Market access rather than travel-time-saved: rethinking transport project benefit-cost analysis

Cameron Gordon
Visiting Professor – ITLS – University of Sydney
Associate Professor, Economics, University of Canberra
Principal Investigator, University Transportation Research Center Region 2, City University of New York

The Basic Question

- Does benefit-cost analysis fully and adequately capture the full ‘efficiency’ effects of proposed transport investments?
- My basic answer is: sometimes yes, but other times, no.
- The overall thrust of my argument will be that B-C analysis is a useful tool, and applied in the ‘traditional’ way is often ‘accurate’, but that traditional methods often need to supplemented or replaced when there are likely to be significant access or agglomeration or other spatial and network impacts.

Outline of the argument

- Framing the issues of ‘standard’ analysis – the ‘social savings debate’
- B-C and three issues: embodied effects, topography and market access/network effects
- Time and space as interdependent rather than independent
- The meaning of access and accessibility
- Agglomeration economies and spatial externality
- When to use ‘standard B-C analysis and when to ‘extend’ it.
- Some possible ways of augmenting B-C analysis

A framing of the issues: Fogel and the ‘Social Savings’ controversy

- Traditional American economic history research held that the coming of the railroads were indispensable in making America a modern industrial power.
- That was the thinking until the economist Robert Fogel wrote his book *Railroads and American Economic Growth* in 1964 (work that he would later win a Nobel Prize for).

The concept of ‘Social Saving’

- Fogel chose a measure of indispensability that he called the “social saving” due to railroads.
- Social saving generally refers to a comparison of the resources saved by a society when it has the use of a technology with the situation that would prevail if it did not have the use of that technology.

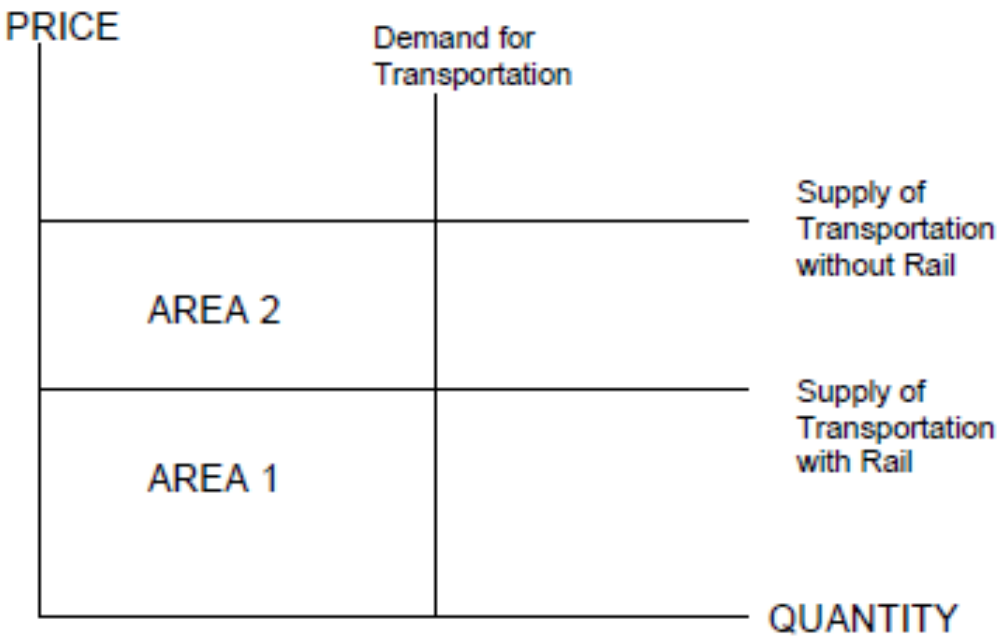


FIGURE 1 Demand for transportation.

Source: Gordon, 2006

- There are two supply curves: the supply of transport available in the presence of railroads and the supply of transportation when railroads are unavailable and when only alternatives, such as canals, can be used.

  The amount spent on transport is equal to the appropriate price times the quantity.

  With railroads available, the amount spent on transport equals Area 1.

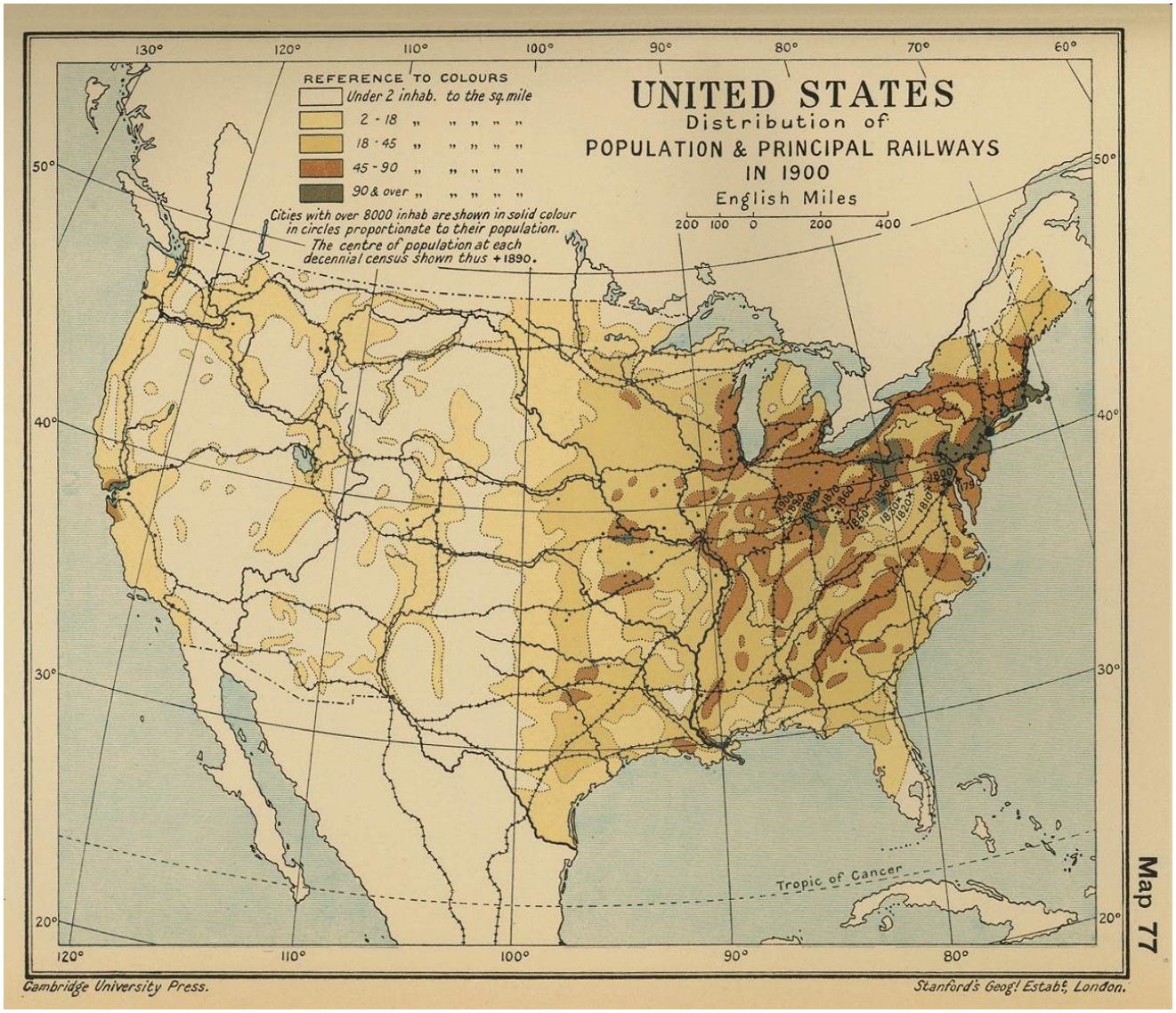
  Without railroads, the amount spent on transport equals Area 1 plus Area 2.

  The social saving, that is, the resources that railroads free up for social uses other than transportation, is equal to the resources that would have been spent on transport without railroads minus the resources that are spent on transport with railroads.
- It should be clear that this difference is Area 2.

The counterfactual

- Fogel looked at the US economy up until 1890 with railroads available and then in 1890 suddenly eliminated them from the economy, with alternative, and more expensive, transport methods such as canals then required.
- He measured the effect on US GNP in that year under several simplifying assumptions (i.e. the social saving).
- Fogel found that the social saving due to the railroads amounted to under 3% of GNP, a significant but not overwhelming figure.
- He argued that by choosing a peak year of agricultural production, when railways would be most utilised, and also by assuming perfectly inelastic demand for transport that he would ‘stack the deck’ in favour of railways so that any estimate he came up with would arguably be on the high side and hence an overestimate of social saving.

This only ‘saved’ 3%?



Back to the future

- It should be clear by now that Fogel was basically doing a benefit-cost analysis of an entire network.
- His assumptions were especially simplistic (to make a point) and he was ‘backcasting’ rather than ‘forecasting’
- But his method was very ‘standard’ and therefore instructive of the strengths and limitations of traditional B-C analysis as are the many objections made to it.

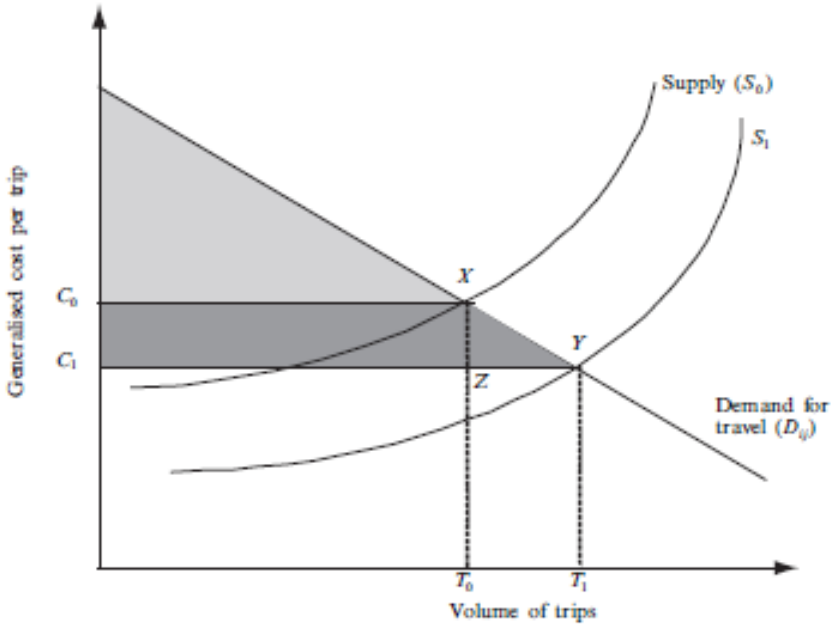


Figure 1. The basic cost–benefit analysis model.

Source: Vickerman, 2007

Basic benefit-cost assumptions

- Demand for travel is derived and a source of disutility.
- The major benefit of transport improvement is the reduction in travel costs (typically the ‘Value of Travel Time Saved’ – VTTS).
- Partial equilibrium (typically)
- Perfect competition
- Instantaneous adjustment across featureless time and space
- Time and space are independent dimensions
- Perfect divisibility of factors (primarily labour and public/private capital)
- Constant returns to scale
- Perfect information
- Rational behaviour
- No ‘strategic behaviour’
- Efficient private markets (no market failure)

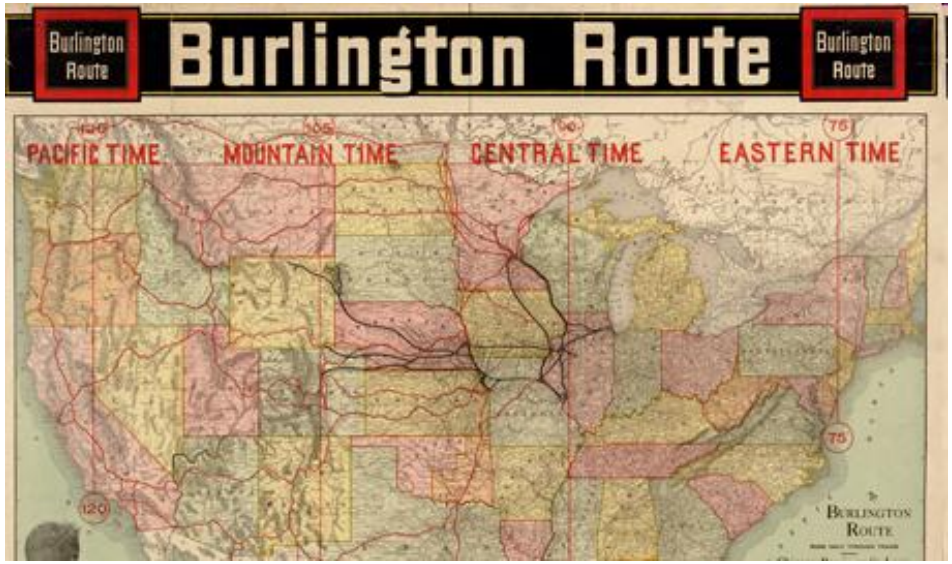
Fogel's backlash…and what it means for current B-C analysis

- Back to Fogel: there were many objections to his method and estimate (and the literature is very worth reviewing because of its empirical, conceptual and theoretical richness and care. Fogel's 1979 review and retort is a good place to start).
- The objections most relevant for the discussion today focus on three main points:
- (1) "Embodied" v "Disembodied" effects
- (2) the role of topography and geography
- (3) Market access and network effects

1. Embodied effects

- Fogel referred to the social saving as a disembodied effect, because any true transport innovation lowers costs and hence results in savings to society.
- However, railroads delivered these savings in a particular form, and the form which savings due to railroads, as opposed to a different transport innovation that might have occurred, are the embodied effects.
- Thus a relatively insignificant social saving might result in otherwise dramatic changes in a society.
- In other words, while total output might have remained fairly similar in the absence of railroads (assuming Fogel's estimate is correct), the economy would nonetheless have looked dramatically different without the specific form of effects that railroads delivered.

- Alfred Chandler and Oliver Williamson both argued that railroads produced not only transportation efficiencies, but organizational and institutional innovations that, when combined with the delivery of speedier and more regular goods and passenger carriage, made a system of mass distribution possible, promoted the growth of mass manufacturing, led to the development of the large, vertically integrated corporation and even established standardised time zones



- Effects such as these are hard, if not impossible, to capture in models that rest on consumption, production, and relative prices alone.
- Yet these may the as important as average net gains or net losses (and will have specific flow-on effects, negative and positive).
- (Chandler 1977; Williamson, 1985)

2. The role of topography and physical space

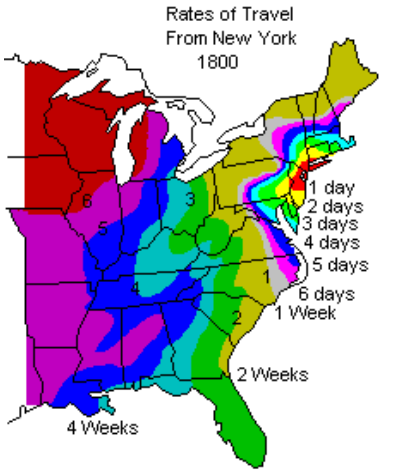
- Social saving due to rails was generally greatest over shorter hauls, where wagons and roads were the main alternative, rather than longer hauls, where waterways were often quite competitive and continued to improve in productivity over the late 19th century.
- Indeed, in this period canals continued to be built in Belgium, France, and Germany.
- Thus, social saving was greatest in countries that relied most on roads—productivity in this sector remained low until the development of the internal combustion engine—and lowest where there were many navigable rivers, terrain well-suited to building canals, and good facilities for coastal trade.
- In general, railways alleviated poor natural endowments. Where these endowments were rich, rails had relatively little effect. Where these endowments were poor, railroads saved society more in terms of transport costs

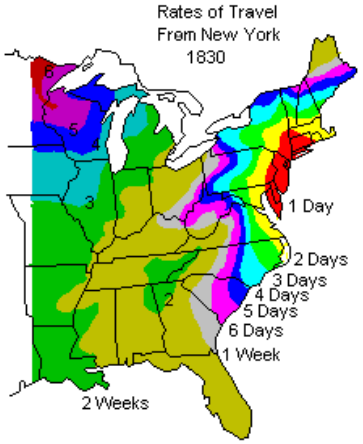
Andrew J. Russell Railway construction in the Green River Valley, Wyoming, near Citadel Rock, 1868. Albumen print

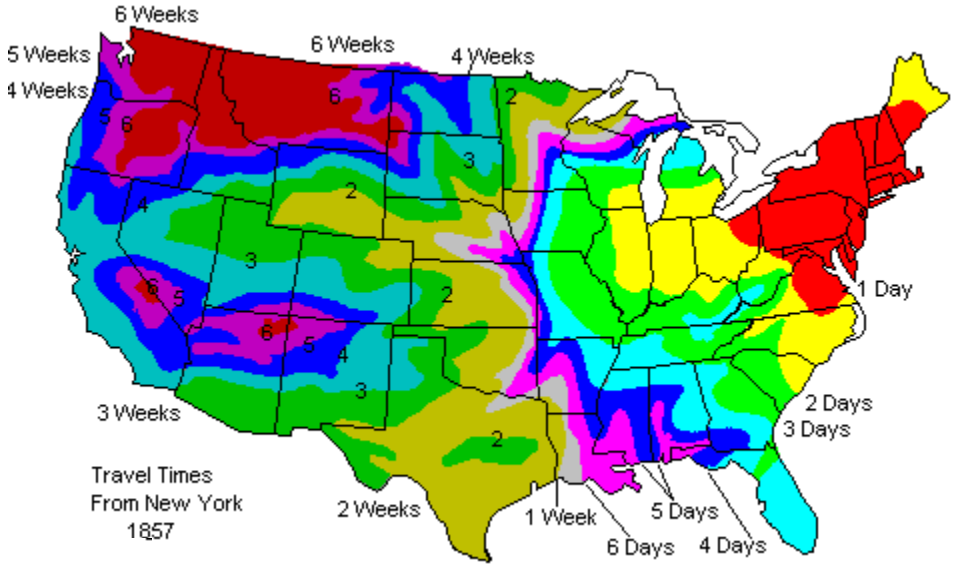
3. Market Access and Network Effects

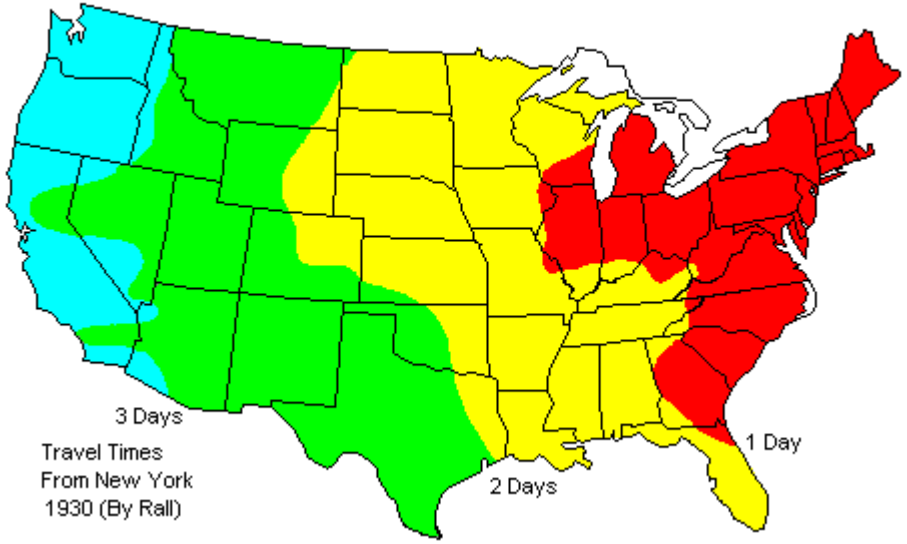
- So embodiment matters…and natural setting matters (arguably the ‘natural’ embodiment in one case, the ‘human’ embodiment in the other) which leads to the issue of how the railways network changed access to markets for both inputs and outputs.
- Fogel’s analysis was very unusual in transport because it explicitly looked at the effects of removal of an entire transport network from an economy which then had to fall back on its previous network (and in this case using a less advanced technology).
- This, of course, is not the question posed in most transport B-C analyses – but it is an important background issue that often needs to be considered in the foreground because networks don’t just reduce travel times…they reshape and reform access.

Did these US transport improvements primarily deliver ‘travel time saved?’





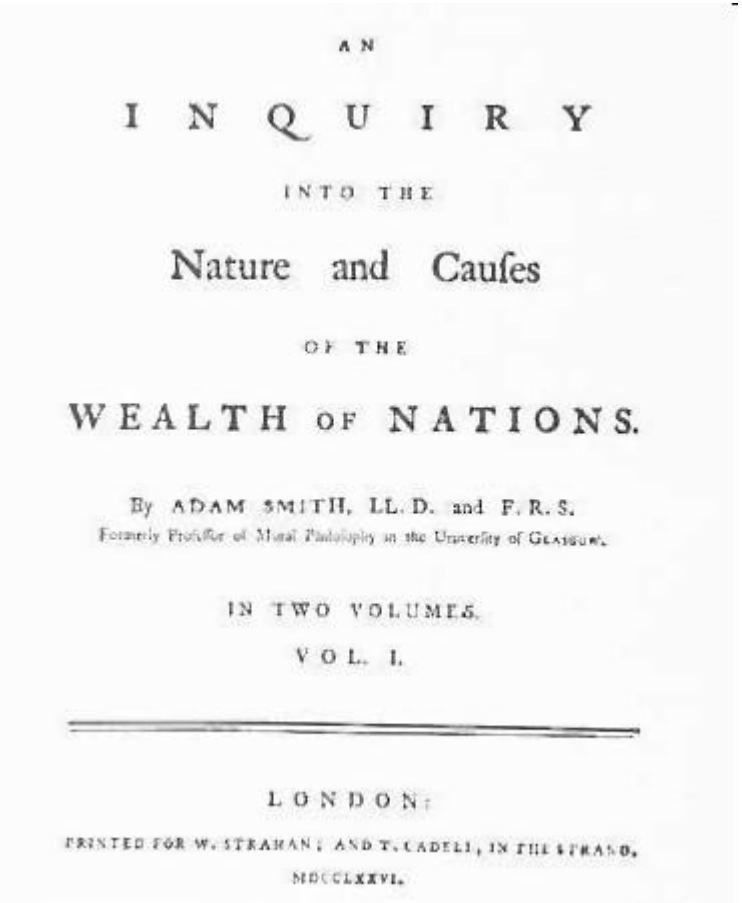




Source: Historical Atlas of the US

‘The extent of the market’

- “As it is the power of exchanging that gives occasion to the division of labour, so the extent of this division must always be limited by the extent of that power, or, in other words, by the extent of the market.”
- (Adam Smith, Chapter 3, The Wealth of Nations)
- In many ways the classical economists and the neo-classical economist Alfred Marshall in particular were thinking about the interaction between economies (internal and external) and the market.
- Marshall himself explicitly brought the spatial dimension into the discussion when in his ‘Principles’ he mentions ‘localized skills’ and ‘local trade secrets’ as key factors in spatial concentration (what later became known as ‘Marshallian Externalities’.



AN

INQUIRY

INTO THE

Nature and Cauſes

OF THE

WEALTH OF NATIONS.

By ADAM SMITH, LL.D. and F.R.S.

Formerly Profeſſor of Moral Philoſophy in the Univerſity of GLASGOW.

IN TWO VOLUMES.

VOL. I.

LONDON:

PRINTED FOR W. STRAHAN; AND T. CADELL, IN THE STRAND.

MDCCLXXVI.

Access and Accessibility

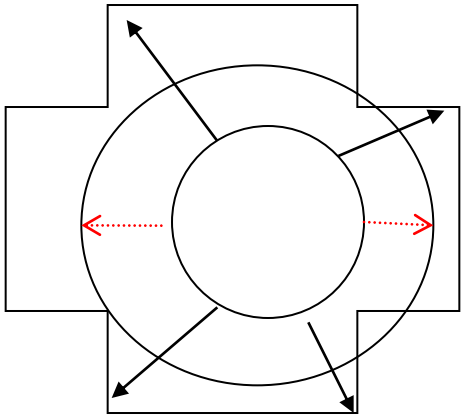
- A key part of market extent is ACCESS and ACCESSIBILITY.
- “Accessibility…has taken on a variety of meanings.
- …the potential of opportunities for interaction (Hansen, 1959),
- the ease with which any land-use activity can be reached from a location using a particular transport system’ (Dalvi and Martin, 1976),
- the freedom of individuals to decide whether or not to participate in different activities (Burns, 1979)
- the benefits provided by a transportation/land-use system’ (Ben-Akiva and Lerman, 1979).

- “Focusing on passenger transport, we define accessibility as the extent to which land-use and transport systems enable (groups of) individuals to reach activities or destinations by means of a (combination of) transport mode(s).”
- “Furthermore, the terms access’ and accessibility’ in the literature are often used indiscriminately. Here, access is used when talking about a person’s perspective, accessibility when using a location’s perspective.”
- (quoted from Geurs and Van Wee, 2004)

Space-time continuum

- One basic implicit assumption underlying ‘standard’ B-C analysis is that time and space are independent of each other.
- In effect the spatial plane remains fixed while transport improvements reduce the amount of time it takes to get to different points on that plane.
- However this Newtonian notion of absolute time and space has been supplanted by Einstein’s ‘space-time continuum’ in which the two are interdependent (a notion elaborated on by Stephen Hawking, amongst others).
- This may seem esoteric, but it has direct relevance for transport investment and service analysis.

The reshaping of space



- Thus one can think of space and time as interdependent.
- Change one and you change the other.
- So in a classic Von Thunen space (the featureless plane), a transport improvement increases the diameter of the market circle.
- In the ‘real world’ with topography and other ‘featured space’ the shifts are more textured and unpredictable (and may, in some cases, lead to loss of access for parts of the original market area).

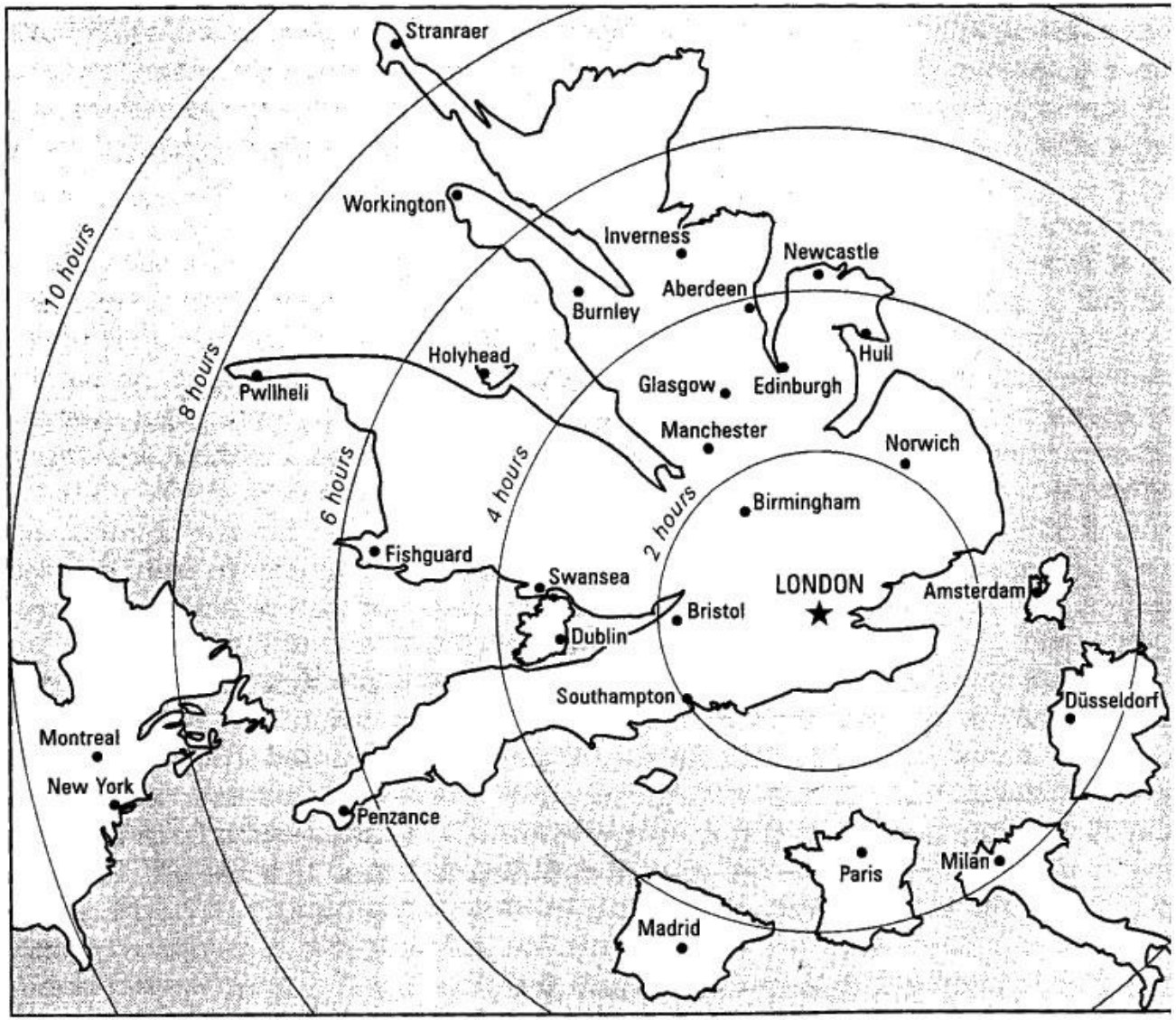


Figure 5.5 The unevenness of time-space convergence
Source: Based on Dicken **and** Lloyd (1981, Figure 2.7)

- Geographers are well aware of this.
- And so are many who study the meaning of ‘access’ and ‘accessibility’
- Note here travel times interact with space to make some points within the UK more ‘distant’ than some international journeys.

Time-space convergence

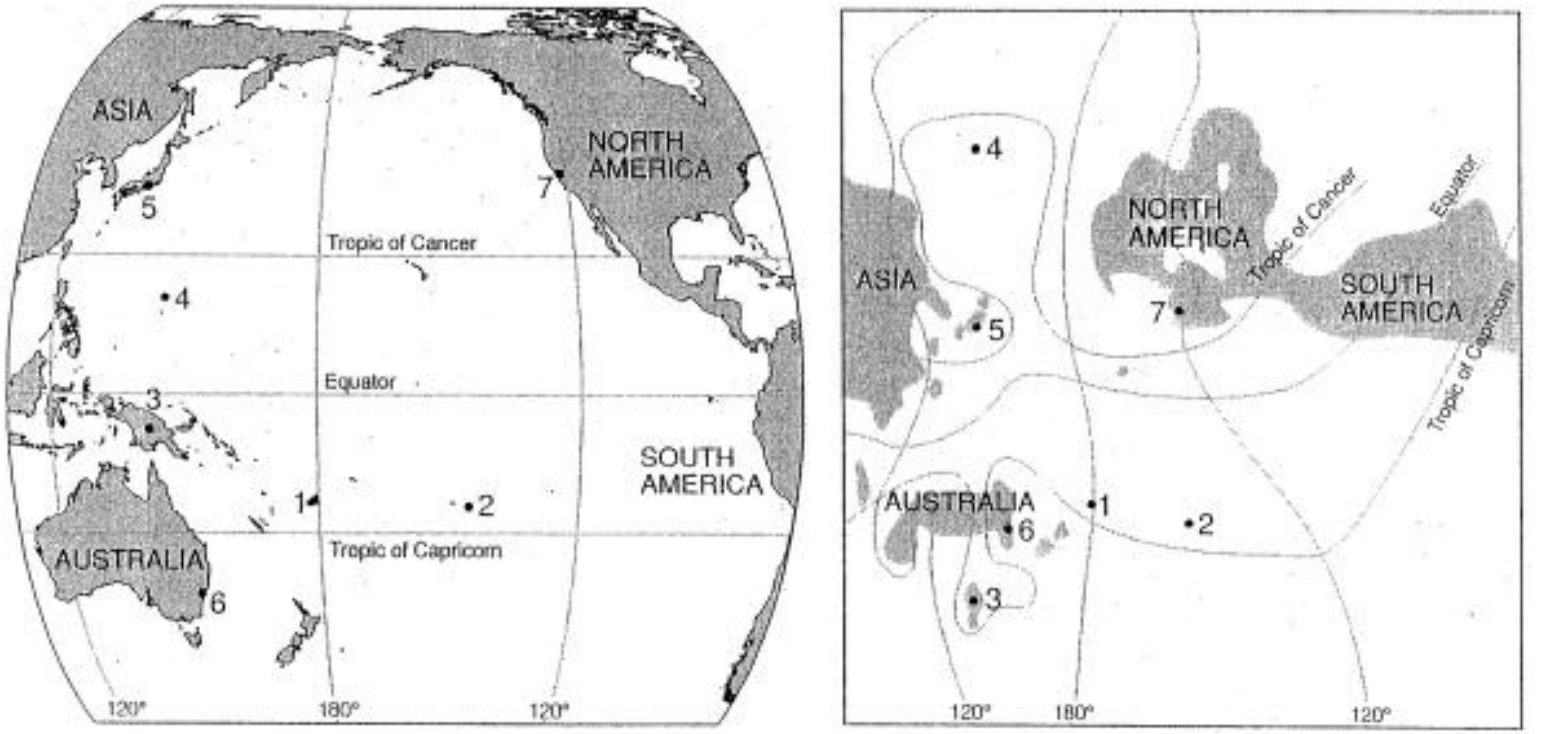


Fig. 4. Time–space map of the Pacific basin based on relative time accessibility by scheduled airline in 1975. (*Source:* Haggett, 1990, Fig. 3.3, adapted by Leyshon. Reproduced with permission.)

R.D. Knowles / Journal of Transport Geography 14 (2006) 407–425

Source: Knowles 2006

- D. Janelle was one of the first to begin to formalise the study of this process of time-space reshaping. To quote from his seminal 1969 article:
- “The premise upon which the study was based was that man adapts the areal structure of his activities in response to changes in transport technology which enable him to travel faster and to have access to larger areas and to more resources.
- Given the assumptions of rationality as developed earlier in this study, the proposed theses seem tenable: 1) that time-space convergence is a significant factor leading towards spatial adaptation, and that 2) spatial reorganization will accelerate most rapidly at those places which stand to benefit most from increasing accessibility.”

- “However, in light of the limitations posed by the assumption of rationality, it is evident that the concept of relative advantage, as applied to the process of spatial reorganization, is not in complete accord with reality. …The decisions of man, however, are often conditioned by his lack of information, or non-maxima goals. In that much of man's areal development is presently directed by various government agencies…rather than by the demand for locational utility, it seems that complete understanding of the process of spatial re-organization may rest upon one's knowledge of the decision-making processes of these agencies.”

Janelle's model

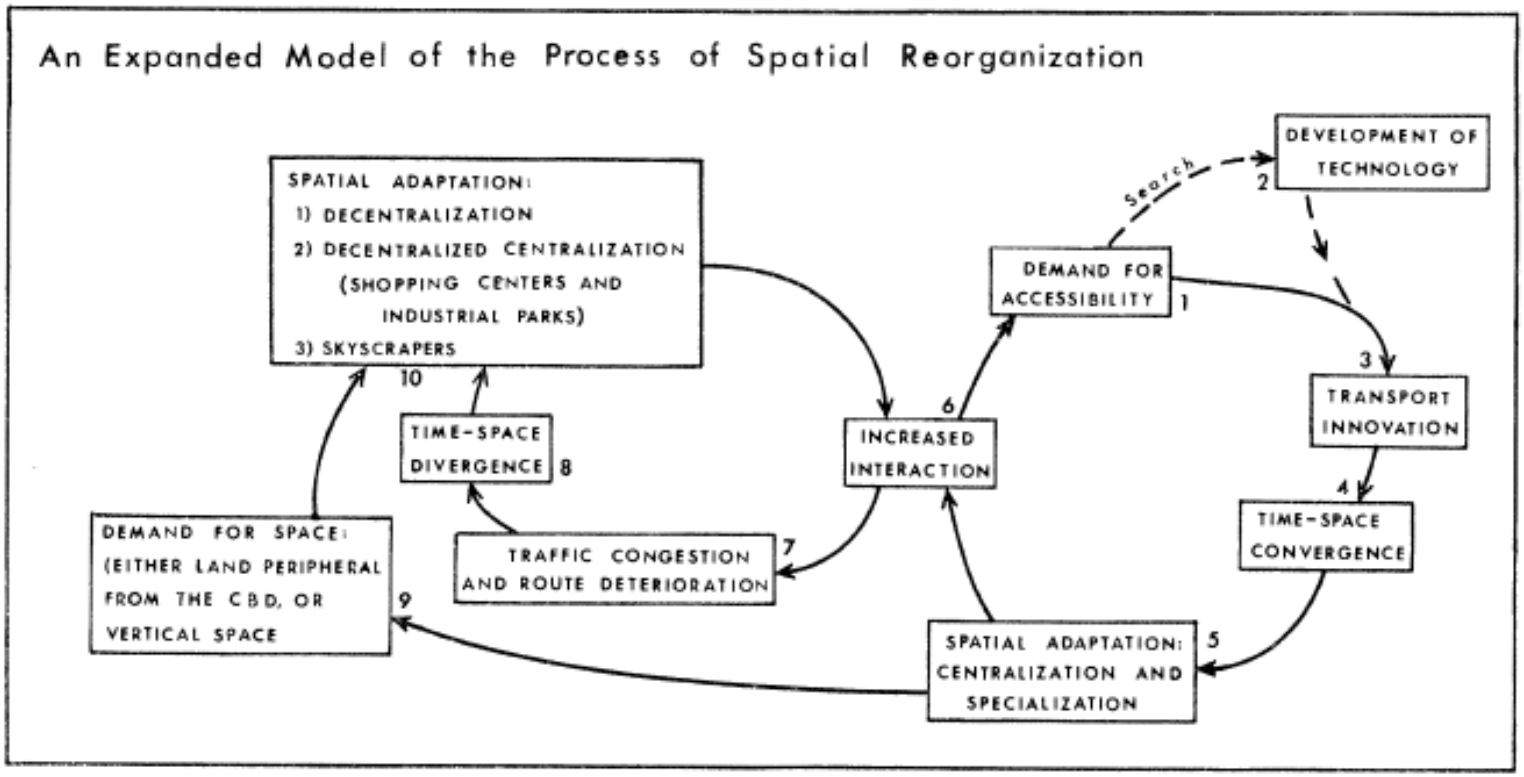


FIG. 5.

Real world complications

- Janelle is not the first to touch upon the importance of institutions in human behaviour.
- He mentions governmental institutions, but equally relevant are private firms, communities and incentive structures in general.
- Information is typically not shared equally across different institutions, motivations (‘objective functions’) typically clash, and ‘strategic behaviour’ is often present.
- Moreover even where institutions are functioning ‘efficiently’, imperfect competition, even something as mild as monopolistic competition, prevails in many factor, intermediate and final goods markets.

Agglomeration economies and externalities

- In fact, many of the economic effects that are associated with spatial density, especially agglomeration economies on the production side (typically broken down into localisation and urbanisation economies) and on the consumption side (e.g. worker utility increased by diverse urban amenities) are now believed to arise from imperfect competition, natural advantages built on by first-mover advantage and historical accident (a la the “New Economic Geography”).
- In particular large network infrastructure investments are indivisible and lumpy, have long time horizons and once built, are largely ‘irreversible’ for long periods of time.
- There is far from consensus here but any approach that abstracts away from these real-world complications may end up missing where the real action (positive or negative) lies.

When to use ‘standard’ B-C analysis

- All of this is not to say that ‘traditional’ methods are inappropriate or inapplicable.
- Although the classic Newtonian system has been shown to be a limited case (relativity being the more general case), the system remains very useful if applied correctly.
- Similarly for B-C analysis.
- The key is to think first about the situation being analysed and whether there are likely to be conditions or effects that may significantly diverge from the assumptions under which B-C analysis functions best.

Conditions under which ‘standard’ B-C is likely to work best

- Projects with no or few likely network effects.
- Relatively ‘small’ projects of relatively short duration in execution.
- Short-lived projects
- Projects designed primarily to speed flow over limited segments with no practical alternative route
- Projects in environments with few institutional distortions or rigidities.
- -- an intuitive ‘screen’ – would a private firm likely be attracted to this project based on a financial rate-of return analysis?

"I have the result of your cost-benefit analysis. You should have retired four years ago."

Conditions under which B-C analysis may need to be adapted

- Projects significantly altering or building an entirely new network
- Very large projects, of long duration in execution and/or very long-lived effects afterwards
- Projects in environments with institutional rigidities and distortions including imperfect competition
- Public-Private Partnerships (at least with respect to key variables such as discount rate since these may well differ between the partners)
- Projects primarily devoted to improving access (e.g. for a particular industry) or accessibility (e.g. making a peripheral location more ‘central’)
- Projects where there are likely to be ‘general equilibrium effects’

How to adapt B-C analysis?

- Of course the Holy Grail for an economist would be to model simultaneously the how the time-space (access) bundles for individuals are changed by different projects, aggregating up as needed.
- There is indeed such a literature – here time-space 'prisms' are shown capturing how a transport improvement increases both time and space choice sets – but in many cases detailed data needed are not available and solutions are not tractable.
- Policymakers also have trouble understanding such approaches even where feasible.

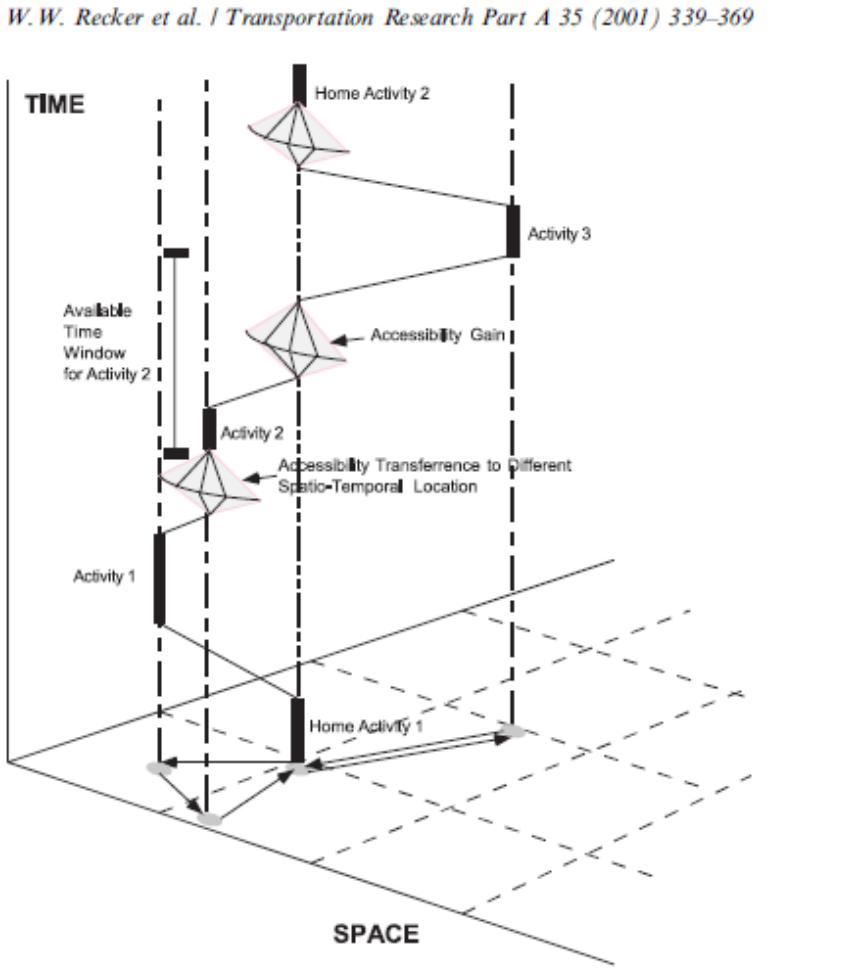


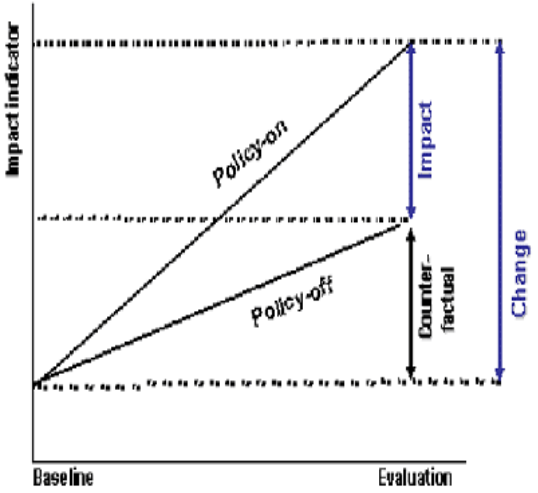
Fig. 2. Example of accessibility gains and transference with more efficient travel behavior.

Simpler adaptations

- Of course simpler adaptations might be used just as well, on a case-by-case basis.
- The Fogel debate is relevant once again.
- Jeffrey Williamson (1975) decided that modelling of dynamic, multisectoral effects was necessary (Fogel's analysis being a static partial equilibrium analysis) and used a computable general equilibrium (CGE) model to get at the broader effects of railways.
- Other historians did not look to models but looked more deeply at the data, in particular focusing on market and cost structures of competing modes (especially canals) to see how much P diverged from MC. They also looked to come up with reasonable transport supply and demand elasticity estimates.
- This information was then used to run sensitivity analyses on key variables to see how social saving estimates changed and what the key drivers of those changes were.
- Much analysis was 'qualitative' in the sense of looking at analogous situations elsewhere (e.g. overseas) to apply to the US situation to increase understanding of underlying dynamics.

Choice of baseline

- Most standard B-C analysis uses a simple ‘with/without’ benchmark to assess net benefits of a project.
- Fogel himself argued that he was not looking for a ‘realistic’ base case but one that would provide an ‘upper bound’ for the social saving estimate.
- This indicates that one can apply ‘standard’ B-C but choose a baseline that best answers the question at hand (e.g. ‘best’ or ‘worst’ case rather than true net gain) or that best reflects facts on the ground (for example, if a project has a very long execution time we may want to choose a baseline that reflects a situation later on, closer to project opening, when there may have been substantial change in the meantime, as opposed to a simple ‘no project now’ versus ‘project exists then’.



Adding metrics

- One may want to add various metrics to a B-C such as distributional analysis, accessibility measures and even maps of space-time contours.
- Showing how trip matrices might change during a project implementation (these are often assumed to be fixed between the start and end of a project) is also often useful, even necessary.
- Of course adding metrics and information to B-C is nothing new but perhaps should be approached more systematically, based on the criteria offered here.

- Returning to Geurs and van Wee's work (op. cit.) analysts can ask themselves how important accessibility gains and losses are likely to be for a given transport project and which of these drivers might be paramount.
- Additional metrics can be chosen accordingly.

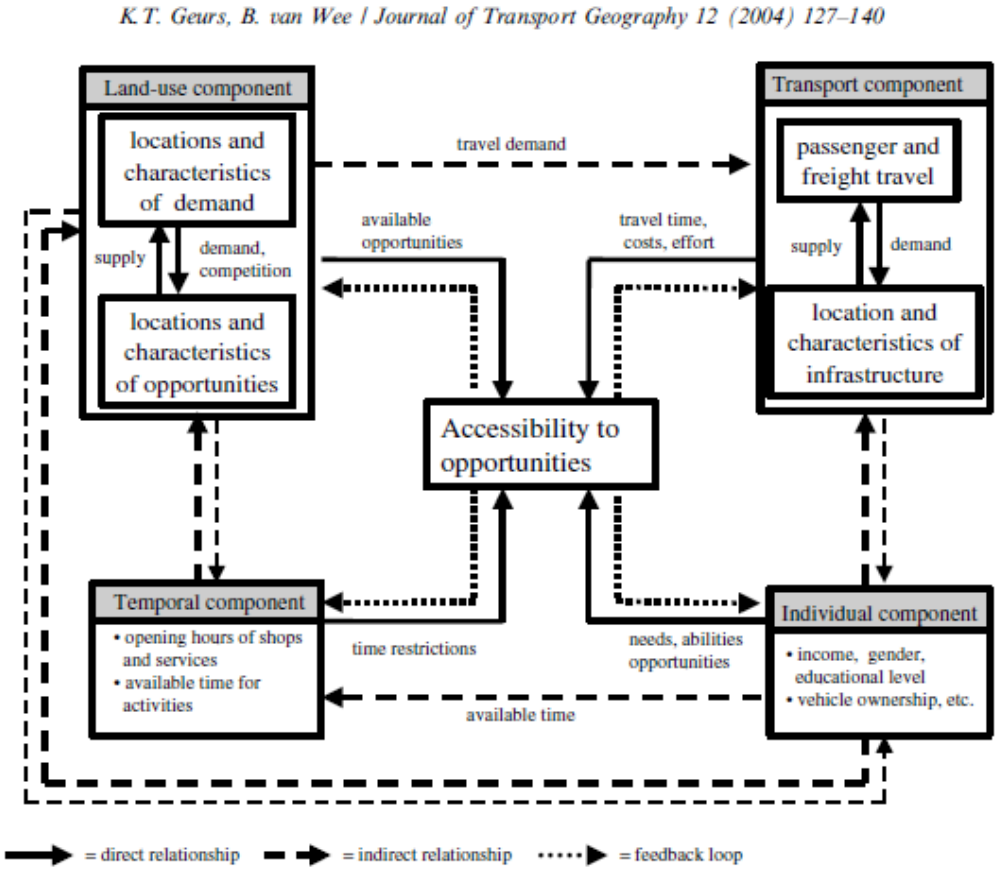


Fig. 1. Relationships between components of accessibility.

Adding relevant analytics

- Finally (and again the Fogel debate is a guide here), one can amplify the traditional analysis by adding relevant analytics.
- For example, Vickerman (2007) shows how B-C should be altered in the case of a simple monopolist, where MC and MB curves provide the market equilibrium, not S and D.
- The change in benefits is now the difference between the two rectangles EFGC'$_1$ minus ABDC'$_0$ which, depending on the relative elasticity of the demand curve and the price/marginal cost mark-up, this could be greater or smaller than the usual benefit C_0XYC_1

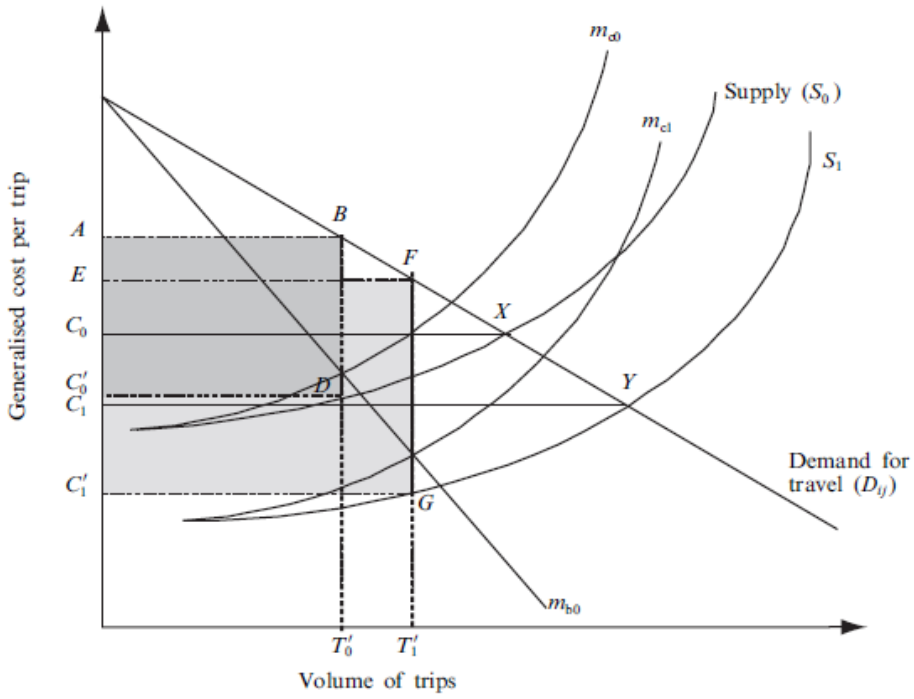


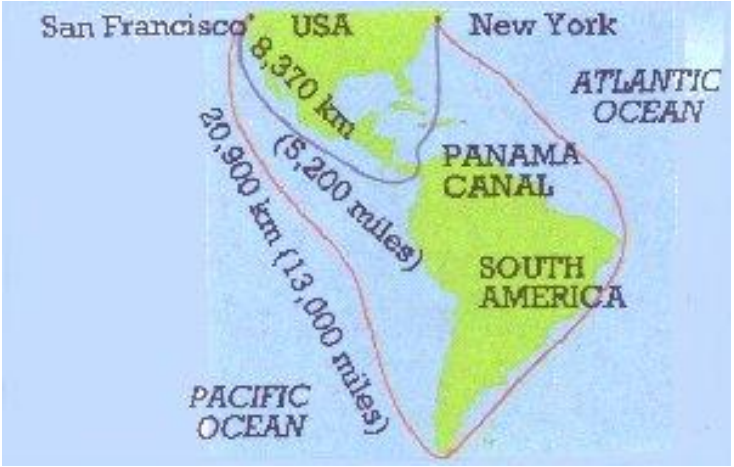
Figure 4. Cost–benefit analysis under imperfect competition.

Conclusion

- We know that transport infrastructure can shift access patterns and that this may often be its primary benefit, even beyond VTTS.
- We also know that many spatial effects occur as a result of transport investments, and many of these might actually be driven by conditions that deviate from standard B-C assumptions.
- ‘Standard’ B-C works well in many instances but its application should not be blind but should rest on a consideration of the likely facts on the ground, with methods adapted or supplemented accordingly.

Euro Tunnel

Panama Canal



References

- Chandler, A. (1977). *The Visible Hand.* University of Chicago Press, Chicago, Ill.
- Dicken, P. and Lloyd, P. (1981). *Modern western society: A geographical perspective on work, home, and well-being,* Harper & Row, New York.
- Fogel, R. (1979). Notes on the Social Saving Controversy. *The Journal of Economic History*, 39, pp 1-54.
- Fogel, R. (1964). *Railroads and American Economic Growth: Essays in Econometric History,* Johns Hopkins University Press, Baltimore, Md..
- 1964.Geurs, K.T. and van Wee, B. (2004). Accessibility evaluation of land-use and transport strategies: review and research directions, *Journal of Transport Geography*, Volume 12, Issue 2, Pages 127-140.
- Gordon, C. (2006). Assessing Adequacy of America's Transportation Policies: Lessons from Debate About the Role of Railroads in Development of the American West. *Transportation Research Record,* Vol. 1966, Pages 96-102.
- Janelle, D. (1969). Spatial reorganisation: a model and concept. *Annals of the Association of American Geographers* 59: 348-64.
- Knowles, R.D. (2006). Transport shaping space: dfferential collapse in time–space. *Journal of Transport Geography 14, 407–425.*
- Recker, W.W., C. Chen, M. G. McNally, (2001). Measuring the impact of efficient household travel decisions on potential travel time savings and accessibility gains, *Transportation Research Part A: Policy and Practice*, Volume 35, Issue 4, Pages 339-369.
- Vickerman, R. (2007). Cost- benefit analysis and large-scale infrastructure projects: state of the art and challenges. *Environment and Planning B: Planning and Design*, Colume 34, 598-610.
- Williamson, J. G. (1975). The Railroads and Midwestern development, 1870–1890. In *Essays on Nineteenth Century Economic History,* (David C.Klingaman and Richard K. Vedder, eds.), Athens, Ohio.
- Williamson, O. (1985). *The Economic Institutions of Capitalism.* Free Press, New York.