Available online at www.sciencedirect.com
SCIENCE DIRECT
Regional Science and Urban Economics 34 (2004) 533–564
www.elsevier.com/locate/econbase

The geographic distribution of production activity in the UK

Michael P. Devereux[a], Rachel Griffith[b], Helen Simpson[c,*]

[a] University of Warwick, Institute for Fiscal Studies and CEPR, Warwick, UK
[b] Institute for Fiscal Studies, University College London, and CEPR, London, UK
[c] Institute for Fiscal Studies, 7 Ridgmount Street, London WC1E 7AE, UK

Received 11 September 2001; received in revised form 30 September 2002; accepted 27 March 2003
Available online 4 September 2003

Abstract

We investigate the geographic concentration and agglomeration of production activity in the UK at the four-digit industry level using a variety of measures. We relate these to comparable patterns in the US and France and find several similarities. We find that conditioning on industrial concentration, the most geographically concentrated industries appear to be relatively low-tech. We find evidence that plant survival rates are higher and both entry and exit rates are lower in more agglomerated industries, but that in some of the most agglomerated industries entry acts to re-enforce agglomeration.

JEL classification: R12; R3
Keywords: Geographic concentration; Agglomeration

1. Introduction

There are many examples of geographically concentrated industries, including the often cited clusters of high-tech firms in Silicon Valley (California), Route 128 (Boston), Cambridge (UK) and Sophia Antipolis (France). But the phenomenon is neither recent, nor restricted to high-tech industries. Other examples abound: the US carpet industry in Dalton, GA; the UK ceramics industry around Stoke-on-Trent, an area known as 'The Potteries'; and the UK lace industry centred in Nottingham.

Understanding how and why these clusters form and persist is an issue of considerable interest both from an academic and policy perspective. A substantial literature has focused

* Corresponding author. Tel.: +44-20-7291-4800; fax: +44-20-7323-4780.
E-mail address: hsimpson@ifs.org.uk (H. Simpson).

0166-0462/$ - see front matter
doi:10.1016/S0166-0462(03)00073-5

on localisation economies, dating back to Marshall (1890).[1] Broadly, Marshall envisaged three types of positive externalities arising for plants located near other similar plants: lower costs arising from close proximity to suppliers and customers, benefits from a flexible and specialist local labour market and benefits that come from knowledge spillovers facilitated by close proximity to other plants. Several recent papers have investigated the evidence for these externalities, which can arise both within and between industries.[2]

The focus of this paper is on examining the extent of such geographic concentration in the UK. We examine whether technology plays an important role in explaining the geographic distribution of production and look at the dynamics of activity in agglomerated industries and regions. Our research builds on two recent papers, Ellison and Glaeser (1997) and Maurel and Sédillot (1999), which have developed measures of geographic concentration and applied these to manufacturing industries in the US and France, respectively. These papers distinguish between geographic concentration arising from unrelated plants locating near to each other and that due to concentration in industrial structure. For example, a single plant monopoly would represent a highly concentrated industry geographically. But the high geographic concentration is not the result of positive externalities between unrelated plants (although high industrial concentration may have arisen endogenously if externalities were sufficiently high for the firms to chose to integrate into a single firm). We use the term 'agglomeration' in this paper to refer to geographic concentration over and above that which would be expected given the extent of industrial concentration in the industry. We compare the overall level of agglomeration and which industries are most and least agglomerated in the UK with the findings in the US and France.

In Section 2, we briefly outline the measures of agglomeration used in this paper. In Section 3, we apply these measures to plant level data in the UK. We examine patterns of geographic concentration and agglomeration at the four-digit industry level and across related four-digit industries. We also investigate, in a simple way, the hypothesis that more high-tech industries tend to be more agglomerated. In Section 4, we compare the patterns of agglomeration in the UK with those in the US and France. In Section 5, we investigate the dynamic properties of agglomeration in the UK. We examine whether there are significant differences between more and less agglomerated industries in patterns of entry and exit and job creation and destruction. We also apply geographic concentration and agglomeration measures to new entrants, to examine whether entry is acting to re-enforce agglomeration. Section 6 summarises and concludes.

2. Indices of agglomeration

A number of measures have been used to investigate the geographic dispersion of plants within industries. For example, Krugman (1991) and Amiti (1998) have used variants of the Gini coefficient to measure geographic concentration. Recent papers by

[1] More recent references include Arrow (1962), Romer (1990) and Krugman (1991). See also Jacobs (1969, 1984).

[2] See, for example, Ellison and Glaeser (1999), Glaeser et al. (1992), Henderson (1994, 1999), Henderson et al. (1995) and Dumais et al. (2002).

Ellison and Glaeser (1997) and Maurel and Sédillot (1999) have proposed indices designed to measure agglomeration—geographic concentration in excess of that which would be expected given industrial concentration. These measures are all based on the distribution of activity over discrete geographic units. By contrast, Duranton and Overman (2002) propose a measure of agglomeration based on a continuous measure of location, the distance between pairs of plants, which therefore avoids aggregating information to the level of discrete spatial units, and allows for an analysis of the geographic scale of industry agglomeration. As here, Duranton and Overman apply their measure to UK data. They additionally test for the statistical significance of measured localisation.

The analysis in this paper primarily uses the index of geographic concentration proposed by Maurel and Sédillot (1999), henceforth MS, although we also compare it with that proposed by Ellison and Glaeser (1997), henceforth EG. We briefly summarise these measures. In a data appendix[3] we also provide our calculations at the four-digit industry level.

Define a variable $u_{ji}=1$ if plant j locates in region i, and $u_{ji}=0$ otherwise. The index proposed by MS is an empirical estimate of γ, defined as

$$\gamma = \mathrm{corr}(u_{ji}, u_{ki}) \text{ for } j \neq k. \tag{1}$$

Within any given industry, γ is treated as fixed across any pair of plants, j and k, and any region, i. Comparing estimates of γ across industries gives an indication of the relative strength of spillovers between plants within each industry. These spillovers may reflect natural advantages to location in the region (such as access to raw materials), or positive externalities arising from close proximity to other plants within the same industry.

MS demonstrate that an empirical estimate of γ can be generated from an estimate of the probability, denoted as p, that plants j and k locate in the same region. To see this, define x_i as the probability that plant j locates in region i: $P(u_{ji}=1)=x_i$. Then the probability of j and k both locating in region i is:

$$P(i,i) = E(u_{ji}u_{ki}) = \mathrm{cov}(u_{ji}, u_{ki}) + E(u_{ji})E(u_{ki}) = \gamma x_i(1-x_i) + x_i^2. \tag{2}$$

Aggregating over M regions:

$$p = \sum_{1=i}^{M} P(i,i) = \gamma\left(1 - \sum_{1=i}^{M} x_i^2\right) + \sum_{1=i}^{M} x_i^2. \tag{3}$$

For the central case, MS define x_i as the proportion of aggregate employment in region i. This implies that a random location process will on average lead to a geographic

[3] Available at http://www.ifs.org.uk/corpact/dgsdata.zip.

distribution of employment matching the aggregate distribution.[4] Given this, MS demonstrate that a simple frequency estimator of p can be found as:[5]

$$\hat{p} = \frac{\sum_{1=i}^{M} s_i^2 - H}{1 - H}, \tag{4}$$

where s_i is the share of total industry employment in region i, and H is the industry Herfindahl index, defined as:

$$H = \sum_{j=1}^{N} z_j^2 \tag{5}$$

and where z_j is the share of plant j in total industry employment, given N plants in the industry. Substituting $\hat{p}$ from Eq. (4) for p in Eq. (3), and solving for the implied estimate of γ, yields the empirical index proposed by MS:[6]

$$\hat{\gamma} = \frac{G - H}{1 - H}. \tag{6}$$

where

$$G = \frac{\sum_{i=1}^{M} s_i^2 - \sum_{i=1}^{M} x_i^2}{1 - \sum_{i=1}^{M} x_i^2}. \tag{7}$$

An industry with high industrial concentration is likely also to have a high geographic concentration; in the extreme, a monopoly consisting of a single plant would be located in a single region. The two components of the index $\hat{\gamma}$ are G—a measure of the raw geographic concentration—and H—reflecting industrial concentra-

[4] Note, however, that the aggregate distribution is itself not random. One could think of different underlying distributions with which to compare, for example, the population or the number of plants.

[5] An alternative approach would be to abstract from plant size altogether, by developing an estimate of p based on the number of plants locating in each region, rather than on the size of employment in each plant. This generates an index that is independent of H.

[6] The Ellison and Glaeser (1997) index is similar, but not identical, to this:

$$\hat{\gamma}_{\mathrm{EG}} = \left\{ \frac{\sum_{i=1}^{M} (s_i^2 - x_i^2)}{1 - \sum_{i=1}^{M} x_i^2} - H \right\} \Big/ (1 - H),$$

where s_i is defined relative to manufacturing, rather than total, employment. See Maurel and Sédillot (1999) for a comparison of the two measures.

tion. Broadly, the index represents the difference between these two, and hence the degree of geographic concentration in excess of that which is due to industrial concentration.

MS demonstrate that $\hat{\gamma}$ is an unbiased estimate of γ. This leads to the interpretation of the index that, when plant location choices are independent, $E(\hat{\gamma})=0$. A value of $\hat{\gamma}$ greater than 0 can therefore be interpreted as implying a localised industry, in the sense that geographic concentration is in excess of that which would be expected given industrial concentration.

3. Patterns of geographic and industrial concentration in the UK

We now turn to an examination of the patterns of geographic and industrial concentration in UK production activity. In the course of presenting these results we discuss a number of issues relating to the appropriate level of industry and regional aggregation. We begin by summarising the data. As a preliminary step, we present evidence of the geographic concentration of the production sector as a whole in the UK. We then present summary measures of geographic concentration and agglomeration at the four-digit industry level. We go on to investigate briefly the relationship between agglomeration and alternative measures designed to summarise the technological intensity of the industry. In the final part of this section, we also investigate co-agglomeration between related four-digit industries.

3.1. The data

The empirical analysis presented below uses plant level data from the UK's Annual Respondents Database (ARD) over the period 1985 to 1992. The ARD contains information on the population of production plants in the UK.[7] This includes the location of the plant (given by the postcode and local authority), the plant's four-digit industrial classification and the number of employees. A broader range of information on output and inputs is available at the establishment level. An establishment can be a single plant or a group of plants (which can be at different addresses) owned by one firm. For the years we consider this information is available for all establishments with over 100 employees and a sample of those with below 100 employees.

Table 1 shows some descriptive statistics for 1992, which is our main year of analysis. Our data includes information on plants in 211 four-digit production industries (using the 1980 SIC classification). These include energy and water supply, extraction and all manufacturing industries.

[7] The ARD contains the population of plants from the Inter-Departmental Business Register (IDBR). Data are stored in two types of files—non-selected and selected data. To construct a dataset of all production plants (or establishments), it is necessary to combine the non-selected and selected data. See Oulton (1997), Griffith (1999), and Barnes and Martin (2002) for a description of the ARD data.

Table 1
Descriptive statistics, 1992

Number of four-digit industries[a]	211
Number of plants:	
in total population	155,849
incorporated, sole proprietors or partnerships, and actively producing	150,695
number of 'firms'[b]	144,404
Average employment per 'firm'	32

[a] We have excluded seven industries due to the small number of plants and the confidential nature of the data. These are (1113) deep coal mines, (1115) manufacture of solid fuels, (1200) coke ovens, (1520) nuclear fuel production, (1620) public gas supply, (2100) extraction and preparation of metalliferous ores, and (4560) fur goods.

[b] This is the number of observations after aggregating plants that are in the same four-digit industry and postcode area and are owned by the same firm. We have dropped plants where the postcode, industry or employment information is missing.

In 1992, there were 155,849 plants. From the population of plants, we restrict ourselves to plants which are part of incorporated companies, are sole-traders or partnerships (therefore excluding publicly owned corporations), plants that are strictly engaged in production activity (rather than distribution or administration), and plants that are active in that year (excluding those that are not yet in production). This leaves 150,695 plants. From the theoretical discussion above it is clear that we are interested in looking at agglomerations of plants that are *not* under common ownership. Therefore, where we observe two plants in the same industry, in the same postcode area that are under common ownership we aggregate them and call them a 'firm'. This leaves us with 144,404 firms or non-related plants.

Average employment in these 'firms' in 1992 was 32 employees. Table 2 shows more detail about the size distribution of firms. Nearly 50% of 'firms' have fewer than five employees, while nearly 90% have fewer than 50 employees. The size distribution of

Table 2
Size distribution of firms[a]

Number of employees in the firm	Percentage of firms	Number of firms (thousands)	Percentage of employment	Number of employees (thousands)
0–1	23.5	34.0	0.7	33.9
2	11.9	17.2	0.8	34.4
3	8.1	11.8	0.8	35.3
4	6.1	8.8	0.8	35.1
5–9	17.4	25.1	3.6	167.5
10–19	11.7	16.9	5.1	231.9
20–49	10.9	15.7	10.5	484.0
50–99	4.6	6.6	10.0	461.2
100–199	2.8	4.1	12.3	565.3
200+	2.9	4.2	55.4	2541.5
Total	100.0	144.4	100.0	4590.2

[a] Firms are defined as in Table 1.

'firms' is similar across regions, although in the South East of England small 'firms' account for a larger proportion of employment than they do in other regions.[8]

3.2. The concentration of total production

Fig. 1 shows the distribution of production employment by postcode area. A postcode area is based on the first two letters of the postcode, for example BS for Bristol.[9] They represent geographic areas that cross local authority and county borders and which are centred around cities or towns, which might be thought of as centres of economic activity. Thus postcode areas correspond more closely to areas of local economic activity than do administrative areas. Fig. 1 shows that the two postcode areas with the highest proportion of total production employment are Central London 5.0% (in the South East of England) Birmingham 4.3% (in the West Midlands). Around 30% of firms are located in the South East of England, and these represent around 23% of total production employment.

Table 3 shows measures of agglomeration calculated for total production activity—γ is the MS index, defined in Eq. (6) and γ_{EG} is the EG index, defined in footnote 6. The measures of raw geographic concentration, G (defined in Eq. (7)), and the industrial concentration measure H (defined in Eq. (5)), are also shown, together with a locational Gini coefficient and concentration index described in Appendix A. For the purposes of examining the aggregate distribution, we construct these measures relative to a uniform distribution; that is, we set $x_i = 1/M$, where M is the total number of regions.

One issue that arises in constructing these measures is what level of regional unit to use for analysis. We consider two levels of administrative regions in the UK, county (65) and local authority (447). Column one uses local authority boundaries to define geographic regions. Column three uses counties. Our preferred regional unit is the postcode area, shown in column two of the table and in Fig. 1, as this corresponds most closely to areas of economic activity.

Table 3 indicates, unsurprisingly, that moving to a larger geographic unit increases the geographic concentration measures G, and consequently γ, γ_{EG}; the concentration index also increases as the number of geographic regions decreases.

3.3. Agglomeration at the industry level

Fig. 2 shows the distribution of γ over the 211 four-digit industries.[10] The overall pattern of agglomeration at the four-digit industry level for 1992 in the UK looks similar to that found in the US and France. The agglomeration measures have similarly skewed

[8] These calculations are based on the UK's 11 administrative regions, 8 for England (South East, East Anglia, South West, West Midlands, East Midlands, Yorkshire and Humberside, North West, and Northern), Scotland, Wales and Northern Ireland.

[9] Each UK postcode identifies an average of 15 individual delivery points. They have four levels. There are 124 areas which have an average of 183,000 delivery points. These are divided into 2900 districts of which there are an average of 21 per area and which have an average of 8197 delivery points within them. These are further broken down into 9000 sectors and within this into units. For example, the postcode GU9 8AQ is in the area GU (Guildford), the district GU9, the sector GU9 8 and the units are identified by GU9 8AQ.

[10] The distribution of γ_{EG}, not shown, is very similar.

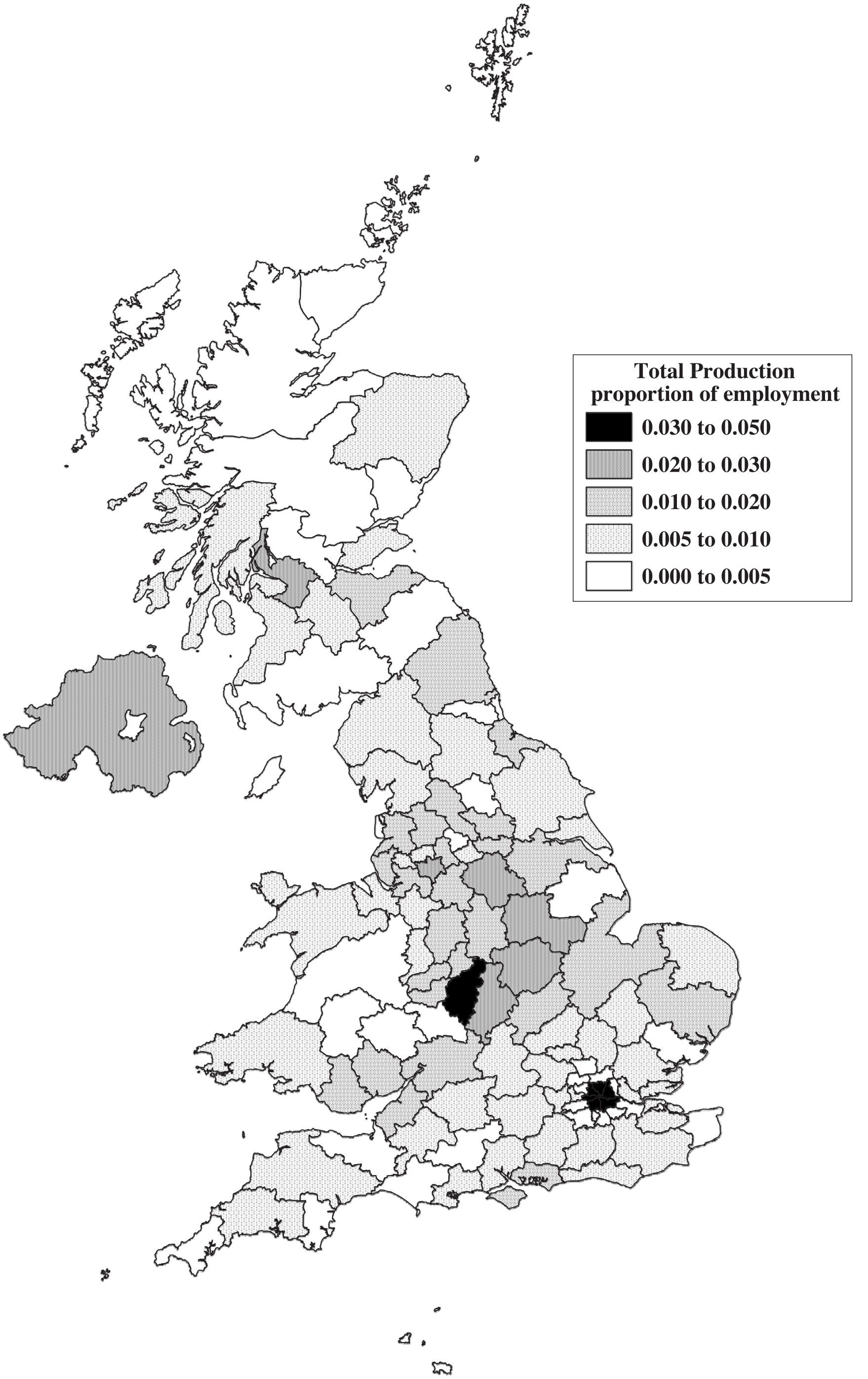

Fig. 1. Geographic distribution of total manufacturing employment.

Table 3
Geographic and industrial concentration measures for total production, 1992

Number of regional units	Local authority, 447	Postcode area, 113	County, 65
G	0.0040	0.0066	0.0147
H	0.0003	0.0003	0.0003
γ	0.004	0.006	0.014
γ_{EG}	0.004	0.006	0.014
Locational Gini	0.482	0.406	0.441
Concentration index	0.218	0.259	0.331

Measures are: G: geographic concentration (Eq. (7)), H: industrial concentration (Eq. (5)), γ: Maurel and Sédillot (1999) agglomeration index (Eq. (6)), γ_{EG}: Ellison and Glaeser (1997) agglomeration index (footnote 6), locational Gini, and concentration index calculated on firms (see Appendix A). The eight central London postcode areas are aggregated to form a single postcode area. Fourteen central London local authorities are aggregated to form a single local authority. Greater London, which covers a larger geographic area, is aggregated to form a single county.

distributions in all three countries (see MS Fig. 1 and EG Fig. 1). The mean value of γ is 0.033. Fifty-nine industries have a value of γ below zero, with half lying below 0.006. The mean value of γ_{EG} for British industries is 0.033 and the median 0.007. Ellison and Glaeser (1997) calculate γ_{EG} across 459 manufacturing industries in the US. They also found the distribution of γ_{EG} to be skewed (their Fig. 1), with a mean value of 0.051 and median value of 0.026. Both EG and MS report the number of industries with γ_{EG} below 0.02 (not very agglomerated), between 0.02 and 0.05 (somewhat agglomerated) and above 0.05 (very agglomerated). Using these definitions we see that in the US 10% of industries are classified as not very agglomerated, in France 50% and in the UK 65%. In the intermediate range EG find 65% of industries for the US, MS find 23% for France and we find 19% for the UK. EG find 25% of industries to be in the high agglomeration range for the US, MS find 27% for France and we find 16% for the UK. These comparisons suggest that British industry is

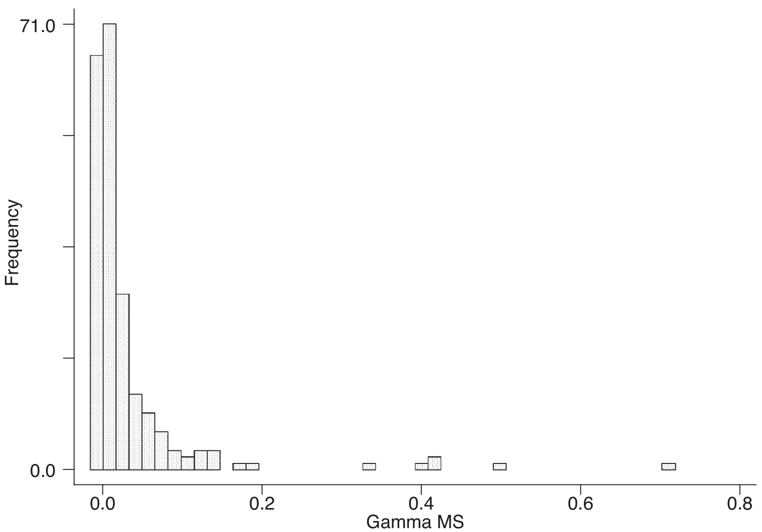

Fig. 2. Distribution of gamma MS agglomeration index.

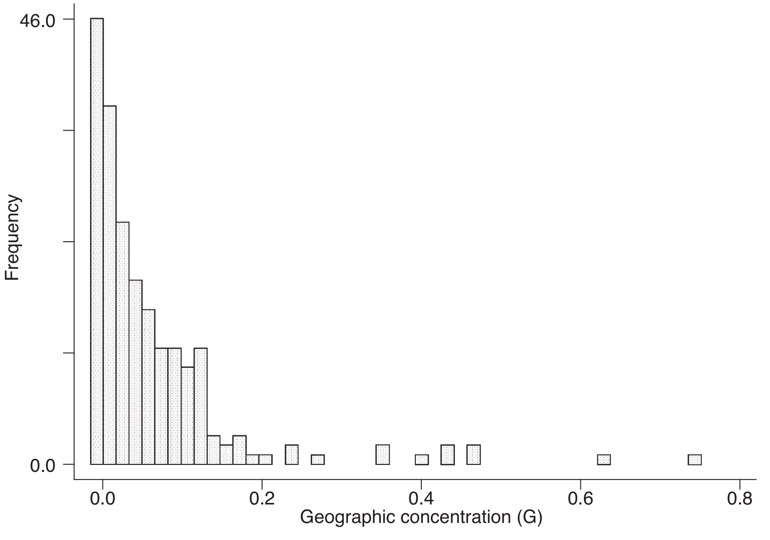

Fig. 3. Distribution of G adjusted geographic concentration measure.

somewhat less agglomerated, on average, than either the US or France. However, care needs to be taken in interpreting these numbers as we use a more disaggregated regional definition, and there are differences in the numbers and sizes of industries across countries.

Figs. 3 and 4 show the distributions of G, and H across four-digit industries. Geographic concentration is considerably less skewed than industrial concentration.

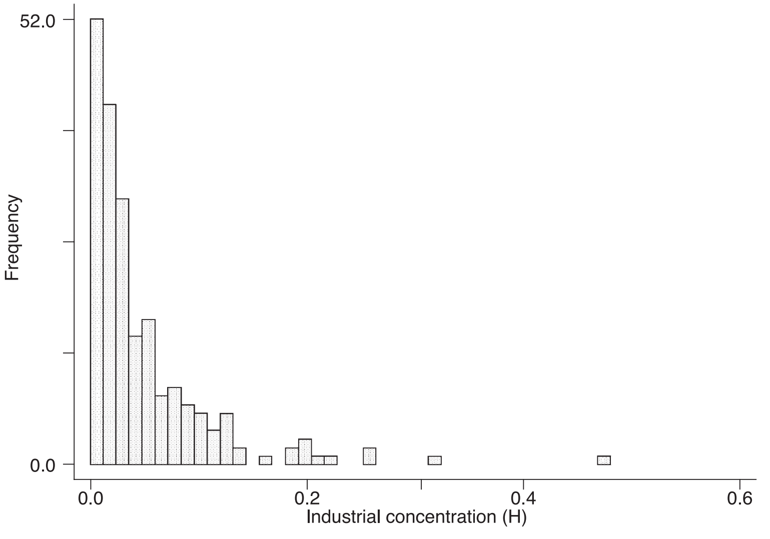

Fig. 4. Distribution of H industrial concentration.

Appendix A presents correlations between the measures and between each measure and the number of firms observed in each industry. The correlation between the agglomeration measures γ and γ_{EG} is positive and high. The rank correlation between these measures is also high. However, as would be expected, the correlation with the number of firms in the industry is in each case low. By contrast there is a strong negative correlation between the locational Ginis (which do not condition on industrial concentration), and the number of firms, and a weaker negative correlation between the concentration index and the number of firms in the industry.

Table 4 summarises the pattern of agglomeration at the four-digit industry level by showing the mean of γ (calculated at the four-digit industry level) by two-digit industry, and the percentage of four-digit industries in each quartile of γ across all four-digit industries (the fourth quartile contains the most agglomerated industries).

Extraction of other minerals (23) and textiles (43) top the table with mean values of γ far in excess of all other two-digit industries. The agglomeration of extraction of minerals (this includes stone, clay, sand, gravel, salt) is clearly driven by the fact that their main

Table 4
Summary of agglomeration in four-digit industries, by two-digit industry

Two-digit industry	Mean γ[a]	Percentage of four-digit industries in quartile (by γ)				Number of four-digit industries
		1 (least)	2	3	4 (most)	
23 Extraction of other minerals	0.172	0	33	33	33	3
43 Textiles	0.149	7	7	13	73	15
31 Other metal goods	0.053	0	29	29	43	14
24 Non-metallic mineral products	0.041	25	42	8	25	12
44 Leather	0.039	0	0	50	50	2
36 Other transport equipment	0.038	17	17	50	17	6
14 Mineral oil processing	0.036	0	0	0	100	2
35 Motor vehicles and parts	0.036	20	0	40	40	5
47 Paper and paper products	0.035	18	18	36	27	11
16 Electricity, gas, other energy	0.033	0	0	0	100	1
45 Footwear and clothing	0.032	0	15	46	38	13
22 Metal manufacturing	0.030	0	14	43	43	7
26 Man-made fibres	0.027	0	0	100	0	1
41 Food, drink and tobacco	0.025	23	23	15	38	13
49 Other manufacturing	0.023	29	43	14	14	7
48 Rubber and plastic	0.011	44	33	11	11	9
32 Mechanical engineering	0.008	23	31	38	8	26
42 Sugar and its by-products	0.007	36	45	9	9	11
34 Electrical and electronic engineering	0.006	53	20	7	20	15
37 Instrument engineering	0.003	50	17	33	0	6
25 Chemicals	0.002	45	25	25	5	20
46 Timber and wooden furniture	− 0.000	33	56	11	0	9
33 Office machinery, data processing equipment	− 0.002	50	0	50	0	2
17 Water supply industry	− 0.012	100	0	0	0	1

Quartile boundaries are by γ (Eq. (6)), 1: (− 0.156, − 0.001), 2: (− 0.001, 0.006), 3: (0.006, 0.027), 4: (0.027, 0.711).

[a] Mean is unweighted.

Table 5
Twenty most agglomerated industries

Four-digit industry	Number of firms	Agglomeration, γ	Geographic concentration, G	Industrial concentration, H
4340 Spinning and weaving of flax*	26	0.711	0.749	0.132
2330 Extraction salt*	5	0.499	0.626	0.253
4350 Jute and polypropylene*	31	0.414	0.474	0.101
2489 Ceramic goods*	744	0.410	0.432	0.037
4395 Lace*	86	0.402	0.432	0.050
3162 Cutlery*	75	0.338	0.412	0.112
3634 Pedal cycles	73	0.191	0.364	0.213
4363 Hosiery*	1341	0.168	0.177	0.010
4910 Jewellery*	1916	0.146	0.151	0.006
3161 Handtools*	324	0.139	0.169	0.035
4752 Periodicals*	2079	0.135	0.143	0.009
4310 Woollen and worsted industry*	508	0.119	0.129	0.011
3523 Caravans	85	0.118	0.155	0.041
4721 Wall coverings	33	0.118	0.199	0.093
4322 Weaving cotton, silk*	267	0.112	0.130	0.020
4831 Plastic coated textile fabric*	18	0.111	0.286	0.197
2235 Other steel forming*	58	0.092	0.124	0.034
4240 Spirit distilling	95	0.091	0.139	0.053
4537 Hats	126	0.082	0.110	0.031
4150 Fish processing	290	0.081	0.120	0.043

Measures are: γ: agglomeration index (Eq. (6)), G: geographic concentration measure (Eq. (7)); H: industrial concentration (Eq. (5)).

* Indicates that the industry was also in the top 20 in 1985.

inputs are physically immobile and geographically concentrated. MS also find extraction industries to be among the most localised in France.

The four-digit industry spinning and weaving of flax (4340) is the most agglomerated industry (see Table 5) and of the 15 four-digit industries within the textiles industry, 6 are amongst the 15 most agglomerated industries and 11 are in the fourth quartile. Textiles industries are found to be highly agglomerated in many countries.[11] In the UK this is a sector in which plants are small (the median number of employees in plants in the textiles sector is 7), a high proportion of workers are unskilled (80% of workers are operatives compared to an average of 70% across all plants and average wages of operatives are 20% below the national average), and where labour market externalities are likely to be among the determinants of agglomeration. Ellison and Glaeser (1999) investigate the extent to which regions' labour market characteristics and endowments of natural resources can explain geographic concentration. They find access to unskilled labour to be the most important factor for the textiles and apparel industries in the US. It appears from our evidence that this could also be true in the UK.

The other notable feature of Table 4 is that the group of industries at the bottom of the table, with low mean γ, contains several high-tech industries, for example office machinery

[11] See Section 4 of this paper, Maurel and Sédillot (1999, Table 2), Ellison and Glaeser (1997, Table 4), Krugman (1991, Appendix D).

and data processing equipment (33), and electrical and electronic engineering (34). In general, it appears that the more technologically advanced industries are less agglomerated. We investigate the relationship between geographic concentration, industrial concentration, agglomeration and technological intensity further in the next section. One reason why more high-tech industries might be less agglomerated is that they are newer. Agglomeration is a dynamic process and geographic concentration in these industries might still be at an early stage. In Section 5, we consider entry patterns in agglomerated and non-agglomerated industries. A second reason that more high-tech industries might be observed to be less agglomerated is that, while technological spillovers are important, developments in communications and transportation mean that geographic proximity is now less important in enabling firms to capitalise on these knowledge spillovers. A third reason is that the externalities may be sufficiently high in these industries that firms have internalised them by merging, so that high industrial concentration accounts for the geographic concentration.

Table 5 shows the 20 most agglomerated four-digit industries as measured by γ. The table shows the number of firms in each industry, the geographic concentration measure G and the industrial concentration measure H. While all of these industries display high geographic concentration relative to the distribution of total production employment it is interesting to note the variation in industrial concentration. For example, ceramic goods (2489), has high geographic concentration and low industrial concentration, whereas pedal cycles (3634) has quite high geographic concentration coupled with high industrial concentration.

Table 6 shows some additional information on the top 20 agglomerated industries. Columns 2 and 3 list the two postcode areas with the highest proportion of industry employment. In some industries the first and second postcodes are adjacent to each other and may thus indicate a larger agglomeration. Examples include lace in Nottingham and Derby (see Fig. 5), hosiery in Leicester and Nottingham (see Fig. 6), and weaving of cotton and silk in Blackburn and Oldham. Columns four and five show the proportion of industry employment in these two postcodes areas. The proportion of industry employment in the top postcode area ranges from 86% in the most agglomerated industry—the spinning and weaving of flax—to 22% in the other steel forming industry. Column 6 shows the total number of firms in the industry, and columns 7 and 8 show the proportion of firms in the two postcode areas. For the top postcode area this ranges from 68% in the cutlery industry to only 5% in the fish processing industry. Finally, columns 9 and 10 show the average firm size in the top postcode area and in all others.

From this table, two types of industries arise, those that have a single agglomeration that contains a large number of firms and those that contain two agglomerations of employment or that have only a few large firms in the most agglomerated region.[12] Examples of the single agglomeration type include ceramic goods in Stoke-on-Trent (see Fig. 7) and periodicals in London. Examples of the second type include jewellery,

[12] The most agglomerated region is defined by the postcode area with the highest proportion of industry employment.

Table 6
Most agglomerated regions, 1992

Four-digit industry	1st postcode area	2nd postcode	Percentage of employment in postcode		Total number firms	Percentage of firms in postcode		Average firm size (employment)	
			1st	2nd		1st	2nd	1st	Other
4340 Spinning and weaving of flax	Northern Ireland	[a]	86.3	[a]	26	57.7	[a]	134	29
2330 Extraction salt	[a]	[a]	[a]	[a]	5	[a]	[a]	[a]	[a]
4350 Jute and polypropylene	Dundee	[a]	67.6	[a]	31	32.3	[a]	204	47
2489 Ceramic goods	Stoke-on-Trent	Derby	66.0	4.1	744	31.3	3.0	120	28
4395 Lace	Nottingham	Derby	63.8	14.2	86	60.5	11.6	36	31
3162 Cutlery	Sheffield	[a]	58.2	[a]	75	68.0	[a]	39	60
3634 Pedal cycles	[a]	Birmingham	[a]	17.7	73	[a]	19.2	[a]	21
4363 Hosiery	Leicester	Nottingham	38.2	17.9	1341	39.7	6.0	40	43
4910 Jewellery	Birmingham	London	33.7	21.3	1916	18.4	25.5	11	5
3161 Handtools	Sheffield	Walsall	40.7	7.7	324	18.8	2.8	51	17
4752 Periodicals	London	Tunbridge Wells	38.4	3.1	2079	29.7	3.0	26	18
4310 Woollen and worsted industry	Bradford	Huddersfield	28.8	17.3	508	17.7	16.3	95	50
3523 Caravans	Hull	Bournemouth	37.1	9.3	85	18.8	5.9	148	58
4721 Wall coverings	Blackburn	[a]	40.1	[a]	33	27.3	[a]	166	93
4322 Weaving cotton, silk	Blackburn	Oldham	32.8	13.0	267	14.6	6.0	111	39
4831 Plastic coated textile fabric	[a]	[a]	[a]	[a]	18	[a]	[a]	[a]	48
2235 Other steel forming	Birmingham	Sheffield	21.6	19.2	58	22.4	10.3	74	78
4240 Spirit distilling	Glasgow	Edinburgh	30.8	12.3	95	13.7	8.4	290	103
4537 Hats	Luton	[a]	29.2	[a]	126	23.8	[a]	26	20
4150 Fish processing	Doncaster	Aberdeen	29.3	14.8	290	4.5	14.8	502	57

[a] Figures cannot be provided for data confidentiality reasons.

which has two relatively similar sized agglomerated regions, one in Birmingham and one in London (see Fig. 8), and fish processing, which has a significant presence in two regions (Doncaster and Aberdeen). In this latter case the two regions are quite different—one (Doncaster) has a small number of firms but the highest proportion of employment, the other (Aberdeen) has a larger number of small firms, which account for a much lower proportion of employment. In this industry average firm size is much higher in the most agglomerated region than in other postcode areas.

Many of these agglomerations date back a number of centuries. Some are referred to in Marshall's (1890) discussion of localised industries. For example, he refers to the location of extractive industries, such as fishing, being driven by access to natural resources, and indeed the two postcode areas containing the highest proportion of employment in the fish processing industry are both located on the coast. He also refers to the localisation of the pottery industry in Staffordshire and the cutlery industry in Sheffield, both of which are still seen to be highly agglomerated today. A map of industrial Britain published between

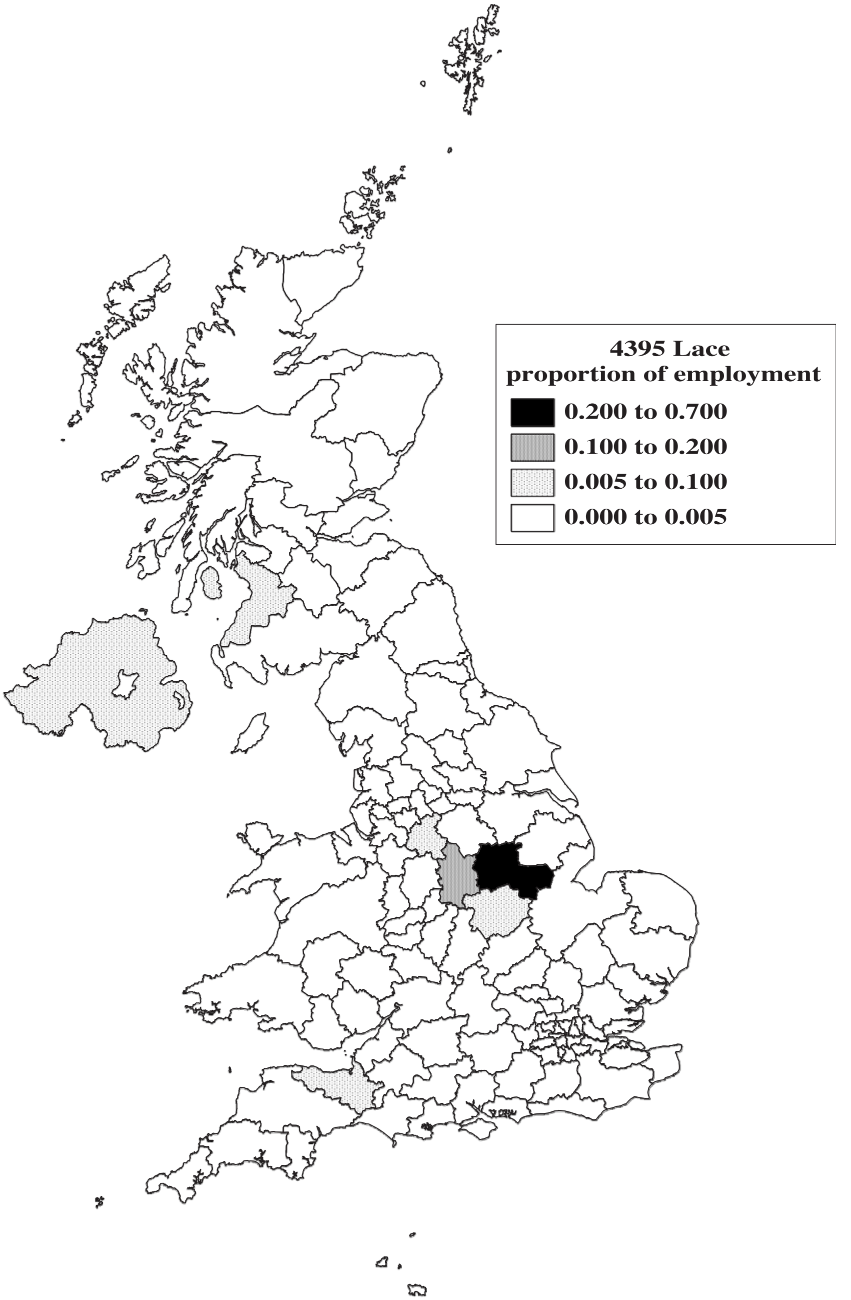

Fig. 5. Lace—geographic distribution of employment.

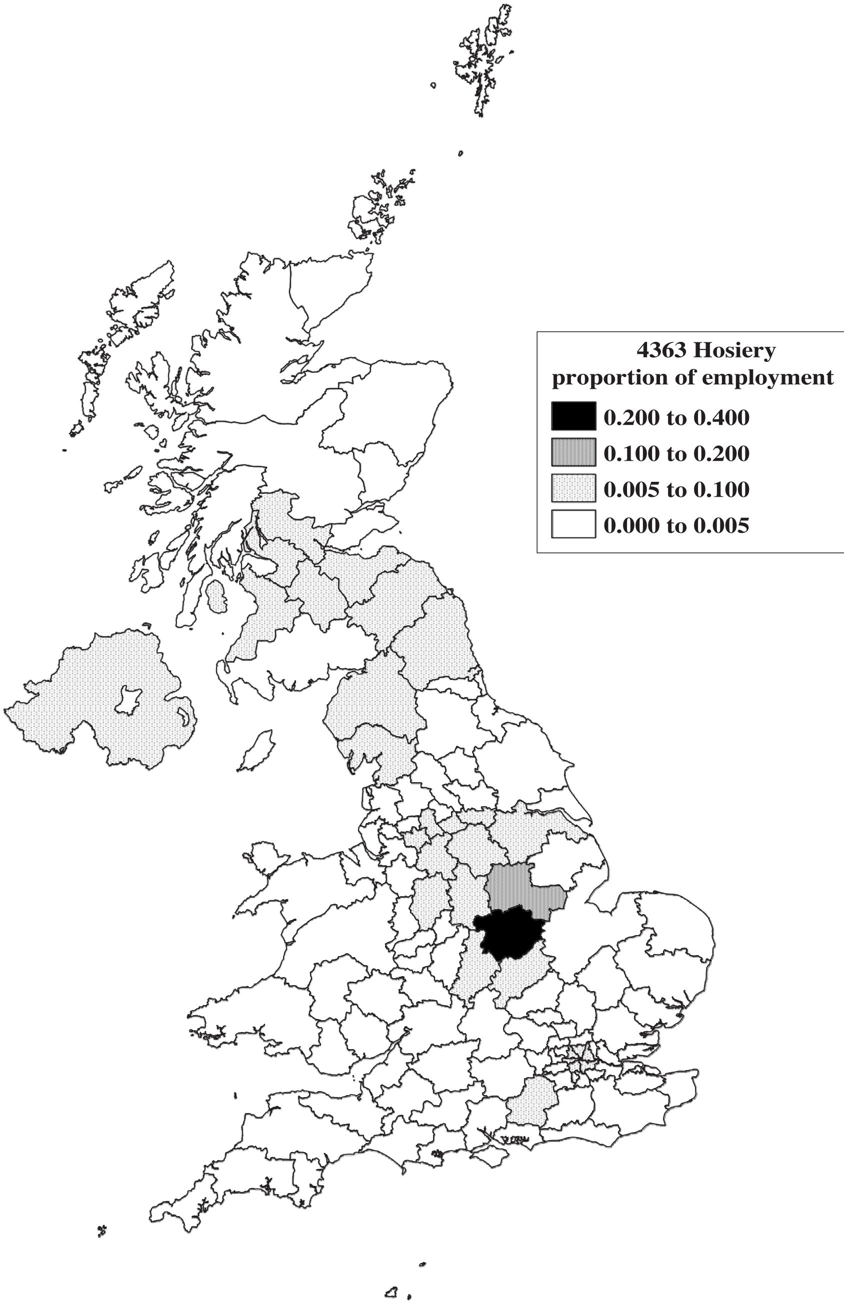

Fig. 6. Hosiery—geographic distribution of employment.

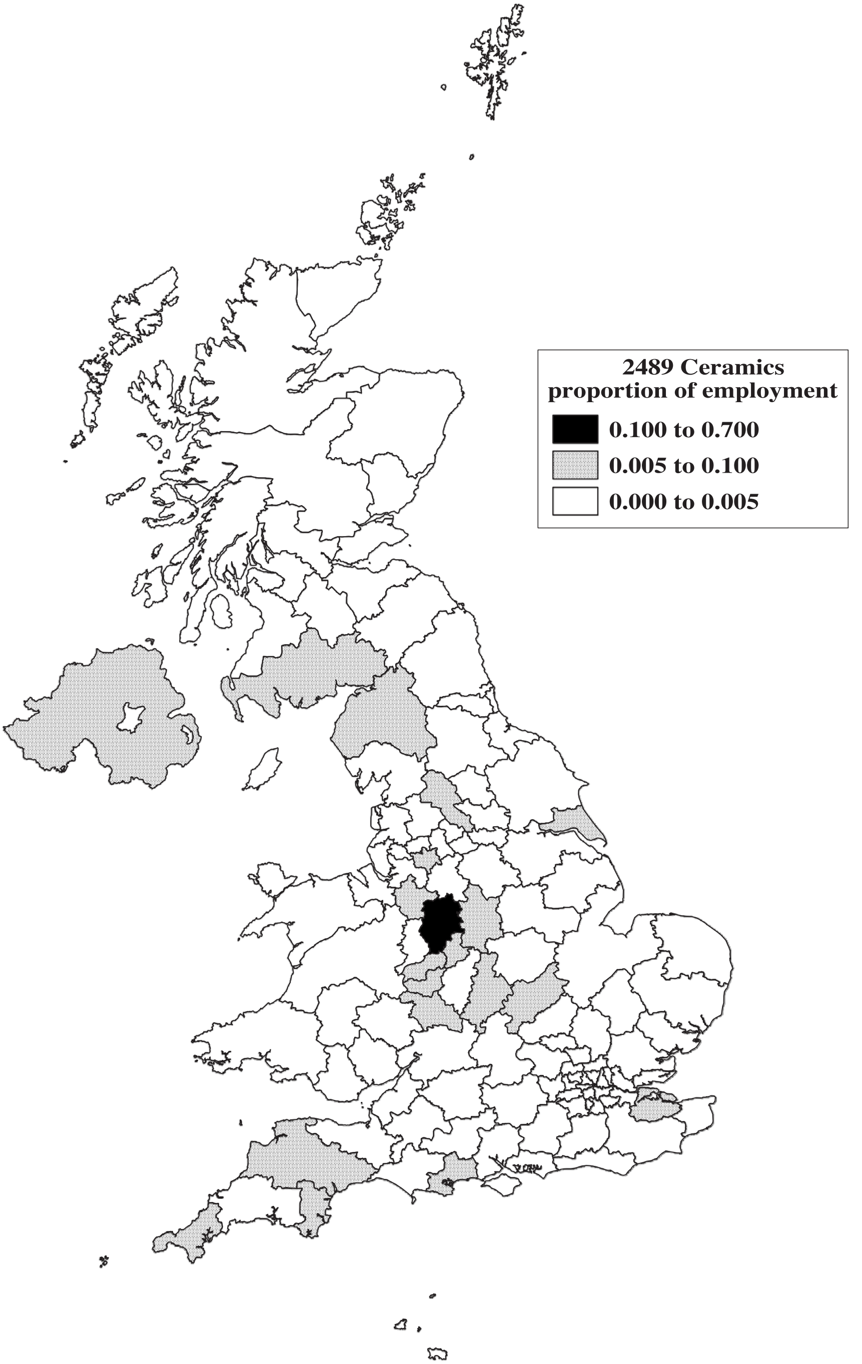

Fig. 7. Ceramics—geographic distribution of employment.

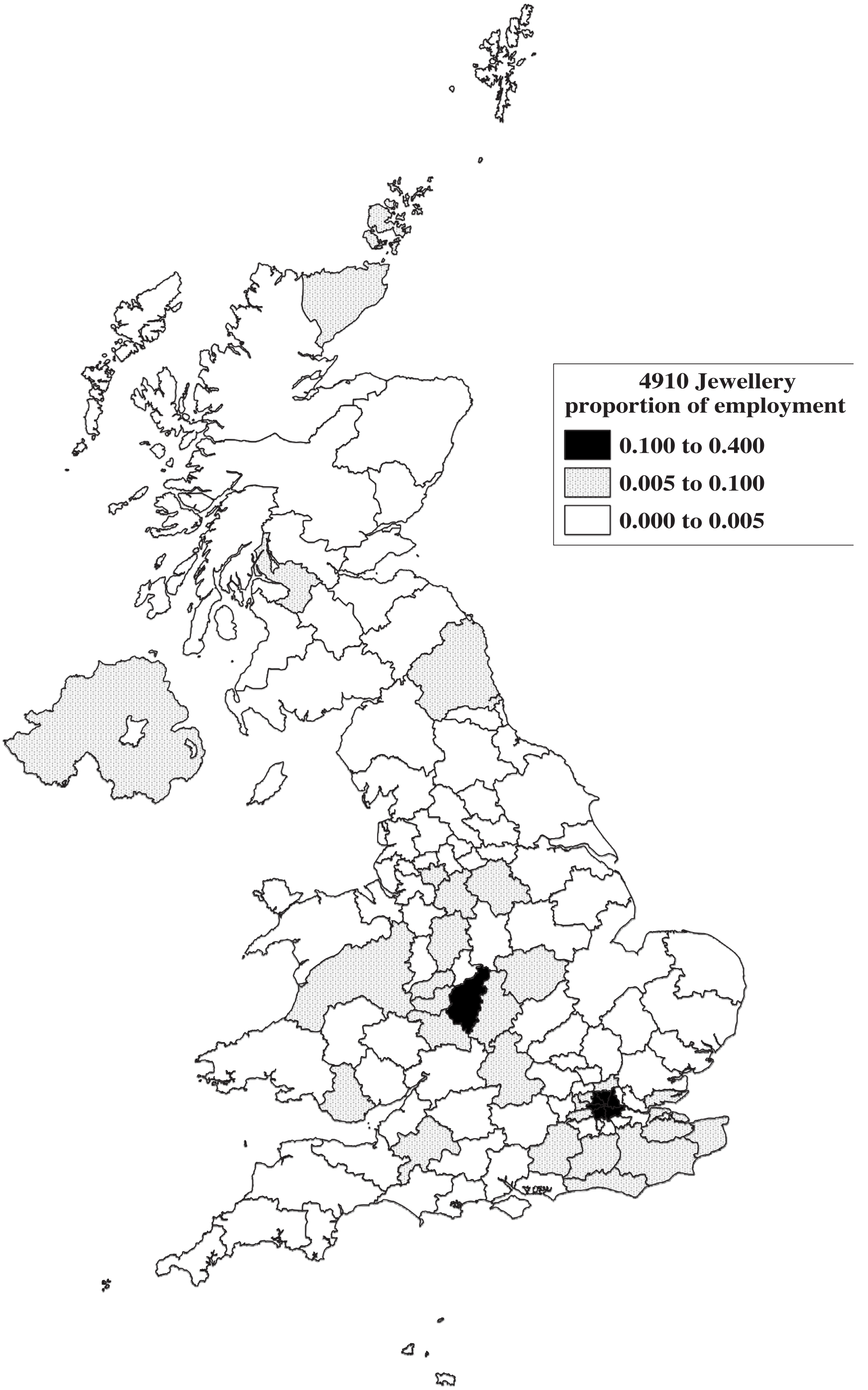

Fig. 8. Jewellery—geographic distribution of employment.

Table 7
Twenty least agglomerated industries

Four-digit industry	Number of firms	Agglomeration, γ	Geographic concentration, G	Industrial concentration, H
4200 Sugar and its by-products	15	− 0.016	0.145	0.159
2515 Synthetic rubber	17	− 0.016	0.210	0.222
2591 Photographic materials and chemicals	62	− 0.015	0.187	0.200
2440 Asbestos goods	39	− 0.015	0.115	0.128
2569 Misc. chemicals for industrial use	85	− 0.013	0.087	0.098
4664 Cork and basketware	33	− 0.013	0.132	0.129
2565 Explosives	40	− 0.012	0.092	0.103
1700 Water supply industry	44	− 0.012	0.090	0.101
4290 Tobacco industry	33	− 0.012	0.102	0.112
3290 Ordinance small arms	106	− 0.011	0.048	0.059
3301 Office machinery	139	− 0.011	0.062	0.072
2420 Cement, lime and plaster	255	− 0.010	0.030	0.040
4811 Rubber tyres	38	− 0.009	0.106	0.114
3435 Electrical equipment for industrial use	401	− 0.007	0.021	0.028
3212 Wheeled tractors	30	− 0.006	0.255	0.260
4833 Plastics floorcoverings	54	− 0.006	0.131	0.136
2513 Fertilisers	105	− 0.006	0.176	0.181
4396 Rope	143	− 0.006	0.058	0.064
2570 Pharmaceutical products	396	− 0.006	0.010	0.016
3441 Telegraph and telephone equipment	531	− 0.006	0.065	0.070

Measures are: γ: agglomeration index (Eq. (6)), G: geographic concentration measure (Eq. (7)); H: industrial concentration (Eq. (5)).

the two World Wars[13] shows jewellery located in Birmingham, lace in Nottingham, cutlery in Sheffield and hosiery in Leicester—locations where we still see high concentrations of firms in these industries.

Using the same data source as this paper, Duranton and Overman (2002) investigate the extent of agglomeration or localisation using a continuous measure of distance, rather than the discrete postcode areas used here. They find that most localisation of activity occurs at distances below 50 km. They too find that the distribution of localisation across industries is skewed, and some of the industries that they find to be most localised (their Table 2) overlap with our findings in Table 5, including cutlery, hosiery and the woollen and worsted industry.[14]

Table 7 shows the 20 least agglomerated industries. For these industries the measures of geographic concentration G and industrial concentration H are very close to each other. This means that, although production is to some extent unequally distributed across regions, compared to the distribution of total production employment, the geographic dispersion of employment is largely explained by industrial concentration. For eight industries in Table 7 the geographic concentration measure G

[13] “The Peoples’ Atlas” published by George Philip & Son for the International Review.

[14] While we find fish processing to be one of the most agglomerated industries, Duranton and Overman find it to be dispersed geographically. This may be because they use a measure based on the number of plants as opposed to an employment-weighted measure, or it may be related to the measure used.

is actually higher than for the hats (4537) industry, which is the least geographically concentrated industry in Table 5, and four are above the median level in Table 5. Other industries, such as electrical equipment for industrial use and pharmaceutical products, simply have low levels of geographic concentration and low levels of industrial concentration.

3.4. The role of technology

It is striking that a number of the least agglomerated industries listed in Table 7 are high-tech, while those listed in Table 5 are relatively low tech. Given the emphasis in the literature on knowledge spillovers as an important source of agglomeration externalities, high-tech industries might be expected to be amongst the most highly agglomerated. Krugman (1991) points out that, although high-tech agglomerations such as Silicon Valley receive much attention, many localised industries are far from high-tech. He also notes that the reason for the formation of high-tech clusters should not simply be assumed to be the presence of positive externalities from knowledge spillovers. Other factors may also be at work.

In order to examine this further, we first investigate the relationship between agglomeration and two measures related to the technological level of the industry: the capital to labour ratio (although this could also pick up returns to scale) and the proportion of the workforce that is skilled (administrative, technical and clerical workers). These measures are available from the ARD data for a large sample of establishments. We construct an unweighted average across plants of each of these variables for each four-digit industry for the years 1985 to 1992. In the top two sections of Table 8 we show how they are correlated with geographic and industrial concentration and agglomeration by regressing G, H and γ on both measures—the capital to labour ratio in the first section of the table and the proportion of skilled workers in the middle section. In each case the first column includes year dummies while the second column also controls for mean differences across two-digit industries. There is a significant positive correlation between capital intensity and geographic concentration G, and a weak correlation with H. This results in there being no statistically significant correlation with the agglomeration measure γ. In the middle section of the table we find that the proportion of skilled workers employed is positively correlated with industrial concentration H, it is not with geographic concentration G, resulting in a significant negative correlation with agglomeration γ. However, this correlation is not present in the presence of two-digit industry effects.

In the final section of the table we correlate the measures of geographic and industrial concentration and agglomeration with OECD indicators of high technology sectors, defined on the basis of research and development intensity (see OECD, 1997). We regress the mean values of G, H and γ (averaged over the period 1985 to 1992 for each four-digit industry), on a series of dummies for the technological intensity of the industry, which vary at the two-/three-digit level. We find that medium–high-technology industries exhibit significantly higher industrial concentration, and significantly lower agglomeration than low-technology industries (the omitted category). These results are consistent with the informal impression gained from Tables 5, 6 and 7 that there is little evidence that high-

Table 8
The role of technology

Dependent variable	Geographic concentration, G		Industrial concentration, H		Agglomeration, γ	
Number of observations	*1688*	*1688*	*1688*	*1688*	*1688*	*1688*
Capital/labour ratio	2.856	1.434	3.442	1.474	− 0.474	0.072
	(1.339)	(0.785)	(1.629)	(0.823)	(0.339)	(0.205)
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Two-digit industry dummies	No	Yes	No	Yes	No	Yes
Proportion of workforce skilled (Administrative, Technical and Clerical)	− 0.020	0.042	0.087	0.027	− 0.117	0.016
	(0.023)	(0.032)	(0.014)	(0.020)	(0.022)	(0.024)
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Two-digit industry dummies	No	Yes	No	Yes	No	Yes
Number of observations		*205*		*205*		*205*
High-technology		− 0.054		− 0.006		− 0.051
		(0.037)		(0.024)		(0.028)
Medium–high-technology		− 0.012		0.024		− 0.037
		(0.016)		(0.011)		(0.012)
Medium–low-technology		− 0.005		0.002		− 0.009
		(0.018)		(0.011)		(0.014)

Standard errors are in brackets. Estimates in top two sections based on a panel of 211 four-digit industries, 1985 to 1992. Capital/labour ratio is defined as capital stock (£m)/number of employees. Estimates in final section based on 205 four-digit industries, where G, H and γ are means across 1985 to 1992. The excluded category is low technology industries. Industries are categorised according to OECD (1997). High-technology sectors (8 four-digit industries): aerospace, office and computing equipment, drugs, radio, TV and communication equipment. Medium–high-technology sectors (71): scientific instruments, motor vehicles, electrical machines excluding communication equipment, chemicals excluding drugs, other transport, non-electrical machinery. Medium–low-technology sectors (52): rubber and plastic products, shipbuilding, other manufacturing, non-ferrous metals, non-metallic mineral products, metal products, petroleum refineries and products, ferrous metals. Low-technology sectors (74): paper and printing, textiles, apparel and leather, food, drink and tobacco and wood products and furniture. Six four-digit industries are not categorised.

tech industries are more geographically concentrated (although they are more industrially concentrated).

These results seem to cast some doubt on the role of knowledge spillovers as being a major driving force in industry agglomeration, at least in the UK. However, while being suggestive, we would not make strong claims for these results. Among other things, there is some evidence that there are smaller clusters of firms in high-tech industries. For example, in the pharmaceutical sector knowledge spillovers may be important within each cluster, but other factors may prevent these clusters being located together in a way that would reveal a highly agglomerated industry. It may therefore be appropriate to investigate the existence of such clusters (for example in the Cambridge area) at a narrower industry definition than four-digit industries.[15]

[15] See, for example, Swann et al. (1998).

3.5. Co-agglomeration

So far, we have considered the location of firms within the same industry. Now we consider the location of firms that are in related industries, on the grounds that externalities may also generate geographic concentrations between industries, for example if two industries are vertically related, or use similar technologies or skills. It is possible to use the measures above to analyse the overall pattern of geographic and industrial concentration for any set of plants, whether they come from one or more industries. But it is also informative to consider the extent to which concentrations arise within and between groups of industries. This can be done using a decomposition of the agglomeration measure calculated over a group of industries.

MS demonstrate that the relationship between γ evaluated at the level of a two-digit industry (denoted as γ_2) and evaluated for the four-digit industries (denoted as γ_j) contained within that two-digit industry is:[16]

$$\gamma_2 = \frac{\sum_{j=1}^{r} \gamma_j \omega_j^2 (1 - H_j)}{1 - \sum_{j=1}^{r} \omega_j^2 H_j} + \frac{C\left(1 - \sum_{j=1}^{r} \omega_j^2\right)}{1 - \sum_{j=1}^{r} \omega_j^2 H_j} \tag{8}$$

where there are r four-digit industries within the two-digit industry, $w_i = T_i / \sum_{j=1}^{r} T_j$ and T_i is total employment in industry i. In this expression, the first term on the right hand side represents the weighted average of γ_j for each of the four-digit industries; this is a summary of the *within* four-digit industry agglomeration. The second term represents the difference between this and the overall agglomeration of the two-digit industry. The term C is equivalent to EG's 'co-agglomeration' index. This measures the degree to which there is agglomeration *between* the four-digit industries.

Table 9 presents estimates of both C (column 1) and γ_2 (column 2) for each two-digit industry. It also shows the proportion of the overall agglomeration (γ_2) accounted for by between-industry agglomeration, measured as the last term in Eq. (8) expressed as a proportion of γ_2 (column 3).[17] Industries are listed in numerical order by SIC code.

There is considerable variation in the two-digit industry agglomeration indices, γ_2. Further, the more agglomerated two-digit industries are typically those with a high co-agglomeration. Thus, for example, mineral oil processing, textiles and leather, all have high values of γ_2, and also all have high co-agglomeration indices. Despite this, the overall agglomeration in these industries is not just driven by co-agglomeration, as indicated in

[16] Clearly, this approach could be used to study co-agglomeration between any group of industries; however, like MS we consider the relationship between an overall two-digit industry and the four-digit industries that comprise it.

[17] Note that the first column of Table 4 presents estimates of the unweighted average of the four-digit agglomeration indices, for each two-digit industry.

Table 9
Co-agglomeration by two-digit industry

Two-digit industry	C	γ_2	Percentage between four-digit industries
14 Mineral oil processing	0.046	0.041	55
22 Metal manufacturing	0.015	0.018	67
23 Extraction of other minerals	− 0.001	0.007	− 6
24 Non-metallic mineral products	0.005	0.024	17
25 Chemicals	0.000	0.000	20
31 Other metal goods	0.018	0.018	83
32 Mechanical engineering	− 0.001	− 0.001	126
33 Office machinery, data processing equipment	0.000	0.005	1
34 Electrical and electronic engineering	− 0.002	− 0.002	99
35 Motor vehicles and parts	0.016	0.022	41
36 Other transport equipment	− 0.004	− 0.003	73
37 Instrument engineering	0.000	0.000	206
41 Food, drink and tobacco	− 0.001	− 0.000	250
42 Sugar and its by-products	− 0.002	− 0.001	151
43 Textiles	0.017	0.036	38
44 Leather	0.023	0.034	33
45 Footwear and clothing	0.008	0.011	63
46 Timber and wooden furniture	− 0.001	− 0.000	249
47 Paper and paper products	0.016	0.018	75
48 Rubber and plastic	− 0.002	− 0.002	92
49 Other manufacturing	0.006	0.008	57

Industries 16 (Electricity, gas, other energy), 17 (Water supply) and 26 (man-made fibres) do not appear because they contain only 1 four-digit industry.

Table 4, and in the last column of Table 9. In fact, for these industries, only relatively small proportions of the high values of γ_2 can be attributed to *between* four-digit industry agglomeration. Other industries—such as mechanical engineering, instrument engineering, and food, drink and tobacco—have much higher proportions (in excess of 100%) attributable to *between* industry agglomeration. For a two-digit sector such as ‘42 Sugar and its by-products’, the low level of agglomeration γ_2 is driven by positive within four-digit industry agglomeration, but negative between industry co-agglomeration, as shown in column 1.

4. International comparisons

Empirical investigations into the extent of agglomeration have also been carried out in other countries, as referred to earlier, and it is interesting to look at the extent to which there are similarities or differences across countries. EG use a US state-industry employment dataset. This means that the US measure is based on a more aggregated regional unit (50 States plus the District of Columbia) than our calculations for the UK (which use 113 postcode areas). MS use French data at the department level (95 departments). In the following tables we have attempted to match the information given in the EG and MS articles with our data. The industry classifications differ across

Table 10
Comparison of γ for UK top 20 agglomerated industries

Four-digit industry	UK		US		France	
	γ	Rank	γ	Rank	γ	Rank
4340 Spinning and weaving of flax	0.711	1	0.28	13		
2330 Extraction salt	0.499	2				
4350 Jute and polypropylene	0.414	3				
2489 Ceramic goods	0.410	4				
4395 Lace	0.402	5				
3162 Cutlery	0.338	6			0.28	19
3634 Pedal cycles	0.191	7				
4363 Hosiery	0.168	8	0.44	3		
			0.40	5		
4910 Jewellery	0.146	9	0.32	8		
			0.30	10		
3161 Handtools	0.139	10				
4752 Periodicals	0.135	11			0.40	10
4310 Woollen and worsted industry	0.119	12			0.44	7
					0.42	9
					0.25	20
3523 Caravans	0.118	13				
4721 Wall coverings	0.118	14				
4322 Weaving cotton, silk	0.112	15				
4831 Plastic coated textile fabric	0.111	16				
2235 Other steel forming	0.092	17				
4240 Spirit distilling	0.091	18	0.48	2		
4537 Hats	0.082	19				
4150 Fish processing	0.081	20				

See Table A3 for industry mapping, regional units used are postcode areas for the UK, States for the US and departments for France.

countries[18] and there is therefore some approximation in this matching. In addition, even where the industry classifications match exactly, firms may engage in quite different activities across countries.

Table 10 shows which of the 20 most agglomerated industries in the UK were also found to be agglomerated in these US and French studies. Four of the most agglomerated UK industries were also identified by EG as being amongst the 15 most agglomerated industries in the US (spinning and weaving of flax, hosiery, jewellery and spirit distilling), while a different three were identified by MS as being amongst the 20 most agglomerated in France (cutlery, periodicals and woollen and worsted). Details of all industries were not available in EG or MS at the most disaggregated level, so we were only able to compare a limited range of industries in this way. A more complete matching and comparison across countries would certainly be interesting.

Table 11 shows how the corresponding British industries compare with the 15 most agglomerated US industries from the EG study. The three British industries that are in

[18] The data used here is based on the 1980 SIC revision for the UK. The more recent 1992 revision has been harmonised across countries.

Table 11
Comparison of γ_{EG} for US top 20 agglomerated industries

US	γ_{EG}	Rank	UK	γ_{EG}	Rank
2371 Fur goods	0.63	1	4560 Fur goods	n/a	n/a
2084 Wines brandy spirits	0.48	2	4240 Spirit distilling	0.099	17
2252 Hosiery not elsewhere classified	0.44	3	4363 Hosiery	0.159	8
3533 Oil and gas field machinery	0.43	4	3254 Construction equipment	0.011	91
2251 Women's hosiery	0.40	5	4363 Hosiery	0.159	8
2273 Carpets and rugs	0.38	6	4384 Pile carpets	0.064	24
			4385 Other carpets	0.050	31
2429 Special product sawmills not elsewhere classified	0.37	7	4610 Sawmilling	0.004	133
3961 Costume jewelry	0.32	8	4910 Jewellery	0.115	15
2895 Carbon black	0.30	9	2516 Dyestuff and pigments	0.035	47
3915 Jewelers' materials lapidary	0.30	10	4910 Jewellery	0.115	15
2874 Phosphatic fertilizers	0.29	11	2513 Fertilisers	0.003	150
2061 Raw cane sugar	0.29	12	4200 Sugar	− 0.036	211
2281 Yarn mills except wool	0.28	13	4340 Spinning and weaving of flax	0.693	1
2034 Dehydrated fruits vegetable soups	0.28	14	4147 Fruit and vegetables	0.044	35
			4239 Misc. foods (inc. soup)	0.004	140
3761 Guided missiles space vehicles	0.25	15	3640 Aerospace equipment	0.003	144

Industry mappings between UK and US industry codes are not exact.

the top 15 both in the UK and US match to five US industries. The fact that some industries are agglomerated in the US, but not in Britain, may be due to the very different nature of the industries. For example, Britain does not have a raw cane sugar industry—the closest match is the sugar industry, which mainly consists of processing. Other industries, where there is closer match, have somewhat different rankings. For example, carpets, which is famously agglomerated in the US, is ranked only within the top 40 industries in the UK.[19]

Finally, Table 12 shows the same comparison with the French top 15 agglomerated industries, based on the γ measure. Four of the French industries match to two British industries ranking in the top 15 in the UK. Many of the same comments apply. For example, made-to-measure clothing (potentially located in Paris) is likely to contain quite different firms than men's and boys' and women's and girls' tailored outerwear, industries that are quite dispersed in Britain.

An interesting line of future research would be to compare the characteristics of these industries across countries to see if this sheds light on the reasons for agglomeration—are there common characteristics across industries, or do the factors driving agglomeration differ? MS report that high-tech industries in France are ranked similarly to those in the US in the extent to which they are agglomerated. However, from the tables above there is clearly some variation across countries in the extent to which more narrowly defined industries exhibit agglomeration. A more complete analysis, for example relating measures

[19] See data in web appendix (http://www.ifs.org.uk/corpact/dgsdata.zip).

Table 12
Comparison of γ for French top 15 agglomerated industries

France	γ	Rank	British	γ	Rank
Extraction of slate	0.88	1	2310 Extraction of stone, clay	0.016	75
Extraction of iron ore	0.88	2	2100 Extraction and preparation of metalliferous ores	n/a	n/a
Made-to-measure clothing	0.80	3	4532 Men's and boys' tailored outerwear	0.007	99
			4533 Women's and girls' tailored outerwear	0.044	36
Extraction of minerals for chemical industry and fertilisers	0.76	4	2396 Extraction of other minerals n.e.s.	0.003	126
Steel pipes and tubes	0.69	5	2220 Steel tubes	0.023	65
Extraction of coal	0.53	6	1113 Deep coal mines	n/a	n/a
Combed wool spinning mills	0.44	7	4310 Woollen and worsted industry	0.119	12
Vehicles hauled by animals	0.42	8	3650 Other vehicles	0.000	149
Wool preparation	0.42	9	4310 Woollen and worsted industry	0.119	12
Periodicals	0.40	10	4752 Periodicals	0.135	11
Watch-making	0.38	11	3740 Clocks, watches	0.004	118
Flat glass	0.37	12	2471 Flat glass	0.001	138
Screw cutting	0.36	13	3137 Bolts, nuts, etc.	0.075	22
Lawn and garden equipment	0.36	14	3286 Other industrial and commercial machinery	– 0.001	161
Carded wool weaving mills	0.34	15	4310 Woollen and worsted industry	0.119	12

Industry mappings between UK and French industry codes are not exact.

of technological intensity to agglomeration measures across industries and countries could prove informative.

5. The dynamics of agglomeration

It is likely that different industries are at different stages of maturity, and therefore that both the rate at which plant births, deaths, expansions and contractions occur and their geographic location can have important effects on the extent of industry agglomeration. One issue of interest is therefore whether the observed patterns of agglomeration are changing over time, and to what extent entry, exit, job creation and job destruction play a role in reinforcing agglomerations. This is interesting for academics in understanding what drives agglomeration, and for policy makers in understanding the implications of agglomeration for policy formation.

Dumais et al. (2002) find that there was little change in the extent of agglomeration in the US over the period 1972 to 1992. As we have already pointed out, there are number of industries in Britain which appear to have been agglomerated for well over a century. However, Dumais et al. also find that underlying the stability, the entry of new firms played a role in reducing the extent of agglomeration while plant closures acted to bolster agglomeration. That is, there was more entry outside rather than inside agglomerations.

In this section, we look at entry and exit rates and job creation and job destruction patterns in agglomerated and non-agglomerated industries in the UK. We calculate entry and exit rates and job creation and destruction rates using the data at the plant level. We classify plants present in period t as entrants (if not present in $t-1$), exitors (if not present in $t+1$), 1-year (not present in $t-1$ or $t+1$), or survivors (present in $t-1$ and $t+1$). Entry, exit, 1 year and survival rates for period t are defined as the number of entrants, exitors, 1-years and survivors, respectively, divided by the total number of plants in period t.

Total new employment is employment in new entrants plus additional employment in plants that increased employment. The job creation rate in period t is total new employment as a ratio of total employment. The total reduction in employment is employment in period t exitors plus reductions in employment in plants that decreased employment. The job destruction rate is total reductions in employment as a ratio of total employment.

These rates are calculated for each year over the period 1985 to 1991. In Table 13, we regress each measure against the agglomeration measure γ in order to examine correlations between the different measures of industry dynamics and agglomeration. In each case the first column includes year dummies and the second column also controls for mean differences across two-digit sectors.

Entry and exit rates are both lower in more agglomerated industries as are the number of 1-year firms. Correspondingly, we see that survival rates are higher in more agglomerated industries. There is some suggestion that job creation rates are lower in more agglomerated industries, but there is no difference in job destruction rates. These findings potentially tie in with the earlier observation that the most agglomerated industries are somewhat more traditional, low-tech industries, and are perhaps at a more stable point in their lifecycle.

We also look specifically at the location of new entrants, to assess whether agglomeration forces are inducing them to enter into agglomerated regions (and thus to reinforce the agglomeration), or whether there is evidence of geographic dispersion among entrants

Table 13
Industry dynamics and agglomeration, 1985–1991

Dependent variable	Entrants		Exitors		One-year	
Number of observations	*1477*	*1477*	*1477*	*1477*	*1477*	*1477*
Agglomeration, γ	−0.128	−0.084	−0.033	−0.035	−0.057	−0.033
	(0.026)	(0.029)	(0.016)	(0.017)	(0.013)	(0.015)
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Two-digit industry dummies	No	Yes	No	Yes	No	Yes
Dependent variable	**Survivors**		**Job creation**		**Job destruction**	
Number of observations	*1477*	*1477*	*1477*	*1477*	*1477*	*1477*
Agglomeration, γ	0.218	0.153	−0.062	−0.037	−0.032	−0.031
	(0.036)	(0.039)	(0.018)	(0.021)	(0.020)	(0.021)
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Two-digit industry dummies	No	Yes	No	Yes	No	Yes

Estimates based on a panel of 211 four-digit industries, 1985 to 1991. Dependent variables are the annual entry rate, exit rate, 1-year rate, survival rate, job creation rate and job destruction rate as defined in the text. Standard errors are in brackets.

Table 14
Twenty most agglomerated industries—geographic distribution of entry 1985–1991

Four-digit industry	Number entrants	Percentage entrants to the top postcode area*	G entrants	γ entrants
4340 Spinning and weaving of flax	26	35	0.235	0.056
2330 Extraction salt	0	–	–	–
4350 Jute and polypropylene	38	16	0.126	0.064
2489 Ceramic goods	1337	27	0.129	0.124
4395 Lace	118	67	0.344	0.321
3162 Cutlery	98	45	0.246	0.224
3634 Pedal cycles	129	3	0.213	0.045
4363 Hosiery	3009	34	0.151	0.148
4910 Jewellery	3959	16	0.109	0.107
3161 Handtools	539	17	0.093	0.085
4752 Periodicals	4397	35	0.155	0.154
4310 Woollen and worsted industry	858	17	0.079	0.070
3523 Caravans	101	18	0.062	0.020
4721 Wall coverings	43	14	0.312	0.008
4322 Weaving cotton, silk	412	10	0.070	0.040
4831 Plastic coated textile fabric	23	0	0.131	– 0.011
2235 Other steel forming	52	23	0.134	0.044
4240 Spirit distilling	106	20	0.088	0.049
4537 Hats	198	16	0.105	0.091
4150 Fish processing	449	4	0.035	0.020

* Defined as the most agglomerated postcode area in 1985. This coincides with that given in Table 6.

in the most agglomerated industries, acting against agglomeration (due, for example, to congestion effects).

Table 14 gives information on the geographic distribution of entry for the 20 most agglomerated industries. It shows the number of entrants over the years 1985–1991, the percentage of those entrants that locate in the most agglomerated region,[20] and the measures of geographic concentration G and agglomeration, γ calculated over entrants only. Entry to many of the most agglomerated industries is also geographically concentrated. In these cases entry is concentrated in the most agglomerated region (the top postcode area). For example, lace and cutlery show both high geographic concentration and agglomeration among entrants and over 40% of entrants locating in the most agglomerated region. On the other hand, in a few industries, such as fish processing, entry appears to be acting against agglomeration, with a low measure of geographic concentration and agglomeration among the entrants, and only 4% of entrants going into the postcode area containing the highest proportion of employment.

We find that, while entry rates overall are lower in agglomerated industries, nonetheless in several agglomerated industries entry acts to reinforce agglomeration. Our time series is not as long as Dumais et al., but on this evidence it looks like, contrary to the US findings,

[20] Defined as the postcode area with the highest proportion of employment in 1985, as listed in Table 6.

the birth of new firms in the UK does not lead to greater geographic mobility of these industries. Again, this would be an interesting area for future research.

6. Conclusions

This paper has investigated the geographic concentration and agglomeration of production industries in the UK at a very disaggregated level both by industrial classification and regional unit of analysis. It has used measures of geographic concentration and agglomeration–which can be interpreted as the 'excess' of geographic concentration over that which would be expected given the industrial concentration of the industry. We use these measures to examine the pattern of production activity in the UK in 1992. As in the US and France, we find a high degree of geographic concentration in some industries. In some cases, a very high measure of geographic concentration can be almost entirely explained by an equally high industrial concentration. However, in other cases, such as ceramics, a high measure of geographic concentration is associated with low industrial concentration. We do not find any evidence to suggest that high-tech industries are more agglomerated, in fact if anything we find the opposite.

We compare our results with those obtained in studies using US and French data. We find a number of similarities in the pattern of agglomeration in the UK, the US and France. Those industries that are most agglomerated appear to be older and relatively low-tech industries. Looking back at historical evidence we see that several of these industries were reported to be agglomerated in Britain over a century ago.

Part of the observed difference in geographic concentration between industries may reflect differences in their stage of development. Analysis of entry, exit, job creation and job destruction rates finds that survival rates appear to be higher and entry rates lower in the more agglomerated industries. But among some of the most agglomerated industries, including those that are well-established in particular locations, new entry appears to be reinforcing the extent of agglomeration. As some of these industries would not be classified as high-tech it may be that local labour market conditions or vertical linkages between sectors are the forces driving agglomeration.

There remain many unanswered questions about what the main factors driving agglomeration are. We hope that the dataset provided in this paper (http://www.ifs.org.uk/corpact/dgsdata.zip) will provide a useful resource to researchers in the field. For example, linking this with detailed data on labour market behaviour would enable researchers to explicitly test for the presence of labour market externalities.

Acknowledgements

The authors would like to thank Stephen Redding, Konrad Stahl, David Stout, John Van Reenen, and two anonymous referees for helpful comments. The analysis contained in this paper was funded by the Leverhulme Trust under grant F/368/I and the ESRC Centre for Microeconomic Analysis of Fiscal Policy at the Institute for Fiscal Studies. This report has been produced under contract to ONS. All errors and omissions remain the responsibility of the authors.

Appendix A

We present a comparison of the main approaches used to measure geographic concentration and agglomeration. As well as the measures described in the text, we consider two further measures that have been used in the literature.

(a) A basic measure of geographic concentration defined as the proportion of plants in an industry in the top 3 regions, denoted as CI:

$$\mathrm{CI} = \sum_{k=1}^{3} s_k \tag{A1}$$

where s_k is the kth region's share of industry output, employment or any other measure of size, and $k = 1 \ldots 3$ are the regions with the three largest shares.

(b) A locational Gini coefficient:

$$L_{\mathrm{A}} = \frac{2}{K^2 \bar{s}} \left[\sum_{k=1}^{K} \lambda_k (s_k - \bar{s}) \right], \tag{A2}$$

where s_k is as defined above, λ_k denotes the position of the region in the ranking of s_k, and $\bar{s}$ is its mean across regions. This is the measure used by Krugman (1991).

In general, $0 \leq L \leq 1$. But for $N < K$, it follows that, $1 - N/K \leq L \leq 1$. An important problem with this measure is that it is sensitive to whether or not regions in which no firms in that industry operate are included. In calculating the Gini coefficient we are faced with the problem of how to deal with industry-regions in which there is no activity (since not all industries have plants in every postcode area). The Gini can be calculated either using only those regions in which there is some activity, or including all of the 113 postcode areas as a possible location, in which case if an industry has no activity in a particular region it is assigned a zero. We take the latter approach. The other measures discussed are not sensitive to this problem. We also present a relative Gini coefficient, L_{R}. This is defined in Eq. (A2), except that s_k is defined relative to the share of the industry's employment in aggregate.

The following tables present estimates of the correlation between these measures and those described in the text (Tables A1 and A2).

Table A1
Correlation between measures

	Number firms	G	γ	γ_{EG}	L_{R}	L_{A}
G	− 0.217					
γ	− 0.070	0.830				
γ_{EG}	− 0.080	0.824	0.996			
L_{R}	− 0.632	0.574	0.290	0.296		
L_{A}	− 0.559	0.633	0.349	0.344	0.966	
CI	− 0.314	0.893	0.612	0.600	0.756	0.834

Measures are: G: geographic concentration measure; γ, γ_{EG}: agglomeration indices; L_{R}: relative locational Gini, L_{A}: absolute locational Gini; CI: concentration index.

Table A2
Spearman rank correlation

	G	γ	γ_{EG}	L_R	L_A
γ [reject independence?]	0.480, [yes]				
γ_{EG} [reject independence?]	0.454, [yes]	0.908, [yes]			
L_R [reject independence?]	0.879, [yes]	0.248, [yes]	0.269, [yes]		
L_A [reject independence?]	0.941, [yes]	0.347, [yes]	0.330, [yes]	0.973, [yes]	
CI [reject independence?]	0.990, [yes]	0.477, [yes]	0.445, [yes]	0.864, [yes]	0.930, [yes]

Measures are: G: geographic concentration measure; γ, γ_{EG}: agglomeration indices; L_R: relative locational Gini, L_A: absolute locational Gini; CI: concentration index.

Table A3
Industry mapping used in Table 10

UK four-digit industry	US	France
4340 Spinning and weaving of flax	2281 Yarn mills except wool	
2330 Extraction salt		
4350 Jute and polypropylene		
2489 Ceramic goods		
4395 Lace		
3162 Cutlery		Cutlery
3634 Pedal cycles		
4363 Hosiery	2252 Hosiery not elsewhere classified and 2251 Women's hosiery	
4910 Jewellery	3961 Costume jewellery and 3915 Jewellers' materials lapidary	
3161 Handtools		
4752 Periodicals		Periodicals
4310 Woollen and worsted industry		Combed wool spinning mills, Wool preparation, Carded wool weaving mills
3523 Caravans		
4721 Wall coverings		
4322 Weaving cotton, silk		
4831 Plastic coated textile fabric		
2235 Other steel forming		
4240 Spirit distilling	2084 Wines brandy, brandy spirits	
4537 Hats		
4150 Fish processing		

References

Amiti, A., 1998. New trade theories and industrial location in the EU: a survey of evidence. Oxford Review of Economic Policy 14 (2) (Summer), 45–53.

Arrow, K., 1962. The economic implications of learning by doing. Review of Economic Studies 29, 155–173.

Barnes, M., Martin, R., 2002. Business data linking: an introduction. Economic Trends 581, 34–41 (April).

Dumais, G., Ellison, G., Glaeser, E., 2002. Geographic concentration as a dynamic process. Review of Economics and Statistics 84 (2), 193–204.

Duranton, G., Overman, H., 2002. Testing for localisation using micro-geographic data. CEPR discussion paper no. 3379.
Ellison, G., Glaeser, E., 1997. Geographic concentration in US manufacturing industries: a dartboard approach. Journal of Political Economy 105 (5), 889–927.
Ellison, G., Glaeser, E., 1999. The geographic concentration of industry: does natural advantage explain agglomeration? American Economic Review 89.2, 311–316.
Glaeser, E., Kallal, H., Scheinkman, J., Shleifer, A., 1992. Growth in cities. Journal of Political Economy 100, 1126–1152.
Griffith, R., 1999. Using the ARD establishment level data: an application to estimating production functions. Economic Journal 109, F416–F442.
Henderson, J.V., 1994. Where does an industry locate? Journal of Urban Economics 35, 83–104.
Henderson, J.V., 1999. Marshall's scale economies. NBER Working Paper, 7358.
Henderson, J.V., Kuncoro, A., Turner, M., 1995. Industrial development in cities. Journal of Political Economy 103, 1067–1090.
Jacobs, J., 1969. The Economy of Cities. Vintage, New York.
Jacobs, J., 1984. Cities and the Wealth of Nations: Principles of Economic Life. Vintage, New York.
Krugman, P.R., 1991. Geography and Trade. MIT Press, Cambridge, USA.
Marshall, A., 1890. Principles of Economics. MacMillan, London.
Maurel, F., Sédillot, B., 1999. A measure of the geographic concentration in French manufacturing industries. Regional Science and Urban Economics 29 (5), 575–604.
OECD, 1997. Revision of the high-technology sector and product classification. STI working papers. 1997/2.
Oulton, N., 1997. The ABI respondents database: a new resource for industrial economics research. Economic Trends (528), 46–57 (November).
Romer, P., 1990. Increasing returns and long-run growth. Journal of Political Economy 94, 1002–1037.
Swann, G.M.P., Prevezer, M., Stout, D., 1998. The Dynamics of Industrial Clustering: International Comparisons in Computing and Biotechnology. Oxford Univ. Press, Oxford.