impacts who found that experience is mixed, with outcomes varying widely across schemes. However, noted that "regeneration or new cluster development is likely to need coordinated actions beyond the transport investment". This review concurs that urban regeneration due to transport investment is not likely to be significant in countries with well-established transport networks and urban areas.

### Scale economies and comparative advantage

Undoubtedly, transport infrastructure may enable scale economies and/or comparative advantages that increase producer surpluses in addition to transport user savings. Where producers have competitive access to more markets, including international markets, they may achieve significant economies of scale and higher surpluses. As illustrated in the Annex, increasing returns to scale may lower average production costs. Also where, following the well-established principle of comparative advantage, access to markets enables producers to specialise in producing the goods that they are relatively efficient at producing, increased specialisation and trade between two or more communities can raise the output and effective incomes in these communities. It give an example of a retail development arising from economy of scale in a more populated area and show that, in this case, the economic surplus is again greater than under the rule of a half measure. Points out that improved freight logistics and supply chains may lead to creation of new firms. Further to this, transport infrastructure may enable multi-plant firms to achieve economies of scale by operating from fewer sites. These changes may in turn raise productivity and hence have economic benefits [5]. These contentions are supported by several empirical studies. The cites examples, mainly in the United States, where improved freight logistics and supply chains have generated developments mainly in rural areas. For Peru, developed an extensive data set to show contend that ports generate significant economies of scale for firms by facilitating exports. In a major work, contended specifically that Indian railways generated major economies of scale and comparative advantages / specialization that led to economic development. And in major review of the economic effects of transport infrastructure, Ferrari et al. (2019, pp. 181-240) cite numerous cases of positive commercial effects. They conclude (p.235) that "there are few doubts that transport infrastructure matters as far as the level of economic activity, productivity, firm creation and trade is concerned". But they also note (p.236) that it is "hotly debated" whether these are growth effects or "simply a reorganisation of economic activity".

### Concluding comments

When production costs fall with economies of scale or specialisation due to comparative advantage, the benefit for firms is the producer surpluses that arise. These may be larger, sometimes much larger, than the transport user savings.

However, these benefits are context specific. They arise mostly with large, non-marginal changes in transport infrastructure and costs. It is also important to allow for displaced activity and not to double count producer surpluses with transport user benefits. Thus, there is no simple, automatic, way to incorporate these benefits into a standard economic appraisal. Rather, the transport benefits and producer surpluses need to be separately modelled, estimated and justified on a case by case basis.

### Dependent residential development

As NZIER (2013) describes, households may adjust over time to changes in transport infrastructure, often by moving to lower density areas. In principle, the impacts of such marginal household moves can be forecast using spatial interaction models and evaluated by the "Rule of a Half" applying to the new trips (see Annex below). If a household did not move at previous transport user costs, the benefit of the household move cannot exceed the change in transport user costs. However, claims are often made that major transport infrastructure will enable substantial residential development close, or related, to the infrastructure with significant economic benefits. In Sydney, the NSW Department of Transport is projecting major benefits for residential redevelopments aligned with the proposed Sydney City South West Metro. In such cases, as described by NZIER (2013), part of the producer surplus of development may be attributed to the transport infrastructure. As we will see, this surplus usually depends on some form of market failure, notably again economies of scale. To start the discussion, two basic points should be made. First, it is critical to distinguish between residential development that occurs, or can occur, independently of the projected transport investment and development that is 'dependent' on the new infrastructure (UK DfT 2018b). Development is a potential benefit only when it is dependent on the new project infrastructure. Second, residential development may involve broadly constant costs of production or scale economies. Consider first the constant cost scenario. Suppose that transport user costs fall by \$5000 a year and generate capital uplift of \$100,000 per housing unit. Where savings are capitalised in this way, it is double counting to count savings in transport user costs and property owner benefits. However, where there are scale economies (public or private) in residential development, there may be an overall surplus from development of new housing without new owners using the new transport infrastructure (or doing so only occasionally). In this case, their benefits are not picked up as new transport user benefits. Net social benefits from residential development occur when dwelling values exceed all costs of development, including land opportunity and development costs, related public infrastructure costs, residential construction, marketing and finance costs. In this case, the net social benefit from new housing (NSBNH) would be:

$$\text{NSBNH} = \text{CBUP} - \text{CVIRP} - \text{HNPM}(\dots\dots\dots) \quad (5)$$

where MPNH denotes market prices of new housing (the gross benefits) and PRIVC and PUBC denote total private and public costs

respectively of housing development. Importantly, "land opportunity costs" include the loss of any housing or other assets on the land that is developed. They should also include any loss of amenity value of public open space and any negative social welfare effects associated with higher urban densities (Winter and Li, 2017). Also, the timing of the housing developments needs to be forecast, compared with development in a base case or other transport options, and the benefits appropriately time discounted.

In principle, as outlined by UK DfT 2018b, the evaluation can be conducted in terms of land value uplift (LVU) rather than in relation to housing values:

$$LVU = )CVIRP - HNPM( \dots\dots\dots$$

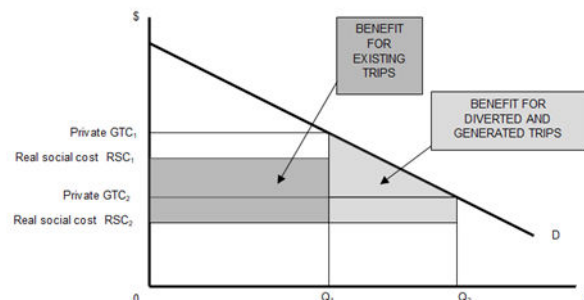
(6)

where MPNH includes the new land values and PRIVC includes initial land values. However, to estimate the change in land values accurately, the analyst needs to estimate both the market prices of new housing and the full private costs including any losses of existing housing assets. Also, public costs must be allowed for. Thus, the house price approach embodied in Equation (5) is more reliable and more transparent than a simple LVU approach and allows for associated public costs. Note also there is no reference here to savings in public infrastructure expenditure elsewhere. This reflects the assumption that any displaced development would be marginal development where the benefits, as reflected in market prices, approximately equalled the sum of private and public costs. In this case, there is no net social gain (or loss) associated with this displaced development. In summary, a WEB may occur when residential developments are dependent on a specific transport infrastructure and would not occur under other transport options. The benefits are also likely to depend on economies of scale in residential production. Absent such scale economies, any increases in house prices would usually reflect use of the new transport infrastructure and be captured in user benefits. Clearly, estimating these benefits is complex. Recognising this complexity, DfT (2016) proposed that any such benefits be included in qualitative terms in the appraisal summary and not in the monetised costs and benefits. This review supports the possibility of dependent housing benefits and the option of estimating them via a full evaluation of alternative land uses and transport options (using Equation 5). But they should be carefully and explicitly described and estimated in a full place-making CBA to support the transport evaluation. This review warns strongly against making simplistic assumptions and add-ons to transport evaluations.

## Annex valuing user benefits and increased output

This annex outlines the basic method for estimating user benefits. The private generalised trip cost (GTC) is the sum of travel time and fares and other out-of-pocket costs, including taxes. The real social cost excludes taxes or charges, such as road tolls, that do not reflect use of resources.

It shows the private GTC and the real social cost (RSC) for a given trip and mode before and after a transport improvement. There are Q1 existing trips and Q2 trips after the improvement. Post-improvement trips include trips diverted from other modes or routes and trips generated by the fall in GTC (Figure 2).



**Figure2:** Benefits of existing, diverted and generated trips.

The benefits to existing trips are the savings in real social cost given by shaded area:  $Q1 \cdot (2CSR - 1CSR)$

Trip makers who divert to a new destination, route or mode are assumed to be willing to pay a price between GTC1 and GTC2. If the demand curve is linear, diverted trip makers would be willing to pay an average price of  $0.5(GTC1 + GTC2)$ . Thus, these benefits are often estimated as  $0.5(Q2 - Q1)(GTC1 - GTC2)$ . This is known as the "Rule of a Half". Where  $GTC2 > RSC2$ , there is an additional benefit =  $(Q2 - Q1)(GTC2 - RSC2)$ . This is a benefit to government and is known as the "resource cost correction", (NZIER 2013, p.7).

The user 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 infrastructure when the cost falls just below GTC1 but other trips would be generated only when the cost falls close to GTC2.

This evaluation model also captures the benefits of increased output when firms produce with constant returns to scale. Suppose that a firm sells 1000 widgets at a price of \$100 and has the following production cost structure, including 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 the existing transport of goods.

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 (assuming no price changes). Thus, some firms may expand output and others may relocate into this market. In such cases, given constant production costs other than transport, the rule of half the savings in transport costs is a realistic measure of the benefit of increased output.

Finally, suppose that there are economies of scale and that, as output increases, other costs fall from \$80 to \$60 per widget. There are then savings (producer surpluses) of \$20 per widget in addition to the direct transport benefits.

## Conclusions

The standard economic appraisal of transport infrastructure includes transport user benefits but may require marginal adjustments for additional economic benefits (WEBs) in a few cases. Claims of large WEBs are generally unjustified. When WEBs are claimed, an economic narrative and explanation is essential rather than applying "assumption laden black-box formulae as has increasingly been the norm". Small agglomeration benefits may occur with actual increases in employment density. However, it needs to be demonstrated that the transport infrastructure will increase employment density. Pending further research, changes in effective density due to lower transport costs are unlikely to have significant productivity effects without changes in actual employment densities. The value of output associated with travel time savings increases with imperfect competition, but this factor is likely to be offset by due allowances for productive work during work trips and for some trip-maker preferences for leisure. Transport improvements may marginally increase labour supply or moves to more productive jobs. These benefits are captured by the rule of a half assessment in a standard evaluation method. There may be small additional benefits from increased tax revenue. Consistent with most of the literature, when transport investment displaces other investment, there are no additional macro-economic benefits. Substantial changes in transport infrastructure may generate producer surpluses in addition to transport user savings where there are high existing transport

barriers or where significant economies of scale occur or comparative advantages are achieved. Transport infrastructure may also generate residential development that would not otherwise occur. But cause and effect need to be shown. Fundamentally, any claims for WEBs should be carefully demonstrated in the context of any proposed new transport infrastructure. It is inappropriate to simply assume that a WEB exists.

## References

1. Chatman, Daniel G., and Robert B. Noland. "Do public transport improvements increase agglomeration economies? A review of literature and an agenda for research." *Transp Rev* 31(2011): 725-742.
2. Donaldson, Dave. "Railroads of the Raj: Estimating the impact of transportation infrastructure." *Am Econ Rev* 108(2018): 899-934.
3. Graham, Daniel J. "Agglomeration, productivity and transport investment." *J Trans Eco Policy (JTEP)* 41(2007): 317-343.
4. Hensher, David A. "Measurement of the valuation of travel time savings." *J Trans Eco Policy (JTEP)* (2001): 71-98.
5. Martincus, Christian Volpe, Jerónimo Carballo, and Ana Cusolito. "Roads, exports and employment: Evidence from a developing country." *J Dev Econ* 125(2017): 21-39.

**How to cite this article:** Abelson P . "A critical review of the wider economic benefits of transport infrastructure." " *Int J Econ Manag* 10 (2021) : 608.