Variable returns to agglomeration and the effect of road traffic congestion

Daniel J. Graham *

Centre for Transport Studies, Imperial College London, London SW7 2AZ, UK

Received 22 June 2006; revised 25 October 2006
Available online 20 December 2006

Abstract

This paper investigates the links between returns to urban density, productivity and road traffic congestion. A generalised translog production-inverse input demand function is estimated to test for the existence of variable returns to agglomeration in manufacturing, construction and service industries. Two separate measures of urban density, in which proximity is represented by straight line distance or by generalised cost, are constructed and included in the translog to identify the effect of road traffic congestion. The results show that for some sectors of the economy diminishing returns to urban density can set in causing the magnitude of agglomeration elasticity to fall as effective densities increase. The generalised cost based measure of agglomeration produces higher elasticities because it captures both time and distance dimensions of density. A comparison of spatial variance in estimates indicates that road traffic congestion plays an important role in explaining diminishing returns for the most highly urbanised locations.

© 2006 Elsevier Inc. All rights reserved.

JEL classification: R12; R41

Keywords: Agglomeration; Productivity; Congestion; Translog

1. Introduction

The theoretical foundations for the existence of agglomeration externalities are well established (see for example Fujita et al. [14], Fujita and Thisse [15], Duranton and Puga [11]). There is also now a good deal of empirical evidence on the extent of agglomeration externalities for
manufacturing industries. For instance, Ciccone and Hall [6] estimate that average labour productivity in manufacturing increases by 6% as the employment density of US states doubles; Ciccone [5] reports an elasticity of average labour productivity with respect to employment density of 0.045 for manufacturing in EU regions; and Rice et al. [30] estimate an elasticity of productivity with respect to economic mass of 0.035 for UK manufacturing. There are a number of other studies, stretching back to the early 1970s, which have estimated agglomeration externalities and have typically reported positive elasticities for manufacturing industries of somewhere between 2% and 10% (for reviews see Rosenthal and Strange [31], Eberts and McMillen [13]).

Despite this evidence, we also know that there are limits to urban growth and that cities can experience declines in population, employment and even output. A number of empirical studies have found evidence of centrifugal forces that can induce industrial dispersion from cities and have also indicated the existence of agglomeration diseconomies (e.g. Beeson [1,2], Carlino [4], Moomaw [27], Hansen [18], Hanson [19]). Urban economic theory does, of course, allow for disparities in the nature of returns to agglomeration and also recognises that for many types of economic activity a highly urbanised location may not be optimal. The new economic geography literature has illustrated these themes by focusing on the interaction between increasing returns, wages and the costs of trade (e.g. Fujita et al. [14], Fujita and Thisse [15]).

Theory tells us that the strength of centripetal forces can be limited or diminished as cities become too large and the processes which give rise to positive externalities consequently become less efficient. Prominent amongst the inefficiencies mentioned in this respect is the existence of congestion in urban transport systems. There is an inherent tendency for externalities of agglomeration and congestion to coincide spatially; the former are thought to be derived from concentration and the latter are a consequence of it. In other words, we expect congestion to be present in those locations where the competition in factor input and output markets is intense (see Coombes and Overman [7]). There, is therefore a very close relationship between urban congestion and agglomeration economies. For this reason, it may be informative to consider the extent to which the benefits of urban density, generated through agglomeration externalities, are constrained by road traffic congestion.

In this paper we provide an empirical investigation of the effect of urban congestion on agglomeration. Our main focus is on determining whether we can identify variable returns from agglomeration and whether any such externalities may be constrained by congestion in cities. The paper estimates nonlinear agglomeration elasticities in a generalised translog production function to test for the existence of variable returns.

Effective density based measures of agglomeration are constructed to capture the scale and proximity of activity that is accessible to any location. Proximity is represented in two different ways: by distance and by the generalised cost of road travel which includes information both about distances and road speeds. Comparing results using densities based on each measure of proximity we are able to investigate the impact of congestion on returns to agglomeration. Results are presented for manufacturing, construction and for seven service sectors: distribution, hotels & catering; transport, storage & communications; real estate; information technology; banking, finance & insurance; business services; and public services (see Appendix A for a description of these sectors).

1 Note that we are concerned in this paper with the relationship between economic productivity and the concentration of industry per se, regardless of the mix of intra-industry and inter-industry densities. Studies that have set out to distinguish the productivity effects of urbanisation and localisation within the same model include Nakamura [28], Henderson [22], Henderson [23] and Graham [16]. Recent studies of localisation are provided by Duranton and Overman [12] and Devereaux et al. [8].
The paper is structured as follows. Section 2 describes the translog production-inverse input demand function which we use as a basic framework for the estimation of agglomeration externalities. Section 3 describes data sources and explains the measures of agglomeration that we construct and how we use these to shed light on the effect of urban road traffic congestion. Empirical results are presented in Section 4 and conclusions are drawn in the final section.

2. The translog production-inverse input demand function

We analyse the relationship between productivity and agglomeration within a production function framework. The translog inverse input demand function proposed by Kim [24,25] provides a very flexible functional form for estimation which allows for a non-homothetic production technology with variable returns to scale (RTS). Agglomeration can be estimated as a Hick’s neutral production function shifter and the quadratic form allows this elasticity to vary across the sample with the level of agglomeration.

The translog approximation to a firm level production function is

$$\log Y = \alpha_0 + \beta U \log U + \frac{1}{2} \beta_{UU} (\log U)^2 + \sum_{i=1}^{n} \alpha_i \log X_i$$

$$+ \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \gamma_{ij} \log X_i \log X_j$$

(1)

where $Y$ is the output level of the firm, $X$ is a vector of factor inputs with elements $X_i (i = 1, \ldots, n)$, and $U$ is some measure of the level of agglomeration experienced by the firm.

The firm faces an expenditure constraint in which input prices ($W_i$) and quantities determine total cost ($C$):

$$\sum_i W_i X_i = C.$$  

(2)

Given competitive inputs markets the maximisation of output subject to the expenditure constraint implies that

$$\frac{\partial Y}{\partial X_i} = \lambda W_i,$$  

(3)

where $\lambda$ is a Lagrange multiplier which is the reciprocal of marginal cost $\partial C/\partial Y$.

From (2) and (3)

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

(4)

and substituting (4) back into (3) after rearrangement yields the inverse input demand equations

$$\frac{W_i}{C} = \frac{\partial Y/\partial X_i}{\sum_i (\partial Y/\partial X_i) X_i}.$$  

(5)

The inverse input demand functions determine prices as functions of quantities as opposed to ordinary demand functions which determine quantities in terms of prices.

Equation (5) can be written in cost share form ($S^C_i$) as

$$S^C_i = \frac{W_i X_i}{C} = \frac{\partial \log Y/\partial \log X_i}{\sum_i \partial \log Y/\partial \log X_i}.$$  

(6)
and therefore differentiation of (1) can be used to obtain the cost share equations
\[
S_i^C = \frac{\alpha_i + \sum_j \gamma_{ij} \log X_j}{\sum_i \alpha_i + \sum_i \sum_j \gamma_{ij} \log X_j}.
\] (7)

The translog parameters can be efficiently estimated by simultaneously estimating (1) and (7) as a nonlinear multivariate regression system. The generalised translog function provides a flexible framework for estimation in which the level of returns to scale can vary across the sample depending on the level of production. Homotheticity and homogeneity represent restricted version of the production technology allowed by Eq. (1).\(^2\) It is desirable that we adopt a flexible functional form for estimation if we wish to be effective in isolating the impact of agglomeration across the sample. Use of more restrictive functional forms, for instance one that estimates a constant level of RTS, can give rise to a less adequate fit and consequently yield less efficient estimates.

Returns to agglomeration are expressed in a quadratic relationship. The effect of agglomeration on production at each point on this quadratic curve can be measured by
\[
\frac{\partial \log Y}{\partial \log U} = \beta_u + \beta_{UU} \log U.
\] (8)

Equation (8) measures the total shift in output that arises from agglomeration. Since we have controlled for input use and the effect of RTS within the production function, the output shift shown in (8) is also a shift in total factor productivity (TFP). Note that non-linear agglomeration effects are accommodated by this elasticity which varies depending on the level of agglomeration. The quadratic specification allows for variable returns including the kind of diminishing returns that might be predicted by theory.

3. Estimating agglomeration and the effect of congestion

To estimate the translog system we need data with detailed geographical and sectoral disaggregation. The data we have found that match these requirements are the firm level data of registered UK companies. Under UK legislation each registered company is required to provide accounting and other data 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 [3]).

FAME record extensive financial data for each company including turnover, a breakdown of costs, and information on the number of employees and on the total assets held by the firms. We use turnover as a measure of output, the number of employees as labour input, and we calculate average wages per employee from total payroll. As a proxy for capital input we use information on total assets. This includes the value of include 'fixed assets’ such as the buildings, plant, machinery and equipment and ‘current assets’ such as stocks and various debts owed by and to the company. We treat total assets as a proxy for capital input (\(K\)) in the sense that they give a measure of the value of the non-labour inputs available.

\(^2\) A homothetic and homogeneous technology is imposed by the restriction \(\sum_j \gamma_{ij} = 0\). Graham and Kim [17] estimate a non-homothetic version of this generalised translog function allowing for factor augmenting agglomeration effects. They show that this model specification can be used to analyse the effect of agglomeration on partial factor productivity, total factor productivity, factor prices and factor demands.
FAME also records information about the location of the company. Since the data are provided by companies and not plants they can include firms that have plants in many locations but that only report aggregated records. To isolate single plants firms from the FAME data we have taken the following steps. First, we have removed firms that record more than one UK trading address. Second, we have removed firms that have a registered office address that is different from their main trading address. Third, we have kept only those firms that do not have a UK or foreign holding or subsidiary company. Fourth, as a further precaution, we have removed large firms from the data because these could record all of their information from one UK registered office address, typically their headquarters, but not provide information about other plants they may have located in different areas of the UK or the World. Our sample is based on firms that have less than 100 employees.

The FAME data are available for a number of years, although the quantity and quality of the data diminishes as we go back in time. We have extracted FAME data over the period 1995 to 2002 for 9 industry groups (see Appendix A).

The principal concern in this paper is with the impact of urban congestion on the relationship between productivity and agglomeration. As discussed above, we can estimate the translog with a measure of agglomeration expressed in quadratic form to identify if there are diminishing returns. However, the presence of diminishing returns itself does not verify the existence of urban congestion. There could be a number of reasons why cities become less efficient as size grows. To identify the impact of urban transport congestion we need to construct measures of agglomeration that contain an implicit transport dimension and that allow us to consider the implications of constraints on the efficiency of travel.

Thus, crucial to our measures of agglomeration is some quantification of the relative ease of accessing urban activity. In other words, we are not just interested in city size or scale but also in the relative proximity of activities, spatially and temporally. This leads naturally to a consideration of densities and in this paper we base our analysis of agglomeration on the concept of effective densities. An effective density measures the amount of ‘activity’ that is accessible from some given location.

For Britain, the most comprehensive spatially disaggregate definition of ‘activity’ we can construct makes use of data on the number of jobs located in the 10,780 wards of Britain. Ward employment data are reported in the Annual Business Inquiry (ABI). To model the proximity of activity, or the nearness of one ward to the next, we can use a measure based on straight line distance calculated using Pythagoras and the ward centroid \( x \) and \( y \) coordinates. Alternatively, we can make use of information on the ward to ward generalised costs of travelling by road. The generalised cost \( g_{ij} \) of road travel by car from ward \( i \) to ward \( j \) is a measure of the total of all the costs faced:

\[
 g_{ij} = p \cdot rd_{ij} + \tau v \left( \frac{rd_{ij}}{s_{ij}} \right) + \sum_c U_c, \tag{9}
\]

where \( p \) is the price or money cost per passenger kilometre and comprises the costs of operating the vehicle, \( rd_{ij} \) is the distance by road between \( i \) and \( j \), \( \tau v \) is the value of in vehicle time, \( s_{ij} \) is the average speed between \( i \) and \( j \), and \( U_c \) is any other relevant user cost.

The generalised cost data we have available to us, supplied by the UK Department for Transport (DfT), assume constant money prices, user costs and values of time. Thus, differences in the generalised cost of travelling from ward \( i \) to ward \( j \), or from ward \( i \) to ward \( k \), reflect only the differences in the relative distances and speeds of travel, not prices or values.
Using these measures of proximity we can define two effective density measures for a firm in industry \( o \) located in ward \( i \) as follows

\[
UD_{io} = \frac{E_i}{r_i} + \sum_{j \neq i} \left( \frac{E_j}{d_{ij}} \right),
\]

\[
UG_{io} = \frac{E_i}{g_i} + \sum_{j \neq i} \left( \frac{E_j}{g_{ij}} \right)
\]

where \( E \) is total employment, \( r_i \) is an approximation of the radius of ward \( i \), and \( d_{ij} \) is the Euclidean distance between \( i \) and \( j \). The effective density measures closely resemble the concept of market potential developed by Harris [21] and used in recent empirical work by Mion [26] and Hanson [20].

Comparing the values of \( UD \) and \( UG \) should enable us to investigate urban road traffic congestion because the essential difference between the two lies in the inclusion, or exclusion, or information about the speed of travel. Our hypotheses are as follows. In large cities, where congestion is present, the ratio of \( UD \) to \( UG \) will tend to be relatively large because while there is a lot of activity concentrated in space, road traffic speeds are low and so the generalised cost of travelling small distances is high. In smaller towns and cities where there is less congestion and consequently higher road speeds the ratio of \( UD \) to \( UG \) will less. In rural areas where traffic moves at free flowing speeds we will expect the ratio of \( UD \) to \( UG \) to be at a minimum.

Essentially, we are hypothesising that if congestion exists then these two measures will tend to diverge as we move towards the upper end of the urban hierarchy. The generalised cost based measure of proximity accounts for the fact that speeds can vary systematically with city size and this provide a sort of upper constraint on the values of urbanisation. The absence of information about travel times or speeds in the \( UD \) variable induces an artificial right-hand skew into the distribution of urbanisation values: it makes locations that suffer from congestion seem more accessible than they really are. In this sense we can argue that if congestion exists, the \( UD \) variable will contain some degree of measurement error.

If it is true that the generalised cost based density variable provides a superior measure of the actual level of agglomeration experienced by firms, then it follows that estimation using a measure based on Euclidean distance will produce a biased estimate of the agglomeration elasticity. We can demonstrate this in the following way. Let the estimating equation for (1) based on the generalised cost (\( UG \)) measure of agglomeration be written in shorthand form as

\[
Y = \alpha + \beta_u UG + \sum_i \beta_i X_i + u
\]

where the logarithmic and quadratic forms of the variables have been suppressed. We assume this to be the ‘true’ model with agglomeration measured to a reasonable degree of accuracy by \( UG \). If instead of estimating (12) we estimate using the variable \( UD \) which contains measurement error we have

\[
Y = \alpha + \beta_u UD + \sum_i \beta_i X_i + v
\]

\[
= \alpha + \beta_u UD + \sum_i \beta_i X_i + \beta_u (UG - UD) + u
\]
where the error term \( v \) in the first row of (13) can be decomposed into two components: the term \( \beta_u(UG - UD) \) and the error \( u \) from the true model. Since \( UD \) will be correlated with this first component of the error term, the estimate of \( \beta_u \) obtained from (13) will be biased, with the direction of the bias depending on the correlation between \( UD \) and \( (UG - UD) \). In the presence of congestion, we expect speeds to fall and therefore generalised costs to grow as the level of the variable \( UD \) increases, and so the \( UG \) variable should grow less than proportionately with \( UD \). Therefore, if congestion does exist \( UD \) should be negatively correlated with \( (UG - UD) \) and consequently the estimates of the agglomeration elasticity based on \( UD \) should be downwards biased.

4. Results

The results section is organised as follows. First, we present results from estimation of a two factor non-homothetic translog system

\[
\log Y = \alpha_0 + \beta_t t + \beta_U \log U + \frac{1}{2} \beta_{UU} (\log U)^2 + \beta_L \log L + \beta_K \log K
\]

\[
+ \frac{1}{2} \gamma_{LL} (\log L)^2 + \frac{1}{2} \gamma_{KK} (\log K)^2 + \gamma_{LK} \log L \log K,
\]

\[
S_L = \frac{\beta_L + \gamma_{LL} \log L + \gamma_{LK} \log K}{\beta_L + \beta_K + (\gamma_{LL} + \gamma_{LK}) \log L + (\gamma_{KK} + \gamma_{LK}) \log K}.
\]

We estimate (14) separately using two effective density measures of agglomeration. First, we use the more conventional measure of agglomeration with proximity based on distance (\( UD \)) to test for the existence of agglomeration externalities in different sectors of the economy and to identify any evidence of diminishing returns. Second, we estimate using the generalised cost (\( UG \)) based agglomeration variable and then compare results. If congestion does exist, we would expect elasticity estimates based on \( UG \) to be higher than those based on \( UD \) and we may also expect to find differences in the estimates that capture the extent of diminishing returns.

Table 1 shows results of the estimation of Eq. (14) using the \( UD \) variable for our nine industry groups. The Wald statistics associated with each regression allow us to reject the hypothesis that homogeneity or linear homogeneity can provide a superior characterisation of the structure of production. Production function parameter estimates are, with few exceptions, significant and of the expected sign. Results relating to the structure of production are summarised in Table 2, calculated at the sample means.

Estimates of RTS from the translog models indicate that firms within our industry groups tend to operate reasonably near constant returns to scale. For manufacturing, the output elasticity of labour is 0.27 and is under half the size of the capital output elasticity. Services tend to be more labour intensive and this is reflected in the output elasticities of labour around or above 0.45 in banking and finance, IT services, business services, and public services.

Our main concern in this paper is on the estimates associated with the agglomeration variables (\( \beta_{UD} \) and \( \beta_{UUD} \)). The quadratic form for agglomeration produced significant estimates for five of the nine industries shown in Table 1. For the remaining four industries, we estimated insignificant quadratic coefficients and found that a log-linear relationship between output and agglomeration produced a better fit. Table 3 summarises the evidence on agglomeration externalities showing, for each industry group, the relevant parameter estimates from the translog and the elasticities of output with respect to agglomeration. Where the quadratic form has been used these elasticities are evaluated using the mean level of agglomeration of each sample.
Table 1
Parameter estimates from the translog models

<table>
<thead>
<tr>
<th></th>
<th>Manufacturing</th>
<th>Construction</th>
<th>Dist., hotels &amp; cater.</th>
<th>Trans., store, &amp; comm.</th>
<th>Real estate</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha_0$</td>
<td>-1.403</td>
<td>(1.802)</td>
<td>-24.849**</td>
<td>-9.602**</td>
<td>-7.234</td>
</tr>
<tr>
<td>$\beta_T$</td>
<td>-0.001</td>
<td>(0.002)</td>
<td>0.037**</td>
<td>-0.091**</td>
<td>0.003</td>
</tr>
<tr>
<td>$\beta_L$</td>
<td>1.124**</td>
<td>(0.018)</td>
<td>0.852**</td>
<td>0.731**</td>
<td>0.938**</td>
</tr>
<tr>
<td>$\beta_K$</td>
<td>0.26**</td>
<td>(0.016)</td>
<td>0.31**</td>
<td>0.178**</td>
<td>0.673**</td>
</tr>
<tr>
<td>$\gamma_{LL}$</td>
<td>0.074**</td>
<td>(0.002)</td>
<td>0.075**</td>
<td>0.063**</td>
<td>0.066**</td>
</tr>
<tr>
<td>$\gamma_{KK}$</td>
<td>0.047**</td>
<td>(0.001)</td>
<td>0.039**</td>
<td>0.055**</td>
<td>0.009**</td>
</tr>
<tr>
<td>$\gamma_{LK}$</td>
<td>-0.076**</td>
<td>(0.001)</td>
<td>-0.062**</td>
<td>-0.052**</td>
<td>-0.064**</td>
</tr>
<tr>
<td>$\beta_{UD}$</td>
<td>1.105**</td>
<td>(0.283)</td>
<td>4.79**</td>
<td>2.488**</td>
<td>1.48*</td>
</tr>
<tr>
<td>$\beta_{UUD}$</td>
<td>-0.086**</td>
<td>(0.022)</td>
<td>-0.369**</td>
<td>-0.188**</td>
<td>-0.095</td>
</tr>
</tbody>
</table>

Sys. $R^2$ 0.676 0.539 0.609 0.536 0.501
Wald 1 5624** 566** 2452** 749** 260**
Wald 2 5626** 566** 2605** 778** 261**
$n$ 13,003 3139 9658 4411 3947
Table 1 (continued)
Parameter estimates from the translog models

<table>
<thead>
<tr>
<th></th>
<th>IT</th>
<th>Banking, finance, ins.</th>
<th>Business services</th>
<th>Public services</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha_0$</td>
<td>$-14.699^{**}$</td>
<td>2.742**</td>
<td>6.338**</td>
<td>4.974**</td>
</tr>
<tr>
<td>$\beta_t$</td>
<td>$-0.025^{**}$</td>
<td>0.059**</td>
<td>0.025**</td>
<td>0.005</td>
</tr>
<tr>
<td>$\beta_L$</td>
<td>1.117**</td>
<td>0.992**</td>
<td>1.265**</td>
<td>0.674**</td>
</tr>
<tr>
<td>$\beta_K$</td>
<td>$-0.114^{**}$</td>
<td>0.385**</td>
<td>$-0.164^{**}$</td>
<td>$-0.063$</td>
</tr>
<tr>
<td>$\gamma_{LL}$</td>
<td>0.005</td>
<td>0.077**</td>
<td>0.077**</td>
<td>0.139**</td>
</tr>
<tr>
<td>$\gamma_{LK}$</td>
<td>0.06**</td>
<td>0.015**</td>
<td>0.074**</td>
<td>0.052**</td>
</tr>
<tr>
<td>$\gamma_{KK}$</td>
<td>$-0.051^{**}$</td>
<td>$-0.052^{**}$</td>
<td>$-0.077^{**}$</td>
<td>$-0.046^{**}$</td>
</tr>
<tr>
<td>$\beta_{UD}$</td>
<td>3.68**</td>
<td>0.251**</td>
<td>0.176**</td>
<td>0.292**</td>
</tr>
<tr>
<td>$\beta_{UUD}$</td>
<td>$-0.285^{**}$</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>

Sys. $R^2$ 0.720 0.533 0.738 0.480

Wald 1 840** 267** 736** 730**
Wald 2 950** 276** 832** 898**
n 5726 2140 3780 2742

Notes. (1) Numbers in parentheses are standard errors.
(2) The $R^2$ value given is the McElroy $R^2$ value for the translog system.
(3) Wald 1 is the test statistic for homogeneity, Wald 2 is the test statistic for linear homogeneity.
* Significant at 0.05.
** Significant at 0.01.
Table 2
Estimated returns to scale and output elasticities calculated using sample means of the data

<table>
<thead>
<tr>
<th></th>
<th>RTS</th>
<th>Output elasticities</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>labour</td>
<td>capital</td>
</tr>
<tr>
<td>Manufacturing</td>
<td>0.956</td>
<td>0.274</td>
<td>0.682</td>
</tr>
<tr>
<td>Construction</td>
<td>0.878</td>
<td>0.200</td>
<td>0.678</td>
</tr>
<tr>
<td>Distributions, hotels &amp; catering</td>
<td>0.976</td>
<td>0.176</td>
<td>0.800</td>
</tr>
<tr>
<td>Transport, storage &amp; communications</td>
<td>0.862</td>
<td>0.232</td>
<td>0.630</td>
</tr>
<tr>
<td>Real estate</td>
<td>1.096</td>
<td>0.398</td>
<td>0.699</td>
</tr>
<tr>
<td>IT</td>
<td>1.030</td>
<td>0.505</td>
<td>0.525</td>
</tr>
<tr>
<td>Banking, finance &amp; insurance</td>
<td>0.912</td>
<td>0.445</td>
<td>0.467</td>
</tr>
<tr>
<td>Business services</td>
<td>1.063</td>
<td>0.452</td>
<td>0.611</td>
</tr>
<tr>
<td>Public services</td>
<td>0.932</td>
<td>0.442</td>
<td>0.491</td>
</tr>
</tbody>
</table>

Table 3
Elasticities of output with respect to agglomeration (UD variable)

<table>
<thead>
<tr>
<th></th>
<th>$\beta_{UD}$</th>
<th>$\beta_{UUD}$</th>
<th>$\epsilon_{UD}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manufacturing</td>
<td>1.105**</td>
<td>-0.086**</td>
<td>0.041</td>
</tr>
<tr>
<td>Construction</td>
<td>4.79**</td>
<td>-0.369**</td>
<td>0.214</td>
</tr>
<tr>
<td>Distr., hotels &amp; catering</td>
<td>2.488**</td>
<td>-0.188**</td>
<td>0.133</td>
</tr>
<tr>
<td>Trans., storage &amp; comm.</td>
<td>1.48*</td>
<td>-0.095</td>
<td>0.274</td>
</tr>
<tr>
<td>Real estate</td>
<td>0.084**</td>
<td>-</td>
<td>0.084</td>
</tr>
<tr>
<td>IT</td>
<td>3.68**</td>
<td>-0.285**</td>
<td>0.089</td>
</tr>
<tr>
<td>Banking, fin. &amp; insurance</td>
<td>0.251**</td>
<td>-</td>
<td>0.251</td>
</tr>
<tr>
<td>Business services</td>
<td>0.176**</td>
<td>-</td>
<td>0.176</td>
</tr>
<tr>
<td>Public services</td>
<td>0.292**</td>
<td>-</td>
<td>0.292</td>
</tr>
</tbody>
</table>

* Significant at 0.05.
** Significant at 0.01.

We estimate positive agglomeration externalities for manufacturing, construction and for each of our seven service industries. The lowest agglomeration elasticity shown in the table is for manufacturing, 0.041. This measures up reasonably well to roughly comparable estimates of 0.06 for the US states (Ciccone and Hall [6]), 0.045 for EU regions (Ciccone [5]), and 0.035 for UK manufacturing (Rice et al. [30]). The largest agglomeration elasticities are for transport storage & communications (0.274), \(^4\) construction (0.214), banking finance & insurance (0.251), business services (0.176), and public services (0.292). The average for the seven service sectors is 0.186, over four times larger in magnitude than for manufacturing. So it seems, on the basis of the figures given in Table 3 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, and business services.

Table 3 also provides some evidence about the existence of variable returns. We estimate negative and significant quadratic parameters ($\beta_{UUD}$) for five sectors. For the remaining four industries we find insignificant quadratic terms but achieve a good fit with a single parameter for

\(^4\) For transport providing firms, the higher elasticities 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. Passenger densities are likely to grow systematically with city size.
the sample as a whole indicating that returns to agglomeration are constant. Figure 1 illustrates the nature of the five quadratic relationships we have estimated by evaluating the agglomeration elasticities for each ward of Britain using the ward agglomeration values.\(^5\) Note that the elasticities vary at each point in the sample and so the plots show the marginal relationship between productivity and agglomeration. Thus, where the elasticities are greater than zero there are increasing returns to agglomeration, where they equal to zero returns to agglomeration are maximised, and where they are less than zero returns are decreasing.

Each of the five sectors shown in Fig. 1 exhibits some degree of diminishing returns to urban density. For manufacturing, IT and construction, returns to agglomeration are maximised (elasticity = 0) at around the same point. In fact, although this looks to be at a relatively low value of agglomeration the distribution has a right-hand skew and for these three industries we actually find positive but diminishing returns over the first nine deciles of agglomeration values. The distribution hotels & catering industry reaches a maximum value of returns to agglomeration slightly later and then decreasing returns thereafter in a small proportion of the most highly urbanised wards. For transport storage & communications we have evidence of diminishing returns but the quadratic function is not maximised using this sample and we do not identify decreasing returns to agglomeration.

To shed light on the impact of road traffic congestion we next introduce the effective density measure based on generalised cost. Re-estimating the translog using the \(UG\) variable we generate a new set of agglomeration elasticities which we denote \(\varepsilon_{UG}\). These are shown in Table 4 and compared in Fig. 2 to the elasticities obtained using effective density measures based on distance. The results of the estimation relating to the structure of production are very similar to those presented in Tables 1 and 2 and so we do not repeat them here.

The overall pattern of agglomeration estimates obtained using each effective density measure is very similar (Fig. 2). Positive and significant agglomeration economies are estimated for all industries shown in the table. However, there are two notable differences between the estimates obtained using the \(UD\) or \(UG\) variables. First, is that with the exception of elasticity estimates for manufacturing and IT, which are very similar in magnitude, the generalised cost based estimates tends to be higher than the elasticities estimated using distance based measures. Second, using the \(UG\) variable we find positive and statistically significant quadratic terms (\(\beta_{UUG}\)) for three service industries: banking finance & insurance, business services, and public services.

From the comparison of estimates we can make two observations relevant to the impact of road traffic congestion on agglomeration externalities. First, as hypothesised in Section 3, we find that the inclusion of travel time information in the measure of effective density produces higher estimates of agglomeration externalities. If we believe that generalised cost densities provide a superior measure of agglomeration in the presence of road traffic congestion, then the implication is that the estimates based on a Euclidean distance measure of density (\(\beta_{UD}\)) are downward biased. We can confirm the direction of this bias by calculating the correlation between \(\log UD\) and the term \((\log UG - \log UD)\) (see Eq. (13)).

Table 5 shows negative correlations for all industries and therefore that the \(UG\) measure of agglomeration grows a slower rate than \(UD\). In other words, the two measures diverge as the level of urbanisation increases, with the \(UD\) variable yielding more extreme high values. This divergence can induce differences of fairly large magnitude in the \(UD\) and \(UG\) agglomeration

---

\(^5\) In making these calculations we assume that the parameters estimates obtained from our samples are applicable over the distribution of all wards.
Fig. 1. Spatial variance in returns to agglomeration.
Fig. 2. A comparison of effective density elasticities based on distance ($\varepsilon_{UD}$) and generalised cost ($\varepsilon_{UG}$).
Table 4
Elasticities of output with respect to agglomeration (UG variable)

<table>
<thead>
<tr>
<th>Industry Group</th>
<th>$\beta_{UG}$</th>
<th>$\beta_{UUG}$</th>
<th>$\varepsilon_{UG}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manufacturing</td>
<td>1.265*</td>
<td>-0.091*</td>
<td>0.042</td>
</tr>
<tr>
<td>Construction</td>
<td>3.379*</td>
<td>-0.233*</td>
<td>0.243</td>
</tr>
<tr>
<td>Distrib., hotels &amp; catering</td>
<td>3.096**</td>
<td>-0.217**</td>
<td>0.151</td>
</tr>
<tr>
<td>Trans., storage &amp; comm.</td>
<td>0.35**</td>
<td>-</td>
<td>0.350</td>
</tr>
<tr>
<td>Real estate</td>
<td>0.114**</td>
<td>-</td>
<td>0.114</td>
</tr>
<tr>
<td>IT</td>
<td>6.175**</td>
<td>-0.446**</td>
<td>0.082</td>
</tr>
<tr>
<td>Banking, fin. &amp; insurance</td>
<td>-3.449*</td>
<td>0.276*</td>
<td>0.329</td>
</tr>
<tr>
<td>Business services</td>
<td>-1.562*</td>
<td>0.13*</td>
<td>0.226</td>
</tr>
<tr>
<td>Public services</td>
<td>-3.02*</td>
<td>0.252*</td>
<td>0.399</td>
</tr>
</tbody>
</table>

* Significant at 0.05.
** Significant at 0.01.

elasticity estimates, particularly for the industries that tend to be most urbanised. For example, the $UG$ estimates are over 30% higher than the $UD$ estimates for real estate, banking finance & insurance, and public services; and are just under 30% higher for business services and transport storage & communications.

The second observation to be made from the comparison of the agglomeration elasticities is that evidence on the nature of diminishing returns differs depending on which measure is used. The elasticities based on straight line distance show evidence of diminishing returns for five of the nine industry groups and constant returns for the remaining four. The elasticities based on the $UG$ variable, on the other hand, show diminishing returns for only four industry groups, constant returns for two, and increasing returns for the remaining three.

These two observations offer evidence consistent with the hypothesis that urban road traffic congestion plays a significant role in ‘constraining’ the benefits of agglomeration, and consequently, that it may serve to reduce achievable levels of urban productivity. The generalised cost based measures of proximity accounts for the fact that speeds can vary systematically with city size. Accordingly, estimates based on $UG$ capture the combined effects of changes in both the distance and speed dimensions of effective density. The $UD$ measure captures changes in spatial but not temporal accessibility giving rise to higher extreme values of urbanisation at the upper bound of the distribution. The fact that the gap between the two measures increases as the level of urbanisation grows, due to the exclusion of travel time information, reflects the impact of road traffic congestion.

If, as our empirical analysis suggests, congestion can contribute to the diminishing returns we observe then the implication is that the productivity benefits of agglomeration could be increased by making appropriate transport interventions. In other words, the fact that the distance based density elasticities may be smaller in the most urbanised locations does not necessarily imply that there cannot be any further advantages from improving accessibility. Rather, our results suggest that an exogenous intervention to reduce the negative externality of congestion could mitigate the effects of diminishing returns: it could shift the productivity-urbanisation curve outwards. Venables [32] provides a theoretical example of this type of relationship showing how a reduction in travel times within a city can induce agglomeration benefits.

We can provide a brief demonstration of these issues here by considering a numerical example of the effect of changing densities. Table 6 shows elasticities averaged for the London agglomeration evaluated using the parameters given in Table 4 and the generalised cost agglomeration
values for the London wards. The table also shows a breakdown of the broad industrial structure of London expressed in employment shares. Weighting the elasticities according to these shares we calculate a total weighted average agglomeration elasticity for all sectors of 0.26.

The London conurbation contains a large volume of employment within a relatively small area (approximately 4 million jobs within 1579 square kilometres). However, many London locations suffer from congestion and speeds can be low, so the generalised cost of travelling one kilometre within London is relatively high with the travel time component comprising around 80% of the total.6

If the volume of employment within London, that is the total number of jobs within the Greater London area, was to increase by 5% then on the basis of the sectoral elasticities and shares given above we would expect a 1.3% increase in productivity, ceteris paribus. In fact, this scenario would actually involve a substantial increase of approximately 200,000 jobs within London and we would therefore expect additional second or third round effects on travel demand, which we cannot predict here, but which would potentially impact on the generalised cost of travel and reduce the net productivity effect.

---

6 For instance, evaluating the average generalised cost of travelling one kilometre by car in London using parameter values from DfT [9], appropriate price indices (ONS [29]), and assuming an average speed of 27 kilometres per hour (DfT [10]) gives a value of approximately £0.65 in 2005 prices. Of this total figure, 80% comprises the time component, 12% the fuel cost, and the remaining 8% the vehicle operating costs.
If, on the other hand, the volume of employment within greater London was to remain constant but travel times were to fall uniformly by five percent, then assuming an 80% share for the time component of generalised cost, this would give rise to a 4% increase in average $U_G$ densities, and given the elasticities and sectoral share shown in Table 6, we would expect a 1.0% increase in productivity, ceteris paribus. Of course, any such improvement in speeds would not be costless and would almost certainly require major capacity expansion to the existing network or the imposition of some comprehensive congestion charging scheme.

These two brief examples are used simply to emphasise that the densities associated with agglomeration have temporal and cost dimensions as well as a physical dimension. The use of effective densities based on distance and generalised cost has provided an interesting empirical perspective on the role that road traffic congestion can play in constraining accessibility. Of course, one crucial assumption we have adopted is that the exclusion of travel time information in the $U_D$ variable introduces measurement error because it make the most highly urbanised locations appear more dense than they really are, certainly in temporal terms. This assumption seems to be supported by our data. However, since not all travel is made by roads, and since a high proportion of travel by urban rail is common in many cities, the effective density of activity could actually be higher than represented by road based proximity. Consequently, it is important to stress that the extent to which error exists in any physical density measure of agglomeration will depend on the particular travel characteristics of the area under consideration.

5. Conclusions

In this paper we have developed an analysis of agglomeration externalities within the framework of a translog production inverse input demand function to test for diminishing return from urban transport congestion. We have constructed two effective density measures of agglomeration that incorporate an explicit transport dimension with proximity modelled using straight line distances or the generalised costs of road travel. Through comparison, these two measures, and estimates based upon them, can help us identify the impact of road traffic congestion.

The results show that there are diminishing and even decreasing returns to agglomeration for some sectors of the economy. However, for the types of activity we expect to find in CBDs and large urban areas; such as real estate, banking & finance, business services, and public services; our estimates show returns to agglomeration that are constant or increasing. The effective density variable when calculated using straight line distances rather than generalised costs because the inclusion of information about travel times constrains the values at the top of the distribution. Use of the generalised costs effective density variable in the translog model produces higher agglomeration elasticities and shows less evidence of diminishing returns because it captures both time and distance dimensions of effective density. Put another way, road traffic congestion serves to reduce the densities of the most urbanised locations. Consequently, it can constrain agglomeration externalities and prove an important factor in the diminishing return to agglomeration that we observe.

Acknowledgments

I would like to thank H. Youn Kim, Lars Rognlien, Gilles Duranton, two anonymous referees, and the journal editor for comments on a previous draft of this paper.
Appendix A. Industry sectors used for estimation

The firm level data are aggregated according to the 1992 SIC in the following ways:

(i) manufacturing (MAN) (SIC 15-40),
(ii) construction (CON) (SIC 45),
(iii) distribution, hotels & catering (DHC) (SIC 50-55),
(vi) transport, storage & communications (TSC) (SIC 60-64),
(v) Real estate (RE) (SIC 70),
(vi) Information technology (SIC 72),
(vii) banking, finance & insurance (BFI) (SIC 65-67),
(viii) business services (BUS) (SIC 741-745),
(ix) public services (PSE) (SIC 75-90).

References

学霸图书馆

www.xuebalib.com

本文献由“学霸图书馆-文献云下载”收集自网络，仅供学习交流使用。

学霸图书馆（www.xuebalib.com）是一个“整合众多图书馆数据库资源，提供一站式文献检索和下载服务”的24小时在线不限IP图书馆。

图书馆致力于便利、促进学习与科研，提供最强文献下载服务。

图书馆导航：

- 图书馆首页
- 文献云下载
- 图书馆入口
- 外文数据库大全
- 疑难文献辅助工具