OECD

International Transport Forum

**JOINT TRANSPORT RESEARCH CENTRE**

# Agglomeration Economies and Transport Investment

**Daniel J. GRAHAM**
**Imperial College**
**University of London**
**London, United Kingdom**

**TABLE OF CONTENTS**

# ABSTRACT

This paper is concerned with the links between agglomeration, productivity and transport investment. If improvements in transport systems give rise to changes in the mass of economic activity accessible to firms, for instance by reducing travel times or the costs of travel, then they can induce positive benefits via agglomeration economies. The paper presents empirical results from an econometric analysis of the relationship between productivity and accessibility to economic activity for different sectors of the UK economy. The results show that agglomeration economies do exist and that they can be substantial, particularly for services. Furthermore, the effect of agglomeration externalities is not trivial when considered in the context of transport appraisal. Initial calculations typically indicate additions to conventional user benefits of 10%-20% arising from increasing returns to economic mass.

## 1. INTRODUCTION

A recent paper by Venables (2007) develops a theoretical model to demonstrate some important links between transport provision and agglomeration. He shows that if, as urban economic theory suggests, there are increasing returns to agglomeration, then transport investments may induce positive productivity benefits by effectively raising accessibility to economic mass. Any such agglomeration externalities can be classed as `wider benefits' of transport investment in the sense that they are typically not captured in a standard cost-benefit appraisal.

To understand the magnitude of the potential `wider benefits' of transport investment we first need quantitative estimates of returns to agglomeration. In other words we require some empirical verification of the existence and magnitude of the relationship between productivity and accessibility to economic mass. Preferably, we want to examine this relationship separately for different sectors of the economy because the benefits derived from agglomeration are unlikely to be uniform across industries.

This paper describes the results of new empirical research on the relationship between agglomeration and productivity for different sectors of the UK economy. It also considers the implications of agglomeration economies for the evaluation of transport investment. The results show that agglomeration economies do exist and that they can be substantial, particularly for services. If transport investments change the densities available to firms, for instance through a reduction in travel times or in the cost of travel, then there are likely to be positive gains from agglomeration.

D. Graham

Furthermore, the effect of agglomeration externalities is not trivial when considered within the framework of transport appraisal. Initial calculations typically indicate additions to conventional user benefits of 10%-20% arising from increasing returns to economic mass.

The paper is structured as follows. Section 2 reviews the literature on agglomeration and productivity and discusses the relationship between transport investment and agglomeration. Section 3 describes the methodology used to estimate agglomeration economies. Empirical results are presented in section 4, including a review of some recent applications of agglomeration benefits within transport appraisal. Conclusions are then drawn in the final section.

## 2. AGGLOMERATION ECONOMIES AND TRANSPORT INVESTMENT

### 2.1. Agglomeration and productivity

The tendency towards concentration or agglomeration is perhaps the most widely observed feature of the spatial organisation of economic activity. It can be discerned across the Globe at a variety of different geographical levels. Agglomeration is evident, for instance, in the existence and growth of cities, in the formation of industrial regions and districts, and in the clustering of like activities within the same neighbourhood of a town or city.

Attempts to explain the microfoundations of agglomeration generally start from the premise that cities and industrial concentrations would not form if there were not some tangible benefits that accrue to firms. The advantages derived through the spatial concentration of economic activities are referred to generically as *agglomeration economies*.

Typically, a distinction is made between agglomeration effects that arise from the scale or density of activity within a particular industry and from those due to urban scale or city size. Economies of industry concentration, termed localization economies, are external to the firm but internal to the industry and are principally thought to be sourced through labour market pooling, the sharing of intermediate inputs, and knowledge sharing or `technological spillovers'. Economies of urban concentration, or urbanization economies, are external to the firm and the industry but internal to the city with benefits arising from the existence of local public goods, the scale of markets, the proximity of input-output sharing, and other kinds of inter-industry interaction.

The theoretical foundations for the existence of agglomeration economies are now well established (see for example Fujita and Thisse 2002; Duranton and Puga 2005). There is also a body of empirical work that has sought to identify these externalities and to quantify their effects on productivity. There are a number of excellent up-to-date surveys of the empirical literature on agglomeration (see in particular Rosehthal and Strange 2004; Eberts and McMillen 1999). This literature has concentrated largely on manufacturing with, until very recently, few published results on the link between agglomeration and service sector productivity. This is almost certainly due the poor quality of service sector data for most countries compared to the manufacturing statistics. Nevertheless, since services now comprise such a large share of many national and urban economies, the emphasis on manufacturing represents a real limitation.

To identify agglomeration economies empirical work typically proceeds by constructing variables that measure the extent of industry and urban concentration, and uses these within a production or cost function framework to estimate effects on productivity. Urbanization is often represented by the total population or total employment of an urban area. Localization is proxied using some measure of local industry scale such as employment. Table 1 provides a summary of some prominent studies of the effects of agglomeration on productivity. It summarises those studies that have produced an actual elasticity estimate of the effects of agglomeration, rather than those that have detected agglomeration effects through the use of dummy variables or other limited variable methods.

Table 1. **Estimates of agglomeration economies from production function analyses**

| *Author* | *unit of analysis* | *independent variable* | *elasticity* |
|---|---|---|---|
| 1  Aaaberg (1973) | Swedish cities | city size (population) | 0.02 |
| 2  Shefer (1973) | US MSAs | RTS at MSA aggregation | 0.2 |
| 3  Sveikauskas (1975) | US MSAs | city size (population) | 0.06 |
| 4  Kawashima (1975) | US MSAs | city size (population) | 0.2 |
| 5  Fogarty and Garofalo (1978) | US MSAs | city size (population) | 0.1 |
| 6  Moomaw (1981) | US MSAs | city size (population) | 0.03 |
| 7  Moomaw (1983) | US MSAs | city size (population) | 0.05 |
| 8  Moomaw (1985) | US MSAs | city size (population) | 0.07 |
| 9  Nakamura (1985) | Japanese Cities | city size (population) | 0.03*a* |
| 10 Tabuchi (1986) | Japanese Cities | city size (population) | 0.04 |
| 11 Louri (1988) | Greek Regions | city size (population) | 0.05 |
| 12 Sveikauskas *et al.* (1988) | US MSAs | city size (population) | 0.01*b* |
| 13 Nakamura (1985) | Japanese Cities | industry size (employment) | 0.05 |
| 14 Henderson (1986) | Brazilian Cities | industry size (employment) | 0.11*c* |
| 15 Henderson (1986) | US MSAs | industry size (employment) | 0.19*d* |
| 16 Henderson (2003) | US MSAs | industry size (no. of plants) | 0.03*e* |
| 17 Ciccone and Hall (1996) | US States | employment density | 0.06 |
| 18 Ciccone (2002) | EU regions | employment density | 0.05 |
| 19 Rice et al. (2006) | GB NUTS 3 | proximity / travel time | 0.04 |

*Notes:* MSA - Metropolitan Statistical Area,
        a - mean value for 14 manufacturing industries,
        b - mean value from 5 model specifications,
        c - mean value for ten industries,
        d - mean value for 9 industries,
        e - mean value for 4 model specifications.


With the exception of studies 17 and 18, which are concerned with effects on total economic productivity, the estimates shown in table 1 are for manufacturing industries. Elasticities describing the strength of localization economies are given in 13, 14, 15, and 16; the remaining estimates show the effect of urbanization economies on productivity.

The estimates of urbanization economies for manufacturing industries shown in table 1 range from 0.01 to 0.20, but the majority of values are under 0.10. This indicates that a doubling of city size is typically associated with an increase in productivity of somewhere between 1% and 10%. The estimates given in the table are all positive although Henderson (1986) and Henderson (2003) do report difficulties in identifying urbanization effects on productivity.

Table 1 shows four estimates of localization economies. Nakamura (1985) estimates the effect of localization economies on the productivity of 20 manufacturing industries. He quotes an unweighted average elasticity of productivity with respect to industry size of 0.05. Henderson (1986) using industry level data for US MSAs and Brazilian cities also find positive localization economies. His estimates for Brazil vary by industry, with a maximum elasticity estimate of 0.20 and a minimum of 0.03, the mean value over 10 industries is 0.11. For US MSAs estimated localization elasticities range from 0.09 to 0.45 with a mean value of 0.19. Henderson (2003) estimates a mean localization elasticity of 0.03.

In addition to studies using MSA population and employment to represent city and industry size there other studies that have incorporated some measures of distance or density into the specification of agglomeration effects. Two papers are particularly interesting in this respect. Ciccone and Hall (1996) derive an equation to estimate the effects of county-level employment density on aggregate state productivity for the United States. They find that over 50% of the variance in aggregate labour productivity across states can be explained by variance in the density of employment and that a doubling of employment density is associated with a 6% increase in average labour productivity (i.e. an elasticity of 0.06). Ciccone (2002) extends the analysis to European data and estimates an elasticity of labour productivity with respect to employment density of 0.045.

Thus, from the empirical literature, we find evidence to support the theory of increasing returns to urban density and of returns to industry size.


## 2.2. Transport investment and agglomeration

It seems intuitively reasonable to suppose a further link between transport provision and the benefits that arise from the spatial concentration of economic activity. Transportation costs are crucial in determining the mass of economic activity (including population) that firms can access. New investments in transport can render a larger scale of activity more accessible by reducing travel times or the costs of travel, giving rise to positive agglomeration benefits. Conversely, where transports systems work inefficiently, or where there are constraints on accessibility, these may inhibit the generation and distribution of agglomeration externalities.

A crucial issue here is that agglomeration economies are *externalities*, that is, they arise as a side effect of the activities of firms which have consequences for the wider economy. This is very important from the point of view of transport appraisal because traditional methods of appraisal based on valuation of travel times do not recognise these types of externalities. For this reason agglomeration effects of transport investment can be classes as *wider economic benefits* because they represent market imperfections that are not accounted for in a standard cost-benefit appraisal.

Venables (2007) formalises this argument and shows that estimates of the elasticity of productivity with respect to agglomeration can be used to shed light on the magnitude of the external benefits of transport improvements. He develops a theoretical model of an urban economy that links productivity to transport investment via effects on city size. His objective is to distinguish real income
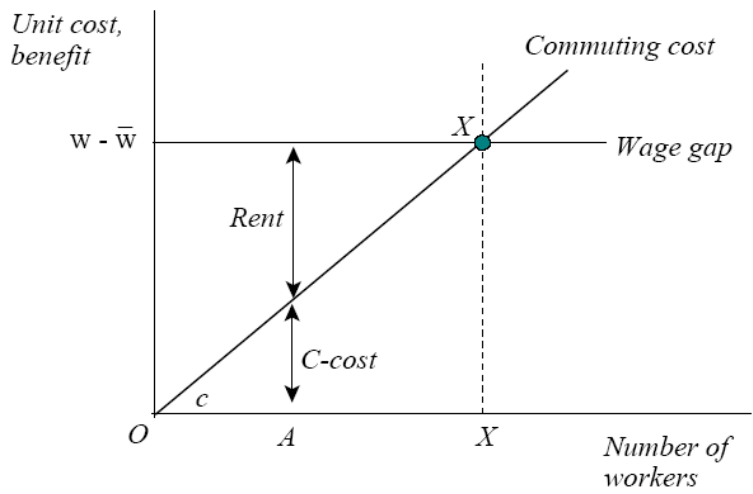
Figure 1a. **Urban equilibrium**



Figure 1b. **Net gains from transport improvement**
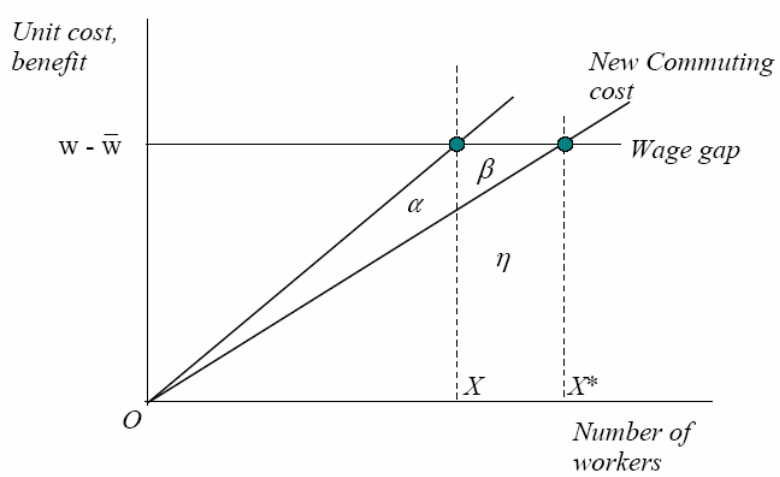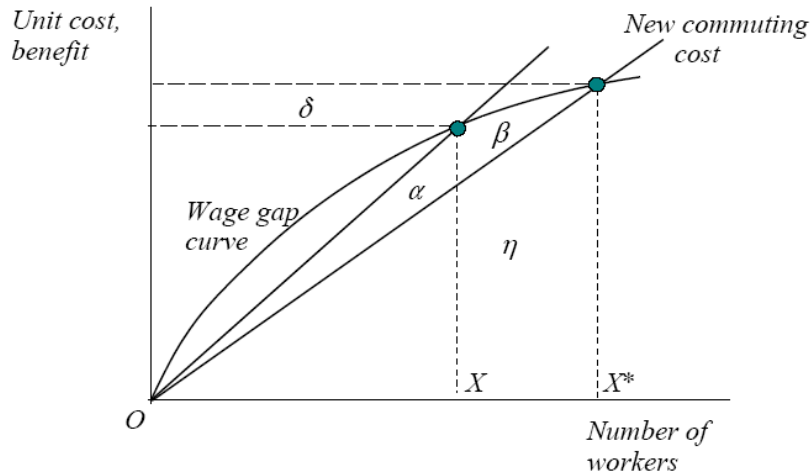


D. Graham

Figure 1c. **Net gains form transport improvement with endogenous productivity**



changes that result from transport investment due to a productivity-city size (agglomeration) effect, from those economic benefits that are captured in standard transport appraisals and which arise from resources saved in commuting and from an increase in urban output.

Venables' paper provides a clear demonstration of the key relevant arguments that link transport and agglomeration. A diagraming representation of the model taken from Venables' paper is given in figure 1.

Figure 1a shows an urban equilibrium in which the size of the city is determine at point $X$, where the wage gap between city workers and non-city workers is taken up in the travel costs of the city worker who is most distant from the CBD.

Figure 1b shows that when a transport improvement is made commuting costs are shifted downwards and consequently the city expands to point $X^*$. The total change in the resources used in commuting is $\eta - \alpha$ , which combined with the change in output ( $\beta + \eta$ ), yields a net benefit from the transport improvement of $\alpha + \beta$ .

In Figure 1c, Venables shows the implications of the existence of a city size-productivity gradient. If larger cities have higher productivity due to agglomeration externalities then the wage gap can be expressed, not as a constant gap, but as a concave curve that increases with city size. Equilibrium is found at the intersection of the commuting cost and wage gap curves. The fact that productivity is non-constant with respect to city size means that the real income gain from a transport improvement is $\alpha + \beta + \delta$ ; where $\delta$ measures the increase in productivity experienced by city workers and is akin to a measure of the elasticity of productivity with respect to city size.

In this way Venables demonstrates that there are external benefits from transport investment related to agglomeration and that these can be quantified using elasticities of productivity with respect to some measure of agglomeration.

# 3. ESTIMATING AGGLOMERATION ECONOMIES

The previous literature has indicated that agglomeration externalities do exist for manufacturing, but the sectoral coverage of existing work is incomplete and the analysis of agglomeration is typically based on data for relatively aggregated industries and spatial areas.

The purpose of the research described in this paper is to estimate a set of agglomeration elasticities that are compatible with the objective of assessing the wider economic benefits of transport investment. We do this for a comprehensive range of industries to find out whether agglomeration externalities really matter across sectors and whether they might be important in assessing the benefits of transport investment[2].

The empirical analysis uses firm level data to represent spatial variance in productivity along with data for small areas to construct measures of the agglomeration `experienced' by firms. The analysis proceeds in four steps. First, we gather data on the production characteristics of firms across a range of sectors. We then use Geographical Information System (GIS) software to identify the location of each firm in an electronic map. In the third step, we overlay on top of our map of firms a framework of small spatial units and use these to construct measures of the agglomeration experienced by each firm in each location. Finally, we use the firm data and the measures of agglomeration within a production function framework to estimate the effect of agglomeration on firm productivity.

There are several advantages to the micro firm-level approach we adopt here rather than the conventional method which uses aggregate spatial areas as the units of observation:

i **Consistency with theory** - the assumptions we use to analyse production behaviour presuppose firms as the basic decision making units, not aggregate spatial areas. Thus, modelling at the firm level provides consistency with the theory we draw upon to analyse productivity.

ii **Compatible measures of agglomeration -** by locating each firm geographically we can capture a high degree of spatial detail in our measures of agglomeration and avoid using data based on large pre-defined geographic units such as administrative areas or metropolitan definitions. Furthermore, using a distance based approach we can include an implicit transport dimension in the measure of agglomeration by considering not just the scale of economic activity within some concentration, but how accessible (proximate) this scale is to each firm.

iii **Flexible representation of production technology -** analysis using production data aggregated over firms require us to assume homogenous technology across those firms and constant returns to scale, restrictions which can give rise to aggregation bias. Firm level data permit the use of more flexible functional forms to represent technology.

D. Graham

**iv Econometric estimation -** in estimating productivity the use of extensive firm level data can help to reduce multicollinearity and provide more identifying variance (e.g. Griliches and Mairesse 1995).

## 3.1. Firm data

The firm data we use for estimation describe the production and cost characteristics of registered UK companies in 2 digit sectors. Under UK legislation each registered company is required to provide accounting and other information about their operations to an executive agency of the Department of Trade and Industry know as Companies House. These data are made available in a commercial software package called Financial Analysis Made Easy (FAME), which is produced jointly by Jordans and Bureau Van Dijk (BVD 2003). The production data are for companies not plants. It is, however, possible to identify and remove multi-plant firms from the sample because they report more than one trading address.

The FAME data record extensive financial information for each firm and are available for a number of years, although the time-series reporting for individual firms is irregular. The basic input data we have on each firm include a measure of capital stock and the number of employees. Capital stock is the value of assets possessed by the firm and includes `fixed assets' such as the depreciated value of buildings, plant, machinery and equipment; `current assets' such as stocks and various debts owed to the company; and `current liabilities' or the amount owed by the company as a result of normal trading. Sales are used as a proxy for output. We also have data on wages and on the total costs of each firm, which includes all direct elements of the cost of the ordinary activities used to produce the firm's output.

## 3.2. Measuring agglomeration

The measure of agglomeration we use is calculated using a *ward* framework because there are extensive economic data available for these areas3, and because they allow for a high level of spatial disaggregation dividing Britain (230,700 square kilometres) into approximately 10,760 units.

Using the ward data we represent agglomeration with an `effective density' measure. This is essentially an accessibility based measure of agglomeration for very small areas. The total effective density (U) of employment that is accessible to any firm located in ward *i* is

$$U_i = \frac{E_i}{\sqrt{A_i / \pi}} + \sum_{j}^{i \neq j} \left( \frac{E_j}{d_{ij}} \right) \tag{1}$$

where $E_i$ is total employment in ward *i*, $A_i$ is the area of ward *i*, $E_j$ is total employment in ward *j*, and $d_{ij}$ is the distance between *i* and *j*. Note that the density effect that arises within the ward in which the firm is actually located (i.e. the first term on the right hand side of equation (1)) is measured by total ward employment divided by a proxy for average ward radius that is calculated assuming that the wards are roughly circular.

It is worth stressing here the implicit transport dimension of (1). Our effective density measure captures the scale and proximity of economic activity that is available in particular locations. We assume that investment in transport will change effective densities because it will alter the relative proximities of activity. Note, that we could also use travel times, or a measure of the generalised cost of travel, as the denominator in equation (1) (e.g. Graham 2007c).

### 3.3. Estimating the link between agglomeration and productivity

As externalities, agglomeration economies are treated as a kind of technology component that serves to shift the firm's production or cost function. For instance, at the firm level a typical specification of the production function would be

$$Y = g(U)f(X) \tag{2}$$

where $Y$ is the output level of the firm, $X$ is a vector of factor inputs, and $g(U)$ is a vector of influences on production that arise from agglomeration economies.

We use firm level data to provide an empirical representation of the production function, allowing us to estimate the effect of agglomeration on firm productivity. Specifically, we a variant of the translog production function which includes a primary production function along with a set of inverse input demand equations, which introduce additional information on costs and factor prices. Using this particular approach we can sketch out a reasonably complete specification of the production technology of firms and analyse some distinct effects of agglomeration on productivity. A description of the translog model used is given in appendix 1. A full demonstration of the model for the estimation of agglomeration economies is provided by Graham and Kim (2007).

## 4. RESULTS

In this section we present estimates of the relationship between agglomeration and productivity for UK industries derived using the production function methodology outlined above. We then review some recent attempts to use these results within appraisal methodology to assess the agglomeration benefits of transport investments. Finally we note some limitations of the approach and suggest some future directions for research.

### 4.1. Production function estimates

The results presented in this sub-section are taken from Graham (2005) and Graham (2006), and the intention here is to provide only an overview of the empirical finding of this previous work. For a full description of methodology, data sources, or other technical aspects of the research the reader should refer to these more detailed reports.

D. Graham

The results are presented for eight industry groups which comprise the following SIC codes:

i.     manufacturing (MAN) (SIC 15-40)
ii.    construction (CON) (SIC 45)
iii.   distribution, hotels & catering (DHC) (SIC 50-55)
iv.    transport, storage & communications (TSC) (SIC 60-64)
v.     Real estate (RE) (SIC70)
vi.    Information technology (SIC 72)
vii.  banking, finance & insurance (BFI) (SIC 65-67)
viii. business services (BUS) (SIC741 -745)

Separate estimates of agglomeration economies from the production function analyses are obtained for each group4. These are expressed as elasticities showing the proportional change in productivity associated with a proportional change in the level of agglomeration. Table 2 below shows estimates of the elasticities of productivity with respect to agglomeration.

Table 2. **Estimated elasticities of productivity with respect to agglomeration**

| industry | elasticity |
|---|---|
| Manufacturing | 0.077 |
| Construction | 0.072 |
| Distribution, hotels & catering | 0.153 |
| Trans, storage & communications | 0.223 |
| Real estate | 0.192 |
| IT | 0.082 |
| Banking, finance & insurance | 0.237 |
| Business services | 0.224 |
| Whole economy | 0.119 |

We estimate positive agglomeration externalities for manufacturing, construction and for each of our six service industries. The lowest agglomeration elasticity shown in the table is for manufacturing (0.077). The largest agglomeration elasticities are for transport storage & communications (0.223)5, banking finance & insurance (0.237), business services (0.224), and real estate (0.192).

The weighted average elasticity for the service sector as a whole, where the weights are based on industry group employment shares, is 0.186. This indicates that a doubling of accessibility to economic mass is associated with an increase in productivity of just under 20%. The service sector elasticity is over twice as large in magnitude than the manufacturing estimate of 8%. So it seems, on the basis of the results given in table 2 that services enjoy higher returns from agglomeration than manufacturing, and particularly the types of activities that we expect to find in CBD locations such as banking finance & insurance, business services, and real estate. Calculating a weighted average elasticity over all industries, gives an estimated elasticity of productivity with respect to agglomeration for the whole economy of 0.119 (12%).

**4.2. Applying the agglomeration elasticities in transport appraisal**

The results given above support the theory of increasing returns to agglomeration across a range of different industries. Proximity to economic mass appears to matter, and for this reason, we may suppose that an increase in effective densities induced through transport investment could have associated productivity benefits via agglomeration. However, we may still wonder about the actual magnitude of these effects in the context of transport appraisal. Would they appear insignificant relative to conventional travel time savings, or could they actually make a real difference to the benefit-cost calculations?

The answers to these questions will ultimately depend on the characteristics of any particular scheme. However, by way of illustration we can draw upon some recent examples of ex ante evaluation that have calculated the agglomeration benefits of certain transport investments for the UK.

The first such evaluation was carried out by the UK Department for Transport (DfT 2005). Using agglomeration elasticities given in Graham (2005), and employing a methodology similar to that suggested by Venables (2007), the DfT have reappraised a proposed London rail scheme called Crossrail to see how these externalities would affect the projected benefits of investment. Table 3 below shows the results of this exercise6. The table shows that inclusion of the urban economic effects, the so called agglomeration benefits, increase the total benefits of the Crossrail project by 25%.

The second recent evaluation of agglomeration benefits has been undertaken by Steer Davies Gleave (SDG) consultants, again using agglomeration estimates given in Graham (2005). They have undertaken a full economic appraisal of various proposed schemes for the Yorkshire & Humberside regionof England. The results relating to estimated agglomeration benefits are shown in table 4. SDGs calculations typically indicate additions to conventional user benefits of somewhere between 10%-20% arising from increasing returns to agglomeration.

The calculations shown in tables 3 and 4 indicate that the inclusion of agglomeration effects could substantially increase the estimated benefits of transport projects. If these are as large as these recent applications show, there are some important implications for those charged with making decisions on transport investment:

Table 3.  **Applying the new appraisal to CrossRail (DfT calculations)**

| *Benefits* | *Welfare ($ million)* |
|---|---|
| Business time savings | 4 847 |
| Commuting time savings | 4 152 |
| Leisure time savings | 3 833 |
| **Total user benefits (conventional)** | **12 832** |
| Agglomeration benefits | 3 094 |
| **Total benefits (new approach)** | **15 926** |

D. Graham

Table 4.  **Appraisal of agglomeration benefits from transport investments**

| Mode | Scheme | Agglomeration |
|------|--------|---------------|
| Road | Leeds to Bradford Improved Highway | 21% |
| Road | Leeds Urban Area Improved Highway | 22% |
| PT | Leeds to Bradford PT Improvements | 15% |
| Bus | Intra Leeds bus subsidy | 11% |
| Road | Leeds to Sheffield Improved Highway | 19% |
| Road | M6 shoulder | 12% |
| Bus | West Yorkshire County bus subsidy | 9% |
| PT | Leeds Urban Area Major PT Investment | 9% |
| Bus | South & West Yorkshire Bus subsidy | 7% |
| Bus | South Yorkshire bus subsidy. | 3% |

- the inclusion of additional information on agglomeration benefits could help inform the prioritisation of schemes for funding allocation.

- the estimation of higher returns to transport could release more public funds for investment.

- identifying impacts on GDP and on welfare could help to assess the trade-offs between scheme objectives.

- the quantification of GDP effects could help support calls for private contributions to infrastructure investment.


**4.3. Limitations of the approach and future research directions**

The research on agglomeration and transport investment has only very recently emerged and there are a number of limitations of the existing approach that require further attention. The empirical identification of agglomeration economies is fraught with difficulties. The actual processes that give rise to these externalities are generally not observed; instead, we use variables reflecting urban or industrial densities to measure the aggregate efficiency gains that we believe are offered by cities and industrial clusters. The measurement and analysis of productive efficiency itself also poses a number of problems, as do the classifications available to describe industrial and functional heterogeneity. In this sub-section, we emphasize some priorities for future research that could address limitations of the existing treatment of agglomeration in transport appraisal.

The first obvious limitation of the existing approach, and of the empirical work presented in this paper, is that it does not actually tell us much about where the productivity benefits of agglomeration come from. The theoretical literature does propose a number of sources of agglomeration benefits, (i.e. labour market benefits, knowledge interactions, and input sharing), but the empirical literature has not yet uncovered the relative magnitude of productivity effects arising from each source. In the context of transport appraisal, this means that we do not know how the sources of agglomeration might relate to transport movements.

This may in fact prove to be an important gap in our knowledge. When transport investments are made they usually affect different types of journey in different ways. Some transport investments will have their greatest impact on business trips, others on commuting, and others perhaps on freight trips.

The extent of the overall agglomeration benefit from a scheme, therefore, will depends on the degree to which agglomeration externalities are currently constrained by transport provision, but also by the extent to which agglomeration is driven by different journey purposes made by different modes.

A second important research theme that requires further attention relates to the geographic scope of agglomeration economies and how we represent this in appraisal. Essentially, the issue here concerns our understanding of the spatial distribution of the agglomeration benefits that might arise from transport spending. For instance, if we make a transport investment in Central London are the agglomeration benefits of this investment available only in the immediate locality of the project, or are they distributed further perhaps through the whole of Central, Inner, or Outer London, or even beyond? This is clearly a very important issue. Whether the productivity benefits of investment via agglomeration affect only a relatively small number of firms or a very large number of firms will radically alter estimates of the magnitude of those benefits. We therefore need to know more about how agglomeration economies diminish with distance from source. Rice et al. (2006) address this theme and it certainly requires more attention in the empirical work on agglomeration and productivity.

Finally, there are a range of other limitations of much of the existing work on agglomeration and productivity that are the subject of ongoing research. These include problems of identification arising from endogeneity and measurement error and the issue of urban functional specialization and the `quality' of inputs7.


## 5. CONCLUSIONS


This paper has considered the link between agglomeration and productivity for sectors of the UK economy. It has developed an `effective density' measure of accessibility to economic mass for small spatial areas which incorporates an implicit transport dimension. The analysis presented above tests the association between productivity and effective density in a firm level translog production function analysis.

The motivation for the study is to identify if there might be any external benefits that arise from the provision of transport infrastructure that are not included in standard transport appraisals. The results show that agglomeration economies do matter and that they can be substantial, particularly for services. We calculate a weighted average agglomeration elasticity of 0.119 for the economy as whole, 0.186 for the service sector and 0.077 for manufacturing. We also find considerable variation across industries in the magnitude of the elasticities.

If transport investment changes the densities available to firms, for instance through a reduction in travel times or in the cost of travel, then there are likely to be positive gains from agglomeration. Having reliable estimates of the economic mass-productivity relationship allows us to quantify these `wider' economic benefits. Some recent applications find that the effect of agglomeration externalities is not trivial when considered in the context of transport appraisal. Initial calculations typically indicate additions to conventional user benefits of 10%-20% arising from increasing returns to economic mass.

**NOTES**

1. Centre for Transport Studies, Imperial College London, London, SW7 2AZ, UK,
   Tel: +4420-7594-6088, Fax: +4420-7594-6107, Email: d.j.graham@imperial.ac.uk

2. In this analysis we concern ourselves with the agglomeration of all economic activity and do not distinguish between urbanization and localization economies. Graham (2007b) uses a similar approach to estimate externalities arising from both sources.

3. We use ward level employment data taken from the Annual Business Inquiry (ABI), the official census of employment in Great Britain.

4. Estimation using a finer industrial disaggregation at the 2 digit level can be found in Graham (2007a).

5. It is interesting that such a high elasticity is estimated for transport services. This result may be indicative of the increasing returns to density which tend to affect transport operators such that unit costs fall as the density of traffic increases. (e.g. Berechman 1993, Graham *et al*. 2003).

6. It is important to emphasise that these calculations have been made by the DfT. The full methodology and a background to Crossrail can be found in DfT 2005.

7. There is evidence to show that functional specialization may vary systematically across the urban hierarchy with larger cities tending to have a higher proportion of firms engaged in specialized functions involving skilled occupations (for instance Duranton and Puga 2005, Rice *et al*. 2006, Combes *et al*. 2007).

# REFERENCES

Aaaberg, Y. (1973). Regional productivity differences in Swedish manufacturing. *Regional and Urban Economics 3*, 131-156.

Berechman, J. (1993). *Public transport economics and deregulation policy*. Amsterdam: North Holland.

BVD (2003). *FAME: UK and Irish company information in an instant*. London: Bureau van Dijk.

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

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

Combes, P., G. Duraton, and L. Gobillon (2007). Spatial wage disparities: Sorting matters! *Journal of Urban Economics (in press)*.

DfT (2005). *TRansport, wider economic benefits and impacts on GDP*. London: HMSO.

Duranton, G. and D. Puga (2005). From sectoral to functional urban specialisation. *Journal of Urban Economics 57*, 343-370.

Eberts, R. and D. McMillen (1999). *Agglomeration economies and urban public infrastructure*, Chapter in HP Cheshire and E S Mills (eds) Handbook of regional and urban economics, Volume III. New York: North Holland.

Fogarty, M. and G. Garofalo (1978). Environmental quality income trade-off functions with policy applications. *paper presented at the Southern Regional Science Association Meeting*, .

Fujita, M. and J. Thisse (2002). *The economics of agglomeration: Cities, industrial location and regional growth*. Cambridge: Cambridge University Press.

Graham, D. J. (2005). *Wider econonmic benefits of transport improvements: link between agglomeration and productivity, Stage 1 Report*. London: DfT.

Graham, D. J. (2006). *Wider econonmic benefits of transport improvements: link between agglomeration and productivity, Stage 2 Report*. London: DfT.

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

Graham, D. J. (2007b). Identifying urbanization and localization externalities in manufacturing and service industries. *Papers in Regional Science (in press)*.

Graham, D. J. (2007c). Variable returns to agglomeration and the effect of road traffic congestion. *Journal of Urban Economics 62*, 103-120.

Graham, D. J., A. Couto, W. Adeney, and S. Glasiter (2003). Economies of scale and density in urban rail transport: effects on productivity. *Transportation Research E 39*, 443-458.

Graham, D. J. and H. Y. Kim (2007). An empirical analytical framework for agglomeration economies. *Annals of Regional Science (in press)*.

Griliches, Z. and J. Mairesse (1995). Production functions: the search for identification. *NBER 5067*.

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

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

Kawashima, T. (1975). Urban agglomeration economies in manufacturing industries. *Papers of the Regional Science Association 34*, 157-175.

Kim, H. (1992). The translog production function and variable returns to scale. *Review of Economics and Statistics 74*, 546-552.

Louri, H. (1988). Urban growth and productivity: the case of Greece. *Urban Studies 25*, 433-438.

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

Moomaw, R. L. (1983). Spatial productivity variations in manufacturing: a critical survey of cross sectional analyses. *International Regional Science Review 8*, 1-22.

Moomaw, R. L. (1985). Firm location and city size: reduced productivity advantages as a factor in the decline of manufacturing in urban areas. *Journal of Urban Economics 17*, 73-89.

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

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

Rosehthal, S. and W. Strange (2004). *Evidence on the nature and sources of agglomeration economies*, Chapter in Henderson JV and Thisse JF (eds) Handbook of Regional and Urban Economics, Volume 4. Amsterdam: Elsevier.

Shefer, D. (1973). Localization economies in SMSA's: a production function analysis. *Journal of Regional Science 13*, 55-64.

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

Sveikauskas, L., J. Gowdy, and M. Funk (1988). Urban productivity: city size or industry size. *Journal of Regional Science 28*, 185-202.

Tabuchi, T. (1986). Urban agglomeration, capital augmenting technology, and labour market equilibrium. *Journal of Urban Economics 20*, 211-228.

Venables, A. J. (2007). Evaluating urban transport improvements: cost-benefit analysis in the presence of agglomeration and income taxation. *Journal of Transport Economics and Policy 41* (2), 173-188.

## APPENDIX 1: THE TRANSLOG PRODUCTION INVERSE INPUT DEMAND MODEL

The generalised translog production function system based on an inverse input demand framework was first proposed by Kim (1992) and has been applied in full for the estimation of agglomeration economies by Graham and Kim (2007).

Let

$$Y = f(X,U) \qquad (3)$$

be the production function for the firm where $Y$ is the output level of the firm, $X$ is a vector of factor inputs with elements $X_i$ ($i = 1,...,n$), and $U$ represents influences on production that arise from agglomeration economies.

If inputs are rented in competitive markets the first-order conditions for output maximisation subject to an expenditure constraint are

$$\frac{\partial Y}{\partial X_i} = \lambda W_i, \qquad (4)$$

where $W_i$ is the price of the $i^{th}$ input, and $\lambda$ is a Lagrange multiplier which is the reciprocal of marginal cost $\partial C / \partial Y$. The expenditure constraint is given by,

$$\sum_i W_i X_i = C, \qquad (5)$$

where $C$ is total cost.

From (4) and (5)

$$\lambda = \frac{\sum_i (\partial Y / \partial X_i) X_i}{C} \qquad (6)$$

and substituting (6) back into (4) after rearrangement yields the inverse input demand equations

$$\frac{W_i}{C} = \frac{\partial Y / \partial X_i}{\sum_i (\partial Y / \partial X_i) X_i} \equiv g_i(X,U). \qquad (7)$$

Note that these inverse input demand functions determine prices as functions of quantities as opposed to ordinary demand functions which determine quantities in terms of prices. Equation (7) can be written in cost share form $\left( S_i^C \right)$ as

$$S_i^C = \frac{W_i X_i}{C} = \frac{\partial \log Y / \partial \log X_i}{\sum_i \partial \log Y / \partial \log X_i}. \tag{8}$$

Taking a translog approximation to equation (3) we have

$$\log Y = \alpha_0 + \sum_{i=1}^{i} \alpha_i \log X_i + \gamma_U \log U + \frac{1}{2} \sum_{i=1}^{i} \sum_{j=1}^{i} \gamma_{ij} \log X_i \log X_j + \sum_i \gamma_{iU} \log X_i \log U + \frac{1}{2} \gamma_{UU} (\log U)^2$$

$$\tag{9}$$

where $\gamma_{ij} = \gamma_{ji}$ $(i \neq j)$, and given (8) appropriate differentiation of (9) yields the cost share equations

$$S_i^C = \frac{\alpha_i + \sum_j \gamma_{ij} \log X_j + \gamma_{iU} \log U}{\sum_i \alpha_i + \sum_i \sum_j \gamma_{ij} \log X_j + \sum_i \gamma_{iU} \log U}. \tag{10}$$

The translog parameters can be efficiently estimated by simultaneously estimating (9) and (10) as a nonlinear multivariate regression system. Since the factor share equations sum to unity, however, estimation of the full system cannot be undertaken because the disturbance covariance matrix is singular and non-diagonal. The singularity problem is addressed by simultaneously estimating the primary translog production function and $n$ - 1 share equations.

Equations (9) and (10) are estimated as a non-linear SURE system using a two stage procedure which first estimate the error covariance matrix using non-linear least squares and then estimates the parameters that minimizes the generalized sum of the squares for the system as a whole. The random errors in each equation are assumed to be distributed independently of the regressors, have expected values of zero and constant variance. The SURE procedure also allows for cross-equation correlation between error terms and the error in each equation can have different variance.

The translog production-inverse demand system provides a generalisation over previous specifications allowing for a non-homothetic production technology in which returns to scale (RTS) and the elasticity of substitution vary with the level of production and with factor proportions. Homotheticity, homogeneity and linear homogeneity each represent restricted versions of the non-homothetic function. From equation (9) RTS are measured by

$$\sum_i \frac{\partial \log Y}{\partial \log X_i} = \sum_i \alpha_i + \sum_i \sum_j \gamma_{ij} \log X_j + \sum_i \gamma_{iU} \log U. \tag{11}$$

Flexibility in RTS across the sample is particularly important for our purposes because we wish to distinguish between scale economies and the effects of agglomeration. To do this effectively we need to ensure as far as possible that the agglomeration term is really distinct; that it does not capture some residual RTS effect arising due to the inadequate fit of a restrictive production function.

**An analytical framework for agglomeration economies**

The system we outline above offers a comprehensive analytical framework for the analysis of agglomeration. The very general specification does not require particularly onerous assumptions about the impact of agglomeration (such as Hick's neutrality) and allows for the possibility of non-linear

effects. We can distinguish three distinct but highly inter-related dimensions of agglomeration that can be identified using our model. First, there are the external returns to agglomeration that influence TFP and the productivity of individual factors and which we will term *productivity effects*. Second, there are *effects on factor prices* which arise as a result of the increase in the productivity of factors induced by agglomeration. Third, there are *effects on factor demands* which follow from the influence that agglomeration has on factor prices. Below we consider each dimension in turn.

*Productivity effects*

The aggregate productivity effect of agglomeration economies is captured by the elasticity of output with respect to agglomeration. Differentiating equation (9) with respect to $U$ we have

$$\frac{\partial \log Y}{\partial \log U} = \gamma_U + \gamma_{UU} \log U + \sum_i \gamma_{iU} \log X_i. \tag{12}$$

Equation (12) measures the total shift in output that arises from agglomeration and thus the effect of agglomeration on TFP. We can decompose this aggregate productivity effect in two parts. First, we have what we will call a *direct* agglomeration effect, which is independent of factor levels and which varies depending on the level of agglomeration ($\gamma_U + \gamma_{UU} \log U$). Note that the quadratic specification allows for non-linear agglomeration effects and thus for the kind of diminishing returns that might be predicted by theory. Second, we have a component that arises through the effect that agglomeration has on the productivity of factor inputs given the volume of factor inputs employed ($\sum_i \gamma_{iU} \log X_i$).

To determine the effect of agglomeration on the productivity of individual factors we use the output elasticities because these are the logarithmic marginal products of each input

$$\frac{\partial \log Y}{\partial \log X_i} = \alpha_i + \gamma_{ij} \log X_j + \sum_i \gamma_{iU} \log U. \tag{13}$$

If $\gamma_{iU} > 0$ then agglomeration is positively associated with the productivity of factor $X_i$, if $\gamma_{iU} < 0$ then agglomeration is negatively associated with the productivity of factor $X_i$. Agglomeration economies are Hick's neutral only if $\gamma_{iU} = 0$.

So in terms of productivity, our framework allows us to identify three types of impact from agglomeration externalities: an aggregate effect on TFP, a 'direct' effect that is independent of factor inputs, and the particular effects on the efficiency of each factor. We can therefore provide some distinction of total productivity effects. This is important because we might expect agglomeration to affect efficiency in different ways. The possibility of accommodating non-linear agglomeration effects, for instance, due to diminishing returns, is also advantageous.

*Price effects*

If there are productivity improvements from agglomeration then these should be capitalized in the prices of factor inputs and these price effects can also be identified within our framework. The inverse input demand equations (7) measure shadow prices or the marginal willingness to pay for inputs by

firms at a pre-determined level of expenditure. In equilibrium the marginal willingness to pay for an input should be equal to price. Denoting $W_i/C$ as $\hat{w}_i$ and rewriting equation (10) in inverse input demand form we have

$$\hat{w}_i = \frac{\alpha_i + \sum_j \gamma_{ij} \log X_j + \gamma_{iU} \log U}{\left(\sum_i \alpha_i + \sum_i \sum_j \gamma_{ij} \log X_j + \sum_i \gamma_{iU} \log U\right) X_i} \cdot \tag{14}$$

Logarithmically differentiating (14) with respect to $U$ gives

$$\frac{\partial \log \hat{w}_i}{\partial \log U} = \frac{\gamma_{iU}}{\left(\partial \log Y / \partial \log X_i\right)} - \frac{\sum_i \gamma_{iU}}{\sum_i \left(\partial \log Y / \partial \log X_i\right)} \cdot \tag{15}$$

Equation (15) identifies the effect of agglomeration on the willingness to pay for input $i$, or its price. This is expressed in terms of the impact of agglomeration on the productivity of factor $i$ ($\gamma_{iU}$) given the contribution of this factor to total output $\left(\partial \log Y / \partial \log X_i\right)$, less the impact of agglomeration on the productivity of all factors given the contribution of all factors to output change.

*Effects on factor demands*

If the price of factor inputs varies systematically with the level of agglomeration then we might expect to find an effect on factor demands. From (14) the inverse price elasticities of each factor are

$$\frac{\partial \log \hat{w}_i}{\partial \log X_i} = \frac{\gamma_{ii}}{\left(\partial \log Y / \partial \log X_i\right)} - \frac{\sum_j \gamma_{ij}}{\sum_i \left(\partial \log Y / \partial \log X_i\right)} - 1, \tag{16}$$

and so using (15) we can determine the effect of agglomeration on factor demands as follows

$$\frac{\partial \log X_i}{\partial \log U} = \left(\frac{\partial \log \hat{w}_i}{\partial \log X_i}\right)^{-1} \cdot \left(\frac{\partial \log \hat{w}_i}{\partial \log U}\right) \cdot \tag{17}$$

Equations (12) to (17) provide a useful comprehensive empirical framework to analyse agglomeration economies. They allow us to derive the effect of agglomeration on TFP, on the efficiency of each individual factor, on factor prices, and on factor demands.