

# Cart or Horse: Transport and Economic Growth

04

**Discussion Paper 2011 • 04**

**Tim LEUNIG**

London School of Economics, UK

This document was produced as background for the 2011 International Transport Forum, on 25-27 May in Leipzig, Germany, on *Transport for Society*. The views expressed in this document do not necessarily reflect those of the member countries of the International Transport Forum.

Further information about the International Transport Forum is available at  
**[www.internationaltransportforum.org](http://www.internationaltransportforum.org)**

# **Cart or Horse: Transport and Economic Growth**

**Discussion Paper No. 2011-4**

**Tim LEUNIG**  
Professor  
London School of Economics  
London, United Kingdom

**April 2011**

## INTERNATIONAL TRANSPORT FORUM

The International Transport Forum at the OECD is an intergovernmental organisation with 52 member countries. It acts as a strategic think tank with the objective of helping shape the transport policy agenda on a global level and ensuring that it contributes to economic growth, environmental protection, social inclusion and the preservation of human life and well-being. The International Transport Forum organizes an annual summit of Ministers along with leading representatives from industry, civil society and academia.

The International Transport Forum was created under a Declaration issued by the Council of Ministers of the ECMT (European Conference of Ministers of Transport) at its Ministerial Session in May 2006 under the legal authority of the Protocol of the ECMT, signed in Brussels on 17 October 1953, and legal instruments of the OECD.

The Members of the Forum are: Albania, Armenia, Australia, Austria, Azerbaijan, Belarus, Belgium, Bosnia-Herzegovina, Bulgaria, Canada, Croatia, the Czech Republic, Denmark, Estonia, Finland, France, FYROM, Georgia, Germany, Greece, Hungary, Iceland, India, Ireland, Italy, Japan, Korea, Latvia, Liechtenstein, Lithuania, Luxembourg, Malta, Mexico, Moldova, Montenegro, Netherlands, New Zealand, Norway, Poland, Portugal, Romania, Russia, Serbia, Slovakia, Slovenia, Spain, Sweden, Switzerland, Turkey, Ukraine, the United Kingdom and the United States.

The International Transport Forum's Research Centre gathers statistics and conducts co-operative research programmes addressing all modes of transport. Its findings are widely disseminated and support policymaking in Member countries as well as contributing to the annual summit.

## DISCUSSION PAPERS

The International Transport Forum's Discussion Paper Series makes economic research, commissioned or carried out at its Research Centre, available to researchers and practitioners. The aim is to contribute to the understanding of the transport sector and to provide inputs to transport policy design. The Discussion Papers are not edited by the International Transport Forum and they reflect the author's opinions alone.

The Discussion Papers can be downloaded from:

[www.internationaltransportforum.org/jtrc/DiscussionPapers/jtrcpapers.html](http://www.internationaltransportforum.org/jtrc/DiscussionPapers/jtrcpapers.html)

The International Transport Forum's website is at: [www.internationaltransportforum.org](http://www.internationaltransportforum.org) or further information on the Discussion Papers and other JTRC activities, please email: [itf.contact@oecd.org](mailto:itf.contact@oecd.org)

**TABLE OF CONTENTS**

EXECUTIVE SUMMARY.....	4
MOTIVATION – TRAVEL AS AN EXPRESSION OF BEING HUMAN.....	7
LESSONS FROM HISTORY.....	8
1. Lessons from History 1: Transport as a new good .....	8
2. Lessons from History 2: Steadily diminishing returns from expansion .....	10
3. Lessons from History 3: Transport and the location of industry .....	11
4. Lessons from History 4: The value of transport relates to underlying economic conditions .....	12
TRANSPORT STATISTICS .....	16
POLICY IMPLICATIONS.....	19
CONCLUSIONS.....	23
BIBLIOGRAPHY .....	24
ANNEX: CONCEPTUALISING TRANSPORT IN THEORETICAL ECONOMIC TERMS .....	27

## EXECUTIVE SUMMARY

This paper argues that transport is more cart than horse, in that transport improvements are not the most important driver of economic growth for most countries. Nevertheless there are circumstances in which transport is particularly important. Big transport breakthroughs – such as replacing walking with railways, or creating a highways network for the first time – do have big effects, but these are unlikely to be seen again in developed economies. Instead transport in developed economies is best seen as having a supporting role. If it is neglected, it can constrain growth, as congestion and unreliable transport systems can exact a heavy price. But as long as the transport system is “good enough”, the returns to greater transport investment will be relatively limited.

The paper is based around a number of historical case studies. We explain why freight railways were worth more in South America than North America (North America had canals which were capable of doing a similar job), or why passenger railways were worth more in the UK than in the US or Russia (towns were too far apart in the US, and Russia was too poor). We look at the rolling out of the US and other highways network and show that transport has diminishing marginal returns – the first bits are very valuable, but after a while additional improvements yield relatively small economic returns. We look too at the role of transport in determining the location of industry – it matters sometimes, but not always, and this literature is not sufficiently developed for us to be able to offer clear predictions for other situations. Finally we look at the role of new transport in old cities, and find that the benefits are greatest when a city is so economically strong that it needs to expand, but that urban transport is unlikely to play an important part in regenerating struggling cities.

We draw a number of policy implications. The first and most important is that the effect of an improvement in transport can only be understood in the context of understanding the existing transport infrastructure, and in the context of the level of economic development that already exists. Building transport too far ahead of demand will not offer good value for money.

Second, we know that speed matters, but the definition of speed is not the top speed by the average speed, door to door, factory gate to destination. That is useful in that there are good reasons why top speeds are unlikely to increase dramatically, whereas average speeds can be increased by eliminating bottlenecks. Congestion is a particular issue, not only because it slows down travel, but because it is unpredictable, which adds a further set of costs as journey times become less reliable. In developed economies these problems are usually worst in major urban cities, where the number of people travelling is high, where the land available for transport is low, and where the cost of improvements – for example tunnels, for roads or metro lines, is phenomenally expensive. We need more intellectual effort to be devoted to solving bottlenecks, and urban bottlenecks are the most challenging.

This report argues that case as follows. It begins with a short section motivating the report, and outlining the extent to which travel has changed over the long sweep of human history. It then has four substantive sections. The first and most extensive section looks at the lessons from history. This section is in turn divided into four sets of case studies. The first looks at when transport is a truly new good, that is, at the value of transport when the mode of transport is

revolutionary, in the sense of hugely reducing the time and or monetary costs associated with travel. The case studies look at the arrival of the railway. Here we show that transport can have a very high social rate of return, but that this is achieved only when society is relatively affluent. The lesson to draw is that when transport “catches up” with the level of development that already exists it is particularly valuable, but when it is “forging ahead” of the general level of development it is much less useful.

This is reinforced by the second case study, which looks at the construction of the highways network. Here we show that the development of the modern multi-lane, graded interchange highway system was initially very valuable, but additions to the network were much less valuable. Here the lesson to draw is that transport investment, like almost everything else, has strongly diminishing returns. The returns to making transport “good enough” are substantial, but once a society’s transport system is “good enough”, then the benefits of improving it further are likely to be slight.

The third historical case study looks at the effect of improvements in transport on the location of economic activity. Again it looks at nineteenth century railways, concentrating on Britain, the US and Spain. Here we find that transport improvements can have a big effect on the location of industry but only when the transport improvements are large and precede major industrial location decisions. In short, it is easy for transport to determine the location of industry if the transport developments occur before the development of industry.

The final case study looks at urban areas. Here we concentrate on the development of two apparently similar light rail schemes in Britain: the London Docklands Light Railway, and the Sheffield Supertram. We show that the reason that the Docklands Light Railway was much more successful than the Sheffield Supertram is that the City of London was “bursting at the seams,” and the DLR allowed a new area – Docklands – to be integrated with the crowded City area. In contrast Sheffield was not bursting at the seams in the same way, and therefore the Supertram did not have a comparable transformative effect.

These historical case studies are followed by a section that looks at the limitations of transport statistics. In many ways we know very little about travel. We know aggregates, but we have much less of a sense of why people move the way that they do. We know, for example, how many vehicle kilometres are travelled, and we can make a reasonable guess as to how many passenger kilometres that represents. But it is hard to classify those journeys: if a Pole working in Britain drives home for Christmas, is the journey home the outwards leg of a leisure journey, or the return leg of a commuting trip. If we cannot classify journeys, it is hard to understand them, and hard to understand how changes in the availability or price of trips will alter behaviour or economic growth.

The problems are not so severe for freight, since freight does not engage in leisure travel. Even here, however, the standard measure of tonne-kilometres is not a good guide to the value or importance of travel. Both the speed of travel and the value of the good being transported also matter.

Nevertheless, despite all of these caveats, we can say that travel increases over time. This is, however, compatible with both the cart before horse and horse before cart scenarios. This is why we have concentrated on the historical case studies, since these are analytically more helpful than approaches based on recent statistics.

The final substantive section looks at the policy implications. Here we draw heavily on the historical case studies, and argue that transport will be a “horse” when the transport improvement is revolutionary, and society relatively affluent, when the transport improvement precedes economic development, or when it allows an area that can expand to do so. In contrast it will be the “cart” when the transport improvement is incremental, or otherwise relatively small, when it is following economic development, or when it is an area in which transport is not obviously constraining development. This second set of circumstances is much more common.

In reality the correct analogy is not horses and carts. Rather transport should be seen, like human capital, financial availability, and so on, as needing to keep pace with development. Where it gets ahead of economic development, transport will have a low return. Where it lags, it can constrain development substantially. The aim should be an appropriate level of infrastructure.

In this context I argue that the biggest issues are congestion and reliability. Both of these are particularly problematic in urban areas, where the transport effect of transport schemes is hardest to model, and where the costs of improvements are extremely expensive. Here the schemes with the highest value for money are likely to be small scale improvements that improve the flows of traffic by all modes, concentrating on the door to door journey times.

The section on policy implications is followed by a short conclusion, and a bibliography. The paper is completed with an substantive appendix that sets out how economists look at transport.

## MOTIVATION – TRAVEL AS AN EXPRESSION OF BEING HUMAN

It is common to say that we travel today as never before. The proportion of the world that has flown in an aeroplane, travelled in a car, is higher than it has ever been. At one level we do travel as never before. In addition, the global economy is integrated as never before, so that goods travel to an unprecedented extent. People in rich countries think nothing of strawberries readily available on every day of the year, and people in all countries think nothing of buying goods mass produced in low cost countries, often thousands of miles away. The movement of people and goods is without parallel.

And yet at another level travel today is commoditised and pales into comparison with the experience of travel of our predecessors. Fifty thousand years ago human beings were fully evolved and for the following forty thousand years they lived as hunter gatherers, moving regularly as part of their existence. Rather than travel being new, it is the concept of a fixed home and place of work that is the thoroughly modern construct. Just as we travel as never before, so we are rooted in a way that would once have been impossible.

Not only did the earliest humans move within their area in search of food and other basic necessities, they also migrated very long distances. Modern DNA studies have confirmed Darwin's idea that all human beings have a common, East African, origin. 60 000 years ago the first people left Africa and, over the following 30 000 years, they walked across all of the earth, populating all regions in which humans can readily exist. Recent advances in genetics reveal that the last millennium of that period saw humans colonise all the islands of the Pacific from Papua New Guinea to Easter Island by canoe (Jones, 2007, p. 143).

The motivations for exploration were often economic. We think of Zhang Qian, the Chinese Imperial envoy who in the 2<sup>nd</sup> Century BC helped established the Silk Road, the Romans exploring – and conquering – that which they could find for plunder, Cabot, Columbus, de Gama and others exploring the world for treasure in the 15<sup>th</sup> century, to the construction of the European Empires into the nineteenth century. Economics was at least a part of such travel. Of course others travelled for the sake of it - from Pytheus' circumnavigation of Britain long before Roman's arrived, through Drake's first circumnavigation of the world, to the adventures of Amundsen, Hillary and so on, who travelled and explored simply because they were able to.

Travel to become rich, and travel as a leisure experience: neither is new, and both appear to be part of what it means to be human. The paper is written in that context.

## LESSONS FROM HISTORY

This section looks in detail at some of the most famous and best studied historical improvements in transport, ranging from the development of canals in nineteenth century America, to the creation of light railway systems in later twentieth century Britain. In each of the four cases we set out the effect of an improvement in transport, before drawing out the lessons.

### 1. Lessons from History 1: Transport as a new good

The coming of the railways was perhaps the defining passenger transport improvement in history. Prior to the railways most people walked, which meant that most people stayed at home. Even the rich travelled by slow and uncomfortable coaches. Once the railway came along, passenger travel was possible in a way that it simply wasn't before. As well as huge improvements in conditions for passengers, freight transport also improved dramatically, not least because railway networks were generally more comprehensive than canals or rivers or coastal shipping, as well as being faster and more reliable.

The standard method of assessing railways is known as the "social savings" methodology, first used by Nobel Laureate Robert Fogel. Details are given in the appendix, but social savings is analogous to total factor productivity growth under tolerable assumptions, and should be seen as equivalent to a rise in GDP in welfare terms.

Since railways were a revolutionary new form of transport, we would expect the returns to transport to be very high. A summary of the results of the better quality studies, for freight, is given below.

Table 1. **Freight social savings: % of GNP**

England and Wales	1865	4.1
USA	1859	3.7
	1890	4.7-10.0
Brazil	1913	18.1
Argentina	1913	26.0

*Sources:* England and Wales: Hawke (1970); USA 1859: Fishlow (1965); USA 1890: Fogel (1964); Argentina: Summerhill (2003); Brazil: Summerhill (2005).

First of all it is worth noting that all of these studies look at railways some years after the railways were first built. As Leunig (2006) has shown, the value of railways in the early days was relatively slight, since the amount of freight carried and the number of passengers using them, was very small. A transport improvement cannot be valuable to society unless it is used.

By the periods to which these studies apply, we find that the countries divide into two groups. On the one hand we have England and Wales and the United States, with relatively low returns, while on the other we have Brazil and Argentina with relatively high returns. The reasoning is that Britain and the United States had good canal networks prior to the construction of the railways, whereas Brazil and Argentina did not. Thus Brazil and Argentina had very high returns from the construction of the railways, whereas Britain and Argentina did not. Railways were a new good in Latin America, but represented only an improved good in Britain and the United States. They were better than canals, but canals had been surprisingly effective and so the gain was limited.

An exception to this pattern is provided by Russia, whose social savings amounted to just 4.6% of GNP in 1907 Metzer (1976). Russia did not, in fact, have a particularly developed set of canals prior to the construction of the railway. Rather the reason for its lower returns is that its level of development in Russia in this era was not sufficient to generate the returns made.

The importance of demand is again emphasised when we look at the social savings for passenger travel in the nineteenth century.

Table 2. **Passenger social savings: % of GNP**

England and Wales	1865	2.8
	1912	6.5
USA	1890	2.6
Brazil	1913	2.1
Russia	1907	1.0

Sources: England and Wales: freight: Leunig (2006),  
USA: Boyd and Walton (1972), Brazil: Summerhill (2005),  
Russia: Metzer (1976).

Here one figure country stands out: England and Wales in 1912. There are two reasons for this. First, the date, 1912, is later than for say the US calculation, and as the two figures for England and Wales demonstrate, the extent of passenger social savings grew over time. Nevertheless, there are two much more important issues. First, Britain and the United States were much richer than Brazil and Russia. Higher disposable incomes translated into potentially higher rates of travel, particularly leisure travel. But Britain was also different to the United States, because Britain is a much smaller country, in which there are many “sensible” train journeys. Even in 1910, London and Manchester, our two biggest cities, were just 4 hours 15 minutes hours apart by train, making the journey much cheaper, and easier than journeys from New York to America’s other large cities – Chicago, Los Angeles or even Boston (Bradshaws Rail Guide 1910, 14:00 departure from London). Britain was rich and the distances between cities were suitable for rail travel. As a result there were large numbers of passenger trains, people travelled, and the returns were high. The story is the same for the United States in the post war era of mass aviation, where the rise of SouthWest airlines had relatively low social returns, owing to the existence of other operators (Leunig and Golson, 2011 work in progress). Again, the issue is whether the transport improvement generates a product that is new and desirable. Given the greater speed of railways over canals, and the greater importance of speed for passengers compared with freight, railways were always a new good to passengers, wherever they were found. But that new good was much more desirable in Britain because Britain was both rich enough and had the relevant spatial setting for railways to be effective. It is likely that this result would be replicated were we to have equivalent studies for say Belgium, the Netherlands, and Germany, other relatively affluent countries with big cities that are not too far apart.

### ***What can we learn?***

This section tells us that transport methods are substitutable to some extent, particularly for freight, where issues of speed, comfort and convenience are generally not as pressing. For this reason, the benefits of investing in one method of transport are generally much lower when another already exists. In this case the value of trains was much lower when canals already existed, but we can easily imagine that the value of high speed trains is lower when conventional trains are already relatively fast, and that the value of an underground metro system is lower if trams running on dedicated tracks already exist.

This section also warns us – again – to remember that the underlying need to travel needs to be borne in mind. If the underlying likely demand for travel is relatively limited, because of fundamental exogenous factors, or levels of income, then the benefits of transport improvements will be relatively limited.

## **2. Lessons from History 2: Steadily diminishing returns from expansion**

In the previous section we looked at the introduction of a major new method of transport, the railway. Although new technologies such as the railway do come along from time to time, most potential transport investments take the form of incremental improvements. This section looks at what is by now the best analysed case of this type of potential investment: the expansion of the US highways network.

In 1989 Aschauer published what was to become a highly influential paper, which estimated the elasticity of output in with respect to US public infrastructure between 1949 and 1985 of between 0.38 and 0.56. This is the era in which the US dramatically expanded and improved the US highways network, which became characterised by highways of four lanes or more, allied to “graded interchanges”, that is where highways crossing each other do so at different levels, so that traffic is not required to slow down or stop at the crossing.

Although those numbers have been comprehensively demolished the paper deserves its iconic status for starting an important debate. Nadiri and Mamuneas (1998) give a good summary of that debate. They argued that the expansion of, and improvements to, the highway network reduced transport costs and increased the profitability of the transport using sector, which in turn crowded *in* private sector investment and so stimulated growth. They found, however, that the beneficial effects fell over time: the elasticity of output with respect to transport investment fell from 0.15 c. 1950 to 0.03 forty years later. This meant that while highways investment accounted for about a third of total private sector TFP growth in the 1950s and 1960s, it accounted for only 4% of the same by the 1980s.

These findings are reinforced by work by Fernald (1999), who found that 1973 represented a watershed. Prior to that date highway construction added about 1.4 percentage points to total TFP growth, while there was no discernable effect after 1973. The intuition is as follows: prior to 1973 the US highways network was still characterised by significant bottlenecks. The gains to the economy from eliminating bottlenecks was significant, and therefore the returns from building 4 lane highways with graded interchanges was large. Since these are relatively cheap to build outside of land-scarce metropolitan areas, those investments made very good economic sense. After 1973 the number of non-metropolitan bottlenecks fell, and so the returns to investments in further highways fell also. We shall return to the issue of metropolitan bottlenecks later.

These findings have been replicated after a fashion for other developed countries as well. Kamps (2004) found that the average elasticity of output with respect to transport spending

across the OECD was 0.2 for the period 1960 to 2001, somewhat above the 0.08 average found by Nadiri and Mamuneas, but of a similar order of magnitude. Within Italy, Destefanis and Sena (2005) find that the highest returns are in areas characterised by poor quality roads initially. This is essentially the cross-sectional version of the US time-series result, namely that when roads are poor the returns to improving them are much higher than when roads of reasonable quality already exist. Combes and Lafourcade find that improvements in road quality explain only a small proportion of the fall in French freight costs between 1978-98, again because the road network was already “good enough” to preclude high returns that occur only when there are bottlenecks to be overcome.

### ***What can we learn?***

These examples serve to remind us of two things. First, investment in transport is very much like investment in other areas: subject to the laws of diminishing returns. Second, investment in transport is not linear.

This combination means that the returns to transport improvements do not necessarily fall steadily as the level of investment increases. The aggregate returns to transport investment function is not smooth and doubly differentiable, in the manner in which traditional investments behave in standard economic textbooks. On the contrary, there are what economists call “network effects”. The benefit of a train line from New York to Omaha is greater if that train line goes on to Sacramento and San Francisco. Were the Panama canal to stop even a metre short of crossing the Panama isthmus, its economic value would be only a small fraction of its potential. The returns to expanding the US highways network were high up until it was completed.

This gives us two results, and a warning. The first result is that completing networks, and attacking bottlenecks are likely to offer very high rates of returns, because then existing infrastructure can be used more effectively. Second, notwithstanding this effect, the returns to transport spending will generally fall as the quantity available rises. Building additional capacity when current capacity is not used intensively is not good economics. The warning comes from the fact that it is not straightforward to distinguish between the two. If a road is not heavily used, it may be that we need to build it further, because it will then connect up more places. Or it may be that we should stop pouring good money after bad. Only good economic modelling and a sound understanding of local economies can help to distinguish between these two different circumstances.

### **3. Lessons from History 3: Transport and the location of industry**

Transport improvements have the potential to change the way that firms do business. We know, from everyday observations of the world that we live in, that firms locate in places that have good transport connections. The relevant question is whether transport improvements can radically reorder the economic geography of a nation.

Here the first important set of papers was by Nicholas Crafts and Abay Mulatu (2006). They set out to use a Midelfart-Knarvik, et al (1999) approach to investigate the effect of transport on the geographical location of the industry during the British industrial revolution. There is a sense in which this is the ultimate test of the ability of transport to lead to industry locating in one place rather than another. This is because the British Industrial Revolution was just that – a revolution that transformed a country that had previously been largely agricultural and so rural, to one that was not only industrialised but also urban. The extent is really quite remarkable: Britain’s 1840 agricultural share of the labour force was matched by France and Germany only after the Second

World War (Crafts, 1985). It is not, therefore, simply that Britain was more urbanised because it was richer, it was also more urbanised for its given level of wealth. Indeed, Britons were 50% more likely to live in a town at any given income level than their neighbours in France and Germany (Crafts and Harley, 2002). Given the rapid pace of industrialisation, the question arises: did transport determine the location of cities and industry?

The answer is in essence no. The railways in particular connected places that were already important, and did not lead to dramatic changes in the overall economic geography of the nation. Of course there were individual railway towns that owed their existence to the railways – Crewe and Swindon, for example – but the railways did not determine the location of the cotton and woollen factories, the breweries, the ship yards and so on. The location of canals had an effect, but even then the location of coal, and access to other raw materials and major ports, was much more important.

This approach has been developed further, both theoretically and empirically by Klein and Crafts (2010). They construct a model that seeks to explain the creation and persistence of the US manufacturing belt a century ago. They find that market potential – for which transport is an important factor – was much more important than factor endowments for the vast majority of industries. In particular, they find that market potential was instrumental in allowing the creation of scale economies, and linkage effects between intermediate industries and final goods industries.

This approach has been replicated for Spain for 1856-1929 by Julio Martinez-Galarraga (2010). He finds that once Spain moved beyond being a primarily agricultural economy, new economic geography, that is, market potential, factors became of first order importance in determining the location of industry within Spain. Again, transport was critical in determining the level of market potential. These effects fell after the First World War.

### ***What can we learn?***

These examples serve to illustrate both the strengths and weaknesses of transport in determining the location of industrial activity. It is the case that transport improvements can increase market potential and so create the potential for industrial concentration, economies of scale and a radically different industrial geography. But for transport to have these effects the population must exist in a relevant area to be connected. In other words transport can connect what is there, but is unlikely to be able to create large market potential in areas that do not have the underlying conditions that allow for success. Even then, we should remember that locations are strongly path dependent, so that transport improvements are more likely to be effective at early stages of development, rather than when locations are already determined. Finally, some transport is “inherent” – for manufacturing mass produced goods access to the coast is always likely to be crucial, hugely privileging coastal areas over those inland. Even in coastal areas some places are likely to be better suited to serving as ports than others. In such cases transport is important in determining the location and even extent of development, but not in ways that can readily be altered by policy makers.

## **4. Lessons from History 4: The value of transport relates to underlying economic conditions**

There are many transport successes around the world, in both the technical and more importantly the economic sense. No-one who has ever visited the Ports of Rotterdam, Shanghai or Singapore can fail to be awed by both their scale and their efficiency. These are not beautiful

places, but they are effective. The same is true for airports such as JFK and Chicago – bustling, busy, cramped and bursting at the seams – and taking millions of people where they want to go.

All of these are symbols of success, and there is a causal dimension. New York and London are world business cities, and their business success is reflected in the busyness of their airports. These cities – like so many others around the world – could not survive without their air links. New York has three major airports, and London has five, with a total of around 1,948 and 2,204 daily flight movements respectively (UK: CAA, December 2010; NYC: PANYNJ, November 2010). Shutting these down, or even heavily restricting them would be hugely costly for these cities competitive positions. They are desirable locations for business because you can get to them, and out of them. This is particularly true for high value migrant workers, for whom the ability to get home is worth a great deal.

Idlewild and Heathrow grew organically with their cities. So did the ports of Rotterdam, Shanghai and Singapore. The link between cause and effect is difficult to ascertain. In this section we look at two examples of transport improvements that clearly had causal effects: the Erie Canal connecting the American Midwest with New England and the London Docklands Light Railway that connects the City with Canary Wharf. Both were successful, we shall argue, in the causal sense. That is to say, economic activity would have been significantly lower without them.

This section seeks to understand why these transport improvements were successful in a comparative framework. In particular it does so by contrasting them with the fortunes of those who tried to learn from them, and copy them. In this case we look also at the fortunes of the Mainline Canal, some 352 km to the south of the Erie, and the Sheffield Supertram, in England's East Midlands. We will show that the underlying economic context of transport improvements matters more than anything else.

### ***Nineteenth Century American canals***

The Erie Canal, completed in 182, runs 584 km from New England's Hudson River to Lake Erie and the American Great Lakes. It connects the Atlantic ports, particularly New York, with the Midwest. Before the Erie Canal the journey from Chicago to New York took six weeks by horse, and as a result costs were high and trade was low.

This means that New England had to be self-sustaining. It had some natural advantages, including water power for industry and timber for fuel and housing. There were good transport connections to other Atlantic ports in the Americas and Europe. But the growth of New England was constrained by low agricultural yields: the soil was thin and the climate difficult.

The Erie Canal had the effect of opening up inter-regional trade, connecting markets in New England with Midwestern grain production. Large loads could be carried for smaller costs. Canals are particularly efficient in this respect, with one animal able to pull fifty times the amount that it could transport by road. The Erie maximized transport capability by connecting existing waterways which were already used for this traffic and minimizing changes in altitude and obstacles such as Niagara Falls. As a result of the canal, the costs of transport fell dramatically, from 30 cents per mile to less than 2 cents (Taylor, 1962). The new economic activity from the canal caused a population surge in the Midwest and helped secure its destination, New York City, as the principal U.S. port.

The canal was developed by the State of New York at a cost of USD 9 million (1825 dollars, about USD 230 m in 2010 money, or USD 156 billion in today's money as a share of GDP). Prior to the canal there had been few projects of this magnitude in the United States. Connecting the

Hudson, rather than New York itself, minimised construction costs. The canal was profitable before it was opened. Four years after construction began the first section opened, collecting USD 1 million in tolls before the entire canal was completed. The construction cost was completely recovered within the first ten years. The Erie Canal's social rate of return, including its financial benefits to developers and the advantage to shippers and consumers was many times greater than the original monetary investment (Atack and Passell, 1994).

The financial and economic success of the Erie led many people to believe that they could replicate its success. The most famous was the Pennsylvania Main Line of Public Works Canal and Road system, completed in 1834. Developed in the wake of the Erie Canal, it ran 599 km from Philadelphia to Pittsburgh. It consisted of two main canals which were connected by a cog railroad over the Appalachian Mountains. In theory this canal should have provided a competitor to the Erie, allowing Midwestern farmers to send their goods to the Atlantic via the Ohio River instead of through Lake Erie.

The economic benefits of the Pennsylvania Main Line were effectively zero. An engineering marvel, it was very expensive to operate. Special articulated barges were necessary to negotiate the sharp turns in the canals. The cost of operating dual infrastructure discouraged shippers from using this route: goods had to be unloaded from barge to railroad freight car and vice versa on each side of the Appalachians. Even at its peak, the quantity of freight shipped over the Main Line system was less than 8% that sent over the Erie Canal (Haite, Mak and Walton, 1975). It was never a financial success.

### ***Late twentieth century British urban transit systems***

The Docklands Light Railway runs from "The City", London's original financial district to Docklands, London's new financial district, sometimes known as Canary Wharf. As the name suggests, the system is based on trains that are significantly lighter not only than mainline network trains, but also than "tube" trains used on London's underground metro system.

London's Docklands was once the world's biggest docks, but the move to containerisation in shipping meant that every single dock closed between 1960 and 1980. The area was characterised by surprisingly high levels of geographical isolation. None of the area was on London's extensive underground system, despite being relatively close to the centre of the city. The Isle of Dogs, an important part of the Docklands area, has only two roads connecting it with the rest of London. With the docks gone, the area became characterised by extensive poverty and poor economic prospects.

The Docklands Light Railway was one part of a package of measures put in place by the British government in the late 1980s in order to regenerate the area. Other parts of the package included the creation of tax exemptions, via "enterprise zone" status, and the creation of the London Docklands Development Corporation, which had extensive rights over planning and issues of eminent domain.

The success of this regeneration project must be seen in the context of the success of the City of London as a financial centre. The late 1980s saw the "Big Bang" financial deregulation that led to large numbers of foreign financial firms wanting a significant presence in London, and existing firms wanted to expand. But the City of London is a relatively small area, and planning constraints – protecting such things as the views of St Paul's Cathedral, mean that it is relatively hard to expand the floor area available. We had therefore a constraint – office space availability – that could potentially be eradicated if entrepreneurs would build new office buildings in the former Docklands area, and if workers were willing to work there. The provision of extension

developments of flats attracted some, particularly younger workers. But banks wanting to move had to be able to persuade but older and more senior workers, with established homes and deep social networks in their communities, that working from Docklands was viable. Here transport – in the form of the Docklands Light Railway, and later the Jubilee line extension, were critical. What matters in such instances is not the physical distance, but the time that a journey takes. The distance was never far in miles, the Docklands Light Railway meant that it was not far in minutes either.

Here the most obvious contrast is with the Sheffield Supertram, which began operating seven years after the Docklands Light Railway. The two networks were similar in scale for many years, both covering around 30 km, but usage was very different: the DLR was used by more than four times as many passengers, and the average passenger travelled more than 50% further. As a result, Dockland's track was used over six times as intensively as the Sheffield Supertram track. The reason is that the Sheffield Supertram did not open up a new area that was previously inaccessible to businesses desperate to expand. As such, it replaced existing journeys by other modes, rather than creating significant numbers of new journeys.

The result is that while the economic effect of the Docklands Light Railway, in conjunction with other measures to ease development, was profound, the economic effect of the Sheffield Supertram is very small. In 1995, when Sheffield's Supertram was completed, Sheffield's GVA was 87.3% of the UK average. By 2004 it was 87.6%, well within the margin of error.

Sheffield is not alone. The Tyne and Wear Metro has the same number of passengers per mile as the Sheffield Supertram, although a much longer network means higher average journey lengths. Even so, Docklands track is used more than twice as heavily as Tyne and Wear Metro track. The Tyne and Wear Metro has also had no discernable effect on regeneration or economic growth. In Sunderland, the area of the Tyne and Wear Metro that is most in need of regeneration GVA grew from was 80.0% of the UK average in 1995 to 80.4% by 2004 (Leunig, 2007).

### ***What can we learn?***

Our two concrete historical comparisons both include a transport improvement that was necessary for the economic development of a region. The American Midwest would not have grown so successfully without the Erie Canal to take the wheat to market. London's Docklands would not have become such a successful financial area had it not been for the transport links that allowed workers to get there in the morning and home at night.

But as the Mainline Canal and Sheffield Supertram both demonstrate, transport per se is not the answer. Transport only works in the context of their being a strong underlying economic proposition. We can generalise this intuition. Let us imagine that there exists a location that can produce a product at a cost of  $c$ , which retails in the final marketplace for a price  $p$ , where  $p$  exceeds  $c$ . Production will occur if transport costs are less than  $p-c$ , and will not occur otherwise. Thus, if a transport improvement takes place that reduces transport costs from greater than  $p-c$  to less than  $p-c$ , the transport improvement will cause an area to engage in new economic activity.

In contrast, if an area does not produce a good for which  $p$  exceeds  $c$ , transport costs are completely irrelevant. If the Midwest could not produce wheat at a lower cost than New England, no amount of canals or other transport improvements would have led to development. Building new transport links today to Yukon or the "badlands" of the American Southwest will not lead to increases in agricultural output, no matter how cheap that transport is. Similarly, if the transport improvement results in transport costs that are higher than  $p-c$ , it will be economically ineffective. This characterises the position of the Mainline Canal.

## TRANSPORT STATISTICS

Having established some underlying economic intuitions, we now move on to look at transport in today's global economy. We begin with some cautionary notes, that set out how little we know.

We know most about the modes of transport that have tickets. It is relatively simple to work out how many people have flown, and how many passenger miles have been travelled. What is not straightforward to understand why people fly. Some journeys are relatively obvious: the vast majority of early morning passengers, on expensive business class tickets, bound for Cincinnati are indeed travelling on business, just as we can be relatively sure that someone on a Saturday charter flight from northern Europe to Spain, in August, is a leisure traveller. But some flights are much more mixed – Europe to Singapore, for example, and some journeys are hard to classify at an intellectual level. As we mentioned earlier, when a Pole working in London flies home for Christmas, should we treat their flights as leisure – a trip to see friends and family in Poland – or should we treat the trip as a commuting trip, since they travel to London simply for work? Statisticians will make a judgment, statistics will be prepared, but we need to be aware that the underlying reality is much more complex than any numbers we have.

Although they also require tickets, train journeys are less well-documented than journeys by air, because many train travellers have season tickets, or “all-day” rider tickets. In these cases we rarely know how many miles they have travelled, or their origin and destination pairs.

Bus journeys are less well-documented still, since as well as the presence of season tickets and all-day passes, many bus systems offer free travel to large numbers of people, such as those who are retired. Even where those journeys are accurately recorded, it is hard to assign economic values to them. There is, after all, no reason to think that the economic analysis of a bus that is full of people who have paid a fair is the same as that for a bus full of people who have paid nothing to travel on it. Additionally, most bus journeys record only the point at which the passenger boarded, and especially in the case of flat fare systems, we have only limited knowledge of the average journey length.

Car journeys are less well-documented again. Since most countries tax road fuel we know how much fuel is being consumed. Most developed countries also require the cars annual mileage to be recorded once a year, as part of safety checks on the car. We have, therefore, a reasonable sense of total road traffic, reinforced in many cases by road sensors that tell us something about road usage at different times of day, and in different places. Again, however, we have much less sense of why people are travelling, save from noting the obvious: people travelling just before the start or after the end of the working day are likely to be commuting, those travelling on Saturdays in the summer holidays are likely to be going on holiday, and those driving trucks are probably not leisure travellers.

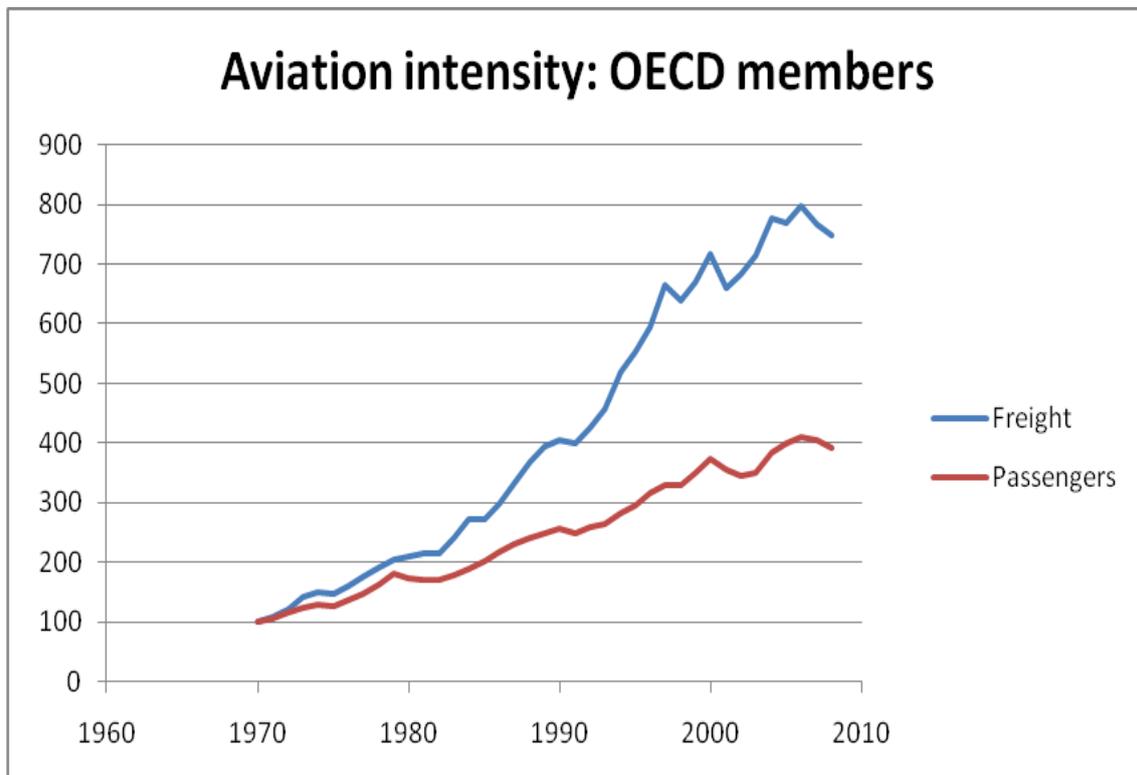
Goods traffic is easier to follow, in two senses. First, goods do not move around as a leisure activity, and therefore we have no need to divide the journeys into business and leisure. Second, goods do not move themselves, and therefore more records are created. Once again, goods shipments by air and rail are better-recorded than those by road, bicycle and carried on foot.

Against that, we have to acknowledge that the standard unit used in analysing freight – the tonne-km – is pretty hopeless economically. A tonne of gold is currently worth around 350 000 times as much as a ton of coal. Those countries that have coal, or other heavy bulky commodities will appear to be more transport intensive than those that have gold, or who produce computer chips, but the importance of transport to each economy may be similar.

Finally, in both the case of international passenger and freight travel it is not intellectually clear to which country we should allocate the transport. If a Brazilian company buy a Swiss machine tool, should the transport of that machine be thought of as increasing the transport intensity of the Brazilian economy, the Swiss economy, the economy of the countries through which it travels, or the economy of the country who own the boat? There is no clear answer to this question. The same is true for passenger transport. Should a Japanese tourist travelling by train from Paris to Rome via Switzerland count towards the transport intensity of Japan, France, Italy, Switzerland or some combination?

These caveats do not mean that we can say nothing, but they do mean that have to be cautious when we analyse what the data can tell us. And at some level we know the single most important fact: those with money travel more, and cause more freight travel. People in rich areas of the world travel more than people in poor areas. People travel more than they did in the past, because they are richer. The same is true for the movement of goods: rich people consume more goods, and are more willing to pay for the transport costs that deliver them.

This is most obvious in aviation, which has the highest income elasticity for people at current levels of income.

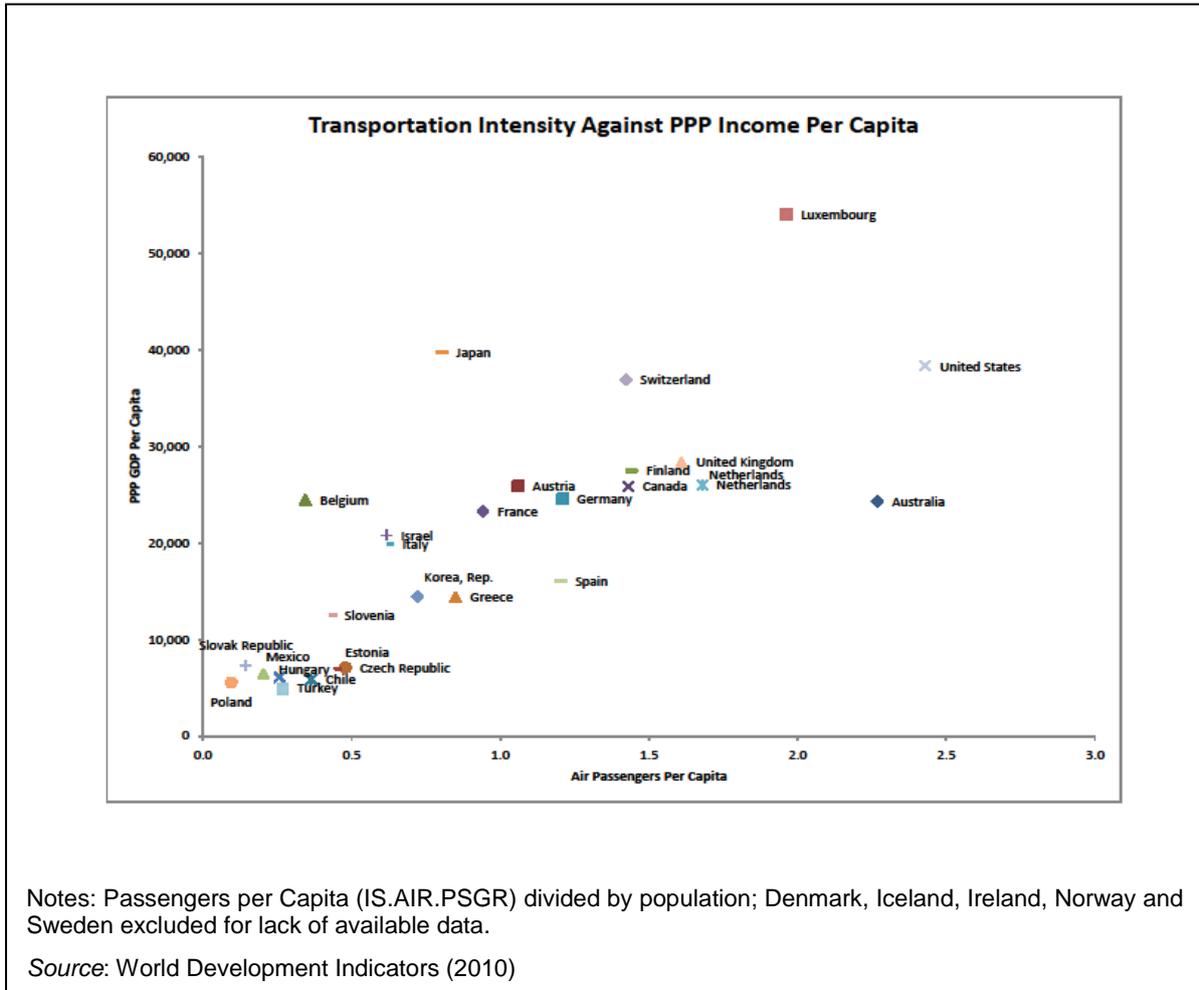


Freight: ton-km per capita

Passengers: Passengers per capita

Source: World Development Indicators (2010)

Clearly, since 1970, the air transport intensity of OECD economies has risen. The compound growth rates are 6.3% and 10.1% for freight and passengers respectively, far outstripping rises in income. The reason that the rate of aviation growth outstrips that of GDP is that the price of aviation has fallen over time relative to that of other goods, and because air travel (and air freighted goods) are income elastic, in that rich people spend disproportionately more of their income on them.



We can see the importance of income in figure two, which looks at air freight intensity in 2006. 2006 is picked because we have better data for this year than for more recent ones, and because it is uncorrupted by the current global economic turmoil. We can see that, as we would expect, rich countries are much more, air traffic intense than poor countries. What matters, above all, are incomes.

## POLICY IMPLICATIONS

What history teaches us is that transport matters when it connects up two places that are synergistic, or when it allows a confined place to grow. The creation of the Silk Road, the discovery of the New World, and connecting the Midwest to the East Coast all come in the former category. Expanding the City of London falls into the second, which is generally a rarer category.

This, then, is the first criteria against which transport proposals need to be evaluated: a simple test which asks, in effect, what is the point of this improvement? The Gravina Island Bridge ([http://en.wikipedia.org/wiki/Gravina\\_Island\\_Bridge](http://en.wikipedia.org/wiki/Gravina_Island_Bridge)) is perhaps the most extreme example of a transport project that fails this test, but there are many others that fail it. Failure is usually a matter of degree, and here the importance of substitutes is critical. If there is already a way to get from A to B, the benefits of building a new transport method from A to B will be much lower than if there is currently no sensible way of getting from A to B. The British proposals to build a high speed line from London to Birmingham need to be considered in the context of the fact that we already have a relatively high speed conventional line from London to Birmingham. In that context the returns are likely to be relatively low. In contrast US high speed train proposals may make more sense because current US trains are relatively slow. Japan's proclivity to build additional roads in areas with low levels of congestion should be seen as likely to offer very poor returns.

Again, returning to history and looking for positive lessons for the future we find three: the importance of speed, the importance of reliability and the importance of cost. Speed is important for most passenger travel, and some portion of freight transport. Here what matters is not the maximum speed at any point in the journey, but the overall speed. This should be hugely encouraging, because the maximum speeds of travel have not been increasing in recent years. A Boeing 777 cruises, for example, at the same speed as an old DC-10, and while a modern car is capable of a much higher top speed than a car of 40 years ago, the maximum speed permitted on the roads has not risen. The top speed of the fastest trains has risen substantially, but these account for a relatively small share of total train miles. We can walk no faster than before.

There are physical reasons for all of this. Energy consumption rises disproportionately in speed, as air resistance rises. We can build faster planes – as Concorde and a huge variety of military planes prove. But fuel consumption increases dramatically as speeds rise, which increases operating costs and reduces range. There will be improvements over time, in that planes can be made more aerodynamic, both by using shapes that are more aerodynamic, and by using materials that have lower levels of friction. There are infrastructure reasons for not increasing the top speeds of vehicles. We can make cars that are capable of 320 km/h, but few people would consider it safe to drive at over 5 km per minute, equivalent to 88 metres per second. Our roads are not that straight, and not that empty. Our reaction times as drivers are simply not good enough, and computer driven cars capable of driving at 320 kph are some way off. Finally, speed dramatically reduces fuel efficiency from around 100 kph.

Instead, improvements in speed are likely to come from looking at the whole journey, rather than at the top speed. This is a harder challenge than completing the US Highways network, but one that is just as important. Most journeys consist of fast bits and slow bits. Even walkers have

to wait to cross roads, or move slowly in busy areas. Cars drive slowly in urban areas. Those who travel by train or plane have to get too or from the station or airport. Those who fly have to pass security and all the other tedium of airports.

There are many ways to increase speeds in this sense. The first is the use of ICT. The newest Satnav systems know not only the distance and official speed limits, but also typical and current traffic flows on the roads. It will not be long before Google Maps direction planner gives you different estimates of the time a journey takes according to the time you leave, and is able to recommend different times for faster journeys. Top of the range satnavs can already recommend different routes at different times of day, based on the experience of other motorists actually in transit. When your author was last driving down the M6 he had his wife check [www.trafficengland.co.uk](http://www.trafficengland.co.uk) from a portable computer to find out the actual speeds ahead to ascertain whether it was worth paying for the toll road (it wasn't on this occasion, as the M6 was flowing well). This sort of technology can pay big dividends: it is much the cheapest way to speed up journeys. Dynamic speed limits – such as limiting speeds on very fast roads so that traffic flows smoothly – is also effective, as the British experience shows.

It is important not only to improve the average whole journey speed, but also to reduce the variation in speeds. The reliability of the journey time matters, as well as the central expectation. If we know that a journey will always take an hour, we can leave an hour for the journey. If it should take an hour, but takes two hours one time in five, then we have to leave two hours if the arrival time is important. Classically this leads people to arrive at airports very early, fearing that a delay will lead them to miss their flight. Notice that there is an asymmetry here: you have to leave the two hours on the way to the meeting, even though there is only a one in five chance that it is necessary to do so, whereas on the way home it is not necessary to allow extra time in case the journey is slow. When there is a one in five chance of an hour's delay, the inbound delay is always one hour, since the extra hour has to be allowed for, whereas the return delay averages 12 minutes. We need, therefore, to consider not only the average delay, but the predictability of the delay, and the nature of the journey.

The same applies to freight. Some freight shipments are not time sensitive, and a 1 in 5 chance of a 1 hour delay adds 12 minutes to the typical journey time. Other products are much more time critical. An express parcels firm has to quote a time for the parcel to be given to them that allows it to reach the airport for sure before the plane takes off. If delays are common, that time quoted has to allow for that delay. The same is true for deliveries to “just in time” production facilities: if the production line will stop for want of a nail, the nail delivery company has to play safe and allow for congestion. In these circumstances increasing the reliability of the transport system, including the reliability of the times taken, is important.

As well as road congestion there are two other important forms of congestion: aeronautical and railway congestion. Aeronautical congestion is made up of runway congestion and airport congestion. Runway congestion means that many plane schedules are slowing down, because airlines cannot be certain of their take off and landing slots. In particular, congested runways mean that planes that arrive early because of good winds are less likely to be able to land promptly, and that stacking is common when anything goes even slightly wrong. Congestion at airports means that it takes a long time for people to get from the edge of the airport to the gate. It is self-evidently nonsensical that law abiding citizens routinely and prudently allow one hour to walk 400 metres from a taxi rank or metro stop at the airport to the gate, because of security. Looking into the number of security lanes available would be the most efficient way to speed up air travel, with small but real benefits. The same is true for passport control on arrival.

Railway congestion is also an issue, and again, as with aviation, the issues are primarily to do with capacity. The most common use of passenger rail the world over is for commuters to reach the central business district from outlying residential suburbs. Here the railways have a particular challenge: almost everyone wants to travel at a particular time, and in a particular direction. The cost of providing additional capacity for everyone to travel in comfort will never be economic, as the rolling stock will only be used at rush hours. For that reason a combination of fares that are higher at peak times, and some congestion, is likely to be optimal. What is noticeable is that governments and transport companies do not look in more detail at the willingness to pay for comfort. Except for the very crude measure of the proportion of people willing to pay to travel first class (very few), we have virtually no sense of the willingness to pay for more capacity, whether that is provided by longer trains, or more frequent trains. For that reason congestion may be either economically too high or too low, or most likely too high in some places and too low in others.

Train congestion is particularly true in terminal stations in city centres, where trains have to cross over each other as they enter and exit stations. This inevitably causes congestion, not least because trains are long, and so the time to cross over a train path is extensive. In the long run the only solution will be to connect stations that are on opposite sides of the city, so that trains can run through, rather than entering and exiting a terminal. That is extraordinarily costly, and takes a long time to implement. The London East-West Crossrail development will open approximately ten years after parliament approved it, and will cost a staggering £135 000 per metre over the entire length, with the cost caused primarily by the 21 km tunnels through the central section.

We need to be aware that within the realm of public transport, there is a significant issue of door to door journey times. Travel by public transport is almost always very slow when measured by this criterion. Go to the journey planner of any major city and enter an origin and destination address (not a station), and calculate the time taken, Find the distance and calculate the average speed. The time to travel between my two offices in London is 22 minutes, an average of 5.5 km/hr, even though both offices are close to tube stations, and even though no change of tube line is required. In terms of promoting agglomeration economies, these times are economically harmful. Again, faster trains are unlikely to be the most cost efficient answer, but instead we need to think about interchanges, times to enter and exit the trains (to reduce dwell times), times to enter and exit the stations, the number of entrances and exits, and so on. None of this is glamorous, but it is what is needed.

The issues for roads are much the same. Outside of urban areas the solutions are simple: repeat the US highways system, that is, build roads that have more than one lane in each direction, and graded interchanges. So long as the right number of lanes are built – which is relatively easy outside of urban areas, since land is generally plentiful – traffic can move at a fairly steady 140 km/h, or perhaps a little more. Such roads have enviable safety records, and, quite simply, work. So long as fuel is taxed, such infrastructure can be paid for by fuel taxes quite effectively.

But as with railways, eliminating congestion in urban areas is not nearly as simple, for lots of reasons. The first is that the congestion is much more deeply embedded. Even if it is cured in one place, it usually stacks up at the next set of lights. The Boston Globe reports that although the “Big Dig” in Boston has increased speeds in the Big Dig area, it has led to more congestion elsewhere so that overall journey times have not fallen (the number of journeys has increased, so the project is not worthless). The second reason that urban areas are difficult to assess is that it is very hard to model behaviour. If people are travelling from Frankfurt to Berlin by car, there are a very limited number of routes that make sense. But there are a large number of plausible routes for any journey within Frankfurt, and the number of different journeys is large.

Although it is difficult, the same principles apply: steadily flowing traffic is the aim. This means eliminating non-graded interchanges. In the past this was achieved with aerial roads, and pedestrian underpasses. Neither work in social terms – aerial roads are ugly, and the road noise travels. Pedestrian underpasses are perceived as unsafe, and are not well used. In contrast short sections of road tunnels can work well, taking through-traffic underneath junctions, and allowing pedestrians to cross at street level. Brussels is a city that does this to a large extent, with the inner urban ring road (R20) serving as an effective example of the role of graded interchanges in promoting more free flowing urban traffic.

Road pricing in urban areas can help, but the willingness to pay is very high, meaning that even large charges can have only a modest effect on transport levels. Transport for London report that traffic levels are back to their pre-congestion charge levels, despite the fact that the charge is now £10 per day – for a journey that would typically be well under 10 km. Thus even with a congestion charge that typically exceeds the cost of fuel by a factor of 5 or more, traffic levels have not fallen, although they are no doubt lower than they would otherwise have been. Furthermore, London's experience shows that although the initial charge leads to a fall in traffic, subsequent rises in the charge have much smaller effects on the number of road users. Almost all those willing to pay £5 prove to be willing to pay £8 and £10 per day. Congestion charging is useful, but it should really be seen as akin to tobacco taxes: much better at raising revenue than cutting consumption. (<http://www.tfl.gov.uk/roadusers/congestioncharging/6723.aspx>, <http://www.tfl.gov.uk/assets/downloads/Ex-post-evaluation-of-quantified-impacts-of-original-scheme-07-June.pdf>, <http://www.tfl.gov.uk/assets/downloads/demand-elasticities-for-car-trips-to-central-london.pdf>)

## CONCLUSIONS

The history of transport is, in some sense, a history of anti-climax. 40 000 years ago human beings set out on the greatest journey ever made, that would take human beings out of Africa and around the world. 500 years ago the inhabited world was being discovered and rediscovered, and continents connected for the first time in any serious manner. Even a hundred years ago we were still involved in races to the North and South Pole, and Everest was yet to be conquered. Today the only areas above water remaining to be conquered are very obscure indeed.

We have also invented all of the major forms of transport that we are likely to invent. The bicycle will be refined, but the biggest restriction is our muscles. The car will be refined, but the basic idea of a roughly family sized powered covered seating area, that can travel from one place to another at a time of the owner's choosing, is well-established and cannot be bettered. The rest is essentially detail: the internal combustion engine may be replaced, and electronics will steadily reduce the demands on the driver, but it will still look like a car, and have the advantages of a car. Trains will continue to travel on rails, ships on water, and planes in the air. Speeds may increase or decrease incrementally, but there will be no breakthroughs of the sort that saw the horse and cart replaced with the railway, or the transatlantic ship with the plane. The private plane in every garage and the jet pack are not going to happen, however much we may once have expected them. The physics does not stack up.

Instead we have a harder challenge: to accept that there are no breakthroughs coming along, and ask ourselves instead what we can do to make what we have work well. To do that we need to understand the intrinsic underlying demand for travel, and that door to door journey times, and price matter. When we concentrate on that, we will be successful. When we forget any one of those, we will not. Transport will not be a horse that powers economies forward in dramatic leaps. But down well it can provide a supportive aid to economic growth, and prevent growth from being curtailed. That is a worthy aim, even if it is not one that will win headlines for anyone in the transport sector.

## BIBLIOGRAPHY

- Aghion, P and M Schankerman (2004). "On the Welfare Effects and Political Economy of Competition-Enhancing Policies," *The Economic Journal*, 114:498, pp. 800-824.
- Aghion, P and P Howitt (1997). *Endogenous Growth Theory*. Boston: MIT Press Books
- Aschauer, D A (1989). "Is public expenditure productive?," *Journal of Monetary Economics*, Elsevier, 23:2, p. 177-200.
- Atack, J and P Passell (1994). *A New Economic View of American History: From Colonial Times to 1940*. New York: Norton.
- Badinger, H (2005). "Growth Effects of Economic Integration: Evidence from the EU Member States," *Review of World Economics*, Springer, 141:1, p.50-78.
- Barro, R J and X Sala-i-Martin (1995). *Economic Growth*. New York: McGraw-Hill.
- Boyd, J and G M Walton (1972). "The Social Savings from Nineteenth Century Rail Passenger Services," *Explorations in Economic History*, Elsevier, 9:1, p.233-254.
- Bradshaws Railway Guide, April 1910. David & Charles Plc, 1968.
- Crafts, N F R (1985). *British Economic Growth during the Industrial Revolution*, Oxford: Clarendon Press.
- Crafts, N F R (2004). "Social Savings as a Measure of the Contribution of the New Technology to Economic Growth," *London School of Economics, Economic History Department Working Paper 06/04*.
- Crafts, N F R and C K Harley (2002). "Precocious British Industrialization: A General Equilibrium Perspective," *UWO Department of Economics Working Paper 200213*, University of Western Ontario.
- Crafts, N and A Mulatu (2006). "How did the Location of Industry Respond to Falling Transport Costs in Britain Before World War I?," *Journal of Economic History*, Cambridge University Press, 66:3, p.575-607.
- Crafts, N and A Mulatu (2005). "What Explains the Location of Industry in Britain, 1871-1931?," *Journal of Economic Geography*, Oxford University Press, 5:4, p.499-518.
- Destefanis, S and V Sena (2005). "Public Capital and Total Factor Productivity: New Evidence from the Italian Regions, 1970-1998," *Regional Studies*, Taylor and Francis Journals, 39:5, p.603-617.

- Fernald, J G (1999). "Roads to Prosperity? Assessing the Link between Public Capital and Productivity," *American Economic Review*, American Economic Association, 89:3, p.619-638.
- Fishlow, A (1965). *American Railroads and the Transformation of the Antebellum Economy*. Cambridge: Harvard University Press.
- Fogel, R (1962). "A Quantitative Approach to the Study of Railroads in American Economic Growth" *Journal of Economic History*, Cambridge University Press, 22:2, p.163-97.
- Fogel, R (1964). *Railroads and American Economic Growth: Essays in Econometric History*. Baltimore: Johns Hopkins Press.
- Haites, E, J Mak, and G Walton (1975). *Western River Transportation: The Era of Early Internal Development 1810-1860*. Baltimore: Johns Hopkins University Press.
- Hawke, G R (1980). *Railways and Economic Growth in England and Wales, 1840-1870*. Oxford: Clarendon Press.
- Hersch, J and J Voth (2009). "Sweet Diversity: Colonial Goods and the Rise of European Living Standards after 1492," *CEPR Working Papers 7386*, CEPR Discussion Papers.
- Hicks, J R (1935). "The Theory of Monopoly," *Econometrica*, Econometric Society, 3, 1-20.
- Jacobs, J (1969). "Strategies for Helping Cities," *The American Economic Review*, The American Economic Association, 59:4, p.652-56.
- Jara Diaz, S R (1986). "On the Relation between Users' Benefits and the Economic Effects of Transportation Activities," *Journal of Regional Science*, Wiley, 26:2, p.379-391.
- Jones, S (2007). *Coral*. New York: Little, Brown.
- Kamps, C (2006). "The Dynamic Effects of Public Capital: VAR Evidence for 22 OECD Countries," *Kiel Working Paper 1224*, Kiel Institute for the World Economy.
- Klein, A and N Crafts (2010). "Making Sense of the Manufacturing Belt: Determinants of U.S. Industrial Location, 1880-2010" *CAGE Online Working Paper Series 04*.
- Leunig, T (2006). "Time is Money: A Re-assessment of the Passenger Social Savings From Victorian British Railways," *The Journal of Economic History*, Cambridge University Press, 66:03, p.635-673.
- Leunig, T (2010). "Social Savings," *Journal of Economic Surveys*, Wiley, 24:5, p.775-800
- Leunig, T and J Swaffield (2007). *Cities Unlimited: Making Urban Regeneration Work*. London: Policy Exchange.
- Lewer, J J and H Van den Berg (2003). "How Large is International Trade's Effect on Economic Growth," *Journal of Economic Surveys*, Blackwell Publishing, 17:3, p.363-396.
- Martinez-Galarraga, J (2010). "The Determinants of Industrial Location in Spain, 1856-1929," *Working Papers in Economics 244*, Universitat de Barcelona.

- Metzer, J (1976). "Railroads in Tsarist Russia: Direct Gains and Implications," *Explorations in Economic History*, Elsevier Press, 13:1, p.85-111.
- Midelfart Knarvik, K H, et al (1999). "What Determines the Economic Geography of Europe?" *CEPR Discussion Paper Series Number 2072*. Centre for Economic Policy Research.
- Nadiri M I and T Mamuneas (1998). "Contribution to Highway Capital to Output and Productivity Growth in the US Economy and Industries," mimeo.
- Rice, P and A J Venables (2004). "Spatial Determinants of Productivity: Analysis for the Regions of Great Britain" *CEP Discussion Paper DP0642*, Centre for Economic Performance, London School of Economics.
- Solow, R (1957). "Technical Change and the Aggregate Production Function," *Review of Economics and Statistics*, MIT Press, 39:3, p.312-320.
- Stigler, G (1956). "Economics Might Serve a More Useful Purpose if they Fought Fires or Termites Instead of Monopoly," *Journal of Political Economy*, University of Chicago Press, p.34.
- Summerhill, W R (2003). *Order against progress: government, foreign investment and railroads in Brazil, 1854-1913*. Stanford: Stanford University Press.
- Summerhill, W R (2005). "Big Social Savings in a Small Laggard Economy: Railroad-Led Growth in Brazil," *The Journal of Economic History*, Cambridge University Press, 65:01, p.72-102.
- Taylor, G R (1962). *The Transportation Revolution 1815-1860*. New York: Holt Rinehart, 1962.
- Tinbergen, J (1942). "Zur Theorie der Langfristigen Wirtschaftsentwicklung" *Weltwirtschaftliches Archiv*, 55, p.511-49)
- Venables, A (1996). "Lecture Notes on International Trade and Imperfect Competition," *Research Institute of Industrial Economics Working Papers 463*, Institute of Industrial Economics.

## ANNEX: CONCEPTUALISING TRANSPORT IN THEORETICAL ECONOMIC TERMS

The contribution of investment and/or technological change to economic growth is a hugely important area for economists and policy makers. There is no single dominant paradigm or model that explains the connection. That being so we present brief outlines of the major approaches, starting with the most traditional, and moving towards the more modern.

### *The Solow growth model and growth accounting*

The modern economic understanding of growth dates to Jan Tinbergen's seminal 1942 paper, 'Zur Theorie der Langfristigen Wirtschaftsentwicklung' (Weltwirtschaftliches Archiv 55: 511-49), which set out, for the first time, the notion that economic growth could be decomposed into those output gains that came about because of an increase in inputs, and those output gains that came about from an improvement in the productivity of those inputs. This intuition was made more explicit and tractable in Robert Solow's perhaps more famous 1957 paper, 'Technical Change and the Aggregate Production Function'. This paper gave economics the famous Solow growth equation, which in a Cobb-Douglas setting is given as:

$$Y = AK^{\alpha}L^{1-\alpha} \quad (1)$$

where Y is output, K is capital, L is labour and A is TFP while  $\alpha$  and  $(1 - \alpha)$  are the elasticities of output with respect to capital and labour, respectively. It follows, therefore, that the growth in labour productivity can be written as:

$$\Delta \ln(Y/L) = \alpha \Delta \ln(K/L) + \Delta \ln A \quad (2)$$

In this context transport can have three effects. First, insofar as transport improvements increase the capital stock, output will rise because of that increase in capital. More formally, output will increase by an amount equal to the growth in the capital stock multiplied by the elasticity of output with respect to transport capital. Second, improvements in transport may increase the productivity of the transport sector itself, that is, the transport sector may be able to supply the same amount of output with a reduced level of inputs. In this case "A", which is often taken to be technological change, would increase. Finally, improvements in transport may lead to improvements in productivity of transport using sectors. Let us imagine a world in which transport is very expensive, so that no product moves more than a few miles. If transport costs then fall dramatically, manufacturing firms can reorganise production into larger more efficient units, and transport their output to distant areas. Productivity in manufacturing – the transport using sector – will rise as a result of the improvements in transport. In this case "A" will also increase, but this time it the "A" is not applied to the transport sector itself, but to the transport-using sector. In addition, the transport using sector may, as a result of the improvements in transport, increase investment levels – for example, mass production may replace small scale artisanal production as scale increases. In this case non-transport K will rise, and output will rise in line with the elasticity of output to non-transport capital. This final set of growth enhancing outcomes from transport improvements are best thought of as externalities, in that they are outside of the transport sector.

As Crafts (2004) notes, the growth accounting definition of the contribution of transport – or any other technology – is too large. As Crafts notes, ‘growth accounting simply addresses the ex-post accounting question “how much did the new technology contribute to growth?” and ignores issues of crowding out.’ (Crafts, 2004, p. 8) Thus, for example, if USD 1 bn is invested in transport capital, and the elasticity of output with respect to transport is 3%, then growth accounting will declare that the transport investment raised GDP by USD 30 m per year. A moment’s thought demonstrates that this emperor has no clothes, or at least not USD 30 m worth of clothes. Had the USD 1 bn not been invested in transport, it could have been invested in other projects, which would also have yielded a return. That return must be subtracted from the USD 30 m to get the “additionality” of transport investment over investments in other projects. It is not straightforward to make this calculation. On the one hand transport is more likely to have the externalities discussed above than other forms of investment. As such it is plausible that although growth accounting estimates will exceed the true value of the transport investment, the true value will be positive. On the other hand, transport improvements are often government led, and may represent vanity projects that have low direct returns and few externalities. In that case the effect of the transport improvement on growth may be negative, since the return from the USD 1 bn in alternative, market-led, investments may have been higher. All we can say for certain is that growth accounting estimates represent an upper bound, and almost certainly a significant over-estimate of the true benefits.

### ***Cost-benefit analysis***

At the opposite extreme conventional cost-benefit analysis traditionally captures only the direct costs and benefits to the transport sector itself. Within this, the benefits are often captured via willingness to pay for improvements. Transport benefits, measured like this, will only equal the economic benefits to society if marginal social benefits equal marginal social costs throughout the rest of the economy, an assumption that de facto implies perfect competition in all transport using sectors (Jara Diaz, 1986). This is not likely to be the case. In an effort to improve traditional cost-benefit analysis governments now try to include the benefits of time saved, even though non-work time is not included in GDP. This area is fraught with difficulty however, since most transport improvements lead to new journeys that would not otherwise have occurred. If the TGV between Paris and Lyon leads people to travel who would not otherwise have done so, how much time have they saved? There is no intellectually satisfying answer to this question. Finally, although governments are increasingly interested in the externalities to other sectors, as given above, these are also hard to model in a conventional cost benefit structure. The British Government has, for example, created “WebTAG” – “Wider economic benefits – transport analysis guidance”, but this should be seen as very much work in progress. This work is publically available at <http://www.dft.gov.uk/webtag/>. For these reasons, cost-benefit analysis is likely to represent a lower bound on the true benefits of transport improvements to the economy, although of course CBA is often produced ex ante by supporters of particular schemes, in which case the results may be biased in the opposite direction. Gourvish has noted, for example, that the British government rarely commissions ex post cost benefit analysis, making it hard to calibrate the over- or under-optimism bias of ex ante studies. The Eddington Report made much the same point.

### ***Social Savings and TFP growth***

An alternative measure of the value of transport was pioneered by Economics Nobel Laureate Robert Fogel, whose pioneering 1962 paper “A Quantitative Approach to the Study of Railroads in American Economic Growth” set out the concept of “social savings”. An up to date summary of the advantages and disadvantages of this approach are set out in Leunig “Social

Savings” (2010). In this context the most interesting result is Foreman-Peck’s demonstration that under competitive conditions, social savings are equal to the rise in Total Factor Productivity. Social savings records the fall in the amount of inputs that are required to supply the ex post rise in output caused by the new technology. TFP growth measures the rise in output possible from a given level of inputs, because of the invention of the new technology. These are intuitively equivalent, and this result was formalised by Foreman-Peck under conditions no less plausible than those that common in economics textbooks.

### ***New Economic Geography***

New Economic Geography is not, in fact, at all new, but instead dates from Marshall’s observations about what were then termed “external economies of scale”. Put simply Marshall observed that costs fell for firms located in the same area as other firms in the same industry. This intuition was later formalised and developed by Arrow and Romer, and these externalities are sometimes referred to as Marshall-Arrow-Romer externalities. Jacobs (1969) extended this analysis by noting that falls in costs could also occur if a firm locates near other firms that are not in the same line of business, but which in some way use the same suppliers, leading to thicker markets or economies of scale in the supplying industry. In some senses all firms rely on some common suppliers – almost every office in the world uses paper and coffee, and every firm accountants. Agglomeration economies do not require people to be in the same industry, although same-sector clusters are likely to offer stronger agglomeration economies than multi-sector clusters, for any given overall size.

In this context transport has two potential effects. First, as mentioned earlier, transport improvements may allow the concentration of production in fewer places by firms that can then use the newly improved transport system to distribute their goods. This can either take the form of one very large firm supplying all the market from one place, or, as is empirically more likely, for firms to cluster together in one area, and supply the whole of the market from there. Thus transport improvements induce the spatial concentration of industry which in turn increases the extent of internal and external economies of scale (Venables 1996). In this context transport improvements a considerable distance from the plant location are important: the ability of Detroit to supply all of America’s cars in Detroit’s heyday depended not only on good transport in the Detroit area, but good transport across America. New economic geography terms places a heavy emphasis on “Market Potential”, which is defined as the amount of demand “close by”, where “close by” is determined in part by transport costs. As transport improves, more distant areas “become close” to a given place. As Midelfart-Knarvik et al (2000) show, this means that transport improvements lead to increased levels of investment and increased clustering.

The second sense in which transport improvements can raise productivity in a new economic geography setting is by facilitating within-area movements. This deepens labour markets, leading to better matching of skills and aptitudes with opportunities, and the development of greater pools of expertise. If a city is well-connected by, for example, an efficient metro system, the city can “become one”, rather than being lots of neighbouring places with only limited labour market integration. It is easy to see how this might be particularly important in areas of high human capital. Rice and Venables (2004) show that population within 80 minutes of a UK city increase productivity in that city, giving a sense of the extent to which cities have significant hinterlands given modern transport systems. Crafts and Leunig (work in progress) find a similar result for Britain in 1900, although of course 80 minutes represented a much smaller distance in that era. This deepening of labour markets would not be included in traditional cost-benefit analysis.

### ***New Growth Economics***

The traditional Solow growth model, as outlined at the start of this paper, decomposes growth in labour productivity into increase in capital and increases in “A”, which is generally taken to be technology, but which would also include improvements in the organisational structure of firms. New growth economics aims to provide microeconomic explanations that are consistent with individually rational agents that can generate long-run, sustained, rises in “A”. These models are often described as “endogenous growth” models.

The particular relevance to transport is that new growth economics models allow for one form of investment to “crowd in” other forms of investments (Barro and Sala-i-Martin, 1995). Thus, to continue with our earlier externality example, a transport improvement that significantly increased market potential can “crowd-in” private sector investment in new factories. The strongest endogenous growth effects would come about if the improvements in transport led not simply to more factories, but to better factories, for example, with higher levels of mechanisation, automation or other productivity enhancing aspects (see Howitt and Aghion for a more general treatment of this intuition).

The final important potential for transport in a new growth economics setting is via the elimination of monopoly power. That monopoly power reduces welfare in a static framework is well known to anyone who has completed even the first two weeks of an introductory economics course. Although economics has not endorse Nobel-laureate George Stigler’s famous response to empirical estimates of the size of these static losses (1956) the profession is increasingly convinced by Nobel-laureate Sir John Hicks argument that “the best of all monopoly profits is a quiet life” (Hicks, 1935). This implies that the problem of monopoly is not the static losses, but that monopolists invest sub-optimal levels of effort into improving productivity and other welfare enhancing measures. Transport improvements can have two potentially beneficial effects. First, by forcing two firms that were previously in de facto separate markets to compete, the more efficient firm will gain sales from the less efficient firm, raising average efficiency and productivity even if neither firm changes their level of efficiency (Aghion and Schankerman, 2004). Second, increases in competition increase the incentives on management not to aim for the quiet life – if they are innovative they can take market share from others, if they are not, they will lose market share to others. Insofar as transport improvements increase the extent to which firms are forced to compete with each other, managerial behaviour will change in a productivity enhancing direction (Aghion et al 1997).

Notice that there is a potential connection between this literature and that on improvements in international trade. Traditionally the modern literature on international trade has seen tariff reductions as instrumental in increasing trade levels, and so generating significant welfare gains (see Lewer and Van den Berg 2003 for a meta analysis and Badinger 2005 for an analysis of post-war European integration). Falls in trade barriers are important, but falls in transport costs have essentially the same effect of integrating economies. For that reason transport economists should take note of findings in the international trade literature.

### ***New Goods***

Improvements in transport can lead to the creation of “new goods”. Here again there are two types of new good: those that are directly related to transport, and those that use transport. The invention of the railway allowed people in London to go on a day trip to the seaside in Brighton for the first time. The distance, 60 miles, was clearly too far to walk for a day on the beach, and even a round trip by stagecoach was barely feasible and would have been a deeply unpleasant

way to spend the day. But by train a day at the seaside became feasible, and relatively quickly the railway companies learned – as SouthWest, Ryanair and Air Asia would relearn a century or so later – if the price is cheap, leisure demand is high. A day at the sea become a common treat for many Londoners (Leunig JEH 2006)

The second type of new good that can be created by transport improvements is the transfer of goods from one part of the world to another. This is most obvious in the case of agricultural commodities – the spice road between Asia and Europe brought all kinds of remarkable items to European consumers for the first time. Today those two staple beverages – tea and coffee – are consumed around the world, often sweetened with sugar. Yet these products were once unheard of, and the value to consumers is clearly considerable (Hersch and Voth, 2009). Smaller improvements in transport can also have big effects: the invention of the refrigerated milk car, for example, allowed the provision of fresh milk in cities, with obviously beneficial effects, particularly for children. Fresher fruit, vegetables and meat also raised nutritional standards.

Finally, falls in transport costs can change the relative price of products in such a way as to make them “new” to some people in the sense that they would not otherwise have been able to afford them. Before sugar was brought back from the New World, it was possible to make it from rose hips, but the cost was fantastically high. New World sugar was not technically a new good, but it was de facto a new good to the vast bulk of the population. The same is effectively true for the effect of the rise of manufactured goods from China and other low cost locations on western consumers: now (almost) everyone in developed countries can afford a very large television, for example.

**International Transport Forum**

2 rue André Pascal

75775 Paris Cedex 16

[itf.contact@oecd.org](mailto:itf.contact@oecd.org)

[www.internationaltransportforum.org](http://www.internationaltransportforum.org)

---