IV. MICROECONOMIC BENEFITS

Microeconomic benefits refer to improvements in the productivity of individual firms due to investments in transportation infrastructure. Depending on market structure, the benefits may accrue to firms themselves as increased profit, be passed along to consumers as lower prices, or some combination of the two. Cost Benefit Analysis is the conventional means of assessing microeconomic benefits. In this section we review these benefits and the extent to which they may be captured in a CBA framework.

The first and most obvious of these benefits is the reduction in inbound and outbound transportation costs. Highway infrastructure improvements – which may include additions of new roads, expansion or improvement of existing roads, and expansion of effective capacity though implementation of ITS – reduce costs for two reasons. First, as the network expands the density of its links increases, making point to point trips less circuitous and thereby reducing distances. (As the network evolves over time and becomes highly connected, however, the marginal impact of link additions on distances decreases.) Second, addition of new roads and capacity expansions on existing roads may decrease congestion and thereby travel times. Since congestion implies that fewer trips can be achieved with the same capital and labor, and since fuel efficiency is lower in congested driving conditions, reductions in congestion translate into reductions in transportation costs. Naturally, induced travel of both freight and passenger vehicles may offset some of the congestion reductions.\(^2\)

The basic analysis by which reductions in transportation costs are captured in a CBA framework is illustrated in Figure 18. Here the demand function DD represents the number of trips that would be made at different costs. At a high cost there are only a few trips made. These are trips whose economic benefits are great enough to justify the cost.

\(^2\) Induced travel does not imply that benefits are negated. Even if all congestion reduction is lost due to increased demand, more transportation services are delivered while holding congestion delay constant.
As the cost declines, less beneficial trips can be economically justified, thus we have the downward sloping demand function. Before the infrastructure improvement, the intersection of the demand function and the horizontal supply function determines the number of trips made \( Q \). The area \( CS \) defined by the difference between \( DD \) and \( SS \) summed up from zero to \( Q \) is the initial level of user benefit provided by the transportation system. This gap between what people are willing to pay and the market price is known as the \textit{consumer surplus}.

**Figure 18. User Benefits**

The effect of an infrastructure improvement is a downward shift in the supply function to \( S' \). This leads to an increase in the number of trips made to \( Q' \) The increase in user benefit has two components: the area \( A \) which represents the reduction in cost enjoyed on all trips that were made prior to the infrastructure improvement and \( B \) which is the user benefit on the incremental trips (\( Q'-Q \)). This theoretical argument provides the rationale for a very straightforward calculation of benefits. The summed areas \( A + B \) are determined by calculating a per-trip reduction in cost – including tolls (where relevant) and factors such as labor, capital depreciation and fuel costs which can be estimated as functions of the change in travel time and distance – and multiplying by the number of trips. Since the number of trips are different before and after the infrastructure
improvement, the user benefit is approximated by the "rule of half": \((S-S')(T+T')\). A general procedure is repeated for personal transportation users as well as freight transportation users and the sum of freight and personal benefits are weighed against the project cost.

A number of refinements on this method are possible. In particular, provision of transportation services may have a number of negative or positive external effects. Air pollution, for example, is a negative external effect that increases with the number of trips made Q. Figure 19 shows a case where the private cost of travel is S is augmented by an external cost e. Because of the upward sloping nature of the S+e functions, the estimated benefit (in this a social benefit) will be lower than the user benefit in Figure 18 by the amount E which represents the extra external cost imposed by the increase in trips. Naturally, if the result of the project is to reduce pollution, a positive adjustment a positive adjustment is needed. In order to implement this adjustment, it must be possible to calculate the net change in external cost in monetary terms and subtract it from the calculated user benefit.

Figure 19. CBA with Externalities

---

3 For a review of CBA applied to transportation projects, see Mackie and Nellthorp (2001).
Where congestion is an issue, the horizontal supply function may be replaced by one with an upward sloping segment to account for the fact that beyond some design capacity travel time is an increasing function of the number of users. The effect of the infrastructure improvement is then to increase the design capacity, thus shifting the upward sloping segment to the right as in Figure 20. The user benefit is defined as in this case as \((P-P')(Q+Q')\). In this case, the scope of the analysis must be fairly broad because congestion on various elements of an infrastructure network tends to be interrelated. For example, if the analysis addresses the addition of a lane to a particular road, it is generally necessary to assess impacts on travel time not only on that road but on all roads that are sufficiently connected to experience positive or negative congestion impacts from the road widening. (Only in the unlikely case that the cost of travel is equal to its marginal cost can the analysis be limited to a single road, Mohring, 1993.) It is also possible that induced trips for personal transportation may negate much of the benefit that might otherwise have accrued to freight users.

**Figure 20. Congestion Effects**

CBA can be extended to address these and a number of other market distortions that result in divergence between private benefits and social benefits. (Venables and Gasiorek, 1999, Part 1.) However recent research indicates that there are a number of benefits from transportation infrastructure whose analysis lies beyond the conventional
supply and demand for transportation services. These include logistics costs effects, facilities consolidations, and other location effects. In what follows we describe these effects and challenges of capturing them in a CBA framework. We also address the issue of imperfect competition, which can result in underestimates of total welfare gains.

**Total Logistics Costs**

The savings arising from reduced freight cost also include non-transportation cost elements such as storage, interest and insurance costs. Suppose a manufacturing firm has to receive a fixed quantity \( m \) of some material input in each one-year production cycle. One of the decisions the firm must make is whether to receive the input in one large order or in a number of smaller shipments. Define *total logistics cost* \( TLC \) as the sum of procurement, carrying, and transportation costs associated with that input. Procurement costs \( P \) will be lowest if one large shipment is received because it will only have to be ordered and processed once. Transportation costs \( T \) will also be lower because it is generally cheaper on a per unit basis to ship large batches of goods. However carrying costs \( C \), which include interest, insurance and storage costs, will be lower if the input is received in small shipments so that the amount held in inventory is minimized. (This is one of the principal benefits of Just In Time systems.)

**Figure 21. Total Logistics Costs**

![Diagram of Total Logistics Costs](image)

Adapted from McCann, 1998
Thus the optimal input shipment size – and the associated optimal level of inventory – depends on the trade-off between procurement and transportation costs on the one hand and carrying costs on the other. A similar argument can be made with respect to the firm's outputs. By delivering goods more frequently and in smaller batches the firm saves on inventory carrying costs by purchasing more (or better) transportation services.

This is illustrated in Figure 21 where the sum of transportation and procurement costs $P+T$ and carrying costs $C$ are graphed against the average size of shipments $B$. The optimal $B$ is found where $TLC = T+P+C$ is at a minimum. When a reduction in transportation cost from $T$ to $T'$ occurs – shown here as a downward shift in the $P+T$ schedule – the effect is a decrease in the optimal $B$. Thus, cheaper freight services leads to reduced storage, insurance and interest costs associated with carrying inventory. (For a theoretical treatment of the effect of transportation cost on optimal shipment size see McCann, 1988.) Of course, maintaining lower inventory increases the risk that the production process will be interrupted because a shipment of some critical input is late. Thus, if freight services are unreliable, greater than optimal inventory must be maintained as a buffer against risk. Improvements in freight services in recent years – brought about in part by application of information and communications technology – have improved the reliability of delivery within narrow time windows, thus allowing producers to reduce inventory and thereby reduce carrying costs. This illustrates that improvement in the quality, as well as reductions of the cost, of freight services yield benefits.

Empirical research on the impacts of improved transportation services on the logistics costs of firms is relatively thin, but a couple of studies provide some interesting insights. A 1995 study by Hickling, Lewis, Brod sought to measure the production benefits of a general reduction in freight travel times for firms in a variety of production sectors. The study employed an intensive interview technique on a small sample of firms, collecting detailed information on logistics practices and asking managers to speculate as to how a reductions in travel time, with corresponding reductions in travel time variances, would translate into reduction in logistics costs.
A compilation of some of their results is provided in Table 7. The first measure is an elasticity of logistics costs with respect to travel time. For example, the elasticity of .548 indicates that a 1% reduction in average freight travel time will lead to a .548% reduction in logistics costs for the surveyed firms in the medical and surgical instrument industry. The numbers in the next two columns indicate, for example, that a general 20% reduction in freight travel times will yield savings equivalent to slightly less that 1% of total sales in the same industry.
Table 7. Measures Of Travel Time Reduction Impacts On Costs (Hickling, Lewis, Brod, 1995)

<table>
<thead>
<tr>
<th>Industry</th>
<th>Logistics cost / travel time elasticity</th>
<th>Logistics cost savings as % of sales</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>20% reduction in travel time</td>
</tr>
<tr>
<td>Retail Food</td>
<td>.055</td>
<td>.04%</td>
</tr>
<tr>
<td>Automotive Parts</td>
<td>.234</td>
<td>.20%</td>
</tr>
<tr>
<td>Telecommunication Equipment</td>
<td>.103</td>
<td>.02%</td>
</tr>
<tr>
<td>Medical / Surgical Instruments</td>
<td>.548</td>
<td>.88%</td>
</tr>
</tbody>
</table>

These results should be regarded with caution, since they are based on a very small sample and they are derived from speculation rather than actual practice on the part of the firms. But the variation across industries – ranging from negligible benefits for food retail to very substantial benefits for medical and surgical instruments – are interesting. The report explains that only a portion of savings are derived from reduced transportation costs, with reductions in inventory carrying costs comprising another significant component of savings. This in part explains why medical and surgical instrument producers, which carry very high value inventories, can reap greater benefits from transportation improvements than the other industries. Thus, incorporating inventory carrying costs calls into question the conventional wisdom that transportation improvements are relatively unimportant for high value added industries.

A very recent study by Shirley and Winston (2001) attempts to draw an empirical link between the provision of road infrastructure and inventory levels. Theory suggests that lower transportation costs and higher reliability allow firms to maintain lower inventories. If costs and reliability are related to the capacity and condition of highway infrastructure, then inventories should be inversely related to highway infrastructure expenditures. Estimating an econometric specification derived from the theory of logistics costs on firm level data from the Census Bureau's Longitudinal Research Database, the authors find empirical support for this hypothesis. Their results indicate,
however, that the marginal inventory reductions for each dollar of infrastructure expenditure are declining over time.

Of course, not all decisions with respect to shipment size reflect the shipping firm's preference to reduce inventories. A survey of British producers of food and drinks found that the most important force leading them to consume more transportation services was pressure from customers (retailers, distributors, etc.) to receive shipments in a more timely fashion so that they could reduce their inventories (McKinnon and Woodburn, 1996). Either way, more and better transportation services lead to greater coordination of production schedules and a system-wide reduction in inventory carrying costs.

**Consolidation of Facilities**

Firms may relocate facilities in response to an infrastructure improvement that alters the spatial distribution of freight accessibility and costs. For example, construction of a circumferential highway (beltway) may attract producers away from urban locations with rail and water access to more peripheral locations along the highway where they can have good access to truck freight services. Generally, such shifts should not be counted as conferring benefits because they increase income in one locale at the expense of another (Forkenbrock and Foster, 1990). There are situations, however, where the locational shift induced by improved freight service leads to increased production efficiencies. In these cases there is a general increase in productivity, which should be attributed to infrastructure improvement.

One such case is where reduced freight costs allow a multifacility firm to concentrate its production on a smaller number of locations in order to take advantage of scale economies. One of the most persistent themes in the theory of industrial location is the trade off between scale economies and transportation costs. Imagine a manufacturing firm that wants to sell its product in a large number of urban markets dispersed throughout a regional or national economy. If its production technology has scale economies, it will minimize production cost by producing in a single location to serve all markets. However it will minimize delivery costs by producing in smaller facilities.
located near important markets. It must therefore choose some optimal configuration of facilities and locations based on the counteracting effects of scale economies and transportation costs. If freight costs decline – whether from improved infrastructure, technical progress, or enhanced productivity in service provision – the firm will have an incentive to close some facilities and expand others, or build one or more large-scale facilities in new locations.\(^4\) In this case there is a transfer of jobs and income from locales where closures occur to locales where expansions occur, but there is also an aggregate efficiency benefit due to scale economies.

Empirical studies show that consolidation of facilities can lead to very substantial savings in logistics costs. For example, a case study of a firm in the medical and surgical products industry with $1.8 Billion in sales in 1990 is provided in Hickling (1995). As Table 8 indicates, this firm was able to consolidate its 16 distribution centers into 6 larger regional centers, resulting in an overall 19% reduction in logistics cost.

### Table 8. Logistics Cost Savings due to Facilities Consolidation, Medical and Surgical Products Case Study

<table>
<thead>
<tr>
<th></th>
<th>Before Consolidation</th>
<th>After Consolidation</th>
<th>Savings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distribution Facilities</td>
<td>16</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Costs ($Millions)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transportation</td>
<td>22</td>
<td>18</td>
<td>18.2%</td>
</tr>
<tr>
<td>Warehousing</td>
<td>9</td>
<td>7</td>
<td>22.2%</td>
</tr>
<tr>
<td>Inventory Carrying</td>
<td>11</td>
<td>9</td>
<td>18.2%</td>
</tr>
<tr>
<td>Total Logistics Costs</td>
<td>42</td>
<td>34</td>
<td>19.0%</td>
</tr>
</tbody>
</table>

Hickling (1995)

\(^4\) Even in the absence of scale economies, there may be benefits from spatial concentration of production. For example, if there is uncertainty in the level of demand in individual markets, concentration of production may have the effect of pooling risk – aggregate demand is more predictable than local demand (Comacho and Persky, 1990.)
A useful way to look at the consolidation effect is as a substitution between production inputs and transportation services (HLB, 2001). Achieving scale economies means that on a per unit of output basis, the firm will be able to purchase fewer inputs of labor, capital, and other inputs. Since the scale economies are achieved by consolidating facilities, however, they come at the expense of more ton-miles of transportation services.
This substitution is depicted in Figure 22 by the isoquant II, which defines different combinations of production inputs and ton miles that can be combined to produce a given amount of output. If transportation costs are reduced, the slope of the iso-outlay line AB is reduced to the line CD. The points of tangency define the optimal combination before (T) and after (T’) a reduction in transportation costs.

In fact, the same sort of logic can be applied to the effect of reduced transportation costs on TLC. As transportation gets cheaper, firms may choose to reduce inventories by increasing the number of inbound and outbound trips. Thus transportation services are substituted for inputs that include working capital, warehouse space and insurance.

*Incorporating Logistical Savings and Consolidations in CBA*

We have described two responses to reduced transportation costs – decrease in shipment size and consolidation of facilities – that allow firms to increase productivity but require increased inputs of transportation services. (In the former case increased services arise as a fixed number of ton miles are distributed over a larger number of smaller shipments, in the latter the number of ton miles actually increases – thus both effects contribute to an increase in the number of vehicle miles). This relationship is
illustrated in Figure 23, where DD is the demand function before the cost reduction and S is the initial constant supply function. When S shifts down to S', there is a short term response represented as an increase from the initial demand level Q to Q'. In the longer run, the firm may adjust its shipment size and consolidate its facilities, which results in a shift to a new demand function D'D' and thus a further increase in transportation services to Q''.

This has significant implications if the reduction in transportation costs is the outcome of an improvement in highway infrastructure. For one thing, it implies that the induced freight travel effect of expanded capacity may be greater than a static demand analysis would predict. Conventional cost-benefit analysis, which seeks to measure the increase in user benefit due to the short run shift from Q to Q', would therefore underestimate the benefits accruing to freight-using firms from highway improvements. Even if it were possible to anticipate the larger increase in demand from Q to Q'', the rule of half approach would still significantly underestimate user benefits. The full estimate must capture the value proportional to F in Figure 23 and add it to the components A and B of conventional CBA. This requires an method to predict the shift in the demand function, rather than just the increase in the quantity demanded.

Figure 23. Shift in Demand for Services
Measurement of this demand enhancement effect presents a tremendous challenge for the implementation of CBA. This problem has been addressed in a number of reports by HLB Decision Economics Inc. (Lewis, 1991; HLB, 1995, 2001). For example, some magnitude of inventory reduction must be defined for a particular level of transport cost reduction arising from an infrastructure improvement. Even more challenging will be estimation of structural shifts such as the consolidation of facilities after infrastructure is improved. Thus, the HLB research provides not only a challenge to conventional CBA methods but also defines the need for extensive empirical research into the interrelationship between the cost and quality of freight transportation services and a whole range of business decisions including inventory policy and the number and location of facilities.

As an illustration of the methods employed, Lewis, 1991 proposes a simple approach to estimating the shift in the demand function. Define the demand function before the improvement in transportation infrastructure as

\[ P_0 = \alpha - \beta Q_0 \]

where \( P \) is an operating cost per vehicle mile and \( Q \) is total demand measured in vehicle miles. Given an estimate of the elasticity of demand

\[ N = \frac{f_Q}{f_P} \frac{\partial P}{\partial Q} \]

then by substitution

\[ \alpha = P_0 + \frac{P_0}{N \sqrt{\Delta Q}} \]

\[ \beta = P_0 / N Q_0 \]

The effect of the infrastructure improvement is measured in terms of a variable \( \Delta Q \) defined as the percent change in transportation services required by the transportation provider to give the same level of service after the improvement. For example, if 5000 vehicle miles were required before the improvement and 4000 are required after the
improvement, $\Delta Q = -2$. If the result of the improvement is a rightward shift in the demand function to

$$P_i = \alpha' - \beta' Q_i$$

the new parameters can be estimated as

$$\alpha' = \frac{P_0 + \frac{-P_0}{N^{\Delta Q}}}{1 + \Delta Q}$$

$$\beta' = \frac{P_0/NQ_0}{(1 + \Delta Q)^2}$$

Using this approach, the additional benefit area $F$ in Figure 22 can be estimated based on observed cost and service demand values and estimates of the price elasticity and the effect of the infrastructure improvement on service provision. This basic framework suffers from a number of strong assumptions – such as that the level of output remains fixed and that logistics costs savings are known a priori – which are addressed in further elaborations of the model (HLB, 2001).

While this approach is an important step, a direct causal link between reduction in transportation costs and the demand shifts should be inferred with a great deal of caution. Given the theoretical arguments and empirical evidence described above, it is tempting to attribute the major logistical transformations described in Part 1 of this report to reductions in transportation costs due to investment in infrastructure. While the macroeconomic studies indicate a link between investment and productivity, there is not much empirical evidence to specifically link changes in logistical practices to transportation costs other than the preliminary work by Shirley and Winston. For example, McKinnon and Woodward's (1996) survey of British firms found only very limited evidence that declining transportation costs were a major driver behind facilities consolidation and declines in average inventories. Consolidation of facilities is part of a

\[5\] This is an appropriate measure only where the infrastructure improvement has the effect of reducing distances between origins and destinations. If the effect were strictly congestion reduction, an alternative measure of demand, such as vehicle service hours, would be appropriate.

\[6\] A complete derivation is provided in appendix D of Lewis, 1991.
long term trend and the move to JIT inventories may arise as much from new control technologies and the heightened awareness of inventory management as from declining transportation costs. Nevertheless, the availability and scope of transportation infrastructure is a key enabling factor behind these trends.

We turn now to two classes of economic benefits that are generally not included even in the most recent work on assessing the benefits of transportation infrastructure, but which may represent important mechanisms by which cheaper and more reliable services are translated into improved productivity for individual forms: location effects and value added effects.

Location Effects

There are a number of ways that improvements in transportation infrastructure may contribute to productivity growth through mechanisms that involve the location choice of the firm. For example, infrastructure investments promote productivity when firms are able to take advantage of agglomeration economies – essentially a class of external scale economies – by locating in large clusters. Agglomeration economies arise for various reasons, including scale economies in the provision of public infrastructure to concentrated centers of demand (urbanization economies), reductions in the cost of transferring intermediate goods among diverse firms linked together in the production chain (juxtaposition economies) and spillovers of knowledge and labor skills that occur when firms in the same industry cluster together (localization economies.) The notion of agglomeration economies has long been used by geographers as a conceptual tool to explain the emergence of cities, and has gained empirical support from recent econometric studies that attribute significant productivity benefits to economic agglomeration (Ciccone and Hall, 1996.)

Naturally, the benefits of agglomeration are at some point offset by the costs of congestion. In this sense major investment in infrastructure for freight movements within, as opposed to between, production centers expand the potential for agglomeration to the extent that they offset growth in congestion (Weisbrod and Treyz, 1998.) These
may be some of the most important benefits that arise from major urban infrastructure improvements such as Boston's Central Artery / Tunnel project (the Big Dig).

Expansion of transportation infrastructure – especially the Interstate Highway system – has expanded the range of possible locations for producers of goods and services. "Greenfield" production sites located at the periphery of metropolitan areas or in rural areas have been sought by many producers to economize on land costs. Thus, just as firms are able to reduce inventory carrying costs by increasing their use of transportation services, firms choosing peripheral locations along highways reduce land costs while consuming more transportation services.

It may seem contradictory to argue that transportation infrastructure promotes productivity on the one hand by allowing firms to cluster together in cities and on the other by allowing firms to spread out into the periphery. But this must be viewed in light of the fact that different firms benefit from different locations. The product life-cycle model (Vernon, 1966) argues that products and the industries that produce them pass through stages in their histories over which their main business requirements evolve. At an early stage of development, acquiring appropriate labor skills, developing product innovations, and market penetration are the key concerns, while at a later stage, implementing process innovations, and economizing on the cost of inputs such as labor and land are the main competitive strategies. The spatial analogue to this argument is that early stage firms do best in urban core locations while late stage firms do best in the periphery. The main point is that a transportation system that provides sufficient capacity and connectivity benefits firms by expanding the range of locations from which they can choose.

As in the case of facilities consolidation, this runs counter to the conventional wisdom whereby location shifts should not be counted among the benefits of highway infrastructure. That argument is correct so far as the shifts simply reflect transfers of identical production activities from one location to another. But the spatial shifts described here are productivity enhancing and in therefore they produce real benefits.
Transportation and Value Added

So far we have discussed the microeconomic benefits of transportation system improvements in terms of cost reduction. Improved infrastructure allows firms to economize on transportation and other logistical costs; production costs due to scale and agglomeration economies; and land costs. Cost reductions translate into productivity enhancements\(^7\) and therefore they can explain much of the impact of transportation infrastructure on productivity growth that has been observed in the macroeconomic literature.

Another way that transportation infrastructure can enhance productivity is by adding value to the output of either the freight using firms or the transportation service provider. Take fresh fish as an example. The best way to add value to a fish is do nothing to it – except get it to the consumer quickly. Fresh fish is worth more than salted, frozen, or otherwise processed fish. Improvements in transportation service that make it possible to get a fish from Maine to St. Louis in less than 24 hours after it is caught are a major source of productivity enhancement. The justification for interpreting this as a case of adding value, rather than just reducing costs, is that the fish can only be produced in one or a few places and may have a scarcity value elsewhere. Thus, transportation makes it possible for the fish producing firm not only to expand markets but to reach markets where its output has a higher value than in its local market. A similar argument can apply to a variety of products that are produced in a limited number of locations because of highly specific skills or resources.

As another example, consider a machine that is used in a production process. The firm that uses the machine will pay more for it if they can be certain that it will never be out of service for more than a few hours. The firm that produces the machine will able to make such a guarantee – and therefore charge a higher price – if they know that it will be possible to ship necessary parts to the machines location if there is ever a breakdown.

\(^7\) HLB 2001 demonstrates the link between cost reduction and productivity growth, section 3.5.3.
Thus the availability of high-quality transportation service increases the value of that machine.

Improved infrastructure can also add value to the services of transportation providers. For example, higher capacity and implementation of ITS infrastructure can help trucking firms deliver goods not only cheaper but within narrower time windows. They can also provide real-time tracking information and a high level of flexibility regarding the volume, location, and frequency of shipments. All of this makes the service more valuable to the client. (In assessing these benefits, care must be taken to avoid double counting – for example, the increase in the value of a trucking firm’s services may be equivalent to the logistics cost savings of its client.)

*Market Structure and CBA*

There is an additional problem that emerges with CBA in the presence of imperfect markets. Figure 24(a) illustrates the case of a provider of transportation services who is a monopolist. The monopolist is able to earn extra profits by providing the service up to the level where its marginal cost (MC) is equal to the marginal revenue (MR), rather than where MC (supply) equals demand (D) as would be expected in a perfectly competitive market. The result is that the total level of services provided is lower than would be the case under perfect competition and the price is higher. Therefore the benefit to consumers (consumer surplus: CS) is also lower. However, there is also a benefit the provider in the form of extra profits. This is known as producer surplus (PS) as shown in the figure. The total welfare benefit from the provision of the service is the sum on CS and PS, which is still lower than CS would be under perfect competition.

Suppose there is an investment in infrastructure, such as a new highway or rail terminal, which reduces MC to MC'. The result is an increase in both CS and PS. Since CBA seeks to measure all welfare gains, no matter to whom they accrue, both increases should be counted in total benefit. Application of the rule of half as in conventional CBA, however, will only capture the change in CS – and in fact may underestimate the
change in CS. Thus CBA is inclined to underestimate benefits of transportation projects if markets for transportation services are imperfect.

**Figure 24(a). Imperfect Competition**

The reason CBA will underestimate CS relates to the way that a reduction in transportation cost translates into a reduction in price. Suppose the producer sets its price by a markup m over marginal cost: \( p = mc(1+m) \). Now suppose the mc is reduced to \( mc' \). The new price is \( mc'(1+m) \). \( p-p' = (mc'-mc)(1+m) \), thus the reduction in price is greater than the reduction in the marginal cost of production.

---

\(^8\) The reason CBA will underestimate CS relates to the way that a reduction in transportation cost translates into a reduction in price. Suppose the producer sets its price by a markup m over marginal cost: \( p = mc(1+m) \). Now suppose the mc is reduced to \( mc' \). The new price is \( mc'(1+m) \). \( p-p' = (mc'-mc)(1+m) \), thus the reduction in price is greater than the reduction in the marginal cost of production.
This sort of market distortion is not limited to the case where there is imperfect competition in the market for transportation services. The same applies if there is a monopolist in a freight consuming goods or service market. Benefits of investments in transportation infrastructure may be assessed from the perspective of derived demand for transportation services. In this case, the supply function in Figure 24 is of a producer whose marginal cost is reduced because of cheaper transportation. Part of the benefit goes to an increase in CS and part to an increase in PS. Again, the rule of half approach will underestimate the combined benefit.

Venables and Gasiorek (1999) have addressed this problem through a series of simulations. Envisioning a situation where an infrastructure improvement has reduced the transportation costs of a monopolist who sells goods into a neighboring region, they calculate first the conventional CBA benefit and then total welfare gain as defined in Figure 24. Their simulations indicate that CBA would have underestimated total benefits by anywhere from 10 to 40%, depending upon the size of the reduction in transport costs, the elasticity of the demand function, the slope of the marginal cost curve, etc.

The results of the simulations tell us two things. First, there is at least a potential for significant underestimation of benefits under imperfect competition. Second, in order
to capture the full benefit we would need knowledge of a variety of parameters (demand elasticities and marginal cost functions in non-transportation markets, for example) that are normally far beyond the scope of project assessment studies. More information and a more comprehensive analytical framework are needed.

The foregoing discussion is set in a context where the firm’s level of production is assumed to be independent of freight transportation costs. An enduring theme in economic geography is that reduced transportation rates make it possible for firms to expand the geographical scope of their markets, thus increasing output and the demand for transportation services. To some extent, market expansion may be viewed as a transfer effect, whereby production in one place is displaced by competition from somewhere else, and should therefore not be included as a benefit of infrastructure investments. Competition over space, however, may lead to regional specialization which yields improvements in aggregate productivity. We turn to this issue in the section on general equilibrium effects below.

V. GENERAL EQUILIBRIUM EFFECTS

In the previous section, we focused on efficiency gains accruing to individual firms as the result of improvements in transportation services, how those gains translate to welfare gains, and how welfare gains can be captured by cost-benefit analysis (CBA). By focusing on the effects on individual firms we have taken a partial equilibrium view. We now shift to the question of how the effects on individual firms redound throughout the entire economy – the general equilibrium view.

Since we are concerned with the issue of assessing economic benefits, a related question is how and whether general equilibrium effects can be captured in the general CBA framework or whether new approaches to evaluation are needed. A couple of notes are in order at the outset. First, it is not fair to say that conventional CBA ignores general equilibrium effects. These effects can be captured in CBA under fairly strong
assumptions about scale economies and competition. (However, it is one thing to say that such effects are captured in theory and quite another to say that they are captured in practice.) Second, while it may be possible to identify types of general equilibrium impacts that are not captured in CBA, there is an important question as to whether the magnitude of these impacts justifies a shift to more complex and less tested analytical frameworks. It has been argued that as one moves to impacts that are less directly connected to the infrastructure improvement in question, the "additionality" to the assessment of benefits becomes too small to worry about (Mackie and Nellthorp, 2001.) A counter argument is that effects may be amplified, rather than dampened, as they spread throughout the economy. In the case of such "cumulative causation," failure to take a general equilibrium view may lead to gross underestimation of both positive and negative economic impacts (Venable and Gasiorek, 1999.)

In the remainder of this section we review the general equilibrium concept of gains from trade as it arises first from the traditional notion of comparative advantage and in the framework of imperfect competition. We end with some speculation about needed empirical study and the development of new assessment methodologies.

**Gains from Trade**

One of the most important economic trends to have occurred over the 20th century is a shift from system of local, regional or national autarky to one of specialization and trade, culminating in the process of economic globalization. Specialization and trade is economically feasible only to the extent that efficiency benefits exceed the cost of interregional shipment and the speed and reliability of shipment make it possible to coordinate production schedules across long distances. Thus, freight cost reductions and quality improvements lead to efficiency gains from trade. These efficiency gains arise for two reasons:

- Each region has a different endowment of natural resources, labor (availability and skills), capital goods, and institutions, that make it highly efficient in some categories of production and less efficient in others. Specialization and trade allows regional
resources to be concentrated in those forms of production for which they are best suited. This is basically David Ricardo’s theory of *comparative advantage*, whose theoretical extensions are explored in the Heckscher – Ohlin – Samuelson framework (Findlay, 1995.)

- Even if regions have similar endowments of resources, there is still a benefit to interregional trade *via* scale economies that are realized as producers target broader markets. Furthermore, production and trade across a national (as opposed to regional) market makes it possible to provide consumers with a broader variety of goods. This is the explanation of gains from trade provided in the *new economic geography* (Fujita, Krugman, and Venables, 1999; Venables and Gusiorek, 1999.)

We address these two bases for specialization and trade in turn and expose the importance of freight transportation in the realization of gains from trade.⁹

*Comparative Advantage*

The theory of comparative advantage has its roots in the 19th century when the British economist David Ricardo argued against the "corn laws" which restricted imports of agricultural commodities into Great Britain. Corn law supporters argued that there could be no possible economic justification for importing grain since British agriculture was at least as efficient as agriculture in countries from which cheap imports originated. Ricardo countered that even if Britain had a small efficiency edge on other countries in agricultural production, it had a very large efficiency advantage in manufacturing production. Since there were fixed amounts of labor and capital resources available in Britain, British agricultural production had a high opportunity cost because it diverted resources from more lucrative manufacturing production. Thus even though Britain had an absolute advantage over its potential trading partners in both agriculture and manufacturing, it had a comparative advantage only in manufacturing. British agriculture was at a comparative disadvantage because of its high opportunity cost. Therefore, if

⁹ An additional reason that is often cited for economic benefits from trade is the increase in competitive pressure which induces firms to increase efficiency and adopt technical innovations (SACTRA, 1999).
foreign grain were imported and resources were transferred from agriculture to manufacturing, both Britain and its trading partners would be better off.

Ricardo's ideas were essentially refined into the Heckscher-Ohlin theorem, which says that if two countries have different relative endowments of capital and labor, they will both benefit if the labor rich country exchanges labor intensive goods for capital intensive goods from the capital rich country. The basic message is the same: all trading partners are better off if they specialize in those things they have the best capabilities to produce and trade with other countries than if they seek to achieve self sufficiency by producing a large variety of goods. An important caveat to this, however, is that gains from trade can only be realized to the extent that they exceed the transportation costs needed to achieve them. Therefore one of the most important benefits of improved transportation infrastructure arises from its role in enabling gains from trade.

Note that this is a *general equilibrium* benefit. It does not arise due to improved productivity in individual production units, but rather from a redistribution of production that leads to higher aggregate productivity. Imagine two regions A and B. Suppose that region A has higher productivity in agriculture than region B and region B has higher productivity in manufacturing than region A. Begin from an initial condition in which both A and B produce sufficient agricultural and manufacturing output to satisfy their own demands. Now if they switch to a situation where A produces all the agricultural goods for both A and B and B produces all the manufactured goods for both A and B, the aggregate productivity of the combined economy of A and B will have increased productivity *even if the productivity of individual production units remains constant*. To the extent that improved transportation infrastructure makes the transition from autarky to trade possible, it yields economic benefits.

The theory of comparative advantage has been the major economic argument in favor of liberalizing international trade. It has provided intellectual ammunition to the proponents of expanded international trade through the GATT/WTO mechanism as well as proponents of regional trading blocks such as the European Union, NAFTA,
MERCOSUR and others. One might assume therefore that economic benefits in the form of gains from trade arise primarily from investments in infrastructure built mainly for international trade: international shipping and air facilities, international bridges, facilities for rapid border clearance etc. In fact, recent experience shows that such infrastructure, along with complementary institutional change, is critical to the success of regional economic integration initiatives (Lakshmanan et al, 2001.) The role of infrastructure in trade creation, however, extends more broadly to the national infrastructure system. For example, a recent study by the Bureau of Transportation shows that 10.4% of U.S. freight movements over domestic road and rail infrastructure are ultimately to support international trade.

More importantly, the notion that transportation infrastructure yields economic benefits that come in the form of gains from trade applies just as well to domestic trade as to international trade – especially in an economy as large and diverse as ours. The United States economy is a multiregional economy comprising a set of distinct but highly integrated economic regions, each with a system of urban centers and resource hinterlands. Over time, these regions have become more highly integrated. For example interregional linkages among the southeast, the west, and the traditional economic core regions have strengthened. The theory of comparative advantage applies just as well here as it does in, say, Europe where diverse regional economies exist within national frontiers. A condition of autarky, where each region is relatively self sufficient, producing most of the goods and services consumed within its borders, is less efficient than a condition of specialization and trade, whereby all regions produce for national and international markets.

The railroads, the inland and coastal waterways, and the National Highway System provide the critical links that make regional specialization and interregional trade possible. Evidence of the increasing specialization over the second half of the twentieth century is found in the fact that ton-miles of freight grew more rapidly than tons of freight, especially in the period for 1960 to 1980 (see Table 6.) Under a system of autarky and regional specialization, the same number of tons might be shipped, but the
average distance of shipment would be much greater in the latter. Thus growth in the average distance over which a ton of freight is shipped is consistent with increasing regional specialization and interregional trade. Evidence of the role of transportation costs in regional specialization is found in a study based on the 1993 Commodity Flow Survey, which showed that interregional movements of commodities is inversely related to transportation cost (Anderson et al., 1998.)

What implication does this have for assessing the benefits of transportation investments? Any project that makes interregional trade easier and cheaper results in improved efficiency (and thereby reduced costs) for those goods that are shipped interregionally. At first blush it might seem that CBA, with an apparently partial equilibrium orientation, will miss this benefit. However, under of conditions of perfect competition the increase in trade will occur up to the point where the cost reduction from specialization and trade is just offset by the transportation cost. Thus, the benefits derived from increased travel (component B in Figure 18) captures the gains from trade.

This is illustrated in Figure 25, which represents the production of the same good in two regions with identical demand functions but different supply functions. The lower supply function in region B means that it has a superior endowment of resources for production of the good in question. Thus, the price in region B is lower than in region A. If trade is allowed to occur, and if transportation were costless, the good will be shipped from B to A up to that point where prices are equalized and excess supply in B is just offset by excess demand in A. If transportation is costly, a somewhat smaller flow would occur and a price difference equal to the transportation costs would remain. If the transportation cost declines, the amount of trade will increase. Thus, increases in the demand for transportation services are good indicators of gains from trade.

This argument is fine from a theoretical perspective, but it suffers somewhat in the implementation. CBA is normally applied ex ante – that is the calculation is made before the infrastructure investment is made. (Otherwise it would be of little use as a decision tool.) This means that the analyst must predict the increase in the demand for
transportation services in order to make the rule of half calculation. Accurate predictions of the full range of economic integrations that may occur after the addition of a major infrastructure improvement is a challenging requirement that is seldom met in applied CBA. Not only would changes in aggregate trade flows need to be predicted, but since this is a relatively long-term phenomenon the timing of the changes must be estimated in order to conduct appropriate discounting. Some attempts have been made to incorporate competitive business expansion impacts into CBA (Weisbrod and Treyz, 1998) but they are rather \textit{ad hoc} and generally only applied to changes in output for a single region or cluster of regions.
Figure 25. Two Region Case
An initial step toward assessing the magnitude of gains from trade attributable to infrastructure investments would be to conduct a series of *ex post* cost benefit analyses where the observed trade impacts are measured some years or decades after an improvement is made. The Commodity Flow Survey could be used to compare changes in trade flows between 1977 and 1993 or 1997. It would be necessary to use multivariate statistical analysis to estimate the component of trade growth that can be attributed to the change, and since it would be difficult to isolate the impact of individual projects the analysis might have to apply to a cluster of investments, such as a set of major regional additions to the interstate highway network.

This *ex post* approach would serve both as a means of assessing the magnitude of the benefits from highway construction that arise due to growth in interregional trade and as a means of comparing benefits calculated after the fact to those found in *ex ante* CBA studies for the same projects.

*The New Economic Geography*

While the theory of comparative advantage is one of the most powerful analytical frameworks ever developed by economists, there are aspects of trade that it does explain very well. Comparative advantage essentially says that gains from trade arise out of diversity across nations or regions in terms of natural resources, capital, human capabilities and institutions. But many of the most important trade relations occur between places that are in fact very similar. For example, the largest bilateral trade relationship in the world is between Canada and the U.S., two countries that are quite similar in terms of most critical economic endowments. Furthermore, the largest share of the goods flowing in both directions -- from the U.S. to Canada and from Canada to the U.S. -- are from the same industry: automobiles and automotive components. Neither the strength of relationship between similarly endowed nations nor the preponderance of *intraindustry* over *interindustry* trade is consistent with the predictions of the theory of comparative advantage.
A new analytical framework has emerged -- called the "new economic geography" -- which addresses just such situations and provides a host of new insights into the spatial configuration of economic activities. Where the theory of comparative advantage is driven by variations in endowments, the new economic geography is driven by *scale economies*.

At the risk of over simplification, consider again the example of region A and region B, who between them must satisfy a certain level of demand for both agricultural and manufacturing output. Assume this time that there is no inherent difference in productivity for either industry between the two regions. In the absence of scale economy, aggregate productivity will be the same under either autarky or specialization. Now assume that scale economies can be realized at the industry level, such that the productivity of agriculture is an increasing function of the amount of output produced in a particular region. The same applies for manufacturing. If this is the case, aggregate productivity will be higher if one region specializes in agriculture and the other specialized in manufacturing – although it doesn’t matter which region specializes in which industry.

Naturally the theory is more complicated than this. It develops a new framework for trade by exploiting the analytical breakthrough of Dixit and Stiglitz (1977) who incorporated scale economies into a general equilibrium model assuming a monopolistically competitive market structure. In this model, product variety is the critical component of competition so that all firms produce distinct but substitutable goods. Consumers' utility functions are defined in such a way that they prefer to consume a variety of goods rather than to concentrate their production on a small number of goods. Thus goods are imperfect substitutes. This means each firm has some degree of monopoly power and can therefore set its price above its marginal cost. The cost structure for each firm includes a fixed component and a constant marginal cost, which results in a downward sloping average cost function indicative of scale economies.
By opening up trade, producers in each region are able to reach broader markets for their unique goods, allowing them to move down their average cost curves and earn greater profits. Naturally this market expansion effect is limited by interregional transportation costs, and any reduction in transportation costs yields increased trade benefits. In the long run, however, more firms -- each providing a unique product variety -- will enter any market where profits are being earned. The presence of more firms shifts the demand curves of all preexisting firms downward until excess profits are exhausted. Thus it is entry of new firms that brings about an equilibrium.

While this view lacks some of the formal elegance of the perfect competition general equilibrium, it has a number of advantages. For one thing, its emphasis on product differentiation as opposed to direct price competition among producers of perfectly substitutable goods is in keeping with the long term trend away from commodity production towards highly differentiated and specialized goods. Growth in consumer utility in recent decades has been due not only to the quantity of goods consumed, but also to the ever increasing variety of goods -- especially consumer electronics and other categories of goods where constant product innovations define the competitive environment. Furthermore, the fastest economic growth has occurred in consumer and producer services, which are also highly differentiated. Thus, while neither perfect competition nor monopolistic competition accurately describes the entire economy, the real world is getting less like the former and more like the latter.

Another advantage is that by adopting imperfect competition and scale economies, this view is able to tackle a whole range of explicitly spatial phenomena, such as agglomeration and persistent regional differences in wages, which mainstream economic theory has largely ignored. In fact, many of the results emerging from the new economic geography are concepts that geographers have espoused for decades but have failed to present in a formal general equilibrium framework.

Finally, the new view provides theoretical underpinnings to a variety of observations found in the empirical literature on international and interregional trade
which do not make much sense from the comparative advantage perspective. For example, the strategy by which some urban regions attain competitive advantage on a global scale by specializing in the production of one or a few high value added commodities, which has been observed by Porter (1990), is consistent with results on agglomeration. Returning to the example of Canada-US trade, a recent study of regionally detailed trade flows found that trade between regions contiguous across the border (such as Ontario and Michigan) tended to be intraindustry rather than interindustry trade. This is consistent with the notion of trade in highly specialized varieties in order to benefit from scale economies, as described in the new economic geography view. Trade between more widely separated regions (with higher transportation costs) tended to be interindustry and therefore more consistent with the comparative advantage view (Brown and Anderson, 1999).

If one accepts the proposition that monopolistic competition is a more realistic theoretical basis for the functioning of many markets, then the following question arises: Is CBA an accurate method for evaluating transportation infrastructure investments when the industries that consume transportation services are monopolistically, rather than perfectly, competitive? This question was recently addressed in a study commissioned for the Standing Advising Committee on Trunk Roads Assessment in the United Kingdom (Venables and Gasiorek, 1999).

The answer to the question is no, for two reasons. The first reason we have already discussed: under imperfectly competitive markets, certain types of benefits are hidden from the rule of half approach to CBA. The second reason is perhaps less evident: in a world of imperfect competition and scale economies, certain general equilibrium effects are too small to be ignored.

To illustrate this, the authors developed a computable general equilibrium model (CGE) based on hypothetical data and parameters incorporating the principals of monopolistic competition as described above. It includes two or more hypothetical regions, each with endowments of two primary inputs: immobile capital and labor,
which is mobile in the long run. Equilibrium wage rates are determined at the regional level. Each region is assumed to include some competitive and some monopolistically competitive industries and input-output parameters describe the linkages among industries in terms of intermediate goods flows. Since firms will enter any monopolistically competitive industry where excess profits are being earned, expansion will occur in industries and regions whose profits expand due to a change in transportation costs.

As in the case of the partial equilibrium simulations described earlier, the CGE results show that total welfare gains from a reduction in transportation costs are greater than would be revealed by CBA. The magnitude of this difference increases with the demand elasticity for transportation services, the degree of market power for transport-using firms, the strength of inter-industry linkages and the strength of agglomeration economies.\(^{10}\) The underestimate identified by the simulations are in the 30% range with some as high as 60%. It is important to bear in mind that these are purely hypothetical estimates from simulations incorporating what the authors believe to be reasonable parameter values. Other observers suggest that the underestimates would be much smaller.\(^{11}\)

Because of the multiregional structure of the simulation, it was possible to identify circumstances where reductions in transportation costs would have differential regional impacts. In the general equilibrium context, these deviations arise for some interesting reasons:

- Symmetric regions: In a simulation where one competitive and one monopolistically competitive industry are both located in two regions with identical resource endowment, CBA was shown to underestimate total welfare gains by roughly the same amount as in the partial equilibrium simulations described earlier. The welfare gain was slightly larger in the general equilibrium

---

\(^{10}\) These general interpretation of Venables and Gasiorek results are found in SACTRA, 1999.

\(^{11}\) SACTRA (1999) quotes Newbery as arguing the effects would be roughly one tenth as high as the simulation results suggest.
case because extra gains by the monopolistically competitive firms were passed on to the perfectly competitive firms via input-output linkages.

- **Agglomeration:** In an economy with two monopolistically competitive industries and two symmetric regions, reduction in transport costs were shown to promote agglomeration whereby each of the two industries became concentrated in only one region. Total benefits are again higher than would be found by conventional CBA and shared equally by the two regions only if the two monopolistically competitive industries are of the same size. If one industry is larger, then the region where that industry agglomerates enjoys greater benefits.

---

12 This result depends on the assumption that within-industry input-output linkages are stronger than between-industry linkages.
Figure 26. Configuration of Three Regions

a)

b)
• **Center and periphery:** A simulation was devised to illustrate the case where one region has higher wages before the reduction in transportation cost occurs. The results indicate that if transportation costs are so high as to result in virtual autarky in the initial case, marginal reductions in transportation cost will cause wages to diverge. At that point, firms are attracted to the more lucrative market, but eventually wage increases in the center region cause shifts in the other direction. Thus, past some critical level of transport costs reductions will result in wage convergence. For this reason, an inverted U-shaped relationship between transportation cost and regional income equality is indicated. So the effect of a transport investment on interregional equity depends on the initial context and may be very difficult to predict.

• **Three region case:** With three equidistant regions as indicated in Figure 26a, there is a reduction in the cost of transportation between regions 1 and 2. The results indicate an increasing share of monopolistically competitive production in 1 and 2 at the expense of the share of three. Also wages increase in 1 and 2 and decline in 3. Thus while overall welfare increases in the three regions combined (at a rate higher than CBA would have captured), the reduction in transportation cost between 1 and 2 causes both a relative and absolute decline in the welfare of region 3.

• **Linear economy:** In figure 26b, the three regions are arranged in a linear fashion so that region 2 is more central and the distance between 1 and 3 is greater than any other interregional distance. In this case region two will always have a higher level of monopolistically competitive production because of its superior accessibility to markets. If the transport cost between 1 and 2 is reduced the effect on region 2 will be small, but region one will benefit in terms of both wages and output at the expense of region 3.
• *Networks and superadditivity:* A further simulation in which the transport costs from 1 to 2 and from 2 to 3 were reduced simultaneously produced an especially important result. The welfare gain from the simultaneous reduction yielded welfare gains that were more than the sum of the welfare gains from reducing the two transportation links independently.

*Implications For Evaluation Methodologies*

The results of the Venables and Gasiorek (1999) simulations have a variety of implications. For one thing, they reinforce the message that in an imperfectly competitive world, conventional CBA analysis tends to underestimate benefits of transportation investments. They also show that specific transport cost reductions can have spatially variable impacts and that reductions that are beneficial in the aggregate can be detrimental to certain places – a result that economic historians will hardly find surprising! More generally, they illustrate the point we made earlier that the benefits accruing from an infrastructure investment depend on the context within which the investment is set. For example, the effect of a transport investment on relative regional incomes depends on their initial values and the initial degree of economic integration. Effects of improving a single link in a multiregional network depend on the spatial configuration of the regions. And the benefits accruing from one infrastructure project may depend on other projects that are being carried out simultaneously.

But what is the implication for the evaluation of individual projects? The importance of context as highlighted in these results suggests that is unlikely that a simple one-size-fits-all framework can identify all the potential benefits and costs of non-marginal infrastructure projects. Is it therefore necessary to abandon CBA and adopt more complex and flexible CGE models as the standard framework?

To answer this, we must be mindful of the fact that both the theoretical and especially the empirical development of these models is in its infancy. CGE models as they exist today have a number of down sides:
• They are complex and highly demanding of data, much of which is not easily available;
• They are generally calibrated by means of a zero degrees of freedom method rather than by means of multivariate statistics, so there is no way to assess how well they fit the data;
• Because they involve so many nonlinear relationships, signs of effects often change on critical values (the inverted-U shaped relationship above) and therefore results are highly volatile and sensitive to small changes in parameter values.

For these reasons, CGE models are probably best used in the kind of illustrative mode in which Venables and Gasiorek used them. There strength is in pointing out the kinds of effects that are likely to be important, rather than in predicting specific outcomes of specific events.

The most prudent course is therefore to continue using CBA -- including the most current methods to adjust for externalities, demand shifts etc. -- as an initial quantitative estimate of the efficiency of any project, but to follow up with more qualitative analyses to determine whether the CBA results are likely to underestimate or overestimate true net benefits. In these more qualitative analyses, the types of issues pointed out in the new economic geography should be given special attention.

**Technology Shifts**

There have been times in history when expanded freight services have made radical changes in the structure of production possible. For example, the development first of canals and later of railroads made it possible for huge areas of the central lowlands of the US to be developed for specialized agriculture serving a national market. To a degree this fits into the standard comparative advantage argument described above, except that rather than a shift from autarky to specialization it involved the creation of new economic regions whose growth was driven from an early stage primarily by export commodities. Furthermore, it involved a fundamental transformation of production technologies, achieving much higher productivities through specialization and large-scale
production. It can be argued that a host of improvements in agricultural technology were induced, at least in part, by the expanded market opportunities made possible by freight improvements.

Another example is the industrial revolution in textile production that occurred on a global scale in the 19th century. In this case, improved freight made it possible to develop a production system that required the movement of materials inputs of cotton from production region (US South, Egypt, India) to production centers in England and New England. Thus, unlike the example above where freight made it possible for a specialized production region to reach broader markets, in this case freight made it possible for widely separated but complementary regions to be integrated into a specialized production system. Again, this story has elements of comparative advantage, but it involves a fundamental shift in technology made possible by improved freight. The key issue is this, while both the Heckher-Ohlin theorem and the new economic geography assume production technologies which are exogenous and fixed, historical examples suggest that new trade opportunities have at times given rise to technological shifts, resulting in an endogenous change in the production technology.

The two historical examples illustrate the role that freight improvements play in fundamental shifts in technology. It is likely that freight improvements provide opportunities for more marginal shifts on an ongoing basis. For example, the interstate system not only allowed producers to seek out locations with lower land costs, but also allowed them to implement more space-intensive technologies to enhance efficiency. Reductions in the cost of global trade due to innovations such as containerization set the stage for a general transformation to global production systems whereby inputs and components are sourced internationally.

It would be very difficult to predict such technological impacts for any specific infrastructure project. It is very possible, however, that this effect plays a significant role in explaining the link between transportation infrastructure investment and productivity growth as observed in the macroeconomic studies described earlier.
While it almost impossible to predict technical shifts, a lot could be learned from \textit{ex post} empirical studies that attempt to identify them. Case studies of industries that have undergone rapid transformations in production technologies or logistical organization could be seek to determine whether technological progress was either enabled or spurred on by new or improved freight transportation options. Such studies need not be limited to goods producing industries, but could also include large scale retail operations which are currently in a phase of rapid technological transformation.

\textbf{Lessons Learned}

Freight transportation continues to play a critical role in the U.S. economy. In recent years this role has been reinforced by qualitative changes in the nature and scope of freight services offered. Not only have the costs of freight services declined, but firms in the freight service sector now offer a broader range of enhanced services allowing freight-using firms more flexibility to restructure their logistical and production activities, and thereby achieve non-transportation cost reductions. This is in large part the outcome of novel applications of IT in the freight service sector and continued public investment in transportation infrastructure. Given its central role in the development of the highway network, FHWA has a critical interest in a better understanding of the role of transport investments in enabling freight service firms to achieve these logistical improvements and related efficiencies, which not only enhance these firms' productivity, but also that of transport using firms in the larger economy.

Prior to the seminal work of Koichi Mera in the 1970’s, assessments of the economic impacts of investments in transportation infrastructure were limited to appraisals of individual projects. Such appraisals provided relatively little insight into the broader role that transportation infrastructure plays in aggregate economic growth and productivity. Despite variations in data, methods, and the magnitude of the effects they uncovered, the sequence of macroeconomic analyses conducted in the U.S. and abroad
over the past twenty five years has identified a persistent, positive influence of transportation investments on aggregate economic performance.

While it is important to know that highway and other transportation investment programs are conferring economic benefits, macroeconomic studies tell us relatively little about the actual mechanisms through which these benefits arise. Policy formulation must address not only the question of whether to invest in infrastructure, but also the question of which among an array of potential projects will yield the greatest economic return. In order to answer this second question, it is necessary to open up the “black box” of macroeconomic studies and attribute economic benefits to specific mechanisms that may vary across projects due to location, network relationships, and other contextual factors.

In this paper we have looked at the underlying mechanisms from two perspectives: the microeconomic (partial equilibrium) perspective and the general equilibrium perspective. From the microeconomic perspective, individual firms benefit from cheaper, better, and faster freight services – benefits that can be captured in a conventional CBA framework. But cheaper and better transportation services may lead to savings in non-transportation inputs as well. For example, presented with lower transportation costs producers may choose to reduce inventories or consolidate facilities, even if it means consuming more transportation services. Important work is underway to incorporate these types of effects into CBA calculations.¹ There are still other possible benefits that, to date, have not been incorporated into the CBA model. For one thing, improved infrastructure increases the locational flexibility of firms, which in turn can lead to a variety of efficiency improvements. Also, there are a number of ways in which improved infrastructure allows firms to add value – this includes both providers of transportation services and freight-consuming producers of other goods and services.

The general equilibrium perspective highlights a different kind of benefits from improved transportation. These benefits arise from economy-wide adjustments and redistributions. The key notion here is gains from trade, whereby aggregate efficiency is enhanced when cheaper or better transportation promotes interregional and international...
specialization and trade. The theory of comparative advantage tells us that producers and consumers are better off when each region specializes in those goods and services it can produce most effectively. High quality, affordable transportation makes this possible. The “new economic geography” shows that in the presence of scale economies and imperfect markets, reduced transportation costs can lead to a host of economic transformations that yield aggregate economic benefits. One of the most important lessons from this emerging line of theory is that the impacts of transportation improvements are context dependent. So, for example, the outcome of a new corridor connecting two regions depends on such things as the state of the pre-existing transportation network; the relative size, wage level, and state of economic development of the two regions; and the degree of type and competition of markets functioning in both regions. Clearly, if this is the case, economic assessments must incorporate a broader range of interrelationships and data than is typical in current practice.

Beyond conventional gains from trade, better transportation can also lead to major shifts in technology that bring about improvements in aggregate efficiency. Specialized commercial agriculture, the industrial revolution, and the globalization of production all represent technological transformations that would not have been possible without non-marginal improvements in transportation systems. Smaller, more incremental technological shifts most probably arise with each successive boost in transportation performance.

With the exception of the direct cost and time savings that are captured in conventional CBA, our ability to measure any of the main categories of benefits described here is relatively poor. Many of the impacts we describe have only been derived from theory or demonstrated by means of hypothetical simulation. Some experts, while conceding that a broad range of indirect benefits may exist, argue that the values of these benefits are probably small and therefore conventional CBA is sufficient. Others counter that once you abandon assumptions of perfect competition and constant returns to scale, indirect effects can be cumulative and large.
Given our degree of uncertainty about many of these benefits, research along two avenues is warranted. The first is the expansion of our analytical toolbox to include a broader range of economic mechanisms. This includes further elaboration of CBA to capture the effects of logistical transformation, productivity-enhancing location shifts, and value-added effects. It also includes the development of more comprehensive frameworks such as CGE (Computable General Equilibrium) models. While the application of these models to transportation analysis is still a nascent field, a lot can be learned from operational models that are currently applied to international trade, tax policy and a number of other fields.

The second avenue of research comprises *ex post* assessment of major infrastructure projects and programs. By means of a more historical perspective we can ask a number of critical questions such as: To what extent did improved infrastructure lead to increased specialization and expansion of markets? How do freight service firms take advantage of improved infrastructure to offer cheaper and better services? What logistical, technological and locational innovations followed in the years after the project’s completion? How did such economic adjustments translate into higher productivities and incomes? Naturally, such assessments must be more than just case studies; they must apply appropriate analytical methods to identify those economic changes that can be attributed to the infrastructure improvement from those that would have occurred without it.

Such a research program is ambitious, but its payoffs may be great. Its goal will be to create both a better understanding of the role of transportation in the economy and better analytical capacity to support more informed decisions about transportation infrastructure decisions in the future, and an increased capacity to serve the Department of Transportation Strategic missions of improving the national economy and national competitiveness.
REFERENCES


Mera, K., 1975. *Income Distribution and Regional Development*, University of Tokyo Press, Tokyo, Chapter IV.


"Infrastructure, Economic Growth and Regional Development: The Case of highly Industrialized Developed Countries," Jönköping, Sweden, June.


SACTRA (Standing Advisory Committee on Truck Road Assessment), 1999. *Transport and the Economy*, United Kingdom, DETR (Department of Environment, Transport and Regions) London.


---

2 See for example Mackie and Nellthorp, 2001