Research in Transportation Economics 23 (2008) 107–115

Contents lists available at ScienceDirect

Research in Transportation Economics

journal homepage: www.elsevier.com/locate/retrec

Transit investment and economic development

Roger Vickerman

Centre for European Regional and Transport Economics, University of Kent, Canterbury, CT2 7NP, UK

ARTICLE INFO

Article history:
Available online 1 March 2009

Keywords:
Investment
Urban land rents
Economic development
Wider economic benefits

ABSTRACT

This paper describes the development of approaches to analysing the links between investment in transit and economic development. It indicates the need to bring together disparate approaches from urban economics and transportation economics to get a full understanding and uses recent results on agglomeration economies to present a more rigorous model of the wider economic benefits of transit investment. Although the evidence increasingly points towards identifiable benefits over and above those captured in conventional transport cost-benefit analysis, the chapter counsels against the assumption of simple rules.

1. Introduction

The role of transport in the urban economy has been the subject of much controversy. In this paper we attempt to sort out a coherent story of the role which transit investment can play drawing on both the classic analyses of this subject and more recent research. Whilst there is no denying that improved transportation will make the wider urban economy function more efficiently, there have been significantly different views on the extent to which there are any wider benefits which cannot be captured simply by those benefits accruing directly to users. Thus one argument would be that, whilst transit investment may have significant impacts on urban form and growth, any investment appraisal exercise can safely ignore these wider effects. On the other hand it is argued that the existence of increasing returns to scale in the activities which use transport will lead to agglomeration effects which are not so captured and hence there is a need to measure wider economic benefits separately. The secondary question is whether such effects are consistent between different urban areas and types of investment, such that a simple multiplier or add-on can account for these wider economic benefits, or that all such effects are case specific.

One of the problems for the development of a consistent analytical approach has been the differences in treatment by urban economists and transport economists arising from their different perspectives and objectives. The urban economist's interest has been in transport as a determinant of land use and urban growth and economic development (e.g. Fujita, 1989). Accessibility has been the key to where within the urban area different activities will locate and hence determines urban form. Lower transport costs imply that the city can grow and hence transport investment can be a determinant of city size. The transport economist's interest is more focused on efficient use of the urban transport infrastructure and the cost which congestion imposes when the infrastructure does not expand with demand (e.g. Small and Verhoef, 2007). Secondly the transport economist has a primary interest in the appraisal of urban transport investments.

Past work by each group has frequently ignored contributions of others leading to a confused view of the interface between the transport system and the urban economy. However there are common elements in their approaches which can provide the basis of a more integrated approach. The key to this is in the evaluation of accessibility. Transport economists have had a long-stranding interest in how transport users value the time savings which arise from transport investment. Urban economists also recognise that the accessibility of a location determines its value, so that as transport improves the implicit rental value of land at any location will rise with consequences for the optimal use of that land. The question is therefore whether transport user benefits are an accurate measure of the wider economic benefits which reflect the use to which the improved transport is put? Or should we use the valuation provided by the changing land values as a better measure of total economic impact. Obviously put this way the two cannot be added together as they will involve double counting. But does either of these measures give a true estimate of any wider economic impacts, except in a world where there is perfect competition in all transport using activities. Once we recognise the existence of increasing returns and imperfect competition, such that prices do not directly reflect marginal costs, changes in transport prices may not pass through smoothly to be reflected in the final prices of activities.

Recent work has improved our understanding of the way in which accessibility affects the performance of firms and, perhaps crucially, of labour markets, enabling us to provide a better account of the way investments in transport will impact on the overall

E-mail address: R.W.Vickerman@kent.ac.uk

0739-8859/$ – see front matter
doi:10.1016/j.retrec.2008.10.007

urban economy (e.g. Graham, 2007; Patacchini & Rice, 2007; Rice & Venables, 2003; Rice, Venables, & Patacchini, 2006). However, the empirical evidence remains problematic for three main reasons. First, work in this area is plagued with questions of endogeneity and causality, essentially whether improved transport leads to improved economic performance or is a consequence of it (see, for example, the discussion in Banister & Berechman, 2000). Secondly, there are conflicting stories between research based on the macro-economic relationship between flows of investment and aggregate output which tend to demonstrate some fairly strong positive links and micro-based estimates which present a more confused picture of the way the improved transport impacts on behaviour (e.g. Vickerman, 2007b). Thirdly, the interrelationships and spillovers between different areas have to be taken into account, the way in which the improvement of transport in one jurisdiction can have both positive (complementary) impacts on adjacent areas and negative (substitution) impacts (e.g. Boarnet, 1998). Individual urban areas cannot be taken in isolation as transport improvements, even to highly localised transport systems, can have profound influences over a very large geographical area.

This will have policy implications. If there are significant negative effects leading to underinvestment in transport infrastructure this could lead to lower growth and congestion whilst over-investment could lead to problems for public budgets and negative externalities associated with over expansion.

In this paper we shall consider these various competing influences in the context of urban transit investments, developing models which allow for increasing returns, imperfect competition and spillovers between areas. The chapter deals in turn with an overview of the links between urban transit and the urban economy, the determinants of land rents and the urban land market, how the urban transport problem relates to this structure before looking in more details at the agglomeration issue. Following the theoretical analysis we shall look for evidence, taken from three broad types of study: macro studies which estimate relationship between macro-economic aggregates; market studies which look for the way transport impacts on individual markets, and particularly the labour market; and micro studies, which focus on examining behaviour within organisations. The chapter concludes with some implications for appraisal methods and for policy.

We face the usual definitional problem of what constitutes urban transit. We have tried to confine the discussion to the consequences of intra-urban travel, or at least travel within a defined urban labour market. However, it is recognised that the growth of mega-cities and the increasing interaction between urban areas makes this a rather fluid concept. Even some international transport has some urban characteristics with the growth of long-distance commuting (weekly migration). To simplify the discussion, however, we will assume in the theoretical sections that we are dealing with a conventional defined and independent urban area and draw primarily on empirical evidence from studies of such areas to illustrate issues.

2. Transport, accessibility and the local economy

Transport has a multiple nature in the urban economy. First and foremost, transport is a derived demand, transport is only required to overcome the spatial separation which is inherent in the urban economy; the level of demand will depend on the level of activity requiring transport. However, transport is also a substitutable input; as transport becomes cheaper firms and households will substitute transport for other (relatively more expensive inputs). Hence both firms and households may move to more peripheral locations (or even right out of a city) to overcome the higher rents of more central locations (and increasingly some of the higher external costs associated with congestion), pollution and crime in such areas. Or conversely they may use improved transport to enable them to enlarge market areas to enhance profitability to pay for the higher land (and labour) costs associated with central locations. Transport can thus become an engine of growth in the city by which a city with better transport can remain competitive as its attempts to grow relative to other cities in the urban system.

Accessibility is a measure of the price of gaining access to markets; for firms this is the market for both outputs and inputs (including labour); for households this is the access to jobs and to the locations for other activities. The potential accessibility of a location depends on an activity function which measures the activities which are to be reached and an impedance function which measures the cost of reaching them. This is conventionally expressed as:

$$A_i = \sum_j W_j^a \exp(-\beta c_{ij}) \tag{1}$$

where A_i is the accessibility of location i, the W_j are the weights associated with the activities available at a range of destinations j, and the c_{ij} are the generalised costs (including all money and time costs) of travel between i and j. Improvements to transport can have three sorts of effect here. Changes on one link will affect the distribution of trips between different modes on that link and also the distribution of all trips between i and the various js. However, the reduction of costs on one link will reduce the overall costs of accessing activities and hence raise overall accessibility, which may increase the aggregate demand for transport from i because it is now more accessible. This third impact is often overlooked in the transport model which tends to assume a fixed volume of trips in the system.

This is perhaps the most useful way of measuring accessibility, following the work of Hansen (1959), since it can be clearly related to the concepts of generalised cost included in most transport models. There are, however, alternative ways of representing accessibility such as the concept of daily accessibility (Tornqvist, 1970) which looks at the population accessible to a location within a certain travel time or the time-space mapping developed by Spiekermann and Wegener (1994) (for a discussion and comparison see Vickerman, Spiekermann, & Wegener, 1999).

Considering an urban area as a whole we need to distinguish two forms of accessibility. External accessibility is that which relates an urban area to other cities and regions. Improving accessibility to other areas can enhance a city's export performance by reducing the costs of accessing markets, but at the same time it lowers the cost of access **from** other cities and hence increase competition for local firms. This same process can apply to labour markets in which improved rail access, for example, to the city can enlarge the relevant labour market for local firms which would have the effect of reducing wage pressure. However, it could also increase demand for resident labour from other cities and hence result in bidding up wages. This is sometimes referred to as the 'two-way road' effect illustrating that improved transportation links operate in two directions (SACTRA, 1999). Which direction has the stronger impact will depend on such factors as the scale and scope economies which existing firms can exploit in each city.

Internal accessibility relates to the performance of the city's own transportation networks and the impact which this has on the overall economic efficiency of the city. As we shall argue in more detail later a better transport system can improve labour supply by reducing the time spent commuting thus enlarging relevant labour market areas and enhancing productivity. The way in which the city's internal transportation network relates to the external network is also important; poor connections between the local transit system and inter-city rail or airports can impact on both mode choice and overall transport costs.

The underlying argument is thus that accessibility is a key determinant of the transport costs faced by all city's activities and this cost of transport has a major potential impact on the economic efficiency of the city. These benefits may accrue just to users, but it is also likely that they may have wider economic effects which we shall need to examine in more detail later. The main role of transport is to provide access to markets and to jobs. If transport is one of the major costs of an urban location, time savings will have a positive value and location will be valued by its accessibility. Hence land rents (the price of location) will be related to the associated transport cost. This lies at the core of the double counting issue: we should either measure cost of access or land rent but not both. However, this assumes that all relevant markets are perfect such that price equals (social) marginal cost? It must be questioned whether spatial markets ever satisfy these conditions?

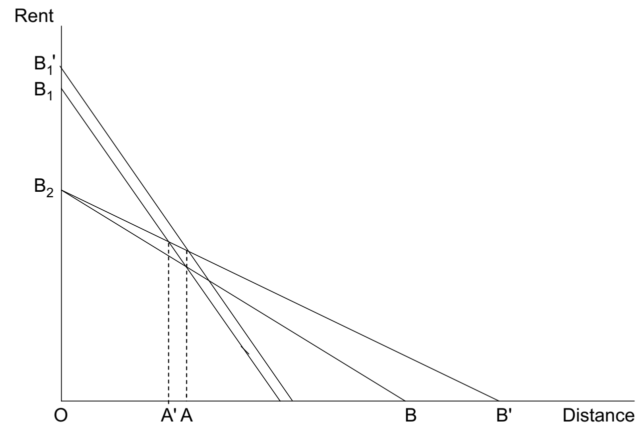

Fig. 2. Transport improvement in the von Thünen model.

3. Transport and land rents

The traditional approach to the relationship between transport improvements and land rents is based on adaptations of von Thünen's theory (1826). Starting with an assumption of perfect markets of fixed size and free entry, von Thünen posed the question as to how land rents will differ for different agricultural products vary within a market area and consequently how land will be allocated between alternative uses.

At the market centre, where transport costs are assumed to be zero, firms will bid land rents up until they equal the profits which can be earned from that activity – i.e. pure economic rent to landowners. At locations away from the market centre the maximum rent which will be bid, the so-called bid-rent, will be reduced by the cost of transport to market centre. This produces an equilibrium location for the individual firm and an equilibrium allocation of land between uses. Thus in Fig. 1 activity 1 has a higher profit at the market centre than activity 2, but also faces higher transport costs, thus the bid-rent curve is steeper. Activity 1 will therefore occupy land up to a distance A from the centre and activity 2 the remainder of the city area. If the allocation of land to activity 1 were too small, the price of activity 1 would rise and hence profits and rents at the market centre and this would result in an increased allocation of land to activity 1.

If the transport costs of activity 2 were now to fall (Fig. 2) this would result in an increased-bid rent at each location away from the market centre. Now activity 2 would secure more land from activity 1 towards the market centre which might result in the price of activity 1 rising as discussed above. The overall consequence of this would be the market area growing in size, hence giving an early example of the way in which more efficient transport would have direct repercussions on markets and settlements.

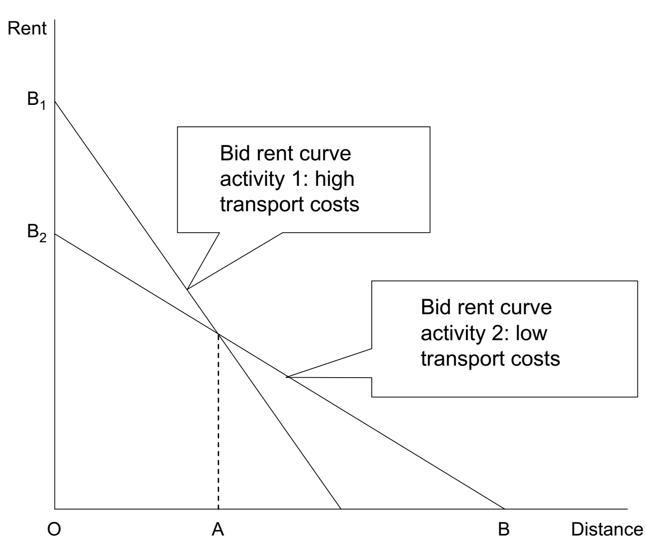

Fig. 1. The von Thünen model.

The von Thünen theory was developed in the context of agricultural markets. The application to industry follows the same basic principles but recognises the need to combine markets for outputs and inputs in different locations. Such an approach (e.g. Weber 1909) shows this as involving a trade-off between the different transport costs of the various activities – the higher the transport cost the closer a firm will locate to that market. Similarly for individuals the largest transport expenditure is likely to be on the work trip and hence the residential location of households will be dominated by access to the workplace.

The fullest translation to the urban context focussed on household access to jobs, but used as a basis the earlier links with land rent. This is best illustrated in the work of Alonso (1964). Assuming a monocentric city with all employment at the CBD, transport costs will be given by distance from CBD. Following the same logic as in the von Thünen model as households move to locations further from the CBD they will incur higher transport costs and thus will be able to spend less on rent. There is however a complication that households may also have preferences over lot size so that there is a trade-off between location (accessibility) and the amount of space they wish to consume. With urban land, unlike the agricultural model of von Thünen, land can also be developed at different densities, this enables more efficient use to be made of land with higher accessibility. Although the aggregate rent of land increases dramatically, the rent for space will not rise so rapidly if density can be increased as well.

The access-space trade-off model developed by Alonso expresses the same concept as the von Thünen model in a simple mathematical format:

Maximise $U = U(x, s, d)$

Subject to $Y = px + r_i s + t_i d$where p, r_i, t_i are prices of goods (x), land (s) and travel measured by distance (d)

This generates the first order condition:

$$\frac{U_d}{U_x} = \frac{1}{p}\left[s \cdot \frac{dr}{dd} + \frac{d(td)}{dd}\right] \tag{2}$$

which tells us that total land costs (rent × space) will decline at the rate at which commuting costs increase. Since s is positive and transport costs must increase with distance, dr/dd must be negative to give the conventional negative marginal rate of substitution.

Graphically the equilibrium of the household is as shown in Fig. 3. The bid-rent curves in this are effectively indifference curves derived from equation (2) which show how different combinations of rent and distance (transport costs to the urban centre) will yield equal satisfaction to the individual or household. Each of these displays a diminishing marginal rate of substitution since a linear bid-rent curve would only generate corner solutions implying only locations at the urban centre or periphery would ever be chosen. Bid-rent curves closer to the origin imply a lower total outlay and hence are preferred. The rent gradient is derived as the envelope of all the household bid-rent curves for a city of a given size. The individual household locates on the rent gradient at a location on the lowest possible bid-rent curve, d_0 in Fig. 3.

The urban land market equilibrium then depends on the allocation of land between all competing uses, recognising that land itself is in (relatively) fixed supply and that as a city grows bigger the amount of land which has to be devoted to transport will also increase in order to transport people to the CBD. The so-called 'New Urban Economics' of the 1970s (see Fujita, 1989, for one of the most complete statements) provides a complete model of land use in the city.

At each location i, for the supply and demand for land for housing to be in balance, the following equation must hold:

$$s_i n_i = 2\pi i h_i \tag{3}$$

where n_i is population at distance i and h_i is the proportion of land allocated to housing at that location. The right hand side of equation (3) measures the amount of land available for housing in an annular ring at distance i (circumference $2\pi i$). The left hand side measures the space required for a population n_i, each occupying a space of s_i. It is assumed that h_i will fall as the centre is approached since more land is required for commuting from the locations at greater distances, it will also fall with city size at any given distance from the centre.

If we aggregate all the demands for land at each distance i from the city centre, assuming that the total supply of land is fixed for a given size of city of maximum radius r and a given total population of N we need to ensure that:

$$N = \int_0^r n_i di \tag{4}$$

i.e. that the allocation to population to each location allows all the population to be accommodated in the city. This will enable us to determine the equilibrium lot size. If the population cannot be accommodated then the urban area may need to grow, i.e. r increases (see Vickerman, 1980 for a fuller development of this model).

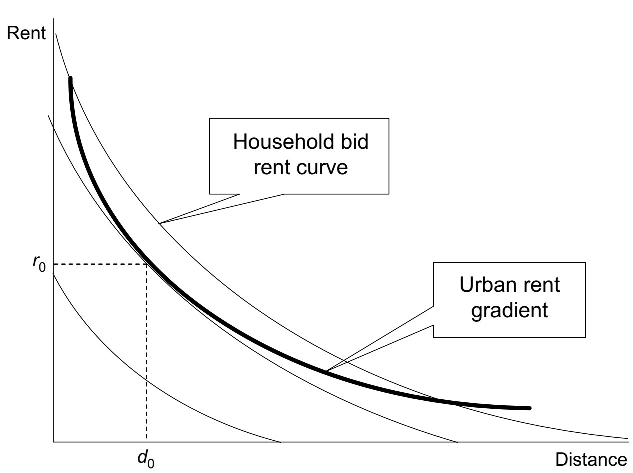

Fig. 3. Equilibrium in the urban land model.

At this equilibrium it is then necessary to determine if the total commuting implied gives rise to congestion? Congestion occurs as the amount of travel to the city centre exceeds the capacity of the land allocated to transport to cope with that amount of commuting. Congestion increases the cost of transport from every location and thus reduces the accessibility of each location and hence the bid-rent at that location to compensate. If the city can invest in more efficient transportation systems (for example economise on land use by installing an underground rail network) this may relax the constraint on growth imposed by the need to allocate more land to transportation. In this case the urban rent and density gradients will show local peaks around the access points to the network (metro stations). If we aggregate the total value of land rents in the city this provides an estimate of the value of the wider benefits of transport; essentially the public good value of transport is measured by the accessibility of all locations as reflected in aggregate land rents.

The urban model allocates land between transport and all other uses. In the monocentric city this provides a limit to city size as flows to the centre exceed capacity. However, the traditional transport model tends to deal with maximising the efficient use of infrastructure for a fixed level of demand. Congestion pricing can ensure the allocation of this capacity between different users with different user values and indicates when capacity is inadequate. Induced demand from expanding capacity (the variable trip matrix) is often ignored in such models (see SACTRA, 1999) and this may be a critical problem.

Bigger cities tend to have more congestion for the reasons outlined above and thus will have increased transport costs. As we have already discussed bigger cities can also invest in more efficient transport systems which will tend to lower marginal transport costs. Bigger cities also have larger markets for the activities located there and this reduces the average costs of producing these activities such that they are able to bear higher transport costs and remain competitive. Indeed these larger markets may well outweigh any extra transport costs such that larger cities are more competitive and can avoid any ultimate constraint on city growth. It is this key element of the role of agglomeration economies to which we now turn.

4. Agglomeration economies and transit investment

The 'New Economic Geography' provides the necessary linkages which enable us to complete the model (see for example Fujita, Krugman, & Venables, 1999; Fujita & Thisse, 2002). Transport costs determine the price of an urban location and hence the **real** wage which can be earned from jobs accessible from that location. This emphasis on the real wage is important as it takes us beyond the simple idea of the value of time savings as a transport benefit and provides the theoretical basis of agglomeration.

The extent to which agglomeration will occur will depend on the interaction of increasing returns in activities, the significance of transport costs and market size. If transport costs are high then there is less scope for agglomeration unless increasing returns are very significant or markets are very large; if transport costs are very low then agglomeration is less necessary. Changes in transport costs between these extremes can lead first to increasing concentration of economic activity and then to a later dispersion. Agglomeration arises because of the extent of linkages in the local economy as these linkages help enlarge the markets not just for final demand activities but for intermediate activities as well. This means that, as activities move towards the larger city costs rise as transport costs within the city will rise because of congestion etc. Land rents will also rise, but so will market size enabling further exploitation of scale economies. Because markets have grown, more activities will be attracted to the city and hence the transport costs embodied in those activities will be reduced as more activities are available locally and do not need to be imported from outside the city.

The critical point to note here is the role of real wages in this process of cumulative causation. Even if nominal wages do not rise, because of the diminishing marginal productivity of labour, since prices of goods will fall real wages may continue to rise. Thus labour will continue to be attracted to the city and contribute to the rising level of demand which will encourage more investment. Similarly, if transport costs fall because the increased prosperity enables further investment in the transit system used for commuting, this enlarges the effective labour market and this will reduce the average unit cost of labour even if wages actually increase.

These changes in labour markets arise for three main reasons: changes in participation rates; increased working hours and moves to more productive jobs. Because workers can now access a wider range of jobs more cheaply (at least in terms of generalised cost) they will be more willing to work. If journey times are reduced workers may be more willing to work longer hours, or if they are not they may nevertheless have an impact through increased expenditure in increased leisure time. Most importantly the enlarged labour market will enable workers to more to more productive jobs within the city, more productive because of the increased market size enabling exploitation of further scale economies. Thus the increased size of the commuting area has impacts on productivity and wage differentials (Venables, 2007).

This can be represented in Fig. 4 (derived from Venables, 2007). Assume the wage differential at the city centre over the periphery is given by B_0; if transport costs are given by C_0 then the limit to the city size (labour market size) is given by L_0. Suppose transport costs now fall to C_1, the labour market will increase in size to L_1' at which point existing city centre workers will gain area α and the new workers will gain area β after incurring extra commuting costs of γ. But now suppose that there are agglomeration effects such that productivity actually rises as the labour market expands for the reasons given above. Now the wage mark up is given not by a fixed amount B_0 but by the wage gap curve W which shows the premium for working in the city increasing with city size, albeit at a decreasing rate. In this case the same reduction in transport costs causes the labour market to expand further, to L_1, with the wage mark up increasing to B_1, and there will be additional benefits to these additional workers of β' (after the additional commuting costs of γ'). However, the extra productivity also raises the mark up to all existing workers such that there are additional benefits of δ. It is this additional amount which is the key to understanding the real total benefits which will arise from transit investment.

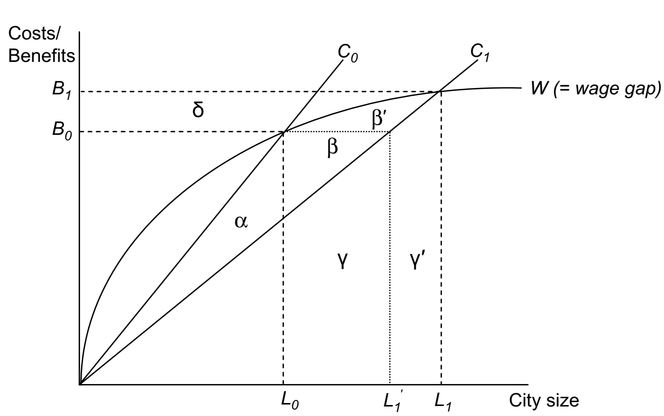

Fig. 4. Agglomeration benefits.

5. Refining the transport model

With this improvement to the urban model we now have a clearer understanding of the complex role which changes in transportation costs may have on the structure and size of the urban economy, operating through scale and agglomeration economies. What, however, of the transport model? It is not sufficient simply to provide a linkage between the traditional transport model and the revised urban model, the transport model itself needs to be reconsidered. The traditional transport model is essentially based on Wardrop's (1952) principles of the least cost assignment of traffic around the network. The generalised cost, which usually includes both direct costs and time costs needs also to allow for choice under charging and the possibility of competition within modes (including roads).

There is a danger that, having overcome the implicit assumption in the transport model that there is perfect competition elsewhere in the economy except in the transport sector, where imperfections arise mainly from externalities, we now forget the increasing significance of the competitive structure in the transport sector. Competition within transport markets has become much more significant with the attempts to privatise, or introduce some form of public–private partnership, in both infrastructure networks and operations.

Competition in transport markets can involve competition between modes and within modes. This requires modification of equation (1) to allow both for multi-modal accessibility and for the existence of more than one operator within a mode. Even in the road mode there is now the prospect of competition with the existence of tolled and untolled roads in a network and the possibility of these being provided by a range of different operators. Within a single network there can also be effective competition between both parallel and serial links in the network. Parallel links are the obvious case of alternative routes between a single origin and destination, but where a journey between these involves travel over the infrastructure or services of different operators there can be serial competition. Here the decision over pricing or service levels by the operator of one of the links will affect the probability of choosing that networks and hence the demand for other links in the network. This is where, for example, agreements on simplified ticketing or interoperability of smart tolling systems can be helpful in reducing the complexity otherwise caused by the fragmentation of networks.

6. Looking for the evidence

We have concentrated thus far on the development of theoretical explanations of the relationship between transport investments and the wider economic impacts they may have in urban areas. We now turn to examine the evidence that exists to substantiate these theoretical insights. Given the link between accessibility and density it is not surprising that much of the early work in this area used data on urban population change to identify the impact of new investments (see for example Davies, 1976). Later work tried to put a more precise economic value on this by looking at the impact on rents. This is only indicative of the impacts, however, and does not capture in full the wider economic effect, only those which are filtered through the commercial and residential land market. We shall review some of the findings of this work later. Finding clear evidence of the wider economic impacts is not straightforward and depends on a number of factors: the geographical scale of the empirical study; the unit of analysis; and the ability to control for other factors which determine urban development.

Broadly speaking we can divide studies into three main types, each of which looks from a different level: those which only look at macro aggregates; those which examine the working of individual markets; those which look in detail at the behavioural responses of individual agents. We look at each of these briefly in turn.

7. Macro studies

Work on the link between levels of investment in transportation infrastructure and the wider economy took off following Aschauer (1989). Although much of the focus of such studies has been on total transportation investment rather than the individual contribution of specific modes and much has been at the national or regional rather than single urban area level they changes the way of thinking about the link. Essentially what was being argued was that public infrastructure could have a positive impact on output, typically through enhancing the productivity of private infrastructure. Whilst later studies showed that Aschauer's initial estimates were much too high because of methodological and econometric problems (see, for example, the review by Gramlich, 1994), and some estimates suggested effects close to zero or even negative, the general consensus was that there was a positive impact. One of the initial objectives was to refute the crowding-out hypothesis that, especially using public money to generate such investment would reduce overall productivity and welfare by reducing private investment. Generally there was seen to be a degree of complementarity between public and private investment, but, crucially, there is no consistency in the relevant elasticities. The impact of investment depends both on the type of investment and the context in which it takes place.

Two sets of conceptual and econometric problems have emerged from this literature. The first and obvious one is that of causality. Does the increased investment lead to enhanced economic performance or does the enhanced performance initiate investment decisions and remove constraints on financing them? Clearly both of these processes are feasible, transportation investment can both be seen to lead economic development and follow it. Indeed transportation investment can often be used simply to boost economic growth through construction without any regard to the longer term consequences. There are also cases where the transportation investment and the accompanying investments are seen as complementary and are planned together as in the opening up of new areas of development where the new transport link is essential to value the development but the development is essential to provide effective demand for the new link. This unlocking argument (see SACTRA, 1999) can be very important in the regeneration of brownfield sites.

The second problem relates to estimation problems arising from spatial autocorrelation and spatial spillovers. A typical way of estimating the effect of infrastructure investment is by using a production function estimated over a cross-section of jurisdictions (or through a Panel approach allowing for both time series and cross-section elements). This causes potential problems of autocorrelation between spatially adjacent areas. More interestingly it also raises the question of the extent to which there are spillovers between areas, and whether these are positive or negative. Work by, for example, Boarnet (1998) using county data in California suggests that both types of effect can be observed where some counties will see a diversion of economic activity as a result of transportation investment in neighbouring counties whilst others may benefit from a positive spillover although the predominant effect was a negative one. Similar mixed results for highway investments were identified by Berechman et al. (2006) and Jiwattanakulpaisarn (2007).

We have to be careful in comparing studies, however since the dependent variable in these econometric studies varies between output, employment and productivity. Arguably the most important variable is productivity here since it is this which relates to the efficiency of the urban system and which identifies the existence of agglomeration effects in which productivity is related to the size of the city.

There are alternatives to aggregate econometric models which try and look also at the way in which the structure of the economy affects the way it responds to new transportation investment. The most used of these are Land Use Transport Interaction (LUTI) Models, but recently there has been increased use of Computable General Equilibrium (CGE) Models. LUTI models allow for multiple markets, each with a specific transport use, which are linked through an input–output structure. Thus any change in transport cost can be fed through the different markets to give an overall impact (Simmonds, 1999). The problem is that the input–output structures used typically have static coefficients such that no allowance is made for the way changing transport provision may change the use of transport (the substitutable inputs issue). Moreover the implicit assumption that transport is a cost means that savings in time will always be valued positively whereas the introduction of congestion related prices will have a negative impact. There is a need to allow for more dynamic behavioural responses.

CGE models tend to suffer from the same reliance on static input–output data but can deal better with assumptions of imperfect competition. Whilst they may be less suitable for evaluating the effect of upgrading individual links they do allow for the evaluation of whole networks with more interaction between markets. Allowing for imperfect competition and these wider effects tends to imply larger total benefits in many cases. CGE modelling has been used mainly in cases of whole network evaluation (e.g. Bröcker, 2000, 2004; Miyagi, 1998, 2001). However, the impacts from such studies are not always unambiguously positive. Elhorst, Oosterhaven, and Rom (2004) found significantly smaller impacts than previous studies on the development of high-speed links in the Netherlands and Oosterhaven and Broersma (2007) identified negative agglomeration effects.

This apparent declining significance of wider benefits from aggregate studies suggest that we might need to explore the way individual markets work at a more local scale to understand whether such wider benefits based on agglomeration effects exist.

8. Market studies

CGE models begin to explore the detail of individual markets which would appear to be important in identifying the way in which agglomeration works. They remain however, constrained by the need to produce overall equilibrium. What are referred to here as market studies are those which explore in greater detail the basis of agglomeration. There are three elements to this: competition effects; agglomeration effects and linkage effects.

Competition effects are the way in which changes in transport provision impact on the degree of competition within different markets. Such effects are ambiguous in their impact. There has been a general assumption that lower transport costs will have pro-competitive effects by reducing the barriers to entry into different spatial markets. The assumption is that where price is equal to marginal cost any changes in transport cost will have to be passed on to consumers because of the competitive pressure. This will be true for firms and for individuals in the labour market where lower commuting costs may lead to wages being squeezed by an increase in supply of labour. However the extent of this will be limited by the existence of imperfect competition and rent seeking behaviour such that where there is already a mark up over marginal costs the reduction in transport costs can be used to reinforce market power.

Agglomeration effects lie at the core of the analysis. Usually such effects are considered to arise in two different ways localisation economies, where there are spillovers between firms in the same sector and urbanisation economies, where there are spillovers between firms in different sectors or where the provision of public infrastructure confers an advantage on all sectors.

Linkage effects are a particular case of these spillovers where the changes in one sector have a direct impact on another sector which lies either backwards or forwards in the production process. The most obvious linkages here are the linkages between changes in a production sector and the labour market. Increased output in a sector benefiting from lower transport costs will have a positive impact on the demand for labour, lower unit labour costs arising from improved supply following transport improvements will impact on the overall costs facing firms employing that labour.

The evidence for these impacts has traditionally been sought in studies which aim to relate productivity to city size. Urban economists have looked at this issue in different ways with similar results to those for the relationship between infrastructure investment and economic performance. The key issue here is whether the evidence points more to urbanisation economies, which are usually measured with respect to total population in the city, or localisation economies, which are measured with respect to industry size (employment). The typical elasticity of productivity with respect to city size is of the order of 0.01–0.20 across a range of studies with a typical value of 0.10 suggesting that there is evidence of modest agglomeration economies (see for example, Sveikauskas, 1975; Moomaw, 1981; Nakamura, 1985; Henderson, 1986). Elasticity with respect to industry size is typically a little larger, with values closer to 0.20 (Nakamura, 1985; Henderson, 1986), although a more recent study by Henderson (2003) has a much lower elasticity of 0.03.

Ciccone and Hall (1996) and Ciccone (2002) provide alternative estimates using measures of employment density which are shown to account for most of the variance in productivity. The elasticity with respect to density is of the order of 0.04–0.06. A more detailed study incorporating more precise spatial measures at a smaller geographical scale by Rice et al. (2006) produces a similar order of elasticity.

Recent work by Graham (2007) for London reaffirms these results for manufacturing industry, but consistently finds elasticities of 0.2 or higher for services. This suggests that for the modern city, highly dependent on skilled labour in financial and business services, there are potentially significant gains which can be generated by changes which enable the labour market to grow. This confirms the expectation of the theoretical model of Fig. 4.

9. Micro studies

Micro studies are defined here as those which look at geographical areas below that of the city or at changes within organisations rather than at the level of the entire market. Micro studies are seen to be important because many of the changes which improvements to transportation will often affect the way that business and households organise their activities. If we do not identify these correctly then we will fail to identify user benefits correctly and will misestimate the total economic impact. This can explain why the same type of investment can have markedly different impacts in different locations. There have, however, been very few ex-post studies of the way in which *n* specific investments have impacted on behaviour and organisation at the firm or household level. The majority of studies have used surrogate markets, such as the housing or commercial property markets, to try and capture these effects, although the majority of tense have tended to assume that it was a much simpler direct link between accessibility and rental values which can be captured as part of the hedonic price of property (Rosen, 1974).

Cervero (1997) surveyed more than forty studies of transit developments in North America which confirmed the expectation that property values would normally increase with proximity to transit. The empirical studies suggest that the typical premium with transit is fairly modest, ten per cent or less. Vessali (1996) estimated an average value of 7 per cent from twenty studies, a figure which is consistent with those of more recent studies such as McMillen and McDonald (2004) and Hess and Almeida (2007). The early study by Dewees (1976) was rather more optimistic. Gatzlaff and Smith (1993) identified insignificant effects for residential properties as did Ryan (2005) for commercial properties. The study by Damm, Lerman, Lerner-Lam, and Young (1980) showed how the property market often adjusts in advance and anticipation of transit developments, to some extent as a precaution against missing subsequent transit induced developments in an area.

The problem with such studies is that property values (rental or capital) are affected by a range of other factors such that it is often difficult to isolate the transit related impact (see, for example, the discussion in Bowes & Ihlanfeldt, 2001). These studies also differ in that some have looked at variations in property values between locations with and without transit access (i.e. between different zones) and others at the variations between locations closer to or further from transit access (i.e. within zones). Changing socio-economic structures, which may themselves be influenced by the transit development, also confuse the impact such that it is not clear how much of this effect can be allocated solely to the transit development (Kahn, 2007). Similarly Duncan (2008) shows that the impact on the price of condominiums was slightly more positive than on single-family homes which he attributes to the pressure towards higher density developments close to stations on new transit routes leading to these benefiting more from changes in accessibility having controlled for other socio-economic factors.

Gibbons and Machin (2005) studied the impact of the Jubilee Line, a new Underground line in London which provided access to the Underground in locations not currently on the network, but giving direct access to locations with new jobs. Their model tried to link the housing market impact more directly with the labour market. This showed an increase in values of an average of 9.3% in areas with new stations and a 1.5% increase in values with every 1 km nearer to a station.

All of these studies imply rational behaviour by households in evaluating changes in accessibility, given the characteristics of the housing and socio-economic characteristics of the individual. What they do not do is explore in detail the way that **household,** as opposed to individual, decisions change in response to changing accessibility to a variety of destinations. This remains a key area for future research.

Studies of the way in businesses respond to new infrastructure through reorganisation of their operation are also rare in the urban context. Research in France of the impact of the TGV high-speed rail network (Plassard & Cointet-Pinell, 1986; Burmeister & Colletis-Wahl, 1996; Klein & Claisse, 1997; SES, 1998) has shown that firms have tended to concentrate towards access to the network rather than between different sized urban areas along the network. This suggests similar responses to the evidence on household movement towards stations. However, there was also evidence of restructuring within organisations which had a presence in cities at both ends of the new link to allow for greater specialisation in function.

10. Implications for appraisal

Moving from the theoretical models and empirical estimates of impacts to a method of appraisal for individual projects poses further problems. The theoretical model give us the basis for a more theoretically correct Cost-Benefit Analysis framework recognising externalities and imperfect competition (Vickerman, 2007a, 2007b). CGE models may be a part of this depending on the scale of projects, but they may not normally be appropriate for estimating the impact of individual links. The relationship between individual links and overall network effect remains an area of some debate (Laird, Nellthorp, & Mackie, 2005).

Wider benefits will include direct user benefits (principally journey time savings), productivity effects, agglomeration effects, competition effects and labour market effects. Data requirements are more demanding and required at a more detailed level than is typical in CGE studies. Table 1 provides an example which is based on the analysis carried out for Crossrail a major new cross-London rail link for which the wider benefits were seen to be critical for justifying the project which eventually was given approval in late 2007.

An essential input to this appraisal was evidence on agglomeration benefits, with estimates taken from work by Graham (2005, 2006). These agglomeration benefits alone add 25 per cent to the direct user benefits. The impacts of increased competition are assumed to be zero in that positive and negative elements will cancel out, but there are positive benefits from the ways that imperfectly competitive firms will enjoy increasing returns to scale. Table 1 also shows that the public finance implications through increased tax revenues dependent on increased employment and higher productivity (and hence profits) were also taken into account and these are a significant potential gain. Since the estimated cost of the project is of the order of £16 billion it can be seen that securing these additional benefits was critical to the appraisal, although this calculation does not include and social or environmental costs and benefits arising directly from the project or indirectly through modal change.

Although the Crossrail case was based on much more rigorous attempts to identify wider economic benefits, such benefits had frequently been used to justify projects previously. Banister and Berechman (2000) report that the wider benefits included as a justification for the earlier Jubilee Line extension in London were around 34% of the direct transport user benefits and this lifted the benefit-cost ratio from less than 1 to a more respectable 1.3.

Even with the improvement in methods of assessing economic impacts and a greater degree of convergence in the values obtained from wider benefit studies, it cannot be assumed that there is a simple rule which can be applied to all such projects. There will continue to be the need for rigorous analysis project by project. However, it has to be recognised that the 'hard' economic evidence arising from the agglomeration benefits associated with increased employment density may not tell the complete story. As suggested by Banister and Berechman (2000), many cities may invest in urban transit, especially rail based systems, as an indicator of their commitment to improved quality. Similarly, many cities seeking inward investors point to their 'address on a network' such as an international airport or station on a high-speed rail line as an indicator of their standing. Such factors will always be difficult to quantify.

Table 1
Estimated welfare impacts of crossrail.

Benefits	Welfare (£mn)
Business time savings	4847
Commuting time savings	4152
Leisure time savings	3833
Total transport user benefits	**12,832**
Agglomeration benefits	3094
Increased competition	0
Imperfect competition	485
Exchequer consequences of increased GDP	3580
Additional benefits	**7159**
Total	**19,991**

Source: Department for Transport (2005).

11. Concluding remarks

Perhaps the key implication of this discussion is that simple rules on the wider benefits of transit improvement for the local, regional or national economy can be dangerous. Investment in transit could seriously damage the health of the local economy, but by the same token, failure to invest in transport could also damage it. There is no universal rule, such that calculations as reported in Table 1 have to be based on the specific evidence for that city. As we have also seen from the property value studies there are substantial variations in revealed results and these can often be city-specific. The greatest danger is that we transport evidence on agglomeration economies from one city to another and apply them without due regard to the industrial structure or market structures of that city.

We have come full circle in the debate on wider benefits. Early work implied that "transport is critical", later work stressed the need to "beware double counting" and more recently we have returned to a belief that "wider benefits are the key". However, we need to beware all simple rules in transport appraisal. The evidence on both the property impacts of transit development and the scope for agglomeration benefits implies on average modest effects of a 10–20 per cent uplift – some impacts appear larger, some smaller, but they depend critically on local circumstances.

There remains much on the research agenda. We need to know more about the role of imperfect competition and the productivity gains from transport. A key to this may come from more research on micro-behavioural evidence. We also need to pursue research on the relative impacts of link versus network effects and on spillovers and jurisdictional competition in transport investment. Finally there is a great need for more ex-post studies, checking whether ex-ante appraisals really capture the claimed benefits revealed after completion, does transport investment really make the difference claimed?

Policy too needs to be developed carefully in the light of this. There is a strong temptation to assume that there will always be added gains from transit development which can justify schemes which might otherwise be rejected on traditional cost-benefit grounds. This can lead to over optimistic claims being made for such gains and pressure for 'transit-oriented development'. Whilst recent work does suggest that there are stronger effects from denser development close to transit developments implying that joint planning of transit and land use may be a more effective way of capitalising on the development, once again it would be wrong to conclude that all such development will produce similar positive results. Frequently this will depend on the extent to which it is the absence of good transportation links, by any mode, is the constraint on economic development. If transportation is not the constraint then no amount of additional investment is likely to create economic benefit.

References

Alonso, W. (1964). *Location and land use*. Cambridge, MA: Harvard University Press.
Aschauer, D. A. (1989). Is public expenditure productive? *Journal of Monetary Economics, 23*, 177–200.
Banister, D., & Berechman, J. (2000). *Transport investment and economic development*. London: UCL Press.
Berechman, J., Ozmen, D., & Ozbay, K. (2006). Empirical analysis of transportation investment and economic development at state, county and municipality levels. *Transportation, 33*, 537–551.
Boarnet, M. G. (1998). Spillovers and the locational effects of pubic infrastructure. *Journal of Regional Science, 38*, 381–400.
Bowes, D., & Ihlanfeldt, K. R. (2001). Identifying the impacts of rail transit stations on residential property values. *Journal of Urban Economics, 50*, 1–25.
Bröcker, J. (2000). Trans-European effects of trans-European networks. In F. Bolle, & M. Carlberg (Eds.), *Advances in behavioural economics*. Heidelberg: Physica.
Bröcker, J. (2004). Computable general equilibrium analysis in transportation economics. In D. A. Hensher, K. J. Button, K. Haynes, & P. Stopher (Eds.), *Handbook of transport geography and spatial systems: Handbooks in transport, Vol. 5*. Oxford: Elsevier.

Burmeister, A., & Colletis-Wahl, K. (1996). TGV et fonctions tertiaires: grand vitesse et entreprises de service à Lille et Valenciennes. *Transports Urbains* 93.

Cervero, R. (1997). *Transit induced accessibility and agglomeration benefits: A land market evaluation*. IURD working paper 691.

Ciccone, A. (2002). Agglomeration effects in Europe. *European Economic Review, 46*, 213–227.

Ciccone, A., & Hall, R. (1996). Productivity and the density of economic activity. *American Economic Review, 86*, 54–70.

Damm, D., Lerman, S. R., Lerner-Lam, E., & Young, J. (1980). Response of real estate values in anticipation of the Washington Metro. *Journal of Transport Economics and Policy, 14*, 315–336.

Davies, G. W. (1976). The effect of a subway on the spatial distribution of population. *Journal of Transport Economics and Policy, 10*, 126–136.

Department for Transport. (2005). *Transport, wider economic benefits, and impacts on GDP*. Technical Paper. London: Department for Transport.

Dewees, D. (1976). The effect of a subway on residential property values in Toronto. *Journal of Urban Economics, 3*, 357–369.

Duncan, M. (2008). *Comparing rail transit capitalization benefits for single family and condominium units in San Diego, CA*. Paper presented at Transportation Research Board Annual Meeting, Washington DC, January.

Elhorst, J.P., Oosterhaven, J., Rom, W.E. (2004). *Integral cost-benefit analysis of Maglev technology under market imperfections*. SOM Report 04C22, University of Groningen (forthcoming in Journal of Transportation and Land-Use).

Fujita, M. (1989). *Urban economic theory: Land use and city size*. Cambridge: Cambridge University Press.

Fujita, M., Krugman, P., & Venables, A. J. (1999). *The spatial economy: Cities, regions and international trade*. Cambridge, MA: MIT Press.

Fujita, M., & Thisse, J. F. (2002). *Economics of agglomeration*. Cambridge: Cambridge University Press.

Gatzlaff, D., & Smith, M. (1993). The impact of the Miami Metrorail on the value of residences near station locations. *Land Economics, 69*, 54–66.

Gibbons, S., & Machin, S. (2005). Valuing rail access using transport innovations. *Journal of Urban Economics, 57*, 148–169.

Graham, D. (2005). *Wider economic benefits of transport improvements: Link between agglomeration and productivity, Stage 1*. Report. London: Department for Transport.

Graham, D. (2006). *Wider economic benefits of transport improvements: Link between agglomeration and productivity, Stage 2*. Report. London: Department for Transport.

Graham, D. (2007). Agglomeration, productivity and transport investment. *Journal of Transport Economics and Policy, 41*, 1–27.

Gramlich, E. (1994). Infrastructure investment: a review essay. *Journal of Economic Literature, 32*, 1176–1196.

Hansen, W. G. (1959). How accessibility shapes land use. *Journal of the American Institute of Planners, 25*, 73–76.

Henderson, J. (1986). Efficiency of resource usage and city size. *Journal of Urban Economics, 19*, 47–70.

Henderson, J. V. (2003). Marshall's scale economies. *Journal of Urban Economics, 53*, 1–28.

Hess, D. B., & Almeida, T. M. (2007). Impact of proximity to light rail rapid transit on station-area property values in Buffalo, New York. *Urban Studies, 44*, 1041–1068.

Jiwattanakulpaisarn, P. (2007). *Granger causality and spatial spillover effects of highway infrastructure on regional economic development: Evidence from an application of spatial filtering in a panel vector autoregressive framework*. Paper to European Regional Science Meeting, Paris, 2007.

Kahn, M. (2007). Gentrification trends in new transit-oriented communities: evidence from 14 cities that expanded and built rail transit systems. *Real Estate Economics* 35.

Klein, O. & Claisse G. (1997). *Le TGV-Atlantique: entre recession et concurrence*, LET, Lyon.

Laird, J., Nellthorp, J., & Mackie, P. J. (2005). Network effects and total economic impact in transport appraisal. *Transport Policy, 12*, 537–544.

McMillen, D., & McDonald, J. (2004). Reaction of house prices to a new rapid transit line: Chicago's Midway Line, 1983–1999. *Real Estate Economics* 32.

Miyagi, T. (1998). A spatial computable general equilibrium approach for measuring multiregional impacts of large scale transportation projects. *Network infrastructure and the urban environment*. Heidelberg: Springer.

Miyagi, T. (2001). *Economic appraisal for multi-regional impacts by a large scale expressway project*. Tinbergen Institute Discussion Paper TI 2001-066/3, Amsterdam: Tinbergen Institute.

Moomaw, R. L. (1981). Productivity and city size: a review of the evidence. *Quarterly Journal of Economics, 96*, 675–688.

Nakamura, R. (1985). Agglomeration economies in urban manufacturing industries: a case of Japanese cities. *Journal of Urban Economics, 17*, 108–124.

Oosterhaven, J., & Broersma, L. (2007). Sector structure and cluster economies: a decomposition of regional labour productivity. *Regional Studies, 41*, 639–659.

Patacchini, E., & Rice, P. (2007). Geography and economic performance: exploratory spatial data analysis for Great Britain. *Regional Studies, 41*, 489–508.

Plassard, F. & Cointet-Pinell O. (1986). *Les effets socio-économique du TGV en Bourgogne et Rhônes Alpes*. DATAR, INRETS, OEST, SNCF.

Rice, P., & Venables, A. (2003). Equilibrium regional disparities: theory and British evidence. *Regional Studies, 37*, 675–686.

Rice, P., Venables, A., & Patacchini, E. (2006). Spatial determinants of productivity: analysis for the regions of Great Britain. *Regional Science and Urban Economics, 36*, 727–752.

Rosen, S. (1974). Hedonic prices and implicit markets: product differentiation in pure competition. *Journal of Political Economy, 82*, 34–55.

Ryan, S. (2005). The value of access to highways and light rail transit: evidence for industrial and office firms. *Urban Studies* 42.

SACTRA (Standing Advisory Committee on Trunk Road Assessment). (1999). *Transport and the economy*. London: Stationery Office.

SES. (1998). *Evaluation de l'impact du TGV Nord-Européen sur la mobilité*. Les Etudes du SES.

Simmonds, D. (David Simmonds Consultancy in collaboration with Marcial Echenique and Partners) (1999). *Review of land use/transport interaction models*. Report to Standing Advisory Committee on Trunk Road Assessment, London: DETR.

Small, K., & Verhoef, E. (2007). *The economics of urban transportation*. London: Routledge.

Spiekermann, K., & Wegener, M. (1994). The shrinking continent: new time-space maps of Europe. *Environment and Planning B, 21*, 653–763.

Sveikauskas, L. (1975). The productivity of cities. *Quarterly Journal of Economics, 89*, 392–413.

Tornqvist, G. (1970). *Contact systems and regional development, lund studies in geography B 35*. Lund: C.W.K. Gleerup.

Venables, A. J. (2007). Evaluating urban transport improvements: cost-benefit analysis in the presence of agglomeration and income taxation. *Journal of Transport Economics and Policy, 41*, 173–188.

Vessali, K. (1996). Land use impacts of rapid transit: a review of the empirical literature. *Berkeley Planning Journal* 11.

Vickerman, R. W. (1980). *Spatial economic behaviour*. London: Macmillan.

Vickerman, R. W. (2007a). Cost-benefit analysis and large-scale infrastructure projects: state of the art and challenges. *Environment and Planning B, 34*, 598–610.

Vickerman, R.W. (2007b). *Recent Evolution of Research into the Wider Economic Benefits of Transport Infrastructure Investments*., Paper to Research Round Table: Macro-, Meso- And Microinfrastructure Planning And Assessment Tools, Paris: European Conference of Ministers of Transport Joint OECD/ECMT Transport Research Centre Paper CEMT/OCDE/JTRC/TR(2007)9.

Vickerman, R. W., Spiekermann, K., & Wegener, M. (1999). Accessibility and regional development in Europe. *Regional Studies, 33*, 1–15.

Wardrop, J. (1952). Some theoretical aspects of road traffic research. *Proceedings of the Institute of Civil Engineers, 1*, 325–378.

Weber, A. (1909). *Über den Standort der Industrien (transl. By Friedrich, C.J. (1929), Alfred Weber's Theory of the Location of Industries)*. Chicago: University of Chicago Press.