TFP growth and the transport sector: a dynamic industry panel analysis

Simon Price
City University and OEF*

October 25, 1999

Abstract

There are strong reasons for believing that infrastructure output may affect the level of output via external effects. An important component of infrastructure which has been singled out in previous research is transport. Much of the existing evidence on infrastructure and growth derives from aggregate data, which presents many problems. We do examine aggregate results, but the main focus is on the results of dynamic panel techniques on disaggregated industry data. We find that transport growth has a small but significantly positive effect on average total factor productivity growth.

Keywords: Growth, total factor productivity, infrastructure, dynamic heterogeneous panels, pooled mean group estimation.

JEL Classifications: C23, H4, O4, R4.

1 Introduction

A pervading theme of recent theories of economic growth\(^1\) is that something is missing from the traditional list of growth factors. One such is the existence of a public infrastructure. In this paper, we concentrate on the role of transport in this context, which is frequently singled out as a key sector within the whole. There are two areas where transport may influence growth (beyond the simple direct contribution to GDP). The first is the contribution the industry may make to ‘knowledge’, as a leading

---

*This work was funded by the Airport Operators Association, the British Air Transport Association and their members, and by the Department for Environment, Transport and the Regions. However, the opinions expressed in the paper are not necessarily those of the above bodies.

\(^1\)There are now several surveys of the new growth literature. Barro and Sala-i-Martin (1995) continues to serve as an excellent introduction.
edge technological sector. The second, and the most relevant, relates to the special ‘infrastructure’ role of the industry. This has an impact on the rest of the economy as an intermediate input like any other, which is not to be disregarded. But the transport infrastructure also has public good and externality-reducing aspects which may imply the sector has a special contribution to make to growth. It is on this which we concentrate here.

We begin in Section 2 by setting out the new ideas underlying economic growth, to define the wider context. With this background, we move on to the existing empirical evidence on infrastructure and growth, in Section 3. In Section 4, we consider the econometric issues raised by the estimation of relationships in levels and in differences. We also consider some special econometric issues raised by the use of a dynamic panel. In Section 5, we describe our data. Section 6 includes the results from aggregate data: Section 7 those from disaggregate data. Our key results are based on total factor productivity growth regressions. Finally, we conclude.

2 Economic growth

2.1 Exogenous growth

The study of growth economics has a history that dates back to the 1950s, but economists have become aware that arguably the most important aspects were being disregarded.

The approach pioneered by Solow (1956) and others was based on the assumption that growth followed from three potential sources. These were: an increase in the capital stock; an increase in the labour supply; and an increase in technical efficiency (‘technical progress’). Given this framework, two things follow. Firstly, knowing the growth of labour and capital and assuming a production function, we can break down growth into its sources with ‘growth accounting’. Technical progress is worked out as what is left - the ‘Solow residual’.\(^2\) Secondly, it follows that if there are diminishing returns to capital, eventually growth will slow down and stick at the level necessary just to keep up with growth in labour supply and technical progress. In a real sense, then, in the long-run growth is fixed, or ‘exogenous’.\(^3\)

There were some clear problems with this paradigm, the foremost being the actual experience of growth in the World. The empirical prediction of the theory was that all countries would eventually grow at the same rate - the convergence hypothesis. Yet this seemed to be refuted by the evidence. Some poor countries remained poor; others caught up and overtook richer ones. It seemed other factors might be driving growth. This sparked off a new literature, aiming to explain this phenomenon. This modern growth literature goes under the general heading of Endogenous Growth Theory.

\(^2\)Solow (1957).

\(^3\)Early models assumed a fixed savings rate. Later work made this a choice variable by examining growth in explicit models of intertemporal optimisation, but the basic result remained.
2.2 Endogenous growth

There are a number of strands to the argument. Firstly, there appeared to be a missing factor of production. Labour supply and capital are important, but there is another, less easily measured factor - human capital, or knowledge. The idea was introduced by Arrow (1962). It was revisited by Romer (1986), who helped to popularise the concept, and added the idea that knowledge has spillover effects (see below). Secondly the role of public capital. The output of at least part of the public sector might also be an important extra factor. In his influential paper, Barro (1990) suggested that government spending, or the public capital stock, might be an extra factor of production. This might have some relevance for transport. Thirdly, the existence of spillover effects. There might well be externalities in operation which act to boost growth. Because no one firm actually chooses these external effects, they are not taken into account by firms, but are still important. Among these, The area usually examined is knowledge, as mentioned above. The key is that knowledge, once acquired, is at least partly a public good. Once something is invented, everyone knows about it. Thus there is an externality at work in its acquisition. This is very relevant to the transport industry, which includes aviation, and is a leading-edge technological industry. The industry creates benefits for output and growth that have effects beyond other industries’ private interests. There are other external spillovers from transport, and these are likely to be the ones of main interest to us. Many studies of public sector capital and growth single out transport as a key example. Transport is clearly a private industry for whose services consumers and firms are happy to pay, but there are still public good aspects of the product. There may be positive externalities from the widespread existence of travel, enabling new ways of trading to emerge; or we may think of transport as simply reducing congestion costs, as in Barro and Sala-i-Martin 1992. Thus transport contributes to the output of firms over and above their private benefit. A larger transport sector leads to a higher level of output than would otherwise seem to follow. Fourthly, there is the conundrum of constant or increasing returns to capital and the convergence puzzle. Convergence requires diminishing returns to capital. This works as increased investment lowers the rate of return on capital at the margin. But if there are constant returns, then as investment proceeds it has no effect on the rate of return, and there is no tendency for growth to diminish. The simplest way to capture this is with the so-called AK model. One way to capture diminishing returns is to write

\[ Y = A K^\alpha \]  

(1)

where \( Y \) is output and \( K \) is capital, and if \( \alpha < 1 \) there is diminishing returns. If \( \alpha = 1 \) we have

\[ Y = A K; \]  

(2)

Hence AK. It is argued that if capital is defined to include human as well as physical capital, this is quite plausible. Equally, if the infrastructure (\( G \), say) grows in line with capital then even if there is diminishing returns to capital then there might be effectively constant returns:

\[ Y = A K^\alpha G^{(1-\alpha)}. \]  

(3)

The importance of this cannot be overstated. What we get in these AK type models is perpetual economic growth, even without technical progress, and this is one of the things that have made these models so exciting. However, for our purposes, the more interesting ideas are the external spillovers from the transport industry.
3 Infrastructure and output

The potential importance of the public sector has led to many attempts to measure its contribution to growth. It should be clear that it is not the ‘public’ part of this which is so crucial, so much as the type of good or service provided. As Barro and Sala-i-Martin (1995) point out, in many countries much of public investment is actually treated as private.\(^4\) As observed above, we have to recognise that there are two contributions which transport makes to output. Firstly, in terms of gross output, transport services are like any other intermediate input. But, secondly, what we are interested in here is the contribution of transport to productivity growth via spillovers - the effect on value added, in other words.

Existing empirical studies fall into two broad categories: those looking at cross-country variations in growth, and those looking at the contribution to output, usually for a single country.\(^5\)

3.1 Growth

Results here have been mixed. Easterly and Rebello (1993) follow the approach pioneered by Barro, and run regressions explaining the rate of growth for many (about 100) countries, with government spending as an explanatory variable. They are able to break down expenditures into different categories. Among their findings, they report (p 431) that ‘[t]ransport and communication investment seem to be consistently positively correlated with growth with a very high coefficient’. Note that this is new investment, and does not necessarily imply anything about the total (rather than marginal) contribution to output. There might also be a confusion of causality, which is a recurrent theme in the literature. Essentially, it may be that growth and infrastructure investment are correlated because economies demand more infrastructure as they grow richer: thus fast growing countries choose to invest more. There is an obvious identification problem. This is sometime referred to as ‘reverse causality’. However, they get even stronger results using instrumental variables, which is intended to control for this. They also report ‘that only transport and communication investment and general government investment are robustly correlated with growth’ (their emphasis). The conclusion seems to be that transport, is important for growth. However, the problem with their results is arguably not that there are no effects but that the coefficient is implausibly large.

3.2 Output

Other evidence comes from analysis of the level of output and related variables. There are two types of analyses that have been undertaken here.

\(^4\)One practical issue is whether to look at the installed capital stock or the output of the relevant sector. For our purposes, while recognising that capacity utilisation can vary, these are very nearly the same. A given capital stock provides a flow of services. If we can measure output, as we can with transport, then that is preferable.

\(^5\)A recent survey of the empirical growth literature is Temple (1999).
3.2.1 Production functions

The initial estimates were derived from estimates of production function. The empirical literature was largely initiated by Aschauer. In his (1989b) paper, he essentially estimates a production function for the US private business sector. He breaks down public capital spending into different types - education, hospitals, and so on. The very clear result is that what he calls ‘a “core” infrastructure of streets, highways, airports, mass transit sewers, water systems, etc, has most explanatory power for productivity’; once again offering strong support for the case that transport is important. He is aware of the reverse causality argument and again uses instrumental variables; however, this may be insufficient. We discuss this below. One result imposing constant returns to scale to all factors, including infrastructure, is

\[ y - k = -2.33 + .001t + .41(n - k) + .40(g - k) + .38 \text{cu} \] (4)

where everything is in logs so that (eg) \( y = \log(Y) \), \( N \) is employment and \( G \) is the government capital stock, excluding the military. He also adds a capacity utilisation term, \( \text{cu} \), to control for cyclical effects. The coefficients are interpretable as the ‘shares’ (although much of public services have no price) and (eg) the marginal product of labour is given by \( .41(Y/N) \). This enables us to calculate the proportion of output that can be ascribed to the sector. Similarly, we can calculate the contribution to growth. Aschauer (1989c) has produced similar results for G7 countries. In a related paper for the US, Aschauer (1989a) looks at the relation between public investment and private. He finds that public investment causes more private investment, by increasing the private rate of return.

On the other hand, Holtz-Eakin (1994) finds that there are no public sector capital productivity effects using US state level panel data. He argues that once state effects are accounted for, there is essentially no relationship. As he points out, this does not imply public capital is unproductive, simply that it does not contribute to productivity growth. However, this striking result seems somewhat odd; an implication is that roads, for example, have no effect on firms’ costs. It may be that the results are subject to bias from unmodelled heterogeneity and dynamics, which are serious albeit largely unrecognised problems in panel estimation; see Haque, Pesaran and Sharma (1999).

There are potential problems with this approach which were not recognised in early work. These stem from three sources. Firstly, non-stationarity and spurious regression. As is now well appreciated, if data are non-stationary (loosely, strongly trended), then there is a severe danger of ‘spurious regression’. As the variance of a non-stationary variable increases with time, unconnected variables may appear to be related. One solution is to take proper account of this by undertaking cointegration analysis. Another is to difference the data. This is appropriate only if the data do not cointegrate, as otherwise the differencing removes relevant long-run information from the sample. Secondly, there may be the problem of causality referred to above. Cointegration also allows us directly to address this. More pertinently, the effects will be second order at the level of aggregation we use in the disaggregated data, as any industry will have a small effect on aggregate transport. Finally, the production function approach ignores information implicit in other variables. In particular, the first order conditions from the firms’ maximisation problem offers a richer structure that may tie down the estimates better, generally via a cost-function approach, discussed below. However, this
is also more demanding in terms of data, and we may also be in danger of rejecting hypothesis because of a misspecification of the cost function.

### 3.2.2 Cost functions

Cost functions have been used in a variety of studies, including an excellent piece of work by Nadiri and Mamuneas (1996). Nadiri and Mamuneas look at the effect of highways on US productivity. In essence, the cost function approach uses the first order conditions from the cost minimisation part of the profit maximisation problem to give expressions for factor demand or factor (cost) shares which are functions of prices. Thus we can use information about factor prices, and treat factor inputs as endogenous variables in estimation. This also offers a richer menu of information about the cost, production and demand structure. However, it should be clear that this approach is not inconsistent with estimating production functions. They argue the production function approach typically gives estimates that are implausibly large. They report that on the basis of their survey, the typical elasticity produced using cost functions is around 0.20. Their own estimates come in at around 0.05 on average, just for highways.\(^6\)

Lynde and Richmond (1993a) have used a cost function approach for the manufacturing sector in the UK.\(^7\) They use total public capital and do not disaggregate. They find the elasticity of output with respect to public capital is around 0.20.

### 3.3 Summary of existing work

Table 1 briefly summarises some results, which have tended to vary somewhat. Aggregation seems to be an important issue; disaggregated studies tend to deliver lower estimates, as do cost functions. The elasticity of output with respect to the public capital figure for the UK from Lynde and Richmond, is about 0.20, compared to Aschauer’s ball-park figure of 0.40 for the US. The best evidence probably comes from disaggregated cost function approaches. Although still subject to potential criticism, the most reliable estimate is likely to be from Nadiri and Mamuneas (1996), who look purely at highway capital. As the table reports, they find an elasticity of around 0.05. This is just for roads. We conclude that there is no consensus in the literature, but early aggregate estimates are thought to be too high. Thus we might expect to find estimates lying between Lynde and Richmond and Nadiri and Mamuneas; in the range 0.05 to 0.20, in other words.

---

\(^6\)There are aspects of their work which give cause for concern. The demand functions use first differencing, which is generally invalid. The pooled time series data method involved pooling some parameters which may be problematic (see Appendix). The equations suffered from serial correlation, usually a sign of misspecification, which was in this case (as is admittedly common in panel studies) ‘corrected’ for in estimation. However, despite these caveats, this is an unusually thorough piece of work.

\(^7\)And also for other countries including the US; Lynde and Richmond (1992, 1993b).
4 Econometric strategy

Our aim is to estimate the parameters of the production functions for the sectors we examine. As output is measured on a value added basis, intermediate inputs net out. Thus any effects we find for transport or for similar public services are affecting total factor productivity.

We operationalise our analysis by estimating log-linear production functions of the form

\[ y_t = \beta_{10} + \beta_{11}t + \beta_{12}k_t + \beta_{13}n_t + \beta_{14}g_t \]  

(5)

where \( g \) is transport or the other relevant variables, which we interpret as approximations to the true production function. In principle, the production function holds at all times, but in practice there are dynamics which require us to use a dynamic specification including lags of the explanatory variables, and possibly measures of capacity utilisation.

4.1 Cointegration, levels and differences

As the data are non-stationary, we need to consider cointegration. A priori, we expect that if one or more cointegrating relationships exist, then either we find a unique relationship (5) or two relationships: (5) and another expressing a demand driven relationship between \( g \) and \( y \), say of the form

\[ g_t = \beta_{20} + \beta_{21}p + \beta_{22}y_t \]  

(6)

where \( p \) is the relative price of transport, \( \beta_{21} \) is the price elasticity and \( \beta_{22} \) is the income elasticity of demand for \( g \). The methodology for estimating and identifying such relationships using the Johansen procedure\(^8\) is now well understood, and will not be rehearsed here.

If we do not find cointegration, this does not necessarily mean we must abandon estimation. In the absence of cointegration, it is legitimate to look for relationships between differenced series. Were there cointegration, this would be a major misspecification; but otherwise it is legitimate. This does raise the issue of why we fail to find cointegration. The answer may be that there are important non-stationary series that we have failed to include in the regression. The most likely candidates are the stocks of human capital and knowledge, both of which are the key concepts in most of the new growth literature, and which are hard, and possibly impossible, to measure.

\(^8\)See, for example, Johansen (1995).
4.2 Dynamic panel data

The use of disaggregated data throws up additional issues. A major advantage is that it helps us to solve the reverse causality issue raised above, and which has plagued empirical growth studies. With disaggregated data, it is simply implausible that a single small sector can influence the entire economy.

The other advantage is the panel dimension, that may allow more efficient estimation. It is not essential, or even desirable, to completely pool the data, as we have enough to run a regression for each industry separately. Indeed, the combination of panel data and dynamics raises some important econometric problems that are often ignored. The panel we are using has dimensions in $T$ and $N$ of roughly equal orders. As static models are rarely adequate for typical time series, dynamic models are usually appropriate. There are profound problems that result from heterogeneity in the model parameters that emerge as soon as a lagged dependent variable is introduced. This problem was forcefully addressed by Pesaran and Smith (1995). Unlike in static models, estimates are inconsistent even in large samples, although there are still problems in static regressions: Haque, Pesaran and Sharma (1999). Happily, in our data set $T$ is sufficiently large to allow individual sectoral estimation. While it is implausible that the dynamic specification is common to all sectors, it is conceivable that at least some of the long-run parameters of the model may be common across the sectors. We can then exploit the cross-sectional dimension to gain more precise estimates of these average long-run parameters, either by averaging the individual sectoral estimates, or by pooling the long-run parameters, if the data allows. This latter approach is known as Pooled Mean Group estimation, and was introduced by Pesaran, Shin and Smith (1999) (PSS).

Setting this out more precisely, the unrestricted specification for the system of Auto Regressive Distributed Lag (ARDL) equations for $t = 1, 2, \ldots, T$ and $i = 1, 2, \ldots, N$ for the dependent variable $y$, which is output in each industry is

$$y_{it} = \sum_{j=1}^{m} \lambda_{ij} y_{i,t-j} + \sum_{j=1}^{n} \delta_{ij} x_{i,t-j} + \mu_i + \varepsilon_{it} \quad (7)$$

where $x_{i,t-j}$ is the $(k \times 1)$ vector of explanatory variables for group $i$ and $\mu_i$ are the fixed effects. In principle the panel can be unbalanced and $m$ and $m$ may vary across industries. (7) can be reparameterised as a VECM system.

$$\Delta y_{it} = \theta_i (y_{i,t-1} - \beta' x_{i,t-1}) + \sum_{j=1}^{m-1} \gamma_{ij} \Delta y_{i,t-j} + \sum_{j=1}^{n-1} \gamma'_{ij} \Delta x_{i,t-j} + \mu_i + \varepsilon_{it} \quad (8)$$

where the $\beta_i$ are the long-run parameters and $\theta_i$ are the error correction parameters. The pooled mean group restriction is that the elements of $\beta$ are common across industries.

$$\Delta y_{it} = \theta_i (y_{i,t-1} - \beta' x_{i,t-1}) + \sum_{j=1}^{m-1} \gamma_{ij} \Delta y_{i,t-j} + \sum_{j=1}^{n-1} \gamma'_{ij} \Delta x_{i,t-j} + \mu_i + \varepsilon_{it} \quad (9)$$

---

9 Quah (1993) has referred to such data sets as ‘data fields’.
10 There is a well-known ‘small $T$’ problem with dynamic panels (Arellano and Bond (1991)), but this is not relevant here as the fixed-effects problem from the initial conditions declines rapidly as $T$ rises.
All the dynamics and the ECM terms are free to vary. Estimation could proceed by OLS, imposing and testing the cross-industry restrictions on $\beta$. However, this will be inefficient as it ignores the contemporaneous residual covariances. A natural estimator is Zellner’s SUR method,\textsuperscript{11} which is a form of feasible GLS. SUR estimation is only possible if $N$ is sufficiently smaller than $T$, as the method involves estimating a large number of error covariance in terms. In our case this condition does hold.\textsuperscript{12}

The industries we examine are very different. So we want to test for homogeneity of the parameters in the model. PSS argue that in panels omitted group specific factors or measurement errors are likely to severely bias the individual industry estimates. This may explain, they suggest, why it is a commonplace in empirical panel to report a failure of the ‘poolability’ tests based on the group parameter restrictions. For example, Baltagi and Griffin (1997, p 308) states that although the poolability test is massively failed ($F(102,396) = 10.99$; critical value about 1.3), ‘like most researchers we proceed to estimate pooled models.’\textsuperscript{13} In particular, as argued by PSS for the cases they examine, we might believe that the long-run parameters are common between industries, in which case efficient estimation is by PMG.\textsuperscript{14} But, arguably, homogeneity is implausible in the current exercise. In either case, we have the possibility of estimating the average effect of transport on productivity over industry as a whole; in all probability, it is best to do this by the mean group method.

In conclusion, it is possible to exploit common cross-industry factors if we can, but if this proves impossible, we are still able to use the individual industry estimates. Even in the latter case, we can calculate the average effect over all industries or over sub-sets of industries, in the knowledge that the estimates this provides are consistent.

5 Data and transport proxies

As our aim is directly to estimate the parameters of production functions, the basic data we require are (log) output $q$, employment $n$, capital stock $k$ and infrastructure proxies. Each of output, employment and the capital stock for the whole non-government economy and manufacturing appear to be non-stationary, and for-

\textsuperscript{11}Zellner (1962).
\textsuperscript{12}We have $T = 73$ in typical regressions, after allowing for lags and $N = 27$. The number of parameters in the covariance matrix is given by $(N^2 + N)/2$ so there are over a thousand degrees of freedom remaining. For cases where this does not hold, which may be typical in macroeconomic exercises using annual data, PSS suggest a maximum likelihood estimator.
\textsuperscript{13}PSS propose a Hausman test. This is based on the result that an estimate of the long-run parameters in the model can be derived from the average (mean group) of the industry regressions. This is consistent even under heterogeneity. However, if the parameters are in fact homogeneous, the PMG estimates are more efficient.

\[ H = \hat{q}'[\text{var}(\hat{q})]^{-1}\hat{q} \sim \chi_k^2 \]

where $\hat{q}$ is a $(k \times 1)$ vector of the difference between the mean group and PMG estimates and $\text{var}(\hat{q})$ is the corresponding covariance matrix. Under the null that the two estimators are consistent but one is efficient, $\text{var}(\hat{q})$ is easily calculated as the difference between the covariance matrices for the two underlying parameter vectors. If the poolability assumption is invalid then the PMG estimates are no longer consistent and we fail the test.

\textsuperscript{14}This has been applied by Lukacs, Pain and Price (1999).
mal ADF tests confirm this. The industry data are for the same variables, over 27 industries. The individual series behave in a varied manner, but generally suggest non-stationarity. Turning to transport, we had a range of proxy variables available. For transport as a whole, there are several possibilities. The basic transport sector output figures, with the total sector worth 5.4\% of GDP, are given in Figure 1. An excluded category is private road use. This is hard to measure, but we have figures for passenger road miles, and these could be translated into an ‘output’ series if we knew the ‘price’ of a passenger km. To construct this, we took an average inter-city fare per km on 20 routes into London in 1995, and then used the rail and total RPI to construct a real series (illustrated in Figure 2). We also used the output from the communications industry as a similar infrastructural industry.

6 Aggregate results

Despite our reservations about the difficulties associated with the use of aggregate data, we began by looking at whole private-sector economy and manufacturing data.

6.1 Levels: cointegration analysis

The methodology we adopted was to begin by estimating an unrestricted VAR and then test for lag length using standard criteria (SBC, AIC). We explored all the proxies, $g$, described above. Apart from these, the variables included were (log) output $q$, employment $n$, capital stock $k$ and communications output. After the lag length was selected, we tested for the number of cointegrating relationships for both non-governmental whole economy and manufacturing industry, including transport, aviation and communications. In essence, there was very little evidence for cointegration. Where we did try to identify long-run relationships, they proved impossible to interpret, either in terms of sensible production functions or a demand for transport or aviation. When we tried to impose identifying restrictions and were able to find convergence in estimation, the restrictions were massively rejected and the parameters nonsensical.

As the Johansen procedure is sensitive to the various choices on lag length, exogeneity, restrictions on constants and trends and cointegrating rank that can be made, we also considered single equation ECM results as an alternative methodology. This is known to be a robust test an estimation method where there is a unique cointegrating

\cite{Greenslade1999}.
relationship.\footnote{For example, see Kremers, Ericsson and Dolado (1992), although they assume exogeneity of the explanatory variables. Zivot (1994) suggests the ECM test is powerful under less restrictive assumptions.} We began by estimating an error correction mechanism:

\[ \Delta q_t = \text{dynamics} - \lambda (q_{t-1} - \beta_1 n_{t-1} - \beta_2 k_{t-1} - \beta_3 g_{t-1} - \beta_4 t) \]  

(10)

for manufacturing and the whole economy. Equation (10) here embodies a long-run relation between output and factor inputs of the form

\[ q = \beta_1 n + \beta_2 k + \beta_3 g + \beta_4 t \]  

(11)

The adjustment coefficient \( \lambda \) offers a test for the existence of cointegration which is distributed asymptotically standard normal for exogenous regressors. Constant returns to scale (CRS) to internal inputs implies \( \beta_1 + \beta_2 = 1. \beta_4 t \) captures technical progress. Unrestricted estimates for the whole economy suggested \( \beta_1 \) exceeded 4 and \( \beta_2 \) was negative; for manufacturing both \( \beta_1 \) and \( \beta_2 \) were negative. In the light of this, we proceeded to impose CRS with respect to private capital and labour, and a capital share of 0.35, in line with aggregate national account figures. Given these restrictions, the results for the whole sample still showed no sign of cointegration in either case.

We therefore conclude there is no evidence for cointegration in this set of variables. This allows us to proceed to an examination of growth rates.

### 6.2 Growth rates

We continued to examine total factor productivity. We ran regressions of the form

\[ \Delta f_t = \alpha_0 + \alpha_1 \Delta f_{t-1} + \alpha_2 \Delta g_{t-1} + \alpha_3 \Delta g_{t-1} \]  

(12)

where \( f \) is (log) total factor productivity, defined so that \( f \equiv y - 0.65n - 0.35k \), for each sector, and \( g \) is the relevant transport proxy. The equation is defined so the long-run impact is estimated through \( \alpha_2 \): allowing for dynamics, the full long-run effect is \( \alpha_2/(1-\alpha_1) \). We were prepared to experiment with the dynamics. In the reported results, \( \Delta g \) is the change in the log of the widest transport aggregate. Estimating this model for the whole (private sector) economy did produce some weak support for our hypothesis: see Table 2. The long-run growth impact of transport as a whole is 0.11, with a \( t \) ratio of 1.51. The diagnostics for this equation are satisfactory.

We found no evidence of a similar relationship for manufacturing over the full sample, but some evidence of a weaker effect over the restricted sample of 1985 to 1997 (Table 3).

These results are weakly suggestive that there is an effect of the type we expect. But clearly, neither of them are strong and the results are not robust. Indeed, there are reasons to be sceptical. Firstly, there is an unresolved problem of causality with this methodology. We did lag transport, but there could still be an identification problem. We therefore instrumented \( \Delta g_{t-1} \), and estimated a larger and more significant effect in the whole economy case; this suggests causality is not the issue. But there is also an unresolved aggregation effect, as we believe that different sectors within the total are likely to have different coefficients - a problem not specific to the effect of
Table 2  
Aggregate growth regression for whole (private) economy, 1975 to 1997

<table>
<thead>
<tr>
<th>Variable</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>t Ratio [Prob]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>0.0010532</td>
<td>0.0010036</td>
<td>1.0494 [.297]</td>
</tr>
<tr>
<td>$\Delta f_{t-1}$</td>
<td>0.44955</td>
<td>0.082498</td>
<td>5.4492 [.000]</td>
</tr>
<tr>
<td>$\Delta g_{t-1}$</td>
<td>0.060606</td>
<td>0.041244</td>
<td>1.4694 [.145]</td>
</tr>
<tr>
<td>$\Delta \Delta g_{t-2}$</td>
<td>0.087911</td>
<td>0.033063</td>
<td>2.6589 [.009]</td>
</tr>
<tr>
<td>$\Delta \Delta g_{t-3}$</td>
<td>0.090894</td>
<td>0.033007</td>
<td>2.7538 [.007]</td>
</tr>
</tbody>
</table>

R-Squared 0.36414  
R-Bar-Squared 0.33422  
Serial Correlation $\chi^2_4 = 3.5624 [.468]$  
Functional Form $\chi^2_1 = 2.5440 [.111]$  
Normality $\chi^2_2 = 4.4906 [.106]$  
Heteroscedasticity $\chi^2_1 = 0.41998 [.517]$  

Table 3  
Aggregate growth regression for manufacturing, 1985 to 1997

<table>
<thead>
<tr>
<th>Variable</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>t Ratio [Prob]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>0.0061013</td>
<td>0.0015826</td>
<td>3.8553 [.000]</td>
</tr>
<tr>
<td>$\Delta g_{t-1}$</td>
<td>0.099867</td>
<td>0.081712</td>
<td>1.2222 [.227]</td>
</tr>
<tr>
<td>$\Delta \Delta g_{t-2}$</td>
<td>0.070637</td>
<td>0.057227</td>
<td>1.2343 [.223]</td>
</tr>
</tbody>
</table>

R-Squared 0.052397  
R-Bar-Squared 0.013719  
Serial Correlation $\chi^2_4 = 1.5403 [.819]$  
Functional Form $\chi^2_1 = 0.054322 [.816]$  
Normality $\chi^2_2 = 3.3043 [.192]$  
Heteroscedasticity $\chi^2_1 = 0.025174 [.874]$  

transport, of course. We conclude this section by noting there is some weak evidence for transport impacting on growth from the aggregate results. We now proceed to the disaggregated estimates, which are much less vulnerable to the simultaneity and aggregation problems.

7 Disaggregate results

We have 27 industrial sectors, with data for output and employment. ONS provided individual sectoral capital stock figures.
7.1 Levels

We began by estimating single equation Cobb-Douglas production functions with a simple, common lag structure, using an ECM specification consistent with cointegration among the variables. Thus, we estimate for each sector an equation along the following lines:

$$\Delta q_{i,t-1} = \beta_0 + \beta_1 \Delta q_{i,t-1} + \beta_2 \Delta q_{i,t-2} - \beta_3 (q_{i,t-1} - \beta_4 k_{i,t-1} - \beta_5 n_{i,t-1} - \beta_6 g_{t-1} - \beta_7 t).$$  (13)

This ECM specification allows one to estimate the long-run parameters directly (e.g., $\beta_6$ is the long-run effect of transport). We estimate such an equation for each industry, then take averages. This will give a consistent estimate of the average effect even if each industry has a different parameter. As discussed above, it is often the case in dynamic data fields that individual coefficients are hard to interpret, but the average is plausible. However, in this case as for the aggregate, they were not, once again suggesting we do not have a cointegrating set. As for the aggregate, we imposed CRS and shares for each sector, using known values from the input-output tables. This did not result in better estimates. For example, for the broad transport aggregate, the average estimate of the output elasticity is -2.96. So this is not evidence for cointegration among these industrial sets.20

7.2 Growth rates

Given these results, we were able to move to equations of form

$$\Delta f_{i,t-1} = \beta_0 + \beta_1 \Delta f_{i,t-1} + \beta_2 \Delta f_{i,t-2} + \beta_3 (1 - \beta_1 - \beta_2) \Delta g_{t-1} + \beta_4 \Delta \Delta g_{t-1} + \beta_5 \Delta \Delta g_{t-2},$$  (14)

These are the sectoral equivalents of the total factor productivity growth equations, defined so that the long-run growth impact (e.g., $\beta_3$) can be read straight off. The factor shares are derived from the input-output tables. The results are presented in Table 4. Only the transport effects are reported here; the other parameters are not of any particular interest, except possibly the constant (see below). We included the components of transport in the analysis, as there may still be identifiable effects from aggregate activity in these sectors.

Not many among the 27 are significant, and the estimates vary widely (although only one of the estimates is significantly negative, and then only at the 10% level). But this is not necessarily a problem. This is a classic result with this kind of data; the aim of the dynamic panel methodology is to try to use these estimates to get a good fix on the average effect. Moreover, despite the fact that we did not play with the dynamic specification, there is not much evidence for autocorrelation (by LM(4) tests) in many of these equations. 4 out of 27 equations fail at the 5% level. This is not a problem in terms of estimating the average coefficient. The average value of the effect of transport on growth for the full 27 is sensible, at 0.131.21

---

20 There is a growing literature on panel tests for cointegration: see Pedroni (1999), McCoskey and Kao (1998) Karlsson and Löthgren (1999) for papers on residual based tests; see Larsson, Lyhagen and Löthgren (1998) for a Johansen panel test. However, in view of the total absence of sensible single equation results we did not implement any of these tests.

21 The meaning of ‘sensible’ is discussed further at the end of this section.
Table 4
Disaggregate growth regressions, 1979 to 1997

<table>
<thead>
<tr>
<th>Sector</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>t Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agriculture, forestry and fishing</td>
<td>0.074816</td>
<td>0.250548</td>
<td>0.30</td>
</tr>
<tr>
<td>Basic metals</td>
<td>0.334041</td>
<td>0.280884</td>
<td>1.19</td>
</tr>
<tr>
<td>Other marketed services</td>
<td>0.100752</td>
<td>0.187684</td>
<td>0.54</td>
</tr>
<tr>
<td>Construction</td>
<td>0.199712</td>
<td>0.161838</td>
<td>1.23</td>
</tr>
<tr>
<td>Chemicals and man made fibres</td>
<td>0.214974</td>
<td>0.164253</td>
<td>1.31</td>
</tr>
<tr>
<td>Communications</td>
<td>0.190601</td>
<td>0.119126</td>
<td>1.60</td>
</tr>
<tr>
<td>Computers and office equipment</td>
<td>0.745594</td>
<td>0.381345</td>
<td>1.96</td>
</tr>
<tr>
<td>Distribution</td>
<td>0.096053</td>
<td>0.105092</td>
<td>0.91</td>
</tr>
<tr>
<td>Extraction</td>
<td>-0.24265</td>
<td>0.40907</td>
<td>-0.59</td>
</tr>
<tr>
<td>Electrical Engineering</td>
<td>0.231319</td>
<td>0.168094</td>
<td>1.38</td>
</tr>
<tr>
<td>Finance</td>
<td>-0.13584</td>
<td>0.148801</td>
<td>-0.91</td>
</tr>
<tr>
<td>Food, beverages and tobacco</td>
<td>0.017661</td>
<td>0.060347</td>
<td>0.29</td>
</tr>
<tr>
<td>Coke, petroleum and nuclear</td>
<td>0.22745</td>
<td>0.357302</td>
<td>0.64</td>
</tr>
<tr>
<td>Mechanical engineering</td>
<td>0.629813</td>
<td>0.151347</td>
<td>4.16</td>
</tr>
<tr>
<td>Metal products</td>
<td>0.3917</td>
<td>0.143275</td>
<td>2.73</td>
</tr>
<tr>
<td>Non-metallic minerals</td>
<td>0.163297</td>
<td>0.159477</td>
<td>1.02</td>
</tr>
<tr>
<td>Motor vehicles and parts</td>
<td>0.776844</td>
<td>0.248725</td>
<td>3.12</td>
</tr>
<tr>
<td>Non-market services</td>
<td>-0.14064</td>
<td>0.074395</td>
<td>-1.89</td>
</tr>
<tr>
<td>Other manufacturing</td>
<td>-0.03486</td>
<td>0.174004</td>
<td>-0.20</td>
</tr>
<tr>
<td>Other means of transport</td>
<td>-0.31521</td>
<td>0.283586</td>
<td>-1.11</td>
</tr>
<tr>
<td>Paper, printing and publishing</td>
<td>0.30357</td>
<td>0.112148</td>
<td>2.71</td>
</tr>
<tr>
<td>Precision and optical instruments</td>
<td>-0.06179</td>
<td>0.234696</td>
<td>-0.26</td>
</tr>
<tr>
<td>Rubber and plastics</td>
<td>0.081601</td>
<td>0.186927</td>
<td>0.44</td>
</tr>
<tr>
<td>Textiles, leather and clothing</td>
<td>0.046344</td>
<td>0.146951</td>
<td>0.32</td>
</tr>
<tr>
<td>Transport</td>
<td>-0.63608</td>
<td>0.84213</td>
<td>-0.76</td>
</tr>
<tr>
<td>Electricity, gas and water</td>
<td>0.256931</td>
<td>0.453318</td>
<td>0.57</td>
</tr>
<tr>
<td>Wood and wood products</td>
<td>0.125355</td>
<td>0.252203</td>
<td>0.50</td>
</tr>
</tbody>
</table>

Estimation method: SUR.
Mean impact: 0.135 (t ratio 2.31).
ways of working out the $t$ ratio; the most robust of these is 2.31 (the other 2.28). So we have a significant average effect from transport broadly defined. A Wald test of the ‘poolability’ restriction strongly rejects ($\chi^2_{26}$ is 76.3, well above the critical value), but this is not a problem, as we are happy to assume divergent elasticities, even if we cannot estimate them precisely. For the record, though, when we do pool this parameter, it takes the value of 0.045 with a low $t$ ratio of 1.3, similar to the un-pooled average. Interestingly, though, the OLS pooled estimate is 0.163 with a $t$ ratio of 3.30, similar to the unpooled average. The results from the other transport proxies are very similar. If we take out aviation, the estimated long run effect is slightly smaller but similar: it is 0.111 ($t$ ratio 1.46). So the aggregate including aviation is rather better determined. Interested in the idea that private road use, which has grown enormously with increased car ownership over our sample, might be important for growth, we added in our proxy for car use. The estimated coefficient fell to 0.105. We then ran a regression replacing the transport aggregate with road use. Intriguingly, far from increasing growth, private car use significantly reduces it. This presumably follows from congestion effects.

All the reported results are based on lagged transport proxies as growth spillover effects are unlikely to feed through contemporaneously. It is also arguable that contemporaneous transport output and factor productivity are procyclical, which might induce a spurious correlation, although the lag structure should mop up some of this. For the record, contemporaneous transport generally enters with a larger, indeed arguably implausibly larger, coefficient. To further control for the cycle, we introduced a series measuring deviations from the Hodrick Prescott filter trend on log GDP, and also growth in GDP. Neither of these made any difference to the results.

So we have a figure for the effect of transport on output (or, equivalently, total private factor productivity) in the region of 0.13. Is this number plausible? In a production function context, we would interpret the estimate as the elasticity of output with respect to transport. The marginal product of transport output can be calculated as $0.13Y_i/G$ where $Y_i$ is the level of output of the $i$th industry and $G$ transport output. There are 27 industries, and as government is about 20% of GDP, the average share of each is roughly 3%. On this basis, the average marginal product of transport can be calculated to be roughly 7.45%. As we are not identifying level effects, it may be better to interpret the figures as marginal contributions to growth. Taking this line, as the transport sector has grown on average at a rate of about 0.7% per quarter, the contribution of transport to growth at the margin has been 0.095%.

8 Conclusions

There are strong reasons for believing that infrastructure output may affect the level of output via external effects on productivity. It is often observed that an important component of infrastructure is transport. In this paper, we have examined this link for the UK. Using aggregate data, we found only weak evidence for the proposition. However, when we used disaggregated industry data we were able to exploit the extra power from the cross-sectional dimension to estimate a significant effect. We used figures for total factor productivity growth (calculating factor shares from the input-output tables) as we did not find a cointegrating relationship with our data set, which
we argue is due to unobservable supply side variables, possibly human capital or knowledge. We then employed appropriate dynamic panel techniques on the disaggregated industry data, and found that transport growth has a small but significant positive effect on total factor productivity growth. Private car ownership appears to have a negative impact, suggesting that increased road congestion offsets the productivity gains that might otherwise flow.

References


