The spatial decay of agglomeration economies: estimates for use in transport appraisal.

Final Report

Daniel J. Graham[*], Stephen Gibbons[**], Ralf Martin[***]

October 2010

[*] Author for correspondence: Centre for Transport Studies, Imperial College London, London, SW7 2AZ, d.j.graham@imperial.ac.uk

[**] Department of Geography and Environment and Centre for Economic Performance, London School of Economics.

[***] Centre for Economic Performance, London School of Economics.

The spatial decay of agglomeration economies: estimates for use in transport appraisal.

Summary

Transport investments can induce positive productivity benefits via agglomeration economies by increasing the scale and efficiency of spatial economic interactions. In assessing the 'agglomeration benefits' of transport investments we need to understand the spatial scale over which these externalities are distributed. This report is concerned with the effect of urban agglomeration on productivity and with how agglomeration externalities diminish with distance from source. It draws on extensive firm level panel data to estimate the effect of access to economic mass on total factor productivity (TFP) for broad sectors of the economy. We adopt a control function approach to addresses potential sources of endogeneity associated with the production function and with the agglomeration-productivity relationship, and to allow for unobserved firm level heterogeneity. A non-linear least squares regression is also used to provide a direct estimate of distance decay. The results show an overall agglomeration effect of 0.04 across all sectors of the economy. For manufacturing and consumer services we estimate an elasticity of 0.02, for construction 0.03, and for business services 0.08. The distance decay parameter is approximately 1.1 for manufacturing, but around 1.8 for consumer and business service sectors and 1.6 for construction. This implies that the effects of agglomeration diminish more rapidly with distance from source for service industries than for manufacturing. But the relative impact of agglomeration on productivity is larger for services than it is for manufacturing. The key results of the research are summarised in the table below.

Summary of empirical results: control function specification, non-linear estimation of alpha.

	sic	agglomeration elasticity	alpha
Manufacturing	15-40	0.024	1.122
Construction	45	0.034	1.562
Consumer services	50-64	0.024	1.818
Business services	65-75	0.083	1.746
Economy (weight aver.)	15-75	0.044	1.659

1. Introduction

Agglomeration economies are positive externalities that arise from the spatial concentration of economic activity. According to urban economic theory, large dense cities and industrial clusters offer tangible economic benefits to firms which are manifest in higher levels of productivity and lower costs. The main sources of agglomeration externalities are thought to arise from improved opportunities for labour market pooling, knowledge interactions, specialisation, the sharing of inputs and outputs, and from the existence of public goods. As the scale and density of urban and industrial agglomerations increase, we expect to find an increase in the external benefits available to firms.

There is a good deal of empirical evidence consistent with this theory showing that cities and industrial concentrations have strong positive statistical association with economic performance of firms. This relationship is found consistently across different empirical contexts and for different industries, although the estimated magnitude of agglomeration externalities depends on the particular circumstances under investigation. Many studies have suggested that agglomeration of economic activity in geographical space can lead to positive spillovers on productivity; i.e. firms that are located in areas with more agglomeration are more productive than they would have been in an area with less economic activity around them.

The existence of agglomeration economies implies that *proximity* is important. But the literature is unclear about what actually counts as proximity, or more specifically, about the geographic scale over which activities could be described as proximate in the sense that they are able to generate these mutual external benefits. Issues concerning the *spatial scope* of productivity effects that arise from agglomeration economies have received little attention in empirical work. Rice et al (2006) provide the most relevant evidence, showing that the relationship between manufacturing wages and economic mass declines sharply with travel time from source. Duranton and Overman (2005) and Rosenthal and Strange (2003), while not concerned with productivity, also indicate that the spatial dimension of agglomeration may be very important. Both papers show that the benefits of industry localization are identified over relatively small spatial scales.

So while we know that there tends to be a positive relationship between productivity and agglomeration, we actually know very little about how this relationship plays out across space. The assumption usually made is that the externalities *decay* with distance from source. But the form any such decay might take, and how this might vary by industry, is largely unknown. Theory can help us to derive ways of representing agglomeration that includes some recognition of the effects of distance, for instance, the widely used *effective density* or *market potential* functions (see Fujita et al 1999, Head & Meyer 2004 and Mion 2004 for a discussion). But it cannot really tell us much about the functional form of these relationships or about the explicit role of distance. The functional form of these distance relationships are usually specified by assumption in theoretical models, and testing of these assumptions is an essentially empirical issue.

The aim of this report is to provide estimates of the geographical range of productivity-related agglomeration economies, in the context of firm-level production function models. Our strategy is to estimate firm production functions that include a total factor productivity term that depends on the local *effective density* of employment at the firm's geographical location (following Graham 2007a, 2007b). We then find the specification of effective density that maximises the power of our models in predicting productivity differences between firms in different locations.

While many studies have found a positive correlation between productivity and agglomeration in the cross section it is not clear that this is evidence for spillovers. The causality might simply go the other way round: more productive firms might locate in denser areas; e.g. they might require more higher skilled employees who are more likely located in urban areas. Another concern is that productivity could be mis measured in ways that systematically interact with geography and agglomeration. There are two issues in particular: Firstly, wages and skills and secondly market power might vary across space. The latter is a problem because in most productivity datasets we measure value rather than volume units of output (Klette Grilliches, Martin). Another issue which has not received much attention in the literature is the question of how persistent any agglomeration spillovers are across space; i.e. how far from the centre of an agglomeration can the spillovers still be felt? This is a key question for any policies aimed at responding to the market failure of spillovers. If these spillovers accrue over a larger area and thereby affect a larger fraction of firms then a case can be made for putting more resources into such policies. This paper responds to all of these issues. Firstly, we develop a framework in the tradition of structural TFP estimation (Olley Pakes, Levinsohn Petrin, Martin) to address both the endogeneity problem of agglomeration as well as the endogeneity of production factors in TFP estimation. Secondly, we examine spatial decay of spillovers by looking at agglomeration at various distance bands. Thirdly, we extend the structural TFP estimation approach to examine variations in market power across space.

The paper is structured as follows. In the next section we discuss the theoretical background to agglomeration economies in production, and consider any theoretical bases for the way distance enters into models of firm production. In the same section we discuss existing empirical work on the specification of market potential and accessibility indices, and assess whether this work sheds light on the specification of effective density in production function estimates. In Section 3, we set out our empirical methods and the data used for estimation. Section 4 presents the results and Section 5 concludes.

2. The theoretical basis and existing empirical evidence

2.1. Theoretical foundations of productivity effects from agglomeration

At their broadest level, agglomeration economies occur when agents benefit from being 'near' to other agents. Nearness can involve physical proximity, but transport and communications play a crucial role because, in most contexts, speed and low costs in transportation and communication provide a direct substitute for physical proximity. In this paper, we are specifically concerned with agglomeration

economies that arise in production[1], and with the role that distance between firms and the sources of these agglomeration economies: workers, other firms, and other facilities that we will discuss further below. It is therefore important to understand what mechanisms are likely to drive production-related agglomeration economies, and we will start with a brief overview of these issues.

Production agglomeration economies usually mean that the productivity of individual firms rises with the overall amount of activity in other "nearby" firms, or with the number of nearby workers or consumers. That is, agglomeration economies arise because of the production benefits of closer connectivity with others. The literature traditionally emphasises three sources of agglomeration economies, roughly following three examples given by Marshall (1890): linkages between intermediate and final goods suppliers, labour market interactions, and knowledge spillovers. Input-output linkages occur because savings on transport costs means that firms benefit from locating close to their suppliers and customers. Larger, denser labour markets may, for example, allow for a finer division of labour or provide greater incentives for workers to invest in skills. Finally, knowledge or human capital spillovers arise when spatially concentrated firms or workers are more easily able to learn from one another than if they were spread out over space.

An alternative taxonomy, which sheds more light on the underlying mechanisms, is provided by Duranton and Puga (2004), who classify the sources of agglomeration economies as 'sharing', 'matching' and 'learning'. 'Sharing' refers to the sharing of indivisible facilities, intermediate suppliers, workers and consumers by firms, which reduces fixed costs, allows specialisation and allows firms to pool risks. 'Matching' benefits are usually discussed in terms of having lots of workers in close proximity to employers, which means it is easier for different types of worker and different types of employer to find each other, and more productive matches occur at a faster rate. 'Learning' refers to the transfer of information, knowledge and skills. Even in world of fast communication technologies, close connections between large groups of people and firms provide more opportunities for learning and more opportunities for face-to-face contact, which tends to facilitate knowledge exchange and transfer of skills. Both the generation of knowledge and its diffusion benefit from these interactions.

Clearly, some of these mechanisms may be more important than others, and the relative importance is likely to be different for different industries. Therefore, it is important to bear these theoretical mechanisms in mind when considering the likely geographical scale over which agglomeration may operate. These theoretical mechanisms should also guide the choice of factors – workers, other firms, population, facilities etc. – that are included in any measure of 'agglomeration', although in practice these factors are closely related to one another and so difficult to distinguish empirically.

Reflecting on the mechanisms described above, it is clear that agglomeration economies depend crucially on the flows of goods, people or information between locations. Therefore, the geographical scope of agglomeration economies will depend on the rate at which these flows decrease with distance. In the next sub-section we consider the empirical implementation of these ideas in the existing literature.

2.2. Empirical implementations

Empirical studies of the economic benefits of agglomeration in production are concerned with estimation of the statistical association between a chosen measure of local economic mass and the productivity of firms in that locality. Firm productivity can be measured by labour productivity, wages or total factor productivity (TFP). TFP measures the quantity of output produced for a given index quantity of all inputs. Issues relating to the measurement of productivity are discussed in detail in Gibbons and Overman (2008) in the context of evaluation of transport improvements, where the case is made that wages and labour productivity are not very useful indicators of the economic benefits of agglomeration because high labour productivity can simply indicate high capital intensity. Total Factor Productivity is the preferred indicator. TFP effects from agglomeration are usually estimated using a production function methodology (see for example the reviews in Rosenthal and Strange 2004, Graham 2007b and Melo et al 2009). Studies typically include a variable representing local economic mass in a standard empirical production function model e.g:

$$r_i = f\left(A_i, k_i, l_i, x_i, \varepsilon_i\right) \tag{1}$$

where A_{it} represents the level of agglomeration at location i at time t, and other production inputs are capital (k_{it}), labour (l_{it}), and other observed (x_{it}) and unobserved (ε_{it}) factors. The left hand side (r_{it}) is typically revenue, or output derived by deflating revenues by price deflators, or value-added (outputs minus material inputs). Inputs can be calculated from direct information inputs such as employment, or from input costs. This kind of production function may be estimated at an aggregate (city, region etc.) or micro level (firm, plant) using statistical regression techniques.

The focus of interest in our empirical work is the construction of the agglomeration term A_i. From the discussion of the micro-foundations of agglomeration (above), it is evident that (at least) two fundamental components need to be considered. Firstly, we need a variable that represents the potential opportunities for a firm to benefit from these agglomeration mechanisms in their locality e.g. employment, population or number of firms. Secondly, we need to define what is meant by 'locality'.

In all existing work, the standard set up is to define A_i as an aggregation of workers, firms or population in the geographical neighbourhood of each firm (i). Let us consider employment as an example. Locality is then either defined using predefined statistical or administrative zones, or, more generally, by aggregating employment with higher weights applied to locations close to firm i, and lower weights to locations further a-field[2]. This type of agglomeration index has the general structure:

$$A_{it} = \sum_{j \neq i} a\left(c_{ijt}\right) z_{jt} \tag{2}$$

where the weights $a\left(c_{ijt}\right)$ are decreasing in the costs or time c_{ijt} incurred in moving between place i and places j, and z_{jt} is the variable (e.g. employment) being aggregated to create the agglomeration index. Weights $a\left(c_{ijt}\right)$ are chosen to apply lower weights to locations j that are further away from location i. Example weighting schemes include: 'cumulative opportunities' weights $a\left(c_{ijt}\right)=1$ if j is within a specified distance of i, zero otherwise; exponential weights $a\left(c_{ijt}\right)=\exp\left(-\alpha c_{ijt}\right)$; logistic weights $a\left(c_{ijt}\right)=\left[1+\exp\left(-\alpha c_{ijt}\right)\right]^{-1}$ or inverse cost weights $a\left(c_{ijt}\right)=c_{ijt}^{\ -\alpha}$.

Graham (2007a) refers to A_{ijt} as *effective density*, defining z_{jt} as postcode-sector-level employment, setting cost (c_{ijt}) as the straight line distance between postcode sectors (d_{ij}) and imposing an inverse distance weighting system (referred to in the literature as 'gravity-based', after Newton). This means that:

$$a\left(c_{ijt}\right)=d_{ijt}^{-\alpha} \text{ and } A_{it}=\sum_{j\neq i} d_{ijt}^{-\alpha} z_{jt} \quad (3)$$

Thus, the effective density measure incorporates the two components that capture the amount of agglomeration experienced by a firm located at a site i: the quantity of employment in another location j (z_{jt}) and the connectedness of site i with site j (d_{ijt}). The parameter α is assumed greater than zero, such that employment at place j has less and less potential influence on a firm at site i as the distance between i and j increases. The larger the value α, the more rapidly the potential effect of employment diminishes with distance d_{ij}. For example, if $\alpha=1$,the weight attached to employment decays inversely with distance (employment 10km away from a firm has 1/10th the effect on effective density as employment 1km away). Graham (2007a) sets $\alpha=1$.

Note that it is not really necessary to specify A_{it} as a single variable. This approach is commonly used because the goal is to estimate a single parameter for the effect of agglomeration on production, for example specifying Equation (1) as $r_{it}=\rho A_{it}+other\ factors$. An alternative approach is to represent agglomeration through employment at several discrete distance (or time) bands e.g. $r_{it}=\sum_{j}\rho_j z_{jt}+other\ factors$. In this second case, a set of estimated parameters (ρ_j) measure how the agglomeration effects decline with distance. We will apply both these specifications in the empirical work below. Comparison of these two representations makes it clear that the agglomeration parameters ρ and distance decay parameter α in effective density measures such as Equation (3) are rather closely related. In fact it is only possible to get separate estimates of ρ and α (i.e. they are only "separately identified") assuming particular functional forms like Equation (3)[3]. Although this distance-decay or gravity functional form is very ad-hoc, it is widely used, and has the advantage that the choice of units for distance used does not matter very much, because the effects are easily rescaled.[4] This does not apply for exponential functional forms, which is one reason for preferring an inverse distance weighting scheme.

2.3. Lessons from the existing literature

The goal of the current research is to provide guidance on appropriate weighting scheme for use in production function based estimates of the effects of agglomeration on productivity, to be applied in transport appraisal. Specifically, we wish to estimate α in the inverse distance-weighted effective density specification in Equation (3).

The reason we model agglomeration using weights that decrease with distance is that interactions between people and firms are expected to decrease as distance increases. The broad reason for this decrease is that interaction and transportation becomes more costly with distance, so flows of goods people and information between places decreases with distance. In other words, decreasing distance weights are simple shorthand for the cost of transporting goods or people over distance, or the higher costs that agents face in interacting with each other over increasing distances. However, theory does not reveal precisely how agglomeration economies should be generated or distributed across space, and does not provide a clear case to support the use of particular functional forms and parameter values to represent proximity. This specific functional relationship is likely to depend on what it is that is being transported, or what kind of interactions we have in mind, and so the appropriate specification for production-related agglomeration indices could well vary by industrial sector. Theory often imposes specific assumptions about the way transport costs enter into models of economic behaviour, and sometimes these assumptions can be crucial to the model tractability and predictions. However, it must be recognised that these are *assumptions* and not *predictions*, and the argument that these assumptions are necessary to solve theoretical models is a poor justification for their use in empirical analysis.

Questions about the rate and functional form of distance decay are essentially empirical questions about the magnitude and nature of transport and connectivity costs. These questions have been addressed before many times in different contexts from ours. In this section we consider what can be learnt from this previous work about the appropriate specification for agglomeration indices, drawing on lessons from four fields a) market access in economic geography; b) transport accessibility analysis; c) the analysis of 'spatial interaction' or flows of goods and people in the trade, commuting and migration literature; and d) the technique of spatial 'surfacing' or smoothing. The following review is by no means exhaustive, but conveys the flavour of work in the field and the main messages that emerge from it.

2.3.1. Market access and new economic geography

The effective density measure in Equation (3) is identical in structure to one form of 'market potential' measure and 'population potential' measure (Harris 1954) that has been widely used in the economic geography, regional science, trade and spatial economics literature. Market potential variables typically use area expenditure, income or sales rather than employment in the numerator of (1), whereas population potential uses residential population. The idea of market potential measures is to create an index of consumer demand, or the market for a product, based on the expenditure, income or number of consumers in neighbouring regions.

A recent resurgence of interest in market potential measures has arisen because of the relevance of market access to New Economic Geography models of geographic concentration. These theoretical models make use of 'iceberg' transport costs after Samuelson (1954). In this set up, transport costs to a region result in a proportional increase in price, relative to price in the producer's home region e.g. a 10% transport cost means that the price at the destination is 1.1 times the producer price. This idea has been extended to allow for continuous distance, by assuming that the price at the destination grows with distance between origin and destination according to an exponential rule. This assumption in turn implies that purchasing power in destination regions decreases with distance from source according to an exponential decay function (e.g. $p_i/p_j = \exp\left(-\alpha d_{ij}\right)$) suggesting an exponential decay weighting structure for 'market potential' indices (Fujita et al 1999). This formulation appears in empirical work that implements market potential measures in NEG models (Hanson 2005, Mion 2005), although there is no a priori reason for believing that the iceberg transport costs assumption is tenable, or that the exponential system is to be preferred over any other distance decay structure. On the contrary, McCann (2005) argues that the exponential iceberg costs assumption – which implies that it costs just as much *per unit* to transport 10,000 units as 10 units, and costs less per mile to transport a unit 10 miles than to transport it 100 miles – sits uncomfortably alongside evidence of economies of scale and distance in transport.

The theoretical foundation of NEG is often invoked to guide other aspects of the construction of market potential indices. For example, Head and Mayer (2004) construct sector-specific market potential indices based on a the sector-specific imports of neighbouring regions, with adjustments for international border effects, where all these components are given structural interpretations based on NEG and are derived from a first stage model of bilateral interregional trade flows. However, the authors *assume* a distance-decay function of the type in Equation (3) when aggregating these components to create market potential. They do estimate rates of distance decay from trade data (see section 2.3.3 below), which they find to be between -0.8 and -2.0, dependent on industry. However, they also find that a simple market potential index of the Harris (1954) form statistically outperforms their more complex theoretically derived index.

On balance then, the theoretical foundation of NEG provides no predictions about the rate at which we can expect agglomeration effects to decline with distance (which depend largely on transport costs), and so does not guide functional form or parameters in the construction of empirical agglomeration indices. Some studies estimate the distance decay structure in prices and spending power implied by transport costs (i.e. prices rise with distance from sources of production) e.g.

Hanson (2005), Mion (2004), Niebuhr (2006). Unfortunately the parameters they estimate are rather difficult to transfer to other contexts.

2.3.2. Accessibility in transport analysis

The accessibility and market potential measures described above are also identical in form to accessibility indices commonly used in transport analysis (e.g. El-Geneidy and Levinson 2006, Vickerman, Spiekermann and Wegener 1999). Transport accessibility indices typically differ from market and population potential indices in that they measure distance or travel times along existing transport networks, rather than straight line or other distances based purely on relative geographical location. Sometimes, these distances and times are converted into generalised transport costs using estimates of the monetary value of travel time, fuel costs etc. derived from elsewhere - see for example Combes and Lafourcade (2005), Graham (2007b).

Accessibility, market potential, population potential or effective density measures based purely on geographical distance create a measure of economic mass at a particular spatial location that depends only on the amount of 'local' employment and how far away that employment is. Superficially, it looks as if effective density is only useful for evaluating the effects of interventions that change the number of workers at a given distance, or move workers closer. However, distance in this framework is simply a proxy for transport costs or time. Given a fixed set of transport infrastructure and a fixed transport policy regime, a distance of, say 10km, has a corresponding (average) travel time or travel cost. To work from a proposed or actual transport improvement to a change in effective density, we need only convert the expected reduction in travel times or travel costs in each direction into an equivalent reduction in distance, under the conditions in which the effective density was calculated. For example, if a proposed transport improvement reduces travel costs to the east of site by 20%, then the new effective density at that site will change in way that is equivalent to moving employment to the east 20% closer (i.e. we need to recalculate effective density with distances in the direction of the transport improvement reduced by 20%).

As discussed above, the more direct way to incorporate transport costs or times into estimates of local economic mass is to base these estimates on existing transport costs or times rather than geographic distances. In this case, local employment counts are aggregated up using a penalty that increases with travel costs or times rather than simple distance[5]. It is then easy to see how to convert a policy-induced change in travel costs or time into a change in accessibility. The drawback of this approach and the reason Graham (2006) uses straight line distances in effective density calculations, rather than network distances, times or costs is that the existing transport network and service is in part dependent on transport demand, which is in turn dependent on the level of economic activity and productivity in a given location. There is thus a risk of inferring that closer connection to employment increases productivity, when it is in fact productivity that has encouraged closer connections through development of the existing transport network. Similar problems arise when trying to calculate international or regional market potential measures using trade flows. There is also an additional computational and data burden in calculating fully specified transport accessibility indices that require information on network distances and/or times.

Construction of indices based on transport costs faces the same problems of functional form and rate of distance decay as indices based on straight-line distances. This strand of literature has fairly little to add in terms of empirical answers to these questions. More useful guidance on rates of distance decay and functional forms comes from the literature that models flows of people, goods and information between places, and the rate at which these flows decline with distance and time between places. We turn now to this literature.

2.3.3. Modelling spatial interaction: Trade, migration, commuting and information

Many fields of investigation model flows of things between predefined sets of origins and destinations. This line of enquiry is relevant to us because the distance, or cost, of moving between each origin point and each destination point is a key variable in theoretical and empirical analysis.

The theoretical foundations of these analyses are varied. In the field of transport, commuting and migration, statistical foundations based on entropy (Wilson 1970) and spatial interaction

(Fotheringham O'Kelley 1989) were common. However, these foundations have largely given way to the utility or profit maximising choice-based framework of the Random Utility Model (e.g. Tversky 1972,), but theoretical approaches alone provide no guidance on how rapidly flows decrease with the distance between origin and destination.

Typical empirical applications include the analysis of trade flows, migration (human and non-human) and commuting. At their most basic, these applications involve estimation of regression models in which flows between locations are explained by a range of characteristics of origin and destination, and the distance between them. Typically, these flows, distance and explanatory variables are transformed to natural logs and models in this form are commonly referred to as 'gravity' models. These aggregate flow models have equivalent micro-level representations, which, for example model the probability that an individual at a given origin makes a given destination choice. In either case, empirical applications can provide us with information about how the connectivity between places is affected by the distance between them i.e. the rate of distance decay. Whilst we cannot provide a comprehensive survey we briefly discuss the key findings relevant to agglomeration economies in production.

Firstly, we expect trade in goods to be an important factor linking firms. A recent meta analysis of gravity models of international trade flows (Disdier and Head 2008) suggests that the elasticity of trade with respect to distance is about -0.9 on average. This result provides some empirical support for the common choice of $\alpha = -1$ in effective density/market potential indices of the type in Equation (3). However, there is no guarantee that this result carries over to intra-national, intra-regional and intra-urban transport of goods. Estimates from intra-national trade flows do, however, produce similar results, e.g. Wolf (2000) estimates a slightly lower figure (0.77-0.81) for intra-US-state trade, and Brown and Anderson (2002) find elasticities of around -0.8 to -1.2 from trade between US states, and between Canadian provinces.

Secondly, commuting links firms together via the labour market. Far fewer recent studies provide helpful estimates of the costs of distance in personal travel. Johanssen Klaesson and Hanson

(2004) consider commutes in Sweden and show (unsurprisingly) that commuting time deters commuting, and that sensitivity to time is higher for intermediate commutes (inside region), than for short (within municipalities) or long (between region) distances. Unfortunately they do not provide the information necessary to translate their parameter estimates into a form that could be compared with the distance-decay parameters in which we are interested. Older work, such as Hansen (1959), Fotheringham (1981) suggests parameter values ranging from 0.5 to 3 depending on trip purpose and study area, but Hansen (1959) cites a figure of -0.9 for work trips, and Fotheringham's own average estimate is -0.9. More recently, de Vries et al 2004 estimate a number of commuting-cost elasticities based on alternative specifications, and conclude that a logistic specification provides the best fit for their data on commuting patterns in Denmark. In their inverse-cost decay specifications, they estimate a cost decay parameter, analogous to the alpha parameter in which we are interested, of around -2. In summary, the literature provides rather a wide range of estimates on the elasticity of commuting with respect to distance or other cost measures, and these estimates are quite context-dependent.

The final channel of connectivity is information and communication. Only one study of which we are aware provides estimates that are directly relevant: Blum and Goldfarb (2006) show that US web site visitors are more likely to visit sites in nearby countries than more distant countries. Even when no on-line purchase is involved the elasticity of visits with respect to distance is over -1, implying that distance imposes costs on connectivity even on-line.

2.3.4. Spatial smoothing of surfacing

Another class of method for estimating distance decay functions and parameters appears as 'surfacing' or 'interpolation' in geographical terminology, 'smoothing' or 'kernel regression' in econometrics and statistics. The terms refer to the process of averaging local population, employment or other spatially distributed data values to create a smooth spatial surface. The general approach is to find the functional form and distance decay parameter(s) that generates a surface that best fits the data from which it is generated. Typically this means maximising the correlation between the original data values and the surface predictions. Coombes and Raybould 2001 carry out this purely data-driven exercise using various GB ward-level census characteristics and an inverse-distance (gravity) weighting system, and arrive at an optimal α parameter of around -1.5. Song (1996) uses a similar methodology for Nevada, but bases the estimation on the effectiveness of smoothed employment counts (accessibility of jobs) in explaining population density in regression models. The theoretical reasoning is that population density at any location should be driven by the accessibility of employment from that location. The conclusion is that an inverse distance weighting function with $\alpha = -1$ is more effective than any other, including exponential and more elaborate functional forms. In other words, the original functional form of Harris (1954 - see above) is as good as any other when assessed under this criterion.

In summary, it is hard to draw any solid conclusions from the diverse strands of literature that have considered these questions of distance decay. Typical estimates suggest that the elasticity of flows of goods and people with respect to distance is in the order of -1, and other methods have come up with similar parameters for rates of distance decay in accessibility indices and market potential measures. However, estimates range quite widely and there is no definitive answer to the question of the most appropriate functional form for these indices. The main feature that the indices need to incorporate is that interaction across space declines with distance, and the specific shape of this decay may not be important. There is no evidence to suggest that a simple inverse-distance function underperforms in this context relative to other more complex measures.

3. Econometric Approach

The equation we want to identify takes the following form

$$\omega_{it} = \rho A_{it} + \varepsilon_{it} \quad (4)$$

where ω_{it} is some productivity measure and A_{it} is a measure of agglomeration as discussed above. Our main productivity measure is total factor productivity TFP which relates output to all factor inputs. We are working with firm level data which allows us to look at firm level revenues but not output quantities. So that, from (1):

$$\omega_{it} = r_{it} - \beta_k k_{it} - \beta_m m_{it} - \beta_l l_{it} \quad (5)$$

where $r_{it}, k_{it}, m_{it}, l_{it}$ are the (log) firm level revenue, capital, intermediates and labour, respectively. This has two implications. Firstly, we are looking at revenue TFP (RTFP) rather than TFP which has implications for our findings that we will discuss in more detail below. Secondly, to assess the impact of agglomeration on RTFP we need firstly unbiased estimates of the β coefficients in equation (5) and secondly of ρ in equation (4).

We employ a number of different approaches for that purpose. Firstly, we simply run regressions of the form

$$r_{it} = \rho A_{it} + \beta_k k_{it} + \beta_m m_{it} + \beta_l l_{it} + \varepsilon_{it} + \eta_{it} \quad (6)$$

In which firm-specific unobserved components are ε_{it} and η_{it}. Below we refer to this as the regression based RTFP estimate. While this provides a good benchmark, it is problematic because both the production factor variables and agglomeration are potentially correlated with ε_{it} i.e. more productive firms are likely to employ more factor inputs and there might be a systematic selection of firms across space according to their productivity.

The simplest way to address the production factor endogeneity is to assume that there is constant returns to scale, perfect competition and all factors adjust immediately to any productivity shocks. In this case unbiased estimates of the β s can be obtained by computing each factors' cost share[6]:

$$S_{xit} = \frac{W_{xt} X_{it}}{R_{it}} \text{ and } \hat{\beta}_x^{shares} = \overline{S}_x \tag{7}$$

Consequently, we get

$$\omega_{it}^{share} = r_{it} - \hat{\beta}_k^{shares} k_{it} - \hat{\beta}_m^{shares} m_{it} - \hat{\beta}_l^{shares} l_{it} \tag{8}$$

which we regress on agglomeration A_{it}.[7]

3.1. *A control function approach*

The cost share approach still does not address the potential endogeneity of A_{it}. Also it requires very restrictive assumptions on firm behaviour. To deal with this we develop the following control function approach to production function estimation (Olley Pakes, Levinsohn Petrin, Martin).

Control function approaches derive conditions under which productivity can be proxied by state variables and control variables, i.e.

$$\omega_{it} = \phi\left(\pi_{it}, z_{it}\right) \tag{9}$$

where $\phi(.)$ is an unknown function that will be determined through a non-parametric procedure, π_{it} is the proxy variable[8] and z_{it} is a vector of firm level state variables[9]. Control function approaches are motivated by an assumption about firm behaviour. For instance, Olley and Pakes (1994) make assumptions about the dynamic optimisation behaviour of firms to show that a firm's investment function is monotone in its unobserved productivity, conditional on various state variables. Under this condition an inverse function of the investment function exists and we can write, and non-parametrically estimate, a function such as equation 9 with investment as proxy variable (instead of net revenue). When using net revenue (revenue minus expenditure on variable production factors) as in 9 we make similar argument using less restrictive assumptions about the short term optimisation behaviour of firms (Martin, 2008).

We start by assuming that agglomeration A_{it} is not a state variable. This assumes that agglomeration has no impact on firm behaviour and firm profits other than through its impact on RTFP. Of course this might be restrictive. Two issues in particular would violate this assumption: Firstly, if agglomeration affects the market power of firms and secondly if agglomeration has an impact factor costs such as wages. We therefore relax this assumption below. However, again it is a good benchmark to maintain the assumption at first. Note that it supports the following very simple strategy to get to an unbiased estimate of ρ. It suffices to run a linear regression of the following equation:

$$\omega_{it}^{share} = \rho A_{it} + \phi_{\varepsilon}\left(\pi_{it}, k_{it}\right) + \eta_{it} \tag{10}$$

where $\phi_{\varepsilon}\left(\pi_{it}, k_{it}\right) = \phi\left(\pi_{it}, k_{it}\right) - \rho A_{it} = \varepsilon_{it}$ by the maintained assumptions and $\phi_{\varepsilon}\left(\pi_{it}, k_{it}\right)$ is approximated by a third order polynomial in its arguments; i.e. net revenue and capital.[10]

3.2. A more flexible control function approach

It might be more realistic to assume that $\omega_{it} = \phi\left(A_{it}, \pi_{it}, k_{it}\right)$, that is, that agglomeration affects not only productivity in a linear way but also has a potentially non-linear impact on the proxy function. Estimation of ρ then proceeds in two stages. Firstly, run the following regression

$$\omega_{it}^{share} + k_{it} = \phi_{\varepsilon}\left(A_{it}, \pi_{it}, k_{it}\right) + \eta_{it} \tag{11}$$

Further assuming that ε_{it} proceeds as Markov process - i.e. $\varepsilon_{it} = \phi_{E}\left(\varepsilon_{it-1}\right) + v_{it}$ we can derive a consistent nonlinear least squares regression of the following equation

$$\omega_{it}^{share} + k_{it} = \beta_k k_{it} + \rho A_{it} + \phi_{E}\left(\hat{\phi}\left(A_{it-1}, \pi_{it-1}, k_{it-1}\right) - \beta_k k_{it-1} - \rho A_{it-1}\right) + v_{it} + \eta_{it} \tag{12}$$

where we approximate with yet another polynomial.

3.3. Allowing for variations in market power

A key motivation for allowing agglomeration to affect $\phi(\ )$ concerns its impact on market power. Note however that equation (12) still maintains constant market power across all firms as

$$\beta_k = \frac{\gamma}{\mu} \tag{13}$$

where μ is the markup of prices over marginal costs. We relax this assumption by assuming that $\frac{1}{\mu}$ varies linearly with agglomeration:

$$\beta_k = \beta_{k0} + \beta_{k1} A_{it}$$

Our second stage regression (12) thus becomes

$$\omega_{it}^{share} + k_{it} = \beta_{k0} + \beta_{k1} A_{it} + \rho A_{it} + \phi_E \left(\hat{\phi}\left(A_{it-1}, \pi_{it-1}, k_{it-1} \right) - \beta_k k_{it-1} - \rho A_{it-1} \right) + v_{it} + \eta_{it} \tag{14}$$

We then interpret a positive value for β_{k1} as evidence that agglomeration reduces market power - i.e. μ smaller for firms in more urban areas.

3.4. Labour productivity

As a benchmark we also compute results where we simply use labour as a dependent variable in (4); i.e. log value added divided by the number of employees: VA_{it} / L_{it}:

3.5. Variations in wages and skill

To account for variations in wages and skill we run all (R)TFP regressions with both a head count measure of labour and using the wage bill as labour input measure. This is somewhat crude as it does not allow us to distinguish between wage and skill effects. Such a distinction is however important as they have different welfare implications. Below we find that any spillover effects (significant and positive ρ) greatly reduce and sometimes disappear when using the wage bill measure. If this is driven by skill variation then the conclusion is correct that there are no positive spillovers. If it is driven by wage variations it simply means that employees, rather than firms manage to appropriate much of the spillover. Consider therefore our two measures as an upper (head count labour measure) and a lower bound (wage measure) of the actual spillovers.

3.6. Computing agglomeration measures and their decay

We use a number of agglomeration measures. All are based on aggregate employment measures at various size bands which we computed for each firm in our sample individually. For instance, employment at size band θ with lower radius $r_{0\theta}$ and upper radius $r_{1\theta}$ is defined as

$$L_{\theta it} = \sum_{j \in \{r_{0\theta i}, r_{1\theta i}\}} L_{jt} \tag{15}$$

where $\{r_{0\theta i}, r_{1\theta i}\}$ is the set of all business units located within the boundaries of size band θ at time t.

Below we experiment with a number of different size band resolutions. Our finest resolution has boundaries at 2.5, 5, 10, 25, 50 and 75 km. We use these bands to compute 3 types of agglomeration measures. Firstly, we include the complete set size band employments as our measure of A_{it} in equation (4):

$$A_{it} = \sum_{\theta} \frac{\rho_\theta}{\rho} \ln L_{\theta it} \tag{16}$$

To study the decay of agglomeration spillovers we can examine how ρ_θ declines as we move to distance bands that are further away.

Secondly, we look at

$$A_{it} = \ln \sum_{\theta} L_{\theta it} d_\theta^{-1} \tag{17}$$

where d_θ is the distance at the mid point of distance band θ. This has been the measure of choice in much of the previous literature. It implicitly assumes that spillovers decay in inverse proportion with distance. This might be restrictive. Our third measure is therefore a more general version of (17):

$$A_{it} = \ln \sum_{\theta} L_{\theta it} d_\theta^{-\alpha} \tag{18}$$

where we allow spillovers to decay at rate α, and where this parameter is estimated using a non linear least squares regression of equation (4).

4. Results

Table 1 below summarises the main findings obtained using the flexible control function specification with non-linear estimation of the distance decay parameters α. This is our preferred specification for the following reasons:

1. It offers a flexible representation of TFP which addresses potential sources of endogeneity within the production function (i.e. unobserved productivity).
2. It allows for an endogenous relationship between agglomeration and productivity.
3. The Markov specification draws on variation over time rather than across firms, and thus effectively eliminates the influence of unobserved firm level heterogeneity.
4. It provides a direct estimate of the distance decay parameter.

Table 1: Summary of empirical results: control function specification, non-linear estimation of alpha.

	sic	agglomeration elasticity	alpha
Manufacturing	15-40	0.024	1.122
Construction	45	0.034	1.562
Consumer services	50-64	0.024	1.818
Business services	65-75	0.083	1.746
Economy (weight aver.)	15-75	0.044	1.659

The table shows an agglomeration elasticity of 0.04 averaged across the four broad sectors of the economy[11]. For manufacturing and consumer services we estimate elasticities of 0.02, for construction 0.03 and for business services 0.08. Thus, we find a pattern of estimates consistent with those reported previously for the UK (e.g. Graham 2007a, 2007b), which show variation across industries in the effect of agglomeration on productivity with business service sectors enjoying the largest effects. However, while the basic pattern across industries is the same, the magnitude of the estimates is substantially smaller. We believe this is largely due to more effective controls for unobserved heterogeneity and a more accurate representation of TFP, as well as differences in industry mix and composition.

Table 1 also shows estimates of the distance decay parameters. For manufacturing the estimate is not far from 1.0, for construction the estimate of alpha is 1.6, and for consumer and business services the estimate is around 1.8. This implies that the effects of agglomeration diminish more rapidly with distance from source for service industries and construction, than for manufacturing. Or in other words, proximity matters more for services and particularly for those that tend to be located in and around Central Business Districts.

The remaining tables in this section show detailed results from regressions for manufacturing firms using the different approaches discussed in section 3 above. These results are useful in demonstrating how different econometric specifications can alter our interpretation of agglomeration economies and their decay over distance. The corresponding tables for all four industry groups are reproduced in the appendix.

Table 2: Regression results with employment bands

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Labour productivity	RFTP	RFTP	RFTP	RFTP	RFTP	RFTP	RFTP	RFTP
	log(VA/L)	Factor Shares	Regression based				Control function		
ln(L) 0 to 25 km	0.031***	0.013***	0.015***	0.002	0.025***	0.003	0.013***	0.005***	0.020***
	(0.006)	(0.003)	(0.003)	(0.003)	(0.005)	(0.005)	(0.002)	(0.002)	(0.005)
ln(L) 25 to 50 km	0.020**	0.001	0.009**	0.006	0.023***	0.017**	0.003	0.000	0.012
	(0.008)	(0.004)	(0.004)	(0.004)	(0.008)	(0.007)	(0.003)	(0.002)	(0.007)
ln(L) 50 to 75 km	0.000	0.007*	-0.002	-0.007**	0.005	-0.005	0.002	-0.001	0.012*
	(0.007)	(0.004)	(0.004)	(0.003)	(0.007)	(0.006)	(0.002)	(0.002)	(0.006)
Controlling for material inputs	no	yes	yes	yes	no	no	yes	yes	no
Headcount labour input	yes	yes	yes	no	yes	no	yes	no	yes
Total Labour costs	no	no	no	yes	no	yes	no	yes	no
Age controls	yes	yes	yes	yes	yes	yes	yes	yes	yes
3 digit sector controls	yes	yes	yes	yes	yes	yes	yes	yes	yes
Coefficients vary at 3 digit sector level	no	yes	yes	yes	yes	yes	yes	yes	yes
obs	27835	27835	27835	27835	27835	27835	27835	27835	27835

Source: Authors' calculations based on ARD data.
Notes: The table shows regressions of various productivity measures on employment at various distance bands. All production function coefficients are allowed to vary at the 3 digit level.

Table 2 shows regressions with 3 distance bands for different productivity measures. Column 1 reports results for labour productivity. We find a positive and significant agglomeration effect for the closest band of 0 to 25km and a somewhat smaller less significant effect for the 25 to 50km band. When moving to factor share based TFP in column 2, and simple regression based TFP in column 3, the 0 to 25km effect becomes smaller but is still significant. A decline when moving from simple labour productivity to TFP is plausible since firms in different areas might be substituting labour for other factor inputs. For example, if wages are higher in more urbanised areas, it is likely that firms will substitute away from labour which would explain the observed differences.

In column 4 of table 2 we use wage costs rather than employment as a measure for labour inputs. This renders all positive agglomeration effects insignificant. As discussed above, this could be driven by two separate issues: Firstly, wages (for the same kind of labour input) vary between different areas and secondly, firms in different areas might be employing different types of labour inputs (skill differences). The former would imply that any effects found when using labour inputs would be indeed a spillover, whereas the latter would suggest that firms in more urbanised areas use different inputs but don't have necessarily higher productivity. Since we expect agglomeration economies to be capitalised in the wage rate the fact that we find no agglomeration economies using wage costs is uninformative about whether the former or the latter effect dominates.

Columns 5 and 6 explore what happens if we do not control for intermediate inputs. This is primarily for reference purposes as many earlier studies relied on firm level datasets without intermediate input information. Note that for regression based TFP this makes a big difference; e.g. the agglomeration coefficient increases from 0.015 to 0.025 between column 3 and 5. This is not surprising as it now measures both the impact of agglomeration on TFP and on intermediate input usage.

Columns 7 to 9 report the results when a control function approach is used. The difference when not controlling for intermediate inputs is less noticeable when using this approach (column 7 vs 9). Again this is to be expected since, if material inputs adjust flexibly to changes in TFP, the control function approach of columns 7 to 9 perfectly controls for the omission of intermediate inputs.

Table 3: Regression results imposing α = 1

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Labour productivity	RFTP	RFTP	RFTP	RFTP	RFTP	RFTP	RFTP	RFTP
	log(VA/L)	Factor Shares	Regression based				Control function		
agglomeration with α=1	0.056***	0.022***	0.026***	0.003	0.053***	0.014***	0.021***	0.005***	0.042***
	(0.005)	(0.003)	(0.002)	(0.002)	(0.005)	(0.004)	(0.002)	(0.001)	(0.004)
Controlling for material inputs	no	yes	yes	yes	no	no	yes	yes	no
Headcount labour input	yes	yes	yes	no	yes	no	yes	no	yes
Total Labour costs	no	no	no	yes	no	yes	no	yes	no
Age controls	yes	yes	yes	yes	yes	yes	yes	yes	yes
3 digit sector controls	yes	yes	yes	yes	yes	yes	yes	yes	yes
Coefficients vary at 3 digit sector level	no	yes	yes	yes	yes	yes	yes	yes	yes
obs	27835	27835	27835	27835	27835	27835	27835	27835	27835

Source: Authors' calculations based on ARD data.
Notes: The table shows regressions of various productivity measures on an agglomeration measure with decay parameter α=1. All production function coefficients are allowed to vary at the 3 digit level.

Table 3 reports results for the agglomeration measure described in equation (17), where the distance decay factor is constrained to equal 1. In our preferred specification, using the control function approach with employment as the measure of labour input (column 7), this leads to an elasticity of agglomeration on TFP of 2.1%.

Table 4: Regressions results when estimating α

	(1)	(2)	(3)	(4)	(5)	(6)
	Labour productivity	RFTP	RFTP	RFTP	RFTP	RFTP
	log(VA/L)	Factor Shares	Regression based		Control function	
Marginal Impact of u	0.064***	0.026***	0.032***	0.067***	0.024***	0.006***
ρ	(0.006)	(0.003)	(0.003)	(0.005)	(0.002)	(0.001)
Decay of agglomeration effects	1.062***	1.077***	1.171***	0.879***	1.122***	1.410***
α	(0.140)	(0.178)	(0.140)	(0.137)	(0.127)	(0.405)
Controlling for material inputs	no	yes	yes	yes	yes	yes
Headcount labour input	yes	yes	yes	no	yes	no
Total Labour costs	no	no	no	yes	no	yes
Age controls	yes	yes	yes	yes	yes	yes
3 digit sector controls	yes	yes	yes	yes	yes	yes
Coefficients vary at 3 digit sector level	no	yes	yes	yes	yes	yes
obs	27835	27835	27835	27835	27835	27835

Source: Authors' calculations based on ARD data.
Notes: The table shows non linear regressions of various productivity measures on an agglomeration measure with undetermined agglomeration effect decay. All production function coefficients are allowed to vary at the 3 digit level.

In table 4 we use a non linear least squares approach (see equation (18)) to estimate not only the marginal impact of agglomeration but also an appropriate decay factor. This is our preferred approach and is consistent with the headline results reported in table 1 and the more detailed results in table A3 of the appendix. Looking again at our preferred specification, this time in column 5, we find a point estimate of 1.122, which is not significantly different from 1, for the distance decay factor. Correspondingly, the agglomeration elasticity of 2.4% is close to that reported in table 2.

Table 5: A more general control function approach

	(1)	(2)	(3)	(4)	(5)	(6)
	Control Function 1		Control Function 2			
ln(L) 0 to 25 km	0.0106***	0.0018*	0.0013***	0.0013***	0.0014***	0.0014***
	(0.0017)	(0.0010)	(0.0004)	(0.0003)	(0.0004)	(0.0004)
ln(L) 0 to 25 km X ln(K)					-0.0013	0.0003
					(0.0012)	(0.0019)
Controlling for material inputs	yes	yes	yes	yes	yes	Yes
Headcount labour input	yes	no	yes	no	yes	no
Total Labour costs	no	yes	no	yes	no	yes
Age controls	yes	yes	yes	yes	yes	yes
3 digit sector controls	yes	yes	yes	yes	yes	yes
Coefficients vary at 3 digit sector level	yes	yes	yes	yes	yes	yes
obs	10199	10199	10199	10199	10199	10199
obs first stage			27835	27835	27835	27835

Source: Authors' calculations based on ARD data.
Notes: The table shows non linear regressions of various productivity measures on an agglomeration measure with undetermined agglomeration effect decay. All production function coefficients are allowed to vary at the 3 digit level.

In Table 5 we relax the rather stringent restrictions on the control function approach. This requires a two stage approach where we identify agglomeration effects from changes over time. With the sample data we are using this creates the additional complication that we lose those observations of firms which are not sampled in consecutive years. For reference, in table 5 we first reproduce the simple control function approach with the smaller sample in columns 1 and 2. We find somewhat smaller agglomeration coefficients on this smaller sample, however they are still significant at 1% for head count labour in column 1 (the equivalent of column 7 in tables 2 and 3 and of column 5 in table 4) and at 10% for wage cost labour in column 2 (equivalent to column 8 in tables 2 and 3 and to column 6 in table 4). Using the advanced control function approach suggests even smaller point estimates of 0.13% with head count labour, which are still strongly different from zero, however. With wage cost labour the effect is remains now unchanged.

The advanced control function approach allows the market power parameter to vary with the degree of agglomeration. When using head count labour in column 5 the negative sign on the interaction between capital and urbanisation suggests that that market power is higher in more urbanised areas. This is consistent with the hypothesis that closeness to customers in urbanised areas gives firms the opportunity to charge higher mark-ups over marginal costs. However, this is not significant. Indeed, when using wage expenditure as measure for labour input in column 6 coefficient becomes positive, yet again insignificant.

Table 6: Regression results with high resolution employment bands

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Labour productivity	RFTP	RFTP	RFTP	RFTP	RFTP	RFTP	RFTP	RFTP
	log(VA/L)	Factor Shares	Regression based				Control function		
ln(L) 0 to 2.5 km	0.009**	0.001	0.002	-0.005**	0.005	-0.006	0.004**	-0.000	0.011***
	(0.005)	(0.002)	(0.002)	(0.002)	(0.004)	(0.004)	(0.002)	(0.001)	(0.004)
ln(L) 2.5 to 5 km	0.010*	0.004	0.006**	0.006**	0.006	0.007	0.003	0.002	0.001
	(0.006)	(0.003)	(0.003)	(0.003)	(0.005)	(0.005)	(0.002)	(0.002)	(0.005)
ln(L) 5 to 10 km	-0.001	0.003	0.001	-0.000	0.001	-0.001	0.002	0.002	0.001
	(0.006)	(0.003)	(0.003)	(0.003)	(0.006)	(0.006)	(0.002)	(0.002)	(0.006)
ln(L) 10 to 25 km	0.009	0.002	0.002	-0.002	0.008	0.001	0.004	-0.000	0.008
	(0.008)	(0.004)	(0.004)	(0.004)	(0.007)	(0.007)	(0.003)	(0.002)	(0.007)
ln(L) 25 to 50 km	0.028***	0.004	0.013***	0.007*	0.028***	0.017**	0.005*	0.001	0.015*
	(0.009)	(0.005)	(0.005)	(0.004)	(0.008)	(0.008)	(0.003)	(0.002)	(0.008)
ln(L) 50 to 75 km	-0.002	0.006	-0.003	-0.007**	0.003	-0.005	0.002	-0.001	0.010*
	(0.007)	(0.004)	(0.004)	(0.003)	(0.007)	(0.006)	(0.002)	(0.002)	(0.006)
Controlling for material inputs	no	yes	yes	yes	no	no	yes	yes	no
Headcount labour input	yes	yes	yes	no	yes	no	yes	no	yes
Total Labour costs	no	no	no	yes	no	yes	no	yes	no
Age controls	yes	yes	yes	yes	yes	yes	yes	yes	yes
3 digit sector controls	yes	yes	yes	yes	yes	yes	yes	yes	yes
Coefficients vary at 3 digit sector level	no	yes	yes	yes	yes	yes	yes	yes	yes
obs	27835	27835	27835	27835	27835	27835	27835	27835	27835

Source: Authors' calculations based on ARD data.
Notes: The table shows regressions of various productivity measures on employment at various distance bands. All production function coefficients are allowed to vary at the 3 digit level.

5. Conclusions

This report has provided estimates of the effect of urban agglomeration on productivity and how agglomeration externalities diminish with distance from source. The analysis is based on extensive firm level panel data which we use to represent total factor productivity (TFP) for broad sectors of the economy. Urban agglomeration is measured in terms of access to economic mass discounted by Euclidean distance. The econometric specification is based on a control function approach which addresses potential sources of endogeneity associated with the production function and the productivity-agglomeration relationship, and which also allows for unobserved firm level heterogeneity. A non-linear least squares approach is used to provide a direct estimate of distance decay.

The results show an overall agglomeration effect of 0.04 across all sectors of the economy. For manufacturing and consumer services we estimate an elasticity of 0.02, for construction 0.03, and for business services 0.08. The distance decay parameter is approximately 1.1 for manufacturing, but around 1.8 for consumer and business service sectors and 1.6 for construction. This implies that the effects of agglomeration diminish more rapidly with distance from source for service industries than for manufacturing. But the relative impact of agglomeration on productivity is larger for services than it is for manufacturing.

References

Ahmed M. El-Geneidy, David M. Levinson (2006) Access to Destinations: Development of Accessibility Measures, Minnesota Department of Transportation

Brown, W.M. and W. P. Anderson, Spatial markets and the potential for economic integration between Canadian and U.S. regions, Papers in Regional Science, 81 (1) 99-120

Combes, P. & M. Lafourcade (2005) Transport costs: measures, determinants, and regional policy implications for France, Journal of Economic Geography 5(3) 319-349.

Coombes , M. and S. Raybould (2001) Public policy and population distribution: developing appropriate indicators of spatial settlement patterns, Environment and Planning C, 19 223-248

De Vries, Jacob J., P. Nijkamp, and P. Rietveld, (2004) Exponential or Power Distance-decay for Commuting? An alternative Specification, Tinbergen Institute Discussion Paper TI2004-097/3

Duranton, G. and H. Overman (2005) Testing for localisation using micro-geographic data. Review of Economic Studies 72, 1077–1106.

Fotheringham, A.S. (1981) Spatial structure and distance-decay parameters, Annals of the Association of American Geographers 71 (3) :425-36

Fujita M, Krugman P and Venables AJ (1999) The spatial economy: Cities, regions and international trade, MIT Press, Cambridge MA.

Gibbons, S. and Overman, H. G. (2008) Productivity in Transport Evaluation Studies, Report to Department of Transport

Graham DJ (2007a) ‘Agglomeration, productivity and transport investment’ Journal of Transport Economics and Policy, 41, 317-343.

Graham DJ (2007b) ‘Variable returns to agglomeration and the effect of road traffic congestion’ Journal of Urban Economics, 62, 103-120.

Hansen, W. G. (1959). How Accessibility Shapes Land Use. Journal of the American Institute of Planners, 25 (2) 73-76

Head, K. and T. Mayer (2004) ‘Market potential and the location of Japanese investment in the European Union’, Review of Economics and Statistics, 86, 959–972.

Melo P, Graham DJ and Noland RB (2009) ‘A meta-analysis of urban agglomeration economies’, Regional Science and Urban Economics, 39, 332-342.

Mion, G. (2004) ‘Spatial externalities and empirical analysis: the case of Italy’, Journal of Urban Economics, 56, 97–118

Rice P, Venables A and Patacchini E (2006) ‘Spatial Determinants of Productivity: Analysis for the Regions of Great Britain’, Regional Science and Urban Economics, 36, 727-752.

Rosenthal, S. and W. Strange (2003) Geography, industrial organisation and agglomeration. Review of Economics and Statistics 85, 377–393

Rosenthal, S. and W. Strange (2004) Evidence on the Nature and Scale of Agglomeration Economies in J. F. Thisse and J. V. Henderson (eds.) Handbook of Urban and Regional Economics, Volume 4. North Holland.

Samuelson, Paul A. (1954) The Transfer Problem and Transport Costs, II: Analysis of Effects of Trade Impediments," Economic Journal 64, pp. 264-289

Skov-Petersen, H. 2001: Estimation of Distance-decay parameters – GIS-based indicators of recreational

accessibility. In Bjørke, J.T. and Tveite, H (eds). ScanGIS 2001. The 8’th Scandinavian Research Conference on Geographical Information Science. Proceedings. Aas, Norway June 24-27 2001

Song, Shunfeng (1996) Some tests of alternative accessibility measures: a population density approach, Land Economics 72 (4) 474-482

Tversky, A., Choice by Elimination, Journal of Mathematical Psychology 9 (1972), 341-367.

Vickerman, Spiekermann and Wegener (1999) Accessibility and Economic Development in Europe, Regional Studies 3(1) 1-15

Wilson A.G. (1970) Entropy in Urban and Regional Modelling, Pion

Appendix: Detailed results.

Table A1: Regression results with employment bands

	manufacturing					
	(1)	(2)	(3)	(4)	(5)	(6)
	Labour productivity log(VA/L)	RFTP Factor Shares	RFTP Regression based	RFTP	RFTP Control function	RFTP
ln(L) 0 to 25 km	0.031***	0.013***	0.015***	0.006**	0.013***	0.012***
	(0.006)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)
ln(L) 25 to 50 km	0.020**	0.001	0.009**	0.007	0.003	0.003
	(0.008)	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)
ln(L) 50 to 75 km	0.000	0.007*	-0.002	-0.005	0.002	0.002
	(0.007)	(0.004)	(0.004)	(0.003)	(0.002)	(0.002)
TTWA skill level				0.204***		0.283***
				(0.039)		(0.030)
Age controls	yes	yes	yes	yes	yes	yes
3 digit sector controls	yes	yes	yes	yes	yes	yes
Coefficients vary at 3 digit sector level	no	yes	yes	yes	yes	yes
obs	27835	27835	27835	27835	27835	27835
	construction					
	(1)	(2)	(3)	(4)	(5)	(6)
	Labour productivity log(VA/L)	RFTP Factor Shares	RFTP Regression based	RFTP	RFTP Control function	RFTP
ln(L) 0 to 25 km	0.046***	0.029***	0.039***	0.026***	0.023***	0.022***
	(0.008)	(0.006)	(0.005)	(0.005)	(0.004)	(0.004)
ln(L) 25 to 50 km	0.009	-0.008	-0.003	0.002	-0.004	-0.002
	(0.011)	(0.009)	(0.007)	(0.006)	(0.006)	(0.006)
ln(L) 50 to 75 km	0.018*	0.006	0.011*	0.009	0.009*	0.010**
	(0.009)	(0.007)	(0.006)	(0.005)	(0.005)	(0.005)
TTWA skill level				0.135**		0.132**
				(0.066)		(0.059)
Age controls	yes	yes	yes	yes	yes	yes
3 digit sector controls	yes	yes	yes	yes	yes	yes
Coefficients vary at 3 digit sector level	no	yes	yes	yes	yes	yes
obs	12044	12044	12044	12044	12044	12044

Source: Authors' calculations based on ARD data.
Notes: The table shows regressions of various productivity measures on employment at various distance bands. All production function coefficients are allowed to vary at the 3 digit level.

	consumer services					
	(1)	(2)	(3)	(4)	(5)	(6)
	Labour productivity log(VA/L)	RFTP Factor Shares	RFTP	RFTP	RFTP	RFTP
			Regression based		Control function	
ln(L) 0 to 25 km	0.066***	0.027***	0.031***	0.020***	0.012***	0.010***
	(0.008)	(0.005)	(0.004)	(0.004)	(0.003)	(0.003)
ln(L) 25 to 50 km	0.007	-0.005	-0.006	-0.004	-0.002	-0.002
	(0.012)	(0.007)	(0.005)	(0.005)	(0.004)	(0.004)
ln(L) 50 to 75 km	0.012	0.000	0.008*	0.010**	0.001	0.003
	(0.010)	(0.006)	(0.005)	(0.004)	(0.004)	(0.004)
TTWA skill level				0.357***		0.321***
				(0.057)		(0.046)
Age controls	yes	yes	yes	yes	yes	yes
3 digit sector controls	yes	yes	yes	yes	yes	yes
Coefficients vary at 3 digit sector level	no	yes	yes	yes	yes	yes
obs	17968	17968	17968	17968	17968	17968

	business services					
	(1)	(2)	(3)	(4)	(5)	(6)
	Labour productivity log(VA/L)	RFTP Factor Shares	RFTP	RFTP	RFTP	RFTP
			Regression based		Control function	
ln(L) 0 to 25 km	0.169***	0.081***	0.101***	0.052***	0.072***	0.064***
	(0.013)	(0.011)	(0.009)	(0.008)	(0.008)	(0.008)
ln(L) 25 to 50 km	-0.046**	-0.024	- 0.041***	-0.025**	-0.013	-0.014
	(0.021)	(0.017)	(0.015)	(0.011)	(0.013)	(0.013)
ln(L) 50 to 75 km	0.039**	0.016	0.033***	0.032***	0.003	0.006
	(0.019)	(0.016)	(0.013)	(0.011)	(0.012)	(0.012)
TTWA skill level				0.370***		0.580***
				(0.130)		(0.140)
Age controls	yes	yes	yes	yes	yes	yes
3 digit sector controls	yes	yes	yes	yes	yes	yes
Coefficients vary at 3 digit sector level	no	yes	yes	yes	yes	yes
obs	8236	8236	8236	8236	8236	8236

Source: Authors' calculations based on ARD data.

Notes: The table shows regressions of various productivity measures on employment at various distance bands. All production function coefficients are allowed to vary at the 3 digit level.

Table A2: Regression results imposing α = 1

	manufacturing						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Labour productivity log(VA/L)	RFTP Factor Shares	RFTP	RTFP	RFTP	RFTP	RFTP
			Regression based			Control function	
agglomeration with α=1	0.056***	0.022***	0.026***	0.011***	0.053***	0.021***	0.019***
	(0.005)	(0.003)	(0.002)	(0.002)	(0.005)	(0.002)	(0.002)
TTWA skill level				0.199***			0.215***
				(0.039)			(0.072)
Controls for intermediate inputs	yes	yes	yes	yes	no	yes	yes
Age controls	yes	yes	yes	yes	yes	yes	yes
3 digit sector controls	yes	yes	yes	yes	yes	yes	yes
Coefficients vary at 3 digit sector level	no	yes	yes	yes	yes	yes	yes
obs	27835	27835	27835	27835	27835	27835	27835

	construction						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Labour productivity log(VA/L)	RFTP Factor Shares	RFTP	RTFP	RFTP	RFTP	RFTP
			Regression based			Control function	
agglomeration with α=1	0.075***	0.031***	0.052***	0.039***	0.084***	0.032***	0.032***
	(0.006)	(0.005)	(0.004)	(0.003)	(0.005)	(0.003)	(0.003)
TTWA skill level				0.109			0.105*
				(0.066)			(0.059)
Controls for intermediate inputs	yes	yes	yes	yes	no	yes	yes
Age controls	yes	yes	yes	yes	yes	yes	yes
3 digit sector controls	yes	yes	yes	yes	yes	yes	yes
Coefficients vary at 3 digit sector level	no	yes	yes	yes	yes	yes	yes
obs	12044	12044	12044	12044	12044	12044	12044

Source: Authors' calculations based on ARD data.
Notes: The table shows regressions of various productivity measures on an agglomeration measure with decay parameter α=1. All production function coefficients are allowed t vary at the 3 digit level.

	consumer services						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Labour productivity log(VA/L)	RFTP Factor Shares	RFTP	RTFP	RFTP	RFTP	RFTP
			Regression based			Control function	
agglomeration with α=1	0.094***	0.029***	0.038***	0.027***	0.094***	0.015***	0.012***
	(0.006)	(0.004)	(0.003)	(0.003)	(0.006)	(0.002)	(0.002)
TTWA skill level				0.328***			0.304***
Controls for intermediate inputs	yes	yes	yes	yes	no	yes	yes
Age controls	yes	yes	yes	yes	yes	yes	yes
3 digit sector controls	yes	yes	yes	yes	yes	yes	yes
Coefficients vary at 3 digit sector level	no	yes	yes	yes	yes	yes	yes
obs	17968	17968	17968	17968	17968	17968	17968

	business services						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Labour productivity log(VA/L)	RFTP Factor Shares	RFTP	RTFP	RFTP	RFTP	RFTP
			Regression based			Control function	
agglomeration with α=1	0.191***	0.089***	0.107***	0.059***	0.141***	0.084***	0.077***
	(0.010)	(0.008)	(0.008)	(0.006)	(0.009)	(0.006)	(0.007)
TTWA skill level				- 0.289***			
				(0.099)			
Controls for intermediate inputs	yes	yes	yes	yes	no	yes	yes
Age controls	yes	yes	yes	yes	yes	yes	yes
3 digit sector controls	yes	yes	yes	yes	yes	yes	yes
Coefficients vary at 3 digit sector level	no	yes	yes	yes	yes	yes	yes
obs	8236	8236	8236	8236	8236	8236	8236

Source: Authors' calculations based on ARD data.
Notes: The table shows regressions of various productivity measures on an agglomeration measure with decay parameter α=1. All production function coefficients are allowed to vary at the 3 digit level.

Table A3: Regressions results when estimating α

	manufacturing			
	(1)	(2)	(3)	(4)
	Labour productivity log(VA/L)	RFTP Factor Shares	RFTP Regression based	RFTP Control function
Marginal Impact of u	0.064***	0.026***	0.032***	0.024***
ρ	(0.006)	(0.003)	(0.003)	(0.002)
Decay of agglomeration effects	1.062***	1.077***	1.131***	1.122***
α	(0.140)	(0.178)	(0.140)	(0.127)
Age controls	yes	yes	yes	yes
3 digit sector controls	yes	yes	yes	yes
Coefficients vary at 3 digit sector level	no	yes	yes	yes
obs	27835	27835	27835	27835
	construction			
	(1)	(2)	(3)	(4)
	Labour productivity log(VA/L)	RFTP Factor Shares	RFTP Regression based	RFTP Control function
Marginal Impact of u	0.083***	0.033***	0.055***	0.034***
ρ	(0.006)	(0.005)	(0.004)	(0.003)
Decay of agglomeration effects	1.158***	1.516***	1.451***	1.562***
α	(0.131)	(0.242)	(0.123)	(0.159)
Age controls	yes	yes	yes	yes
3 digit sector controls	yes	yes	yes	yes
Coefficients vary at 3 digit sector level	no	yes	yes	yes
obs	12044	12044	12044	12044

Source: Authors' calculations based on ARD data.
Notes: The table shows non linear regressions of various productivity measures on an agglomeration measure with undetermined agglomeration effect decay. All production function coefficients are allowed to vary at the 3 digit level.

	consumer services			
	(1)	(2)	(3)	(4)
	Labour productivity log(VA/L)	RFTP Factor Shares	RFTP Regression based	RFTP Control function
Marginal Impact of u	0.103***	0.030***	0.043***	0.024***
ρ	(0.006)	(0.004)	(0.003)	(0.003)
Decay of agglomeration effects	1.195***	1.610***	1.337***	1.818***
α	(0.102)	(0.204)	(0.120)	(0.190)
Age controls	yes	yes	yes	yes
3 digit sector controls	yes	yes	yes	yes
Coefficients vary at 3 digit sector level	no	yes	yes	yes
obs	17968	17968	17968	17968

	business services			
	(1)	(2)	(3)	(4)
	Labour productivity log(VA/L)	RFTP Factor Shares	RFTP Regression based	RFTP Control function
Marginal Impact of u	0.196***	0.089***	0.113***	0.083***
ρ	(0.012)	(0.009)	(0.008)	(0.007)
Decay of agglomeration effects	1.477***	1.560***	1.599***	1.746***
α	(0.097)	(0.173)	(0.119)	(0.144)
Age controls	yes	yes	yes	yes
3 digit sector controls	yes	yes	yes	yes
Coefficients vary at 3 digit sector level	no	yes	yes	yes
obs	8236	8236	8236	8236

Source: Authors' calculations based on ARD data.
Notes: The table shows non linear regressions of various productivity measures on an agglomeration measure with undetermined agglomeration effect decay. All production function coefficients are allowed to vary at the 3 digit level.