Journal of Economic Geography **12** (2012) pp. 1167–1189
Advance Access Published on 4 November 2011 doi:10.1093/jeg/lbr031

Agglomeration economies and regional intangible assets: an empirical investigation

*Michael J. Artis**, *Ernest Miguelez***,† *and Rosina Moreno****

*Department of Economics, University of Swansea & CEPR, Singleton Park, SA2 8PP, Swansea, UK
**AQR-IREA, Department of Econometrics, Statistics and Spanish Economy, University of Barcelona, Av. Diagonal 690, 08034 Barcelona, Spain
***AQR-IREA, Department of Econometrics, Statistics and Spanish Economy, University of Barcelona, Av. Diagonal 690, 08034 Barcelona, Spain.
†Corresponding author: Mr. Ernest Miguelez, Department of Econometrics, Statistics and Spanish Economy, University of Barcelona, Av. Diagonal 690, 08034 Barcelona, Spain. *email* <emiguelez@ub.edu>

Abstract

This article examines the impact of agglomeration economies on economic outcomes and is specifically concerned with analysing the effect on standard regressions of including variables that measure intangible assets that might boost productivity as well. Thus, in assessing the relationship between agglomeration and productivity, we consider endogeneity issues arising from the omission of these variables, together with problems of simultaneity and missing cross-regional covariates. We examine each of these features and consider the extent to which they represent a source of bias in the agglomeration elasticity if they are not controlled for. While our results indicate that agglomeration has a marked, positive influence on productivity, these estimates are reduced dramatically when spatial dependence and other, hitherto omitted, variables proxying intangible assets are controlled for.

Keywords: agglomeration economies, intangible assets, endogeneity, interregional externalities
JEL classifications: C21, J24, R10, R11, R12
Date submitted: 28 August 2009 **Date accepted:** 15 September 2011

1. Introduction

The geographic concentration of firms has been the subject of much analysis in economics. As early as 1890, Marshall forwarded the idea that by locating close to other firms a firm can benefit from external economies, taking advantage of the division of labour and input and information exchange. The impact of these agglomeration economies on productivity has come under examination in recent decades (Henderson, 1986; Rauch, 1993; Ciccone, 2002, to mention just a few), in the belief that the density of economic activity is a source of enhanced productivity gains due to the effect of spatial externalities, which are generated by the Marshallian forces and which result in increasing returns.

A key question posed in the literature examining the impact of density on productivity is concerned with the fact that local endowments, in their broadest sense, do differ across regions, which affects productivity, independently or otherwise of the presence of agglomeration economies. As Combes (2011) indicates one would like to control for private and public capital local endowments, physical geography and for the quality of local institutions and the level of available technology. In addition to these theoretical concerns, it has to be acknowledged that econometric estimates may well suffer from missing variables in the specifications that seek to measure the effect of agglomeration economies on productivity and which fail to include these local endowments. If such variables are omitted from the regression, the impact of density is overestimated since it captures effects that do not directly reflect the impact of agglomeration.

This article, in examining the impact of agglomeration economies on economic outcomes, is specifically concerned with analysing the effect of including variables that measure intangible assets that might boost productivity in the standard agglomeration economies' regressions. We adhere to the hypothesis that the mere location of individuals and firms within a given geographic location is not the only source of aggregate increasing returns. Rather, it is our belief that the qualitative characteristics of each region are also important in explaining economic outcomes. Hence, starting from Ciccone's (2002) basic model, we include various modifications so as to control for a wider range of private returns as well as to allow for a broader variety of social returns or externalities within the region. Failing to control for the spatial distribution of these qualitative, intangible features would mean that some regions might appear to be more productive than they really are—even if no externalities are directly derived from the agglomeration of economic activity, but simply because their inhabitants are endowed with these specific qualitative characteristics. We recognize that, to some extent, the externalities attributable to these features might be a partial manifestation of agglomeration effects on productivity, and may, or may not, be correlated. However, the different proxies suggested might have a separate, significant effect that needs to be isolated. Thus, even if agglomeration remains important, controlling for these variables could make a significant difference for the policy maker who seeks to improve the productivity of firms and workers in a given region.

Apart from endogeneity problems due to the omission of relevant variables related to intangible assets, additional concerns with respect to other sources of endogeneity, such as reverse causality, should be considered. It might be the case that the concentration of employees leads to better economic outcomes or, on the contrary, that better economic outcomes attract more workers to live in a given region due to higher wages. In other words, reverse causality arises when people and firms choose their location according to their returns, directly linked to local productivity. In this case, productivity shocks unobserved by the econometrician but observed by agents are correlated to density due to endogenous location choices. If the latter occurs, estimation by OLS will yield inconsistent estimates. To deal with this problem, we conduct our estimation using GMM.

Moreover, the presence of externalities across regions might well bring about additional endogeneity problems, which must be accounted for. In this article, we carefully consider the potential presence of such interregional externalities and, if identified, various empirical approaches are adopted. First, the interregional effects of intangible assets are investigated as a source of spatial dependence across regions by including the average of these assets in the nearest regions. Second, spatial econometric techniques are used if the first approach does not succeed in eliminating spatial dependence problems.

In short, the main goal of our study is to bring together endogeneity issues attributable to the missing relevant variables related to local endowments of qualitative features, the simultaneity of the effect of agglomeration, cross-regional externalities, and, by extension, the endogeneity of those additional missing variables related to intangible endowments. We, therefore, explore stage by stage the extent to which these features represent a source of bias in the agglomeration elasticity if they are not controlled.

A further novelty presented by our study is the measurement of regional productivity figures. Hence, a new data set is used, namely, gross value added (GVA) per job filled (Wosnitza and Walker, 2008). It has the advantage of avoiding a number of the measurement errors that have afflicted other productivity data sets. Additionally, as a proxy for the agglomeration of economic activity, our study uses a concept developed by Rice et al. (2006), that of 'economic mass'. The NUTS3[1] regions of Great Britain are used in our empirical investigation.

Our results do suggest that agglomeration economies are significant in determining productivity, although our estimates of their size are somewhat reduced when the intangible asset endowments, which characterize the knowledge-based economy are introduced, and they are dramatically diminished when across-region externalities are controlled for.

The article is organized as follows: Section 2 reviews the recent empirical literature on agglomeration economies; Section 3 presents our model and discusses some data issues; Section 4 outlines the main results, while Section 5 concludes.

2. Background

The effects of agglomeration economies have been widely studied in the relevant literature. The earliest empirical contributions can be traced back to Sveikauskas (1975), Segal (1976), Moomaw (1981), Henderson (1986), Sveikauskas et al. (1988), Beeson and Husted (1989), and Rauch (1993), among others. Interesting contributions have also been made by Ciccone and Hall (1996) and Ciccone (2002) for, among other things, their focus on endogeneity and reverse causality issues. Recently, in Combes (2011) and Combes et al. (2011), the authors undertake a general discussion as to whether the relationship between urban productivity and city population is causal, examining the main sources of bias in the proper identification of agglomeration effects. The contributions by Ciccone and Hall (1996) and Ciccone (2002) are discussed in detail here as they represent our point of departure for our empirical assessment of agglomeration effects.

Broadly speaking, much of this literature highlights the idea that the density of economic activity is a source of enhanced productivity gains due to the effect of spatial externalities leading to increasing returns within regions. Three main sources have been forwarded to explain why improved aggregate economic results may be derived from the agglomeration of economic activity—already raised by Marshall (1890). First, easier access to suppliers and customers, in the presence of transportation costs that rise with distance. Second, the presence of thicker and larger input markets—that is markets

1 NUTS corresponds to the French acronym for "*nomenclature d'unités territoriales statistiques*" and refers to administrative divisions within Europe devised for statistical purposes.

which are more efficient in terms of market matching and provide a large and diverse supply of certain inputs (Rosenthal and Strange, 2004), which could be characterized by strong-scale economies in input production. Finally, the concentration of economic activity results in more intensive and more frequent knowledge spillovers, given that firms can learn from each other. More recently, additional sources have been forwarded, including natural advantages, home–market effects (Hanson, 2005), consumption opportunities (Glaeser et al., 2001) and rent seeking (Ades and Glaeser, 1995).

However, Duranton and Puga (2004) note that these three main sources cannot constitute the basis for understanding the theoretical mechanisms underlying agglomerations. Thus, they distinguish three types of micro-founded mechanisms, namely, *sharing*, *matching* and *learning*. They identify *sharing* mechanisms which 'deal with sharing indivisible facilities, sharing the gains from the wider variety of input suppliers that can be sustained by a larger final goods industry, sharing the gains from the narrower specialization that can be sustained with larger production, and sharing risks' (Duranton and Puga, 2004, pp. 2–3). Agglomeration also enhances the opportunities for *matching* between actors in a given space. *Learning* refers to the mechanisms by which knowledge is generated and diffused. However, as they stress, the results observed when analysing the effects of agglomeration economies are the same whatever the operating mechanisms.

According to Ciccone and Hall (1996), density is crucial for explaining the variation in productivity since they report that the doubling of employment density leads to a 6% increase in average labour productivity. Ciccone (2002) expanded the scope of his previous study by estimating the agglomeration effects for the NUTS3 regions of France, Germany, Italy, Spain and the UK with a model in which the concentration of production is the main source of agglomeration economies. This study points to the presence of substantial agglomeration effects in Europe, with estimated elasticities of $\sim$4.5%, which do not differ significantly across countries. In fact, the empirical literature concerned with the effect of agglomeration economies on economic performance has grown enormously in recent years, and a number of useful surveys (Rosenthal and Strange, 2004; Duranton, 2008) can be consulted. In broad terms, the majority of studies obtain elasticities between 0.01 and 0.10, using different proxies for agglomeration and for economic outputs, and at both the aggregate and plant level. Thus, the doubling of a city's or region's size leads to an increase in productivity of between 1% and 10% (Graham, 2007).

Interesting recent studies include, among others, those by Cingano and Schivardi (2004) and Combes et al. (2008), who stress the importance of human capital—the latter focusing their attention on the endogenous nature of human capital. Panel data techniques and dynamics are suggested in Blien et al. (2006), Brülhart and Mathys (2008) and Brülhart and Sbergami (2009). The role of diseconomies when dealing with agglomeration effects on economic outcomes are highlighted in Graham (2007) and Brülhart and Sbergami (2009), with the former identifying large differences in the estimated agglomeration effect according to the economic sector under analysis—from elasticities $\sim$0.04 for manufacturing sectors to values of 0.18 for certain service sectors. Finally, Baptista (2003), Fingleton (2003) and Rice et al. (2006) are interesting references for the British case.

The present inquiry builds on these recent studies and extends the literature by bringing together several endogeneity issues that may well bias the agglomeration effect if not taken into proper consideration.

Despite the recent trend in the literature toward the use of microeconomic data[2] for measuring the agglomeration effect, our aim here is to focus on aggregated data. We do so because we explicitly seek to assess how the local intangible assets of a region in which a given firm is located—which, to some extent, are 'in the air' and, as such, are difficult to assign to specific individual economic agents—determine productivity. Thus, from an input perspective, we believe that the list of intangible assets we are taking on board can be better seen as a property of the group rather than a property of the individual. If positive spillovers or strategic complementarities exist between individuals that possess these intangibles, the estimated aggregate elasticity will incorporate both the individual level response and the effects stemming from social interactions (Glaeser et al., 2010). Second of all, from an output perspective, we find our GVA-aggregated measure more appealing to appraise the general level of regional well-being, rather than wages or salaries since we believe that we should also account for profits, for instance. Finally, from a methodological viewpoint, our aim is to bring together in the same empirical framework several endogeneity issues. On the one hand, due to omission of relevant variables related to interactions across different regions, and on the other hand, endogeneity issues due to system feedbacks of some r.h.s. variables. Surprisingly, the analysis of other endogenous variables in the r.h.s. of the equation within the spatial literature has been rather neglected so far (Fingleton and Le Gallo, 2008b). For all these reasons, we still find the aggregate analysis valuable.

As stressed in the introductory section, a number of hitherto missing intangible assets are considered here. First, the literature has widely stressed the role played by human skills in determining regional economic outcomes (Ciccone and Cingano, 2003; Moretti, 2004; Combes et al., 2008). This hypothesis relies, first, on the assumption that, despite equal technologies across regions, there exist differences between areas in terms of the ability of individuals to make that technology productive (Fingleton, 2003); and, second, that human capital spillovers increase aggregate productivity beyond the effect of this capital on an individual's productivity (Moretti, 2004).[3] However, human capital can be acquired both from the educational system and while working. Therefore, the occupational composition of the region is also important (Ciccone and Cingano, 2003) when analysing the regional impact of human capital.

Second, the differential access enjoyed by each region to knowledge can explain productivity differentials across regions as well (Fingleton, 2003). In fact, the access to knowledge capital is rooted in the so-called theories of endogenous economic growth. We hypothesize that private returns to knowledge and knowledge externalities arise both from knowledge inputs—that is, R&D efforts and the number of employees working in high-technology industrial sectors—and from knowledge outputs such as patents.

Third, as Audretsch (2003) and Rosenthal and Strange (2004) suggest, the entrepreneurial or business culture of a region can also boost economic performance. Indeed, in HM Treasury (2001), we find that entrepreneurial activity is regarded as a key driver of productivity growth in the economy. The creation and enlargement of firms is associated with the introduction of new technologies, innovative production

2 A non-exhaustive list of such studies includes, among others, Glaeser and Maré (2001), Wheaton and Lewis (2001), Combes et al. (2008) and Baldwin et al. (2010).

3 See Moretti (2004) for a detailed review of theories and empirical studies on human capital and human capital externalities.

processes and increased competitive pressure on the other firms in a given market, providing them with strong incentives to further innovate and adopt new technologies (Glaeser et al., 1992). Thus, it is important to consider not only the amount of new entrepreneurial projects set up in a given region, but also the overall growth of firms during the whole period, in order to take into account not only the business culture of the region, but also its success.

Following the literature surveyed above, here we limit our intangible endowments to those of human capital, knowledge and entrepreneurial culture.[4] The logic behind our approach is that by not properly controlling for the uneven spatial distribution of these endowments, some regions might appear to be more productive even in the absence of a local agglomeration effects, simply because of the fact that they host more skilled people, knowledge and entrepreneurship (Combes, 2011).

Local endowments, in this broad sense, do differ across regions, and this affects productivity independently, or otherwise, of the presence of agglomeration economies (op. cit.). This separate effect must be isolated if we aim at drawing the correct policy recommendations. Indeed, if these sources of productivity are not controlled for, the estimation of the agglomeration effect could be biased upward.

3. Methodology and some data issues

3.1. The model

We start from the approach described in Ciccone (2002):

$$y = Q_s f(l \cdot H, k;\ Y_s, A_s) = Q_s((l \cdot H)^{\beta} k^{1-\beta})^{\alpha} \left(\frac{Y_s}{A_s}\right)^{(\lambda-1)/\lambda}, \tag{1}$$

where y stands for output per hectare, l for the number of workers per hectare, H for the average level of human capital, k for the amount of physical capital per hectare, all in region s; Q_s is the region's total factor productivity (TFP) index; and Y_s and A_s denote total production and total number of hectares in the region respectively; α captures returns to capital and labour per hectare, β is a distribution parameter and $(\lambda - 1)/\lambda$ is the parameter which captures spatial externalities arising from the concentration of economic activity—in this case, density of production (Y_s/A_s). Here, we introduce a number of modifications for empirical testing. Basically, we consider that this specification fails to represent a wide variety of individual and social returns that might also generate economic outcomes, thereby resulting in an omitted variables problem (Bode, 2004). Thus, our hypothesis is that the size of the estimated elasticity of the concentration of economic activity as a predictor of spatial productivity differentials may not be robust to the consideration of local intangible endowments. Given these arguments, we propose a model in which other sources of private and social returns of the regions are considered, i.e. looking beyond the concentration of economic activity and the externalities that this entails.

4 We are concerned about the omission of other kinds of intangible asset, such as relational capital, social capital, territorial capital, cognitive capital, intellectual capital and the like. We assume, however, that the three types of intangible assets we introduce are taking into account to a certain extent the possible effects of these unidentified intangible assets on productivity.

Thus, our set of intangible assets (discussed in the previous section) enters the production function and directly affects the TFP index of each region—Q_s, in order to capture a greater variety of private returns and social externalities (see Bode, 2004, for a similar approach). These considerations lead us to new TFP measure such as[5]

$$Q_s = Q_s(Q, H_s, O_s, RD_s, MAN_s, PAT_s, E_s, S_s), \tag{2}$$

where Q are the determinants of TFP which do not differ at the NUTS3 level. H_s and O_s are educational and occupational human capital indicators, respectively, RD_s is an indicator of knowledge efforts, MAN_s is an indicator of high-tech manufacturing knowledge and PAT_s is an indicator of knowledge outputs; E_s is an entrepreneurship capital indicator and S_s is an entrepreneurship success indicator, all of them within the region s (see Appendix for a description of the variables). Thus, returning to Equation (1), the final model would be

$$y = Q_s(\cdot) \cdot (l^{\beta} k^{1-\beta})^{\alpha} \left(\frac{Y_s}{A_s}\right)^{(\lambda-1)/\lambda}, \tag{3}$$

where $Q_s(\cdot)$ is the total factor productivity index affected by a wider range of private and social returns as in Equation (2). In order to be able to estimate this function, we can transform it into an aggregate regional production function of the form

$$Y_s = y \cdot A_s = A_s Q_s(\cdot) \left(\left(\frac{L_s}{A_s}\right)^{\beta} \left(\frac{K_s}{A_s}\right)^{1-\beta} \right)^{\alpha} \left(\frac{Y_s}{A_s}\right)^{(\lambda-1)/\lambda}, \tag{4}$$

where output, labour and capital (Y_s, L_s, K_s) correspond to their levels in each region rather than in each hectare. Rearranging and solving for labour productivity, yields

$$\frac{Y_s}{L_s} = \left(\frac{L_s}{A_s}\right)^{\alpha\lambda-1} Q_s(\cdot)^{\lambda} \left(\left(\frac{K_s}{L_s}\right)^{1-\beta} \right)^{\alpha\lambda}, \tag{5}$$

At low levels of regional disaggregation, data on the quantity of physical capital do not exist. To cope with this weakness, we follow Ciccone (2002) and assume that the rental price of capital is the same within every NUTS1 region. Hence, from Equation (1), we are able to derived the capital–demand function, $K_s = (\alpha(1-\beta)/c)Y_s$, where c is the rental price of capital in each larger region. Thus,

$$\frac{Y_s}{L_s} = \left(\frac{L_s}{A_s}\right)^{\alpha\lambda-1} Q_s(\cdot)^{\lambda} \left(\left(\frac{\alpha(1-\beta)}{c} Y_s\right)^{1-\beta} L_s^{-(1-\beta)} \right)^{\alpha\lambda}, \tag{6}$$

$$\left(\frac{Y_s}{L_s}\right) = \left(\frac{L_s}{A_s}\right)^{\theta} \Omega_s Q_s(\cdot)^{\omega}, \tag{7}$$

where $\theta_i = \frac{\alpha\lambda-1}{1-\alpha\lambda(1-\beta)}$ and measures the net effect of regional employment density on regional productivity; $\Omega_s = (\alpha(1-\beta)/c)^{1/1-(1-\beta)\alpha\lambda}$ and is a constant which only depends on the rental price of capital in a larger region and $\omega = \lambda/1 - \alpha\lambda(1-\beta)$. Taking logs,

5 Note that in (2), private and social returns to intangible endowments are observationally equivalent.

and assuming that the productivity term, $Q_s(\cdot)$, enters in a logarithmic form, yields

$$\begin{aligned}\log\left(\frac{Y_s}{L_s}\right) = {} & \log\Omega + \theta_1(\log \text{Agglomeration}_{s1}) + \theta_2(\log \text{Agglomeration}_{s2}) \\ & + \phi_0 \log Q + \phi_1 \log H_s + \phi_2 \log O_s + \phi_3 \log \text{RD}_s + \phi_4 \log \text{MAN}_s \\ & + \phi_5 \log PAT_s + \phi_6 \log E_s + \phi_7 \log S_s + \varepsilon_s\end{aligned} \tag{8}$$

where ε_s is a random error term. In Equation (8), we allow the model to include two measures of agglomeration among its covariates so that we might go some way to exploring the spatial scope of this effect—for a description of the variables used see the next section. Regional dummies are also included to capture differences in exogenous TFP not explained in the model ($\phi_0 \log Q$)—which are assumed to be marginal, and, in particular, $\log\Omega$, because differences in physical capital or its rental price can be captured by allowing for spatial fixed effects for larger regions (Ciccone, 2002). Thus, a dummy for larger regions (NUTS1) replaces $\phi_0 \log Q + \log\Omega$. Next, $\phi_j = \omega \cdot \delta_j$, and δ_j are the elasticities of TFP with respect to its determinants, where $j = 1, ..., 7$ for the coefficients of the seven indicators of intangible assets.

3.2. Data

Our empirical model is estimated using data for 119 British NUTS3 regions. Productivity is defined as GVA per filled job for the period 2001–2005. In spite of the appropriateness of exploiting the temporal dimension of the data through panel data techniques, no proper deflator at the NUTS3 level for GVA per job filled is available, so productivity figures can not be expressed in real terms in our case. Thus, we are unable to isolate volume changes from price changes and so averages of the 5 years' productivity figures are used instead. These averages are also used for the explanatory variables. Studies in the literature have typically used either wages and earnings, or GVA per head or employee to proxy regional productivity. However, productivity measures should ideally include more than just wages or salaries, and allow, for instance, for profits also. Wosnitza and Walker (2008), following the OECD methodology, decompose GVA per head in British regions into four elements, namely, productivity (actually GVA per job filled, which is calculated on a workplace basis rather than on a residence basis), employment rate, commuting rate and activity rate. Taking GVA per job on a workplace basis as our measure of productivity allows us to avoid some of the potential distortions of GVA per head, particularly in cities that receive a significant number of commuters, or have low economic activity rates.[6] We also believe this measure to be superior to work-place-based GVA per employee figures, because a person, for instance, may be performing more than one job and consequently biasing these productivity figures upwards (i.e. there are more jobs than employees).

6 Variables such as GVA or GDP, for instance, are usually estimated at workplaces while people are counted where they are born, so GVA per capita tends to be overestimated if the region excludes dormitory areas (Cheshire and Magrini, 2009). This is what makes this dataset so valuable.

To proxy for the concentration of economic activity, we use the concept of 'economic mass', coined by Rice et al. (2006). This measure is based on the total employment of a given area located within a series of driving time bands around the centre of each NUTS3 area.[7] Thus, we do not understand agglomeration as population per hectare within a given administrative region, but rather as employment within a band, or isochrone, of a certain number of minutes' driving time by car. According to the authors, this measure is an economically more meaningful proxy for agglomeration than the more traditional measure of employment density in the own or neighbouring regions. British NUTS3 areas are sufficiently small,[8] with boundaries determined administratively rather than economically, and so travel time bands capture the effective potential employment (or jobs filled in our case) available for each area much better. Further, by including more than one travel time band, we can capture not only own area effects, but also cross-region effects, so we are able to assess the scope of the agglomeration effect as well.[9]

Intangible assets are not easily defined and measured, due primarily to a lack of consensus as to what they exactly constitute. What is more, they tend to be a multidimensional concept, which we seek to reflect in our proxies. Information regarding the construction of each variable and the data sources drawn upon are given in the Appendix Table A1. In principle, we assume these variables to be exogenous since they predate our period of analysis (2001–2005); these data, in fact, pertain to the period 1996–2000. However, the processes determining spatial variations in productivity presumably evolve slowly, perhaps exhibiting strong time trends and, as such, these variables might themselves be determined in part by productivity. Therefore, predating variables would not be sufficient in order to lessen endogeneity. We address this endogeneity issue in subsection 4.2.

Table 1 sets out the variables used in this study and provides information as to how they vary across the UK regions. Across-region differences are marked as illustrated by our dependent variable, which ranges from £22,761 per job filled in the Scottish Borders region to £46,594 in Inner London, West. Differences between the regions are also high for the explanatory variables, especially for the concentration of employment, applied patents and employment in R&D.

Table 2 shows the correlation matrix. Other than the correlation of 0.7 between occupational human capital and VAT registrations, the correlation among the independent variables is, in general, sufficiently small and collinearity does not pose any problem in our estimation.

7 Data on travel times (and distances as well) were calculated using Microsoft Autoroute 2002. We are very grateful to Patricia Rice and Anthony Venables for providing us with these data. To adapt our data to the travel time data provided by Rice and Venables, the regions of Eilean Siar (Western Isles), Orkney Islands, and Shetland Islands have been excluded. Moreover, the following areas have been aggregated: East Cumbria and West Cumbria; South and West Derbyshire and East Derbyshire; North Nottinghamshire and South Nottinghamshire; Isle of Anglesey and Gwynedd; Caithness, Sutherland and Ross and Cromarty, Inverness and Nairn and Moray, Badenoch and Strathspey, Lochaber, Sky, Lochalsh and Argyll and the Islands.

8 To give an idea of the size distribution of the regions, note that the average area is 190,514 sq km and the average population is 1,251,878 inhabitants.

9 As Rice et al. (2006) mention, the ideal situation would be to include several time bands of no more than 20 min each, although this would introduce serious collinearity problems in the estimation. In our study, therefore, we include two travel time bands of 60 min each. As such two parameters, θ_{0-60} and θ_{60-120}, are estimated in our regressions.

Table 1. Summary statistics

	Observations	Mean	Coefficient of variation	Min	Max
GVA per filled job	119	29,785	0.136	22,761	46,594
Employment 60 min	119	1,251,878	0.965	51,342	6,120,282
Employment 60–120 min	119	4,827,812	0.704	0	1.26e + 07
Educational human capital	119	0.96	0.162	0.66	1.48
Occupational human capital	119	24.24	0.184	11.53	39.63
Employment in RD and computers	119	0.79	0.846	0.2	4.3
High-tech manufacturing employment	119	1.17	0.501	0.08	2.84
Applied patents	119	407	1.107	25	3247
VAT registrations	119	2.73	0.430	1.23	12.37
Cumulative Annual Growth Rate (CAGR) VAT registrations	119	1.64	0.623	−0.34	4.92

Notes: Summary statistics are calculated using the raw variables before any logarithmic transformation.

Table 2. Correlation matrix

	1	2	3	4	5	6	7	8	9
1. ln(employment 0–60 min)	1								
2. ln(employment 60–120 min)	0.41	1							
3. Educational HK	0.19	−0.05	1						
4. Occupational HK	0.24	0.41	0.41	1					
5. Empl. RD&IT	0.32	0.32	0.48	0.57	1				
6. High-tech employment	0.10	0.15	−0.32	−0.20	−0.06	1			
7. ln(applied patents)	0.19	0.26	0.47	0.52	0.57	−0.07	1		
8. ln(VAT registrations)	0.13	0.22	0.51	0.71	0.55	−0.36	0.45	1	
9. CAGR VAT registrations	0.42	0.42	0.35	0.49	0.56	0.03	0.36	0.55	1

Notes: The correlation matrix has been calculated having first log-transformed the variables indicated.

4. Results

4.1. Omitted variables: the role of IA

Table 3 shows the OLS estimates of model (8). We report, initially [Column (i)], the estimates of the effect of agglomeration on productivity, using only the educational human capital location quotient as a control, as is the case in much of the literature reviewed in Section 2. In the next column, we show the effects of including the additional variables suggested by the model discussed in Section 3 [Column (ii)]—the intangible assets.

In line with Ciccone (2002), we assume that the capital-income share, $\alpha(1-\beta)$, is 0.3, while the income share of land, $(1-\alpha)$, is 0.015. The agglomeration parameter within the first 60-min travel time band, θ_{0-60}, is, according to our estimates of the restricted model, 0.059. To obtain an approximation of the elasticity of production density on

Table 3. White-robust OLS and GMM estimates

	OLS		GMM		
	(i)	(ii)	(iii)	(iv)	(v)
ln(employment within 0–60 min)	0.059***	0.042***	0.056***	0.039***	0.039***
	(0.008)	*(0.008)*	*(0.010)*	*(0.008)*	*(0.009)*
ln(employment within 60–120 min)	0.015***	0.009***	0.017***	0.010***	0.009***
	(0.004)	*(0.003)*	*(0.004)*	*(0.002)*	*(0.003)*
Educational HK	0.333***	0.167**	0.334***	0.166**	0.236**
	(0.065)	*(0.080)*	*(0.063)*	*(0.073)*	*(0.095)*
Occupational HK		−0.002		−0.001	0.002
		(0.003)		*(0.003)*	*(0.002)*
Employment RD&IT		0.048***		0.050***	0.050***
		(0.014)		*(0.013)*	*(0.016)*
High-tech manufacturing employment		0.056***		0.056***	0.066***
		(0.013)		*(0.012)*	*(0.014)*
ln(applied patents by inventor)		0.015		0.013	0.007
		(0.011)		*(0.010)*	*(0.013)*
ln(VAT registrations)		0.079*		0.078**	0.041
		(0.044)		*(0.040)*	*(0.068)*
CAGR VAT registrations		0.020*		0.021**	0.016
		(0.011)		*(0.010)*	*(0.014)*
Constant	8.950***	9.203***	8.965***	9.231***	9.175***
	(0.121)	*(0.117)*	*(0.115)*	*(0.108)*	*(0.119)*
NUTS1 dummies	Yes	Yes	Yes	Yes	Yes
Sample size	119	119	119	119	119
Adj. R^2	0.616	0.739	0.615	0.748	0.728
Joint test for intangibles [F-test$_{7,\ 99}$ and Wald test $\chi^2_{(7)}$]		14.61		121.18	128.30
P-value		(*0.000*)		(*0.000*)	*0.000*
Moran's *I*	3.801	3.550			
P-value	(*0.000*)	(*0.000*)			
Hansen *J* statistic			0.803	0.858	1.097
P-value			*0.669*	*0.651*	*0.578*
ln(Empl. 60 mn)—*Partial* R^2			*0.778*	*0.751*	
ln(Empl. 60 mn)—*Shea* R^2			*0.734*	*0.732*	
ln(Empl. 60 mn) —*First stage F-stat*			*53.43*	*49.13*	
ln(Empl. 60–120 mn)—*Partial* R^2			*0.973*	*0.968*	
ln(Empl. 60–120 mn)—*Shea* R^2			*0.917*	*0.944*	
ln(Empl. 60–120 mn)—*First stage F-stat*			*1804.41*	*1402.15*	

Notes: OLS and GMM estimates with several levels of significance: ***1%, **5%, *10%. In the case of the parameters, White-robust standard errors are presented in italics and parentheses below each associated parameter. Moran's *I* test for the residuals of the OLS estimations is provided, rejecting the null and, therefore, indicating that they remain spatially autocorrelated. Each test presents its *P*-value in italics and parentheses below. The variables expressed in percentages and location quotients are not log-transformed in order to facilitate the interpretation of their coefficient. Hansen *J* statistics for mutual consistency of the available instruments are provided [Columns (iii) to (v)] and we cannot reject the null hypothesis that the excluded instruments are valid and uncorrelated with the error term, so there are no overidentification problems. Instrument validities in Column (v) are not reported for reasons of space, although they are available upon request from the authors.

Dep. Var.: lnGVA per job filled.

total output, we use the fact that $\lambda - 1/\lambda = 1 - \alpha + \alpha(1 - \beta)\theta_i/1 + \theta_i$, so the estimated parameter implies results for the coefficient that captures spatial externalities in Ciccone's (2002) model of 5.3% for our sample.

When the extended model is estimated [Column (ii)], the adjusted R^2 increases by 0.12, so that the specification explains a larger proportion of variance than the restricted specification. Moreover, the implied elasticity of the density of production is 4.07%, about 77% of that in Column (i). For the case of the second travel time band, 60–120 min, the parameter is also dramatically reduced.

Interestingly, the majority of the variables included in our model are significant and present the expected sign. However, the occupational human capital indicator does not have a significant impact on productivity, which we attribute to social and institutional factors, and to labour market segmentations within high-performing regions, since people in those regions may demand that low-productivity services be located inside. Knowledge outputs, i.e. applied patents according to their inventor's region of residence, are not significant either.[10] We chose to re-run the same regression as in Column (ii), but now eliminating these two non-significant intangible assets, but the results (available from the authors upon request) remained unchanged. Likewise, an *F*-test for the joint significance of the parameters accompanying the intangible proxies clearly rejects the null hypothesis.

In short, although the estimated agglomeration effect, θ, and the implied production density parameter are somewhat smaller when intangible assets are included in the model, agglomeration economies still matter, although their impact—in quantitative terms—and their scope—in distance terms—are estimated to be lower and shorter, respectively.

Several additional concerns should be borne in mind at this juncture. On the one hand, the concentration of economic activity and employment could suffer from reverse causality with productivity, since workers could tend to concentrate where economic outcomes and consequently wages are higher. On the other hand, other sources of externalities apart from those related to the concentration of employment may arise not only within a given region, but also across neighbouring regions. In the subsections that follow, we take each of these considerations into account.

4.2. Agglomeration and productivity: a simultaneous relationship

A principal concern when assessing the robustness of the relationship between the concentration of economic activity and productivity is that of a possible 'two-way causation', i.e. are cities highly productive because they are big and dense or, rather, are cities big and dense because they are highly productive? To cope with this concern, we use GMM estimation techniques. In doing so, we use two instruments, in order that we might perform overidentification tests as well. Thus, in line with Rice et al. (2006), as our first instrument, we use the population in 1801 in the regions whose respective

10 Earlier versions of this study included interactions between educational human capital and the three dimensions of knowledge capital among the covariates. However, they were eliminated in the final draft for reasons of space (results are available from the authors upon request). When the total elasticities evaluated at the sample mean were calculated, together with the standard errors using the delta method (Serfling, 1980), we encountered a strong complementarity relationship between educational human capital and applied patents. This last variable not only increased its value considerably, but it also became strongly significant.

centres lie within two travel time bands. As the authors noted, the validity of this instrument lies in the assumption that the patterns that determined settlement at the beginning of the 19th century are not correlated with current levels of productivity, apart from its influence through current population and employment concentration. Further, following Ciccone's (2002) suggestions, as our second instrument, we use the total land area of the regions whose respective centres are located within each of our two isochrones. As Ciccone is at pains to stress, current administrative boundaries were often drawn up in order to balance out the population size of each region, so that the area can be used as an instrument if the original sources of population concentration (adhering to mainly geographical explanations) affect productivity solely through agglomeration.

In Columns (iii) and (iv) of Table 3, we repeat the estimations of Columns (i) and (ii) respectively, but instrumenting our main explanatory variables—i.e. employment within each isochrone—using the aforementioned instruments. As shown in the bottom panel of Table 3, the first stage F-statistics for the joint significance of the instruments are larger than 10, which is usually considered a good threshold whereby the instruments cannot be judged as weak, while the partial R^2 of the first regression is high. Moreover, Shea's partial R^2 test results (which takes account of the collinearity between instruments—see Shea, 1997) are also shown since in models with multiple endogenous variables the first-stage F-statistic and the usual partial R^2 statistic of the first stage are not sufficiently informative. Should the partial R^2 statistics present large values and Shea's R^2 statistics be small, the instruments would lack sufficient relevance to explain all the endogenous regressors (Baum et al., 2003). As can be seen, the differences between the two measures are almost negligible.

The results of these estimations, and the conclusions that can be drawn from them, are similar to those above: we record a reduction (both in quantitative and distance terms) in the agglomeration effect when controlling for intangible assets, while these assets are important in fostering productivity—both jointly and individually. It is worth noting that the estimated coefficient of the agglomeration effect is somewhat lower when instrumented, suggesting that the parameter was somewhat upwardly biased in the OLS estimation and that the GMM estimation was necessary.

Next, the variables referring to the intangible assets predate our period of analysis (2001–2005). These data in fact pertain to the period 1996–2000, and, as such, we believe that any endogeneity biases due to simultaneity will be considerably diminished. However, we are aware that the processes determining spatial variations in productivity could evolve slowly, possibly exhibiting strong time trends, and, as such, the variables for the intangible assets could themselves be determined in part by productivity. Given the time-persistent feature of the productivity measure, it is worthwhile ensuring that no endogeneity problems remain. To do this, we use the three-group method, as described in Fingleton (2003). Although it was believed to cope with measurement errors (Kennedy, 1992), we assume that by instrumenting these previously lagged variables, any endogeneity problems will be solved. The three-group method involves sorting the variables and dividing them into three groups of equal size, taking the value 1 if the observation is in the highest third of the variable, 0 if it is in the middle and -1 if the value is in the lowest third of the regressor. Column (v) of Table 3 repeats the GMM estimations, but now all the covariates are instrumented. It is worth noting that few differences were found, other than an increase in the estimated parameter for occupational human capital—although not enough to make it significant.

Additionally, proxies for entrepreneurship capital were no longer significant. Additionally, tests for the joint significance of the intangibles reject the null. Instrument validity measures—not reported here—including partial R^2 and F-tests of the first stage are both quite high, although, contrary to what is shown in the columns of Table 3 above, the differences between partial R^2 and Shea's R^2 are markedly greater for some of the variables. It should be borne in mind that these instruments are thought to deal better with measurement error and, therefore, our results should be taken with caution. Note, importantly, that the Hansen statistic [reported at the bottom of Column (v)] for mutual consistency of the instruments does not reject, by large, the null hypothesis that the excluded instruments are valid and uncorrelated with the error term.

4.3. Spatial structure of productivity

As discussed above, both tangible and intangible assets generate private returns as well as social externalities between agents within regions. When the sender and the receiver of these externalities do not coincide in the same region (especially in the case of small regions, such as the NUTS3 considered here), we can expect a correlation (i.e. spatial dependence) between the productivity in a given region and that of its neighbours. We, therefore, need to take this dependence into account in our model estimation since otherwise the estimates of the relationship between agglomeration (both of employees and intangible endowments) and GVA per job filled would be either biased or inefficient.

To check for spatial dependence, we need to define a measure of proximity, which can be summarized in a $n \times n$ matrix of spatial weights, W.[11] We define the elements of such a matrix, w_{ij}, so that $w_{ij} = \exp(-0.01\, d_{ij})$, d_{ij} being the travel time by car between the centres of region i and region j.[12] A cut-off of 120 min is introduced since interdependencies greater than a 2-h travel time should be negligible. Table 4 shows the values of Moran's I test used to check for the presence of spatial autocorrelation in the distribution of our dependent variable, GVA per job filled, employing our main W matrix as well as various definitions of proximity—including contiguity, physical distance and variations of time travel-dependent measures. Although we find some variation across the various values of the test when using different weight matrices, the null hypothesis of random distribution of the variable is clearly rejected, so we conclude that spatial correlation is significant in the geographical distribution of GVA per job filled.

This spatial dependence, found at an exploratory level in the distribution of productivity, could appear within the regressions conducted herein. As can be seen in Table 3, Moran's I test for spatially autocorrelated residuals after the OLS estimates

11 The most common definition of proximity is that of first order physical contiguity, that is, if two regions share the same administrative border, $w_{ij} = 1$ and $w_{ij} = 0$ otherwise. Other contiguity criteria have been defined in the literature and include commercial exchanges (Cabrer-Borrás and Serrano-Domingo, 2007) and technological proximity (Moreno et al., 2005). We focus our attention here on another definition of contiguity, which is somewhat more relevant for our purposes. Thus, as Patacchini and Rice (2007) stress, travel times between regions are a more economically meaningful measure of proximity than physical contiguity or physical distance.

12 From among various options we use a distance decay of 0.01, since it shows the highest pseudo-R^2 after the FGS2SLS estimations (p.$-R^2$ 0.856 for 0.01; p. $-R^2$ 0.804 for 0.02; p. $-R^2$ 0.774 for 0.03; p. $-R^2$ 0.792 for 0.04; p. $-R^2$ 0.643 for 0.05; p. $-R^2$ 0.733 for 0.08; p. $-R^2$ 0.765 for 0.1).

Table 4. Global spatial autocorrelation tests

	W1	W2	W3	W4	W5	W6
Moran's *I*						
ln(GVA filled job)	12.994	6.598	5.800	6.858	7.318	11.117
P-value	*0.000*	*0.000*	*0.000*	*0.000*	*0.000*	*0.000*

Notes: W1: main matrix [$w_{ij}=\exp(-0.01d_{ij})$, d_{ij} being the travel time by car between the centres of region *i* and region *j*]; W2: row-standardized contiguity binary matrix, with value equal to 1 in the case of sharing borders and 0 otherwise; W3: row-standardized binary matrix where the weight is 1 if a centre of a region is within a 0-60 min travel time band, and 0 otherwise; W4: row-standardized binary matrix where the weight is 1 if a centre of a region is within a 0–90 min travel time band, and 0 otherwise; W5: row-standardized binary matrix where the weight is 1 if a centre of a region is within a 0–120 min travel time band, and 0 otherwise; W6: $w=1/m$, where m = miles between each regional centre.

seems to indicate that spatial autocorrelation is still present in the residuals of our estimation, so that our model does not fully consider the spatial structure of productivity.

In consequence, we first include in our empirical model an average of the endowment of each intangible asset in the neighbouring regions (i.e. the intangibles' 'spatial lag') to pick up the interregional externalities. In this way, we are able to analyse whether the inclusion of the spatial lag[13] of the intangibles eliminates the spatial autocorrelation of the residuals. As shown in Columns (i) and (ii) of Table 5, despite the fact that most of these interregional externalities arising from intangible endowments are significant in the regression, spatial autocorrelation remains after their introduction (Moran's *I* test is still significant).

Second, adhering to the idea of attempting to pick up the spatial structure of productivity through our specification, in Table 6 we consider narrower economic mass bands. It might be the case that the spatial autocorrelation in the residuals could be eliminated by slightly modifying the model specification, and considering these narrower economic mass bands. Previous OLS and GMM estimations were, therefore, repeated by including three 40-min isochrones—within 40, 40–80-, 80–120-min travel time. Once again, what we seek is to determine whether spatially meaningful variables can eliminate the spatial correlation of the residuals. The second and third travel time bands are significant. However, as shown in Columns (i) and (ii) of Table 6, the introduction of such fine bands does not seem to eliminate the spatial correlation of the regression residuals since Moran's *I* test remains significant.[14]

All in all, neither the consideration of the endowments of intangible assets in the neighbouring regions nor the narrower bands for agglomeration economies seems to eliminate the spatial autocorrelation present in our regressions. Thus, in estimating our

13 Our main weights matrix, W, is used to calculate all the spatial lags of the regressors and the dependent variable throughout unless otherwise started.

14 We also repeated the whole set of estimations (not reported here, but available upon request), though on this occasion we excluded the two non-significant explanatory variables, i.e., occupational human capital and applied patents, so that these variables might not account spuriously for variations in the dependent variable. The results and conclusions remained virtually unchanged.

Table 5. Spatial structure of productivity: intangible assets in the neighbouring regions

	OLS (i)	GMM (ii)
ln(employment within 0–60 min)	0.049***	0.037***
	(*0.015*)	(*0.013*)
ln(employment within 60–120 min)	0.013***	0.013***
	(*0.005*)	(*0.004*)
Educational HK	0.173**	0.169**
	(*0.078*)	(*0.069*)
Occupational HK	−0.002	−0.002
	(*0.003*)	(*0.003*)
Empl. RD&IT	0.043***	0.043***
	(*0.016*)	(*0.014*)
High-tech manufacturing employment	0.058***	0.057***
	(*0.014*)	(*0.012*)
ln(applied patents by inventor)	0.012	0.010
	(*0.013*)	(*0.011*)
ln(VAT registrations)	0.105**	0.101**
	(*0.048*)	(*0.042*)
CAGR VAT registrations	0.016	0.017*
	(*0.011*)	(*0.009*)
Educational HK in neighbouring regions	−0.054	−0.042
	(*0.042*)	(*0.035*)
Occupational HK in neighbouring regions	−0.001	−0.002
	(*0.002*)	(*0.002*)
Empl. RD&IT in neighbouring regions	0.002	0.001
	(*0.011*)	(*0.009*)
High-tech manufacturing employment in neighbouring regions	0.015*	0.015*
	(*0.009*)	(*0.008*)
ln(applied patents by inventor) in neighbouring regions	0.004	0.004
	(*0.010*)	(*0.009*)
ln(VAT registrations) in neighbouring regions	0.065	0.061*
	(*0.040*)	(*0.035*)
CAGR VAT registrations in neighbouring regions	−0.012	−0.010
	(*0.012*)	(*0.010*)
Constant	9.132***	9.272***
	(*0.187*)	(*0.162*)
NUTS1 dummies	Yes	Yes
Sample size	119	119
R^2	0.809	0.806
Adjusted R^2	0.754	0.7508
NUTS1 dummies	Yes	Yes
Moran's I	4.323	
P-value	*0.000*	
Hansen J statistic		2.132
P-value		*0.3444*
ln(Empl. 60 min)—partial R^2		*0.5542*
ln(Empl. 60 min)—shea R^2		*0.5361*
ln(Empl. 60 min)—first-stage F-stat		*22.23*
ln(Empl. 60–120 min)—partial R^2		*0.9601*
ln(Empl. 60–120 min)—shea R^2		*0.9289*
ln(Empl. 60–120 min)—first-stage F-stat		*892.68*

Notes: OLS and GMM estimates with several levels of significance: ***1%, **5%, *10%. In the case of the parameters, White-robust standard errors are presented in italics and parentheses below each associated parameter. Moran's I test for the residuals of the OLS estimations is provided, rejecting the null and, therefore, indicating that they remain spatially autocorrelated. The variables expressed in percentages and location quotients are not log transformed in order to facilitate the interpretation of their coefficient. Hansen J statistic for mutual consistency of the available instruments is provided [Column (ii)] and we cannot reject the null hypothesis that the excluded instruments are valid and uncorrelated with the error term, so there are no overidentification problems.
Dep. Var.: lnGVA per job filled.

Table 6. Spatial structure of productivity: narrower bands for proxying agglomeration economies

	OLS		GMM	
	(i)	(ii)	(iii)	(iv)
First isochrone	0.047***	0.035***	0.061***	0.039***
	(0.012)	*(0.011)*	*(0.014)*	*(0.010)*
Second isochrone	0.006*	0.001	0.004	0.002
	(0.003)	*(0.003)*	*(0.003)*	*(0.003)*
Third isochrone	0.017**	0.011**	0.016**	0.010***
	(0.007)	*(0.005)*	*(0.007)*	*(0.004)*
Educational human capital	0.316***	0.163**	0.314***	0.187**
	(0.065)	*(0.081)*	*(0.064)*	*(0.074)*
Occupational human capital		−0.002		−0.003
		(0.004)		*(0.004)*
Employment in RD and computers		0.055***		0.066***
		(0.017)		*(0.015)*
High-tech manufacturing employment		0.058***		0.061***
		(0.015)		*(0.013)*
ln(applied patents by inventor)		0.015		0.010
		(0.012)		*(0.011)*
ln(VAT registrations)		0.058		0.065
		(0.046)		*(0.041)*
CAGR VAT registrations		0.030**		0.024**
		(0.014)		*(0.012)*
Constant	9.076***	9.295***	8.935***	9.265***
	(0.145)	*(0.155)*	*(0.152)*	*(0.136)*
NUTS1 dummies	Yes	Yes	Yes	Yes
Sample size	119	119	119	119
Adj. R^2	0.542	0.701	0.523	0.692
Moran's I	4.013	3.678		
P-value	*0.000*	*0.000*		
Hansen J statistic			9.586	13.470
P-value			*0.0224*	*0.0037*
ln(Empl. 40 min)—partial R^2			*0.7332*	*0.7267*
ln(Empl. 40 min)—shea R^2			*0.7257*	*0.7269*
ln(Empl. 40 min)—first-stage F-stat			*31.75*	*37.40*
ln(Empl. 40–80 min)—partial R^2			*0.9810*	*0.9788*
ln(Empl. 40–80 min)—shea R^2			*0.9561*	*0.9618*
ln(Empl. 40–80 min)—first-stage F-stat			*892.41*	*751.20*
ln(Empl. 80–120 min)—partial R^2			*0.9669*	*0.9486*
ln(Empl. 80–120 min)—shea R^2			*0.9427*	*0.9486*
ln(Empl. 80–120 min)—first-stage F-stat			*194.30*	*185.49*

Notes: OLS and GMM estimates with several levels of significance: ***1%, **5%, *10%. In the case of the parameters, white-robust standard errors are presented in italics and parentheses below each associated parameter. Moran's I test for the residuals of the OLS estimations is provided, rejecting the null and, therefore, indicating that they remain spatially autocorrelated. The variables expressed in percentages and location quotients are not log transformed in order to facilitate the interpretation of their coefficient. Hansen J statistics for mutual consistency of the available instruments are provided [Columns (iii) and (iv)] and we cannot reject the null hypothesis that the excluded instruments are valid and uncorrelated with the error term, so there are no overidentification problems.
Dep. Var.: lnGVA per job filled.

models we would like to control, at least econometrically, for this dependence, so that the results remain uninfluenced by it.

To move in this direction, we therefore need to choose an appropriate estimation method.[15] Most studies in the literature (see, for example, Rice et al., 2006) use maximum likelihood (ML) procedures. However, their reliability and feasibility require specific distributional assumptions (Kelejian and Prucha, 1998) and when there are other endogenous variables in the right-hand side of the model it is difficult, if not impossible, to implement (Fingleton and Le Gallo, 2008).

Thus, we adopt the feasible generalized spatial two-stage least-squares estimator (FGS2SLS) proposed by Kelejian and Prucha (1998), which includes the spatial lag of the endogenous variable (in our case, the average of the productivity in the nearest regions). It is somewhat modified here, however, in order to control for endogeneity problems arising from the reverse causality of the agglomeration variable. To implement this modification, we follow the procedure described in Fingleton (2003) when estimating agglomeration economies for Great Britain. This procedure is consistent, but not efficient if additional spatial correlation remains in the disturbance term. In this instance, we would need to estimate a spatial autoregressive parameter in the error term of the equation, the latter already including the spatial lag of the endogeneous variable (in line with Kelejian and Prucha, 1999).

Our model's estimation results including the average of productivity in the nearest regions [spatial lag model, Column (i) in Table 7] indicate that the spatial lag of the endogenous variable matters, although its value is small. Moreover, Moran's I test for 2SLS[16] indicates that there is no remaining residual spatial autocorrelation at a 5% confidence level (although this is not the case at the 10% level). So, by means of a robustness check, in Column (ii) of Table 7, we show the results with the inclusion of a spatial lag both in the dependent variable and in the error term. The most striking aspects of these estimations are, basically, that the parameters accompanying proxies for intangible capital assets remain, on the whole, and present similar values to those in Table 3. Additionally, the spatial lag is significant at a 5% level and with a value of 0.001. Likewise, the elasticity of the agglomeration effect falls to 0.024, whereas the parameter for the second isochrone is no longer significant.

To sum up, from Table 7, we can conclude that the externalities arising from neighbouring regions—captured through a spatial lag of the dependent variable—matter, although the magnitude of their effects is very small (0.1%). Besides, increasing returns arising from agglomeration economies are markedly reduced when spatial autocorrelation is allowed for and are significant only for distances below 60-min travelling time by car. However, the small value of the spatial lag coefficient and the residual spatial autocorrelation that remains, albeit only slightly, after the first step of the FGS2SLS lead us to think that the spatial lag does not account for all the externalities across regions. However, we do at least controlling econometrically that their presence does not make our estimates of the impact of density and intangible assets unreliable.

15 Ordinary least squares would not be an appropriate technique, resulting in inconsistency or inefficiency depending upon the kind of spatial autocorrelation in question.

16 A Moran's I test for 2SLS residuals (distributed as a standard normal), as proposed by Anselin and Kelejian (1997) is performed, since the usual Moran's I based on OLS residuals, where all the explanatory variables are exogenous, would not be appropriate here.

Table 7. Spatial structure of productivity: FGS2SLS estimates

	(i) First stage of FGS2SLS	(ii) FGS2SLS with spatial autoregressive error term
W·lnGVA filled job	0.001*	0.001***
	(0.000)	*(0.000)*
ln(employment within 0–60 min)	0.021*	0.024*
	(0.012)	*(0.012)*
ln(employment within 60–120 min)	0.008	0.003
	(0.005)	*(0.006)*
Educational human capital	0.178***	0.144**
	(0.066)	*(0.067)*
Occupational human capital	−0.002	0.003
	(0.003)	*(0.003)*
Employment in RD and computers	0.044***	0.042***
	(0.014)	*(0.013)*
High-tech manufacturing employment	0.053***	0.038***
	(0.012)	*(0.012)*
ln(applied patents by inventor)	0.016*	0.010
	(0.009)	*(0.009)*
ln(VAT registrations)	0.069*	0.037
	(0.036)	*(0.035)*
CAGR VAT registrations	0.019**	0.021**
	(0.009)	*(0.009)*
Constant	9.474***	9.491***
	(0.172)	*(0.178)*
NUTS1 dummies	Yes	Yes
Sample size	119	119
Pseudo-R^2	0.856	0.856
Joint test for intangibles [Wald test$\chi^2_{(7)}$]	47.85	90.54
P-value	*0.0000*	*0.000*
Sargan statistic	24.757	25.538
P-value	*0.3630*	*0.323*
| Moran's *I* z-statistic |	1.68	
P-value	*0.095*	
Lambda		0.561
P-value		*0.000*

Notes: FGS22SLS estimates with several levels of significance: ***1%, **5%, *10%. Standard errors are presented in italics and parentheses below each associated parameter. Sargan statistics for mutual consistency of the available instruments are provided and we cannot reject the null hypothesis that the excluded instruments are valid and uncorrelated with the error term, so there are no overidentification problems. Instrument validities are not reported for reasons of space, although they are available upon request from the authors. The Pseudo-R^2 is calculated as the ratio of the variance of the fitted values of the dependent variable over the variance of the dependent variable.
Dep. Var.: lnGVA per job filled.

5. Conclusions

In the study conducted here, basing our analysis on Ciccone's (2002) model, we examine the hypothesis that regions are endowed with certain kinds of intangible assets that have come to characterize the knowledge-based economy, and which are simultaneously sources of private and social returns. In contrast with previous studies, we have taken these qualitative features into account when estimating the aggregate effect of agglomeration economies on the economic performance of regions so as not to introduce an upward bias in our parameter estimations. Further, we have hypothesized that strong social returns derived from several sources, both tangible and intangible, cross the administrative borders of regions and need to be taken into account in our estimation.

The main conclusions to be drawn from applying our methodological approach to the available data sets are as follows: agglomeration economies—as measured here—matter when explaining differences in economic performance across regions; however, their importance in quantitative terms and their extension become somewhat constrained when variables proxying intangible assets—knowledge, human capital and entrepreneurial culture—are included in our estimations. More specifically, most of the variables proxying these intangible assets are significant and present the expected sign. Our results are consistent even when explicitly treating 'two-way causation' problems between productivity and agglomeration.

What is more, the explanatory power of intangible assets in this framework remains, on the whole, unaffected when externalities across regions are considered in the model. However, the coefficients for agglomeration economies are somewhat reduced, albeit that they remain significant.

As for the policy implications of our findings, it would seem that, to a certain degree, improvements in local/regional transportation infrastructure that reduce the length of business and commuting journeys might boost labour productivity by means of increasing returns derived from transportation cost reductions, shared inputs and knowledge spillovers. Thus, as has been stressed elsewhere (Graham, 2007), investments in such infrastructure should be made. However, at the same time, the accumulation of certain kinds of intangible endowments in a given region is of great importance, and so low-density, non-metropolitan areas should also profit from the concentration of these intangible assets. Policies that address this issue are, therefore, of great relevance.

Acknowledgements

Part of this study was completed while Michael J. Artis was Director of the Manchester Regional Economics Centre, part of the Institute for Political and Economic Governance (IPEG), at the University of Manchester and Ernest Miguelez was a visiting scholar. The use of IPEG's facilities is gratefully acknowledged. The authors would also like to thank the following for their helpful comments: Christian Catalini, Vicente Royuela, Marianne Sensier, participants at the IREA Annual Conference and Workshop (Barcelona, 3 February 2009), participants at the III World Spatial Econometric Conference (Barcelona, 8–10 July 2009), participants at the 24th Annual Congress of the European Economic Association (Barcelona, 23–27 August 2009), the Editor (Henry Overman) and three anonymous referees.

Funding

E. M. and R. M. acknowledge financial support from the *Ministerio de Ciencia e Innovación* (ECO2008-05314). Ernest Miguelez also acknowledges support from the *Ministerio de Educación* (AP2007-00792) and Rosina Moreno from the 7th FP project: Intangible Assets and Regional Economic Growth (FP7-SSH-2007-1-216813).

References

Ades, A. F., Glaeser, E. L. (1995) Trade and circuses: explaining urban giants. *Quarterly Journal of Economics*, 110: 195–227.

Anselin, L., Kelejian, H. H. (1997) Testing for spatial error autocorrelation in the presence of endogenous regressors. *International Regional Science Review*, 20: 153–182.

Audretsch, D. B. (2003) Entrepreneurship: a Survey of the literature, Brussels: Enterprise Directorate General, European Commission. Enterprise Papers No 14.

Baldwin, J. R., Brown, W. M., Rigby, D. L. (2010) Agglomeration Economies: Microdata Panel Estimates from Canadian Manufacturing Economic Analysis. *Journal of Regional Science*, 50: 915–934.

Baptista, R. (2003) Productivity and the density of regional clusters. In J. Bröcker, D. Dohse, R. Soltwedel (eds) *Innovation Clusters and Interregional Competition*. Berlin: Springer.

Baum, C., Schaffer, M., Stillman, S. (2003) Instrumental variables and GMM: estimation and testing. *The Stata Journal*, 3: 1–31.

Beeson, P. E., Husted, S. (1989) Patterns and determinants of productive efficiency in state manufacturing. *Journal of Regional Science*, 29: 15–28.4.

Blien, U., Suedekum, J., Wolf, K. (2006) Local employment growth in West Germany:a dynamic panel approach. *Labour Economics*, 13: 445–458.

Bode, E. (2004) *Agglomeration Externalities in Germany*. ERSA Conference Papers, European Regional Science Association, Paper No ersa04p120, Vienna.

Brülhart, M., Mathys, N. A. (2008) Sectoral agglomeration economies in a panel of European regions. *Regional Science and Urban Economics*, 38: 348–362.

Brülhart, M., Sbergami, F. (2009) Agglomeration and growth: cross-country evidence. *Journal of Urban Economics*, 65: 48–63.

Cabrer-Borrás, B., Serrano-Domingo, G. (2007) Innovation and R&D spillover effects in Spanish regions: a spatial approach. *Research Policy*, 36: 1357–1371.

Cheshire, P., Magrini, S. (2009) Urban growth drivers in a Europe of sticky people and implicit boundaries. *Journal of Economic Geography*, 9: 85–115.

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

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

Ciccone, A., Cingano, F. (2003) Skills and clusters. In J. Bröcker, D. Dohse, R. Soltwedel (eds) *Innovation Clusters and Interregional Competition*. Berlin: Springer.

Cingano, F., Schivardi, F. (2004) Identifying the sources of local productivity growth. *Journal of the European Economic Association*, 2: 720–742.

Combes, P.-P. (2011) The empirics of economic geography: how to draw policy implications? *Review of World Economics*, 147: 567–592.

Combes, P.-P., Duranton, G., Gobillon, L. (2008) Spatial wage disparities: sorting matters!. *Journal of Urban Economics*, 63: 723–742.

Combes, P.-P., Duranton, G., Gobillon, L. (2011) The identification of agglomeration economies. *Journal of Economic Geography*, 11: 253–266.

Duranton, G. (2008) From cities to productivity and growth in developing countries. *The Canadian Journal of Economics*, 41: 689–736.

Duranton, G., Puga, D. (2004) Micro-foundations of urban agglomeration economies. In J. V. Henderson, J. F. Thisse (eds) *Handbook of Regional and Urban Economics*, vol. 4. North-Holland: Elsevier.

Fingleton, B. (2003) Increasing returns: evidence from local wage rates in Great Britain. *Oxford Economic Papers*, 55: 716–739.

Fingleton, B., Le Gallo, J. (2008) Estimating spatial models with endogenous variables, a spatial lag and spatially dependent disturbances: finite sample properties. *Papers in Regional Science*, 87: 319–339.

Glaeser, E., Kallal, H. D., Scheinkman, J. A., Shleifer, A. (1992) Growth in cities. *Journal of Political Economy*, 100: 1126–1152.

Glaeser, E., Kolko, J., Saiz, A. (2001) Consumer city. *Journal of Economic Geography*, 1: 27–50.

Glaeser, E., Maré, D. C. (2001) Cities and skills. *Journal of Labour Economics*, 19: 316–342.

Glaeser, E. L., Sacerdote, B. I., Scheinkman, J. A. (2010) The social multiplier. *Journal of the European Economic Association*, 1: 345–353.

Graham, D. J. (2007) Agglomeration, productivity and transport investment. *Journal of Transport Economics and Policy*, 41: 317–343.

Hanson, G. H. (2005) Market potential, increasing returns and geographic concentration. *Journal of International Economics*, 67: 1–24.

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

HM Treasury, (2001) *Productivity in the UK: 3-The Regional Dimension*. London: Department of Trade and Industry.

Kelejian, H. H., Prucha, I. R. (1998) A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. *Journal of Real Estate Finance and Economics*, 17: 99–121.

Kelejian, H. H., Prucha, I. R. (1999) A generalized moments estimator for the autoregressive parameter in a spatial model. *International Economic Review*, 40: 509–533.

Kennedy, P. (1992) *A Guide to Econometrics*, 3rd edn. Oxford: Blackwell.

Marshall, A. (1890) *Principles of Economics*. London: Macmillan.

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

Moreno, R., Paci, R., Usai, S. (2005) Geographical and sectoral clusters of innovation in Europe. *Annals of Regional Science*, 39: 715–739.

Moretti, E. (2004) Human capital externalities in cities. In V. Henderson, J. Thisse (eds) *Handbook of Urban and Regional Economics*, vol. 4, pp. 2243–2291.

Patacchini, E., Rice, P. (2007) Geography and economic performance: exploratory spatial data analysis for Great Britain. *Regional Studies*, 41: 489–508.

Rauch, J. E. (1993) Productivity gains from geographic concentration of human capital: evidence from the cities. *Journal of Urban Economics*, 34: 380–400.

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

Rosenthal, S. S., Strange, W. C. (2004) Evidence on the nature and sources of agglomeration economies. In V. Henderson, J. Thisse (eds) *Handbook of Urban and Regional Economics*, vol. 4. North-Holland: Elsevier.

Segal, D. (1976) Are there returns to scale in city size? *Review of Economics and Statistics*, 58: 339–350.

Serfling, R. J. (1980) *Approximation Theorems of Mathematical Statistics*. New York: Wiley, cooperation.

Shea, J. (1997) Instrument relevance in multivariate linear models. A simple measure. *Review of Economics and Statistics*, 79: 348–352.

Sveikauskas, L. A. (1975) The productivity of cities. *The Quarterly Journal of Economics*, 89: 393–413.

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

Wheaton, W. C., Lewis, M. J. (2001) Urban wages and labor market agglomerations. *Journal of Urban Economics*, 51: 542–562.

Wosnitza, B., Walker, M. (2008) Regional economic indicators with a focus on differences in sub-regional economic performances. *Economic and Labour Market Review*, 2: 40–53.

Appendix A

Table A1. Variables and data construction

Variable	Proxy	Dates	Source
Productivity	GVA per job filled	Average 2001–2005	Wosnitza and Walker (2008).
'Economic mass'	Sum of the jobs filled within all the regions whose centre is located within two 60-min travel-time bands starting from the centre of each region.	Average 2001–2005	Wosnitza and Walker (2008) for the jobs data and data acknowledged to Patricia Rice and Anthony Venables.
Educational human capital	Location quotient[a] of the percentage of economically active population with a first and higher degree; nursing and teaching qualifications (NVQ4) or with A-levels; GNVQ Higher level, or Advanced certificate of Vocational Education (NVQ3)	Average 1999–2001	NOMIS database, collected by the Office of National Statistics (ONS)
Occupational human capital	Percentage of economically active population who are enrolled in occupations such as corporate managers, managers/proprietors in agriculture/services, science and technology professionals, health professionals, teaching and research professionals, and business and public service professionals	Average 1999–2001	NOMIS database, collected by the Office of National Statistics (ONS)
Employment in RD and IT	Location quotient for each area giving the workforce specialization in computing and related activities and in research and development	Average 1996–2000	NOMIS database
High-tech manufacturing employment	Location quotient for each area giving the workforce specialization in chemicals and man-made fibres; machinery and equipment; optical and electrical equipment; and transport equipment	Average 1996–2000	NOMIS database
Applied patents by inventor	Patents applied for in a given region, regionalizing them according to the household of the inventor who has registered the patent to the European Patent Office, using the OECD database[b]	Average 1996–2000	OECD REGPAT database, May 2008
Entrepreneurship culture	VAT registrations per head	Average 1996–2000	NOMIS database
Entrepreneurship success	CAGR of VAT registrations	Average 1996–2000	NOMIS database
Area	Sum of the square kilometres within all the regions whose centre is located within two 60-min travel-time bands each starting from the centre of each region.		ONS
Population in 1801	Sum of the 1801 population within all the regions which centre is located within two travel-time bands of 60 min starting from the centre of each region.	1801	'Britain through time'. Great Britain Historical Geographical Information System. University of Portsmouth. Department of Geography.

[a]The regional share over the national share.
[b]By collecting data on applied patents in this way, we seek to avoid the bias introduced by the accumulation of patents in regions where the headquarters of several firms are located.