

Wren, Colin

Conference Paper

Geographic Concentration, Observational Equivalence and the Source of Industrial Agglomeration

50th Congress of the European Regional Science Association: "Sustainable Regional Growth and Development in the Creative Knowledge Economy", 19-23 August 2010, Jönköping, Sweden

Provided in Cooperation with:

European Regional Science Association (ERSA)

Suggested Citation: Wren, Colin (2010) : Geographic Concentration, Observational Equivalence and the Source of Industrial Agglomeration, 50th Congress of the European Regional Science Association: "Sustainable Regional Growth and Development in the Creative Knowledge Economy", 19-23 August 2010, Jönköping, Sweden, European Regional Science Association (ERSA), Louvain-la-Neuve

This Version is available at:

<https://hdl.handle.net/10419/119285>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Geographic Concentration, Observational Equivalence and the Source of Industrial Agglomeration

Colin Wren

Economics, Business School, Newcastle University,
Newcastle upon Tyne, NE1 7RU, England, UK.

Abstract

Ellison and Glaeser's (1997) index of geographical concentration distinguishes between natural advantages and spillovers as a source of industrial agglomeration, but the well-known 'observational equivalence' means little is known about the relative importance of these. This paper uses the difference in the temporal scope of the agglomeration source to decompose the index, and sets out a methodology for measuring each of these using the frequency estimator approach of Maurel and Sédillot (1999). When applied to a dataset on foreign investment it shows spillovers decay and on average extend over three to five years. An implication is that the geographic concentration index will mainly reflect natural advantages, revealing comparatively little about spillovers.

Key words: geographic concentration, index, industrial agglomeration, spillovers.

JEL classification codes: R12, R30 and L00.

Contact: Prof. Colin Wren, c.m.wren@ncl.ac.uk .

Acknowledgements: The research is undertaken as part of the UK Spatial Economic Research Centre and supported by the Economic and Social Research Council. The author is grateful to comments from Jonathan Jones, Stephen McDonald and Francis Kiraly, but the usual disclaimer applies.

1: Introduction

The geographic concentration of activity within industries is widely observed at different spatial and industrial scales in many countries (e.g. Barrios and Strobl, 2004), and the reasons underlying this are of considerable interest (e.g. Krugman, 1991; Rosenthal and Strange, 2003; Brülhart and Traeger, 2004). While explanations rely on agglomeration economies, such as the externalities of the new growth theories and inter-firm linkages of the new economic geography (Neary, 2001; Döring and Schnellenbach, 2006), traditional notions of comparative advantage continue to play a role (Kim, 1999).¹ Indeed, some attribute at least half of geographic concentration to factors that they describe as ‘natural advantages’, i.e. access to input and output markets (Ellison and Glaeser, 1999).²

To measure the geographic concentration of an industry, Ellison and Glaeser (1997) propose an index that satisfies many of the criteria for a good index (Combes and Overman, 2004), being independent of industry scale and the plant size distribution.³ The index is derived from a choice-theoretic framework, and it models the processes leading to concentration arising from either natural advantages or ‘spillovers’, where the latter refers to technology and knowledge externalities, as well as the gains from inter-firm trade (Ellison and Glaeser, 1994). However, it is well-known that the index is subject to an ‘observational equivalence’, since natural advantages and spillovers contribute to it in

¹ Agglomeration economies include technological and pecuniary external effects (Scitovsky, 1952), with antecedents in the static localisation economies of Marshall, comprising labour-market interactions, inter-firm linkages and knowledge spillovers. Quigley (1998) considers the micro-foundations of agglomeration economies, and Duranton and Puga (2004) characterise them as matching, sharing and learning processes.

² Natural advantages occur “when location decisions are influenced by factors which can be regarded as giving a “natural” advantage to certain of the geographic areas” (Ellison and Glaeser, 1994, p. 8). They include natural phenomena, such as the availability of raw materials, proximity to coast or climate, as well as factors that are relatively immobile, e.g. infrastructure, large markets and a skills base.

³ The index is much used, although other measures exist, such as location quotients (Holmes and Stevens, 2002), entropy measures (Brülhart and Traeger, 2004) and indices of raw geographical concentration, but which are sometimes based on the Ellison and Glaeser index, e.g. Dumais *et al* (2002).

an identical fashion. As a result of this, little is known about the relative importance of natural advantages and spillovers to industrial agglomeration.⁴

The index of geographic concentration is calculated for an industry at a point in time and is static in nature, whereas the processes that underlie the index are dynamic, and importantly occur over different time frames. Thus, natural advantages persist and in empirical work are often treated as time invariant fixed effects, whereas spillovers flow between firms locating in different time periods and are likely to have a more transitory effect on location (Henderson, 1997). The usual explanation for this is that knowledge decays (Rosenthal and Strange, 2004), although there are other reasons why firms will move sooner rather than later to exploit a locational advantage. These include the spatial diffusion of technology, the impatient of firms and competitive pressures.

The purpose of this paper is to exploit this difference in temporal scope between natural advantages and spillovers to decompose the geographic concentration index, and thereby relax the observational equivalence. This is based on the Maurel and Sédillot (1999) version of the index, which has the same functional form to Ellison and Glaeser, and interprets the index as a combinatorial problem that involves counting the number of possible job pairings in different plants occurring within areas relative to those that occur both within and across areas. Crucially, such a decomposition allows the paper to set out a methodology for measuring the components due to natural advantages and spillovers that is based on the frequency estimator approach. This is applied to a dataset on foreign investment location decisions in Great Britain over 1985-05 and finds good support.

Natural advantages are broadly constant over time, but spillovers decay and on average

⁴ Research focuses on agglomeration economies (Rosenthal and Strange, 2004), such that when terms are included in econometric work to pick-up natural advantages they are often treated as uninteresting controls (e.g. Rosenthal and Strange, 2001; Barrios *et al*, 2005). There are methodological difficulties in capturing natural advantages, as they are fixed over time, there may be more natural advantages than there are cross-sectional units, and in any case it may be difficult to specify them all. Ellison and Glaeser (1999) include 16 terms to measure natural advantages across US states, which reduce the mean geographic concentration index by 20%, but Ellison and Glaeser reckon that a full set of such terms would reduce it by at least 50%.

extend over about 3 to 5 years, suggesting that firms move sooner rather than later. An important implication of the paper is that the geographic concentration index will tend to pick-up effects due to natural advantages, and so reveal little about spillovers

In the next section, a brief account is given of the indices of Ellison and Glaeser and Maurel and Sédillot, which is important to the subsequent analysis. Section 3 sets out the index decomposition, and Section 4 proposes the methodology for measuring the components of the index due to natural advantages and spillovers. This is applied to the location data in Section 5, and conclusions are drawn in Section 6.

2: Geographic Concentration Indices

The geographic concentration of an industry results in Ellison and Glaeser (1997) from a sequence of profit-maximising location decisions by individual plants. Agglomeration arises due to ‘natural advantages’, which are common to all firms in an industry, or due to ‘spillovers’, which are available from location close to other firms in the industry.

They are represented by parameters γ^{na} and γ^s respectively, where $\gamma^{na}, \gamma^s \in [0, 1]$. In addition there are idiosyncratic plant-specific effects that counter-balance these.

As a simplified account of the Ellison and Glaeser model, for the case of natural advantages, let π_r denote the probability that a plant locates in region r ($= 1, \dots, R$), and let x_r denote the employment share of all industries located in region r .⁵ Then, there are two restrictions placed on the first and second moments of π_r :

$$E(\pi_r) = x_r \quad \text{and} \quad \text{var}(\pi_r) = \gamma^{na} x_r (1 - x_r) \quad (1a, 1b)$$

⁵ For ease, we refer to regions rather than areas throughout, but the approach extends to any spatial unit. Profitability is chosen by nature, e.g. resource endowments, and in Ellison and Glaeser the probability π_r is the average profitability from locating in region r relative to the profitability of all regions.

In (1a) industry location on average reproduces the employment pattern of all industries, while the variance in (1b) captures natural advantages. When $\gamma^{na} = 0$ these advantages have no effect on location and each plant locates with a probability x_r , so that $\text{var}(\pi_r) = 0$. However, when $\gamma^{na} = 1$ natural advantages overwhelm the plant-specific effects and the region with the best endowments attracts all plants, so $\text{var}(\pi_r) = x_r(1 - x_r)$.

Spillovers are modelled by Ellison and Glaeser (1997) as a Bernoulli random event, which is equal to one with a probability of γ^s . Letting u_{ir} be an indicator variable, such that $u_{ir} = 1$ if and only if plant i locates in region r , then to reproduce the pattern of aggregate employment it is assumed $E(u_{ir} | \pi) = \pi_r$.⁶ A plant locates in a single region only, i.e. $\text{var}(u_{ir} | \pi) = \pi_r(1 - \pi_r)$, while $\text{corr}(u_{ir}, u_{jr}) = \gamma^s$ ($i \neq j$), so that:

$$\text{cov}(u_{ir}, u_{jr}) = \gamma^s \pi_r (1 - \pi_r). \quad (2)$$

Between any pair of plants the spillovers are of an “all or nothing” variety, occurring only within the same region but independent of distance. The spillovers are symmetric and transitive between plants, which importantly mean that the resulting distribution of plants is independent of the order in which they locate, no matter how far apart in time. Ellison and Glaeser (1997) refer to this as a rational expectations equilibrium.

Based on the above, Ellison and Glaeser (1997) derive their index γ as:

$$\gamma = \frac{\frac{G}{(1 - \sum_r x_r^2)} - H}{1 - H}, \quad (3)$$

where G is the raw geographic concentration of the industry, H is the Herfindahl index of the industry plant-size distribution. The index permits comparison of the geographic

⁶ Using (1a), Ellison and Glaeser (1994) jump straight to the relationship $E(u_{ir} | \pi) = x_r$.

concentration of industries. It is zero if jobs are only as concentrated as expected based on industrial concentration and the location of all industry across regions. However, observational equivalence arises because the index is symmetric in γ^{na} and γ^s :

$$\gamma \equiv \gamma^{na} + \gamma^s - \gamma^{na} \gamma^s. \quad (4)$$

Maurel and Sédillot (1999) focus on the probability that two plants i and j locate in the same region r , i.e. $E(u_{ir}, u_{jr})$, from which they derive an identical expression to (3). This is useful as it enables a particular interpretation to be given to the index. From the definition of covariance, using (1a) and (2) with γ^s replaced by γ they get:⁷

$$E(u_{ir}, u_{jr}) = \text{cov}(u_{ir}, u_{jr}) + E(u_{ir})E(u_{jr}) = \gamma x_r (1 - x_r) + x_r^2. \quad (5)$$

Summing this across regions, using $\sum_r x_r = 1$, the probability p that a pair of plants locates in any same region is therefore:

$$p = \gamma (1 - \sum_r x_r^2) + \sum_r x_r^2. \quad (6)$$

Maurel and Sédillot (1999) propose a frequency estimator of the probability p , on which much more is said below. When weighted by the plant employment size it gives:

$$\hat{p} = \frac{\sum_r s_r^2 - H}{1 - H}. \quad (7)$$

Substituting this for p in (6) yields a geographic concentration index, but with the same functional form as (3). However, whereas $G = \sum_r (s_r - x_r)^2$ in the Ellison and Glaeser index, where s_r is the regional shares of industry employment, in the Maurel and Sédillot

⁷ Since natural advantages are common to all plants in an industry, and location depends on π_r then the second equality in (5) can also be derived from (1a) and (1b), where $\text{cov}(u_{ir}, u_{jr}) = \text{var}(u_{ir}) = \text{var}(\pi_r)$.

version of (3) raw geographic concentration is $G = \sum_r s_r^2 - x_r^2$. Maurel and Sédillot

show that the expectation of the difference in the G terms is in fact zero.

3: Decomposition of the Index

To decompose the index, we consider a plant i of fixed scale choosing to locate in region r in one of $\tau = 1, 2, \dots, T$ time periods, and write:⁸

$$u_{ir} \equiv u_{ir1} + u_{ir2} + u_{ir3} + \dots + u_{irT}. \quad (8)$$

If plant i locates in region r at time t then $u_{irt} = 1$, in which case $u_{ir\tau} = 0$ for all $\tau \neq t$. In this case, $u_{ir} = 1$, but $u_{ir} = 0$ if region r is not settled, but which is as before. Then, for two series of random variable, by the Law of Addition (Mood *et al*, 1974):

$$\text{cov}(u_{ir}, u_{jr}) = \text{cov}\left(\sum_{\tau=1}^{\tau=T} u_{ir\tau}, \sum_{\tau=1}^{\tau=T} u_{jr\tau}\right) = \sum_{\tau=1}^{\tau=T} \sum_{\tau=1}^{\tau=T} \text{cov}(u_{ir\tau}, u_{jr\tau}). \quad (9)$$

Hence, by (8) and (9) the first equality in (5) can be written as:

$$E(u_{ir}, u_{jr}) = \sum_{\tau=1}^{\tau=T} \sum_{\tau=1}^{\tau=T} \text{cov}(u_{ir\tau}, u_{jr\tau}) + E(u_{ir})E(u_{jr}), \quad (10)$$

where by definition:

$$\text{cov}(u_{ir\tau}, u_{jr\tau}) = \text{corr}(u_{ir\tau}, u_{jr\tau}) (\text{var } u_{ir\tau})^{1/2} (\text{var } u_{jr\tau})^{1/2}. \quad (11)$$

The correlation term includes the time subscript, and the form of this is now considered.

Like elsewhere, it is supposed natural advantages and spillovers are each of an “all or

⁸ Specifically, it means plants do not grow, contract or close over time, which considered further below.

nothing” variety, but whereas natural advantages persist over time, several reasons were advanced above why spillovers have a limited temporal scope, i.e. obsolescence, spatial diffusion, impatience and competitive pressure. Ultimately this is an empirical matter, and it is examined below, but the limited available evidence is that these spillovers peak between 2 and 5 years for different industries (Henderson, 1997).⁹

Formally, it is assumed spillovers flow between plants and that they are of a constant strength over q periods ($1 \leq q < T$), but otherwise zero. Letting $\nu = 1, 2, \dots, T$, then the correlation coefficient in (11) is specified as:

$$\text{corr}(u_{ir\tau}, u_{jr\nu}) = \begin{cases} \gamma^{na} & \text{if } |\tau - \nu| > q \\ \gamma^{na} + \gamma^s & \text{if } |\tau - \nu| \leq q \end{cases} \quad (12)$$

Natural advantages affect location in every period, but spillovers affect the location of other firms only up to q periods after a plant has located, including plants that co-locate. This allows for the possibility that spillovers are transitive, i.e. either direct or indirect. For the purpose of tractability, a simple sum of γ^{na} and γ^s is taken, but consistent with the interpretation of these as correlation coefficients, then $\gamma^{na} + \gamma^s \in [-1, 1]$.

Substituting (12) into (11) into (10), setting $\text{var } u_{ir\tau} = \text{var } u_{jr\tau} = \pi_r (1 - \pi_r)$, using (1a) and then summing across regions, the probability p that a pair of plants locate in any same region is now given by:

$$p = [\gamma^{na} T^2 + \gamma^s \{q(2T - q - 1) + T\}] (1 - \sum_r x_r^2) + \sum_r x_r^2. \quad (13)$$

⁹ It is for the effect on productivity from own-industry employment. I know of no other studies that have specifically addressed this issue, and none in relation to the index. Glaeser *et al* (1992) and Henderson *et al* (1995) find that the characteristics of a city impact on growth over 20 years, but these are invariant and they can be construed as natural advantages (Henderson, 1997).

There are T^2 covariance terms in this, of which γ^{na} appears in every one and γ^s appears in $q(2T - q - 1) + T$.¹⁰ When no account is taken of the time period in which the plants locate, p is given by (6), but now with γ replaced by γT^2 . Using this to substitute for p in (13), then after simplification the index decomposition is:

$$\gamma = \gamma^{na} + \left\{ \frac{q(2T - q - 1) + T}{T^2} \right\} \gamma^s. \quad (14)$$

The left-hand side is the Maurel and Sédillot index of geographic concentration, so that (14) decomposes this into those parts due to natural advantages and spillovers. It differs from the Ellison and Glaeser decomposition in (4), as these now have a different temporal scope. If spillovers extend across T periods, i.e. $q = T$, then (14) reduces to $\gamma = \gamma^{na} + \gamma^s$, and if instead we write $\text{corr}(u_{i\tau}, u_{j\nu}) = \gamma^s + \gamma^{na} - \gamma^s \gamma^{na}$ for $|\tau - \nu| \leq q$ in (12), then when $q = T$ it reduces to the Ellison and Glaeser decomposition in (4).

The decomposition of the index has an important implication since in practice q is likely to be small relative to T , in which case γ is approximately equal to γ^{na} (the term in curly brackets in (14) is a function of q/T and $1/T$, and is small when q/T is small). For example, if $q = 1$ and $T = 20$ then (14) gives $\gamma = \gamma^{na} + 0.145 \gamma^s$, so that $\gamma \approx \gamma^{na}$. Thus, while the index γ may give a good account of geographic concentration, it reveals little about spillovers. This is because the distribution of plants at any time reflects location decisions taken over many decades, but if spillovers extend over short periods then most of the correlations making-up the index will in fact be due to natural advantages.

¹⁰ If plants i and j each locate randomly across T periods, then both locate in the same period in T of the T^2 possibilities, while they locate within q periods of each other in a further $2\{(T-1) + (T-2) + (T-3) + \dots + (T-q)\}$ periods. These sum to $q(2T - q - 1) + T$. For example, if $q = 1$ there are $3T - 2$ observations on γ^s , of which T relate to the same period, $T - 1$ go from i to j between a single period and $T - 1$ from j to i .

4: The Source of Agglomeration

The index decomposition in (14) relaxes the observational equivalence, but to give it power it is necessary to be able to determine γ^{na} and γ^s . For this we apply an empirical methodology based on the frequency estimator approach of Maurel and Sédillot (1999). To avoid onerous notation, attention focuses on the plant location decision, which is equivalent to considering job location, but in which plants are of the same size. Writing the Herfindahl index as $H = (1 + c^2) / n$, where c is the coefficient of variation of plant size and n the number of plants (Clarke, 1985), it is equivalent to setting $c = 0$, so that $H = 1 / n$ in (3). The extension to job location decision is considered in Appendix A.

Given that the plants are of the same fixed size, s_r now refers to the share of the number of plants in the industry located in region r , where $\sum_r s_r = 1$, and x_r refers to the share of the number of plants across all industries in region r , where $\sum_r x_r = 1$. Hence, if n_r denotes the number of plants in region r , where $\sum_r n_r = n$, then $s_r = n_r / n$.

Now, for a given distribution of plants across regions r , the frequency estimator of Maurel and Sédillot (1999) is equal to the number of all possible plant pairings within regions expressed as a fraction of all possible pairings both within and across regions. It can be written in a combinatorial form as follows, where the numerator is the number of plant pairings within R regions and the denominator is the total number of pairings:

$$\hat{p} = \frac{{}^{n_1}C_2 + {}^{n_2}C_2 + \dots + {}^{n_R}C_2}{{}^nC_2} \equiv \frac{\sum_{r=1}^{r=R} n_r (n_r - 1)}{n (n - 1)} \equiv \frac{\sum_r s_r^2 - 1/n}{1 - 1/n}. \quad (15)$$

This simplifies to give the second identity in (15), and dividing through by n^2 gives the final term. Of course, the final term is just \hat{p} in (7) with H replaced by $1 / n$, so that the

geographic concentration index can be interpreted as a counting exercise.¹¹ Since the pairings may be due to natural advantages or spillovers (or both) the frequency estimator approach offers a way of determining the source of agglomeration.

Empirical Methodology

To decompose the index empirically, the approach is to first estimate γ^{na} based on the plant pairings over which spillovers are assumed *not* to occur, i.e. $|\tau - \nu| > q$ in (12), and then to estimate γ^s from (14), representing an exact decomposition of γ . Letting $n_{r\tau}$ denote the number of plants locating in r at time τ , where $n_\tau = \sum_r n_{r\tau}$, then conditional on q , the frequency estimator for the probability of a natural advantage p_q^{na} is:

$$\hat{p}_q^{na} = \frac{\sum_{r=1}^{r=R} \left\{ n_r (n_r - 1) - \sum_{t=q+1}^{t=T} \left[2 n_{rt} \sum_{\tau=1}^{\tau=q} n_{rt-\tau} + n_{rt} (n_{rt} - 1) \right] \right\}}{n (n - 1) - \sum_{t=q+1}^{t=T} \left[2 n_t \sum_{\tau=1}^{\tau=q} n_{t-\tau} + n_t (n_t - 1) \right]}. \quad (16)$$

This is derived in Appendix A, where it is generalised to the job location decision. The numerator of (16) sums all the plant pairings occurring within regions, but deducts the pairings in which spillovers are assumed to be present, i.e. plants that locate within q time periods of one another. The denominator calculates on the same basis the plant pairings that occur both within and across regions. The probability \hat{p}_q^{na} lies in the unit interval, such that when all plants locate in a single region then $\hat{p}_q^{na} = 1$.

Given \hat{p}_q^{na} in (16), it is possible to empirically decompose the index. First, for $|\tau$

- $\nu| > q$, (6) leads to $\hat{p}_q^{na} = \hat{\gamma}_q^{na} \left(1 - \sum_r x_r^2 \right) + \sum_r x_r^2$, which after rearrangement gives:

¹¹ To give the correct number of pairings the numerators and denominators of the last two terms in (15) should be divided by 2, but these cancel. If all plants locate in a single region then $\hat{p} = 1$, and substituting it for p in (6) gives $\gamma = 1$. If plants locate according to the pattern of all industry, i.e. $s_r = x_r$, then this gives $\gamma = 1 / (1 - n)$, which is approximately zero for large n .

$$\hat{\gamma}_q^{na} = \frac{\hat{p}_q^{na} - \sum_r x_r^2}{1 - \sum_r x_r^2}. \quad (17)$$

Hence, using (16) this gives $\hat{\gamma}_q^{na}$. Second, substituting (17) into (14) gives $\hat{\gamma}_q^s$ as:

$$\hat{\gamma}_q^s = \left(\frac{T^2}{q(2T - q - 1) + T} \right) (\hat{\gamma} - \hat{\gamma}_q^{na}). \quad (18)$$

Both of these depend on the period q over which the spillovers are assumed to occur, and by construction they exactly decompose the Maurel and Sédillot index. Intuitively, $\hat{\gamma}_q^{na}$ is based on all the plant pairings that are sufficiently far apart in time that they cannot be construed as including spillovers, while from the overall value of the index it is possible to determine $\hat{\gamma}_q^s$. If plants locate across regions in the short-run the same as in the long-run then there is no difference between $\hat{\gamma}$ and $\hat{\gamma}_q^{na}$, and $\hat{\gamma}_q^s$ is zero, which is irrespective of q . However, if spillovers are present for some q , $\hat{\gamma}$ and $\hat{\gamma}_q^{na}$ differ and $\hat{\gamma}_q^s$ is non-zero. In this case, $\hat{\gamma}_q^s$ potentially depends on the location of all industry, x_r .

The above methodology is used below to empirically decompose the geographic concentration index, allowing the time period q over which the spillovers are assumed to exist to vary. Since natural advantages persist then $\hat{\gamma}_q^{na}$ should be robust to the value of q , but if spillovers decay then $\hat{\gamma}_q^s$ should decrease as q increases. The approach allows the contribution of natural advantages and spillovers to be quantified, but in the case of the latter it is important to be clear about exactly what is measured. It is the effect of plant location on the location of other plants over the subsequent q periods, including the current period (i.e. more generally job location on other jobs). Like Ellison and Glaeser

(1997) the spillovers are intra-industry only, both direct and indirect, and include market effects and externalities, which in the latter case are MAR externalities.

Numerical Exploration

Before the empirics, the methodology is first subjected to a simple numerical exploration to gain confidence. For simplicity, there are four regions, four time periods and sixteen plants, i.e. $R = 4$, $T = 4$ and $n = 16$, while spillovers extend over a single period, so $q = 1$. Different location patterns are exhibited in Figure 1, in which four plants locate in each time period. In (a) to (d) two regions only are settled equally, in (g) and (h) four regions are settled equally, while (e) and (f) are intermediate. Table 1 calculates the frequency estimates, \hat{p} and \hat{p}_1^{na} , and also calculates $\hat{\gamma}$, $\hat{\gamma}^{na}$ and $\hat{\gamma}^s$ for each of these. This is under two scenarios about how all industry locates (not shown), i.e. either equally across all four regions ($\sum_r x_r^2 = 0.25$) or equally across two regions only ($\sum_r x_r^2 = 0.50$).

The rather stark nature of the simulations means the estimates can be high, while since the number of time periods T is small the weight in (14) is relatively large at 0.625, but interest is in the relative magnitude and sign of these estimates. Overall, a plausible pattern is obtained and several points can be highlighted from Table 1. First, (a) to (d) suggest $\hat{\gamma}^s$ is greatest when a large number of plants locate in a region subsequent to a large number of plants in the same region in the preceding period, as in (d). If plants locate in the same region in every period then this tends to be attributed to natural advantages and $\hat{\gamma}^s$ is small, as in (b). Second, as regards (e) to (h), $\hat{\gamma}^s$ is greater in (f) than (e), and greater again in (g), which are each plausible. If plants locate according to the pattern of all industry in (h), i.e. $\sum_r x_r^2 = 0.25$, then there are neither spillovers nor natural advantages, which clearly is a desirable feature. Finally, $\hat{\gamma}^s$ is insensitive to the location of all industry x_r , but which is mainly reflected in $\hat{\gamma}$ and $\hat{\gamma}^{na}$.

5: Application to Foreign Investment

The decomposition of the index is now applied to a dataset on location. The purpose is to examine whether the approach yields plausible estimates of $\hat{\gamma}^{na}$ and $\hat{\gamma}^s$, especially as q is allowed to vary. In so doing, the plant dynamics are relevant, but which is in effect assumed away in Ellison and Glaeser and above by supposing that the scales are fixed, so that the plants neither expand, contract nor exit.¹² This is no longer satisfactory when confronted with the data, although it could be argued that it is not a serious issue, as Dumais *et al* (2002) find that the combination of births, closures, expansions and contractions act together to maintain concentration over time. In our case the data are for plant births only (i.e. investments), so expansions and contractions are not relevant, while to deal with the exits there are basically two approaches.

The first is simply to calculate the index on a static basis for a single cross-section of plants at time T , and to base the estimates of γ^{na} and γ^s on the dates when each of the included plants entered. This shadows the approach of Ellison and Glaeser, but the drawback is that any plant that existed over the period but exited prior to time T will not be taken into account, which potentially biases $\hat{\gamma}^{na}$ and $\hat{\gamma}^s$. The second approach, that is adopted here, mirrors the dynamics of the process and bases the index on all the births over T periods, including those that exited by time T , and decomposes this. It means $\hat{\gamma}^{na}$ and $\hat{\gamma}^s$ are better determined, which reflects our interest, but it implies $\hat{\gamma}$ is in some sense hypothetical. Kim (1999) and Dumais *et al* (2002) both find that the index is stable over long time periods, while Rosenthal and Strange (2001) calculate the index for births only and find that it does not differ too much from the usual static measure.

¹² For example, if the geographical concentration of an industry at some point in time arises from spillovers that occurred in the past, then in determining the source of agglomeration it is relevant whether the plants expanded, contracted or even exited in the period up to the time that the index is calculated.

Data

The data give information by year on about 6,500 investments by foreign-owned plants across the regions of Great Britain over 1985-2005. Foreign investment is highly mobile, for which spillovers are likely to be important (Blomström and Kokko, 1998), while the data are much used to examine location (Alegria, 2009; Dimitropoulou *et al*, 2006; Jones and Wren, 2010), where details can be found.¹³ The data were gathered on a consistent basis over a long time period, and they include the smaller plants that may be absent from the UK production ‘census’, which samples smaller plants. Interest is in whether foreign investment reflects the pattern of previous foreign investment. These investments tend to be larger in size, which gives the index a good economic meaning.

To maximise the number of observations the regions that are used are large and the industries are broad, but the purpose is to assess the methodology, which can then be applied to other datasets. The areas are the Government Office regions for Great Britain, defined at the *Eurostat* NUTS I level, where London is part of the South East region. The concentration index is comparable at different levels of aggregation (Ellison and Glaeser, 1997). In the case of spillovers, Rosenthal and Strange (2001) find that they are more important for smaller areas, but Lamorgese and Ottaviano (2002) argue that pecuniary external effects occur over larger areas. In the case of natural advantages it is reasonable that these are invariant over a 20-year time span, not least as the UK ‘regional problem’ has persisted. Industries are for manufacturing at the 2-digit level, although where data necessitates sometimes aggregated to form homogeneous groupings, or where data

¹³ The data were supplied by the UK’s main national investment agency, *UK Trade and Industry*, and give information on all known investments carried out by foreign-plants in Great Britain. Comparison with the published aggregate FDI stock data from the production census shows no significant difference by region. Data take a broad definition of FDI, including new plant entrants, acquisitions and re-investments (Wren and Jones, 2009), where each is potentially mobile and accounts for about a third of the total cases.

permits disaggregated to identify particular activities. Co-agglomeration economies are not considered, but this is mitigated by the broad nature of the industries.

Results

There are 22 industries, giving an average of 294 observations on location per industry. This is about 43,000 pairings in each case (i.e. $^{294}C_2$), which may be assigned to natural advantages or spillovers according to (12). The industries are defined in Appendix B, along with the frequency estimates of \hat{p} and \hat{p}^{na} . Table 2 presents the results for $\hat{\gamma}$ and its decomposition $\hat{\gamma}_q^s$ and $\hat{\gamma}_q^{na}$ for each industry. It reports the coefficients as percentages for values of q equal to 1, 2, 3 and 5 years. Since regional policy operated in favour of foreign investment throughout the period, and biased manufacturing location (Jones and Wren, 2010), x_r is calculated from the data for all manufacturing

Overall, the results are encouraging. For the average across 22 industries $\hat{\gamma}^{na}$ is robust to q , lying between 2.41% and 2.72%, while $\hat{\gamma}^s$ decreases monotonically as q increases, from 4.00% over 1 year to 1.90% over 5 years. In order to interpret these, Ellison and Glaeser (1997) regard a value of the index of less than 2% as being not very localized and a value of more than 5% as being highly localized, and these can be applied to results in Table 2. The industries are geographically concentrated on average, i.e. $\hat{\gamma} = 3.27\%$, and spillovers extend over a period of about 3 to 5 years, although there are some strong differences between industries that are worthy of consideration.

Spillovers occur in about half the industries, which not only include relatively labour-intensive activities, such as textiles, leather, publishing and furniture, but capital-intensive activities, such as refining, chemicals, pharmaceuticals, basic metals, metal products and other transport. As q increases spillovers generally decline monotonically,

although electronic components is an exception, and seven industries record a value of $\hat{\gamma}^s$ at or above 2% over a 5-year time horizon. Of course, what is measured is not whether a location 5 years ago affects current location, but whether all locations in the previous 5 years affects current location, which might bias the results towards longer horizons, but either way the effects appear to be reasonably long-lasting. Natural advantages occur in about half the industries, but not necessarily in the same industries as spillovers.

The results support the methodology, and suggest that foreign investment tends to be geographically concentrated within manufacturing industries, with natural advantages and spillovers having a similar effect over a 3-year time horizon, although it is only over this period that it is consistent with the conjecture of Ellison and Glaeser (1999). They also confirm that the index will reveal comparatively little about spillovers. Thus, across the industries in Table 2, the correlation coefficient between $\hat{\gamma}$ and $\hat{\gamma}_q^{na}$ is 0.96, 0.94, 0.96 and 0.96 for $q = 1, 2, 3$ and 5 respectively, but between $\hat{\gamma}$ and $\hat{\gamma}_q^s$ it is only 0.34, 0.32, 0.34 and 0.41. This is because the spillovers have a limited temporal scope, which for any given industry is regardless of the magnitudes of $\hat{\gamma}_q^{na}$ and $\hat{\gamma}_q^s$.

Some Robustness Tests

Before concluding, several tests were undertaken of the methodology. First, longer-time horizons for spillovers were considered, and Appendix C reports the results with q equal to 7 and 10 years (the results for 3 and 5 years are reproduced from Table 1). These suggest spillovers can extend over a seven-year time horizon, but that they are generally exhausted over ten years. Indeed, for some industries $\hat{\gamma}_{10}^s$ is negative and large, but this arises because when q is large the number of pairings due solely to natural advantages is small, so that $\hat{\gamma}_{10}^{na}$ is poorly determined. It suggests data may limit the calculations over

longer time horizons, in which case $\hat{\gamma}_q^{na}$ falls out of line with previous estimates, although spillovers are generally exhausted by then.

Second, the weight on γ^s in (14) is based on the proportion of periods over which the spillovers are assumed to occur, but it takes no account of the fact that plants arrive unevenly over time, so that there may be either a large or small number of observations on these. As an alternative, rather than (18), spillovers were determined using:

$$\hat{\gamma}_q^s = \left(\frac{n(n-1)}{\sum_{t=q+1}^{t=T} \left[2 n_t \sum_{\tau=1}^{\tau=q} n_{t-\tau} + n_t (n_t - 1) \right]} \right) (\hat{\gamma} - \hat{\gamma}_q^{na}). \quad (19)$$

The denominator of the weight in (19) is the number of observations on plant pairings over q periods, while the numerator gives the total number of pairings. Using this makes no difference to either $\hat{\gamma}$ or $\hat{\gamma}_q^{na}$, so that the results for $\hat{\gamma}_q^s$ only are presented in Appendix D based on (19), along with (18) for the purpose of comparison. However, the two sets of estimates are virtually identical, so that the same conclusions emerge.

6: Conclusions

Geographical concentration is the propensity of economic activity to locate unevenly across geographic space, and it is measured by the Ellison and Glaeser index. The index distinguishes between the agglomeration source, i.e. natural advantages in the form of access to input and output markets, and spillovers that are technological or pecuniary in nature, but a difficulty is that there is an ‘observational equivalence’, as these contribute to the index in an identical manner. The contribution of this paper is to decompose the geographic concentration index and to offer an empirical methodology that relaxes the

observational equivalence. This utilises the difference in the temporal scope, as whereas natural advantages persist in a location it is argued spillovers have a transitory effect on subsequent location, which allows the separate effects to be identified.

The approach is applied to a dataset on the location decisions of about 6,500 foreign-owned manufacturing investments across the regions of Great Britain, and yields plausible and interesting results. Overall, it finds that the factors that determine locations within short time spans of one another differ from those that determine locations over longer time horizons, and provides convincing evidence for the existence of spillovers. While natural advantages appear to have a broadly constant location effect over different time horizons, spillovers decay and on average occur over a time period of about 3 to 5 years, although for some industries it is longer than this. On average, over a 3-year time horizon they have a similar effect on location as natural advantages.

The methodology and results of this paper have several important implications, of which perhaps a key finding is that the geographic concentration index will primarily capture location effects due to natural advantages. This is because spillovers have a limited temporal scope, so that rather than contributing to the index in an identical way, most of the correlations making at the index will be due to natural advantages. This in turn has implications for empirical work that uses the index in an attempt to determine the source of agglomeration, e.g. whether the spillovers are pecuniary or technological in nature. It may be that the low coefficients of determination found in econometric work do not just reflect the difficulty of specifying a full set of natural advantages, but that it may also be because the index gives a low weight to spillovers.

Figure 1: Simulation of Plant Location Patterns

(a)

	r_1	r_2	r_3	r_4
t_1	4	0	0	0
t_2	0	4	0	0
t_3	4	0	0	0
t_4	0	4	0	0

(b)

	r_1	r_2	r_3	r_4
t_1	2	2	0	0
t_2	2	2	0	0
t_3	2	2	0	0
t_4	2	2	0	0

(c)

	r_1	r_2	r_3	r_4
t_1	4	0	0	0
t_2	0	4	0	0
t_3	0	4	0	0
t_4	4	0	0	0

(d)

	r_1	r_2	r_3	r_4
t_1	4	0	0	0
t_2	4	0	0	0
t_3	0	4	0	0
t_4	0	4	0	0

(e)

	r_1	r_2	r_3	r_4
t_1	2	2	0	0
t_2	2	2	0	0
t_3	0	0	2	2
t_4	2	2	0	0

(f)

	r_1	r_2	r_3	r_4
t_1	2	2	0	0
t_2	2	2	0	0
t_3	2	2	0	0
t_4	0	0	2	2

(g)

	r_1	r_2	r_3	r_4
t_1	2	2	0	0
t_2	2	2	0	0
t_3	0	0	2	2
t_4	0	0	2	2

(h)

	r_1	r_2	r_3	r_4
t_1	1	1	1	1
t_2	1	1	1	1
t_3	1	1	1	1
t_4	1	1	1	1

Note: Each cell gives number of the 16 plants locating across the 4 regions and 4 time periods.

Table 1: Numerical Explorations

	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
\hat{p}	0.47	0.47	0.47	0.47	0.27	0.27	0.20	0.20
\hat{p}_1^{na}	0.70	0.48	0.41	0.11	0.21	0.14	0.03	0.22
$\sum_r x_r^2 = 0.25:$								
$\hat{\gamma}$	0.29	0.29	0.29	0.29	0.02	0.02	-0.07	-0.07
$\hat{\gamma}^{na}$	0.60	0.31	0.21	-0.19	-0.05	-0.14	-0.30	-0.04
$\hat{\gamma}^s$	-0.50	-0.03	0.13	0.77	0.11	0.26	0.37	-0.05
$\sum_r x_r^2 = 0.50:$								
$\hat{\gamma}$	-0.07	-0.07	-0.07	-0.07	-0.47	-0.47	-0.60	-0.60
$\hat{\gamma}^{na}$	0.41	-0.04	-0.19	-0.78	-0.58	-0.71	-0.94	-0.56
$\hat{\gamma}^s$	-0.77	-0.05	0.19	1.00	0.18	0.38	0.54	-0.06

Notes: Decomposition of the geographic concentration index based on location patterns in (a) to (h) of Figure 1. \hat{p} and \hat{p}_1^{na} are given by (7) and (16) with $q = 1$, $\hat{\gamma}$ by (3) and $\hat{\gamma}^{na}$ and $\hat{\gamma}^s$ by (17) and (18). $\sum_r x_r^2$ gives the assumed plant location pattern for all industry.

Table 2: Decomposition of the Index

Manufacturing Industry	$\hat{\gamma}$	$\hat{\gamma}_q^s$				$\hat{\gamma}_q^{na}$			
		1 yr	2 yr	3 yr	5 yr	1 yr	2 yr	3 yr	5 yr
Food, Beverages and Tobacco	-0.67	0.39	0.24	0.32	-0.16	-0.72	-0.72	-0.77	-0.60
Textiles and Textile Products	2.72	6.90	6.20	6.43	5.87	1.77	1.33	0.75	0.05
Leather and Leather Products	8.29	31.76	22.42	12.77	5.71	3.89	3.25	4.38	5.69
Wood and Wood Products	1.56	1.05	1.90	1.58	0.38	1.42	1.14	1.08	1.39
Pulp, Paper and Paper Products	5.44	0.85	1.32	1.48	0.22	5.32	5.14	4.98	5.34
Publishing and Printing	10.43	4.37	1.58	-0.31	0.99	9.83	10.07	10.52	9.97
Coke, Refined Petroleum Products	2.16	2.83	3.06	0.17	0.91	1.77	1.48	2.11	1.75
Chemicals	1.36	2.31	2.30	2.10	1.96	1.04	0.84	0.72	0.47
Pharmaceuticals	4.66	3.74	3.70	3.63	4.29	4.14	3.83	3.55	2.70
Rubber and Plastic Products	0.28	0.78	0.84	0.62	-0.02	0.17	0.09	0.09	0.29
Mineral Products	1.54	0.07	0.22	0.55	1.99	1.53	1.49	1.37	0.64
Basic Metals	2.93	4.15	3.74	2.81	2.37	2.36	2.09	2.07	1.85
Metal Products	2.73	0.98	0.68	0.08	0.07	2.59	2.58	2.71	2.70
Machinery	0.92	0.70	0.54	0.58	0.04	0.82	0.80	0.74	0.90
Office Machinery	4.23	2.26	1.77	2.24	1.82	3.92	3.83	3.54	3.40
Electrical Machinery	0.62	1.76	0.97	0.50	-0.19	0.37	0.41	0.47	0.71
Electronic Components	10.84	1.70	1.86	3.47	4.56	10.60	10.42	9.77	8.76
TV and Radio	1.36	2.16	1.74	1.15	1.56	1.06	0.97	1.01	0.65
Medical and Optical Instruments	0.59	0.94	0.54	0.81	0.93	0.46	0.47	0.35	0.17
Motor Vehicles	6.23	0.32	0.32	0.54	0.69	6.18	6.16	6.06	5.91
Other Transport	0.61	8.98	5.55	3.02	3.02	-0.63	-0.64	-0.32	-0.77
Furniture and Leisure Goods	3.15	8.99	8.14	6.89	4.72	1.90	1.32	1.03	1.00
Mean	3.27	4.00	3.17	2.34	1.90	2.72	2.56	2.56	2.41

Note: Decomposition of index in (14), based on (17) and (18) and frequency estimates in Appendix B. Index calculated for $q = 1, 2, 3$ and 5 , where index expressed as a percentage. x_r is the regional share across all manufacturing industries.

Appendix A: Derivation of the Frequency Estimators for p^{na}

To derive the estimator of p^{na} for *plant* location, we first note that the total number of plant pairings within regions is $\sum_{r=1}^{r=R} n_r (n_r - 1)/2$, and that between and within regions it is $n (n - 1)/2$, so that the ratio of these gives (15).

Initially, suppose spillovers extend over a single period only, i.e. $q = 1$, so that they occur between $\tau = t - 1$ and t . To get p_1^{na} then from each of the above we deduct the relevant number of plant pairings occurring over a single time period and within the same period of location, i.e. time t . The number of pairings between and within times $t - 1$ and t for region r is $n_{rt} + n_{rt-1} C_2$, but excluding those at $t - 1$ it is $n_{rt} + n_{rt-1} C_2 - n_{rt-1} C_2$, i.e.

$$\frac{(n_{rt} + n_{rt-1})(n_{rt} + n_{rt-1} - 1)}{2} - \frac{n_{rt-1}(n_{rt-1} - 1)}{2} = n_{rt} n_{rt-1} + \frac{n_{rt}(n_{rt} - 1)}{2}.$$

Hence, for all regions the number of the pairings is $\sum_{r=1}^{r=R} [n_{rt} n_{rt-1} + n_{rt}(n_{rt} - 1)/2]$, and

over the T periods it is $\sum_{t=2}^{t=T} \sum_{r=1}^{r=R} [n_{rt} n_{rt-1} + n_{rt}(n_{rt} - 1)/2]$. By the same reasoning the deduction from the denominator is $\sum_{t=2}^{t=T} [n_t n_{t-1} + n_t(n_t - 1)/2]$, giving:

$$\hat{p}_1^{na} = \frac{\sum_{r=1}^{r=R} \left\{ n_r (n_r - 1) - \sum_{t=2}^{t=T} [2 n_{rt} n_{rt-1} + n_{rt} (n_{rt} - 1)] \right\}}{n (n - 1) - \sum_{t=2}^{t=T} [2 n_t n_{t-1} + n_t (n_t - 1)]}. \quad (A1)$$

If spillovers extend over q periods, then to capture the number of pairings between

plants at time t and those over the pervious q periods, substitute $\sum_{\tau=1}^{\tau=q} n_{rt-\tau}$ for n_{rt-1} in

(A1), which gives \hat{p}_q^{na} in (16) of the text.

To get the estimators for the geographic concentration of *jobs* between plants (i.e. in effect allowing the plant size to vary), then to avoid excessive notation we note that the number of job pairings within regions is $\left(\sum_{r=1}^{r=R} s_r^2 - H\right)/2$, and that between and within regions it is $(1 - H)/2$, where s_r is the share of region r in total employment, and the ratio of these gives the frequency estimator in (7). To get \hat{p}_1^{na} it can be shown that the number of job pairings between plants over a single time period and within the same period of location is $\sum_{t=2}^{t=T} \sum_{r=1}^{r=R} [s_{rt} s_{rt-1} + s_{rt} (s_{rt} - 1)/2]$. Further, between and within regions it is $\sum_{t=2}^{t=T} [s_t s_{t-1} + s_t (s_t - 1)/2]$. These now give:

$$\hat{p}_1^{na} = \frac{\sum_{r=1}^{r=R} \left\{ s_r^2 - \sum_{t=2}^{t=T} [2 s_{rt} s_{rt-1} + s_{rt} (s_{rt} - 1)] \right\} - H}{1 - \sum_{t=2}^{t=T} [2 s_t s_{t-1} + s_t (s_t - 1)] - H}. \quad (\text{A2})$$

This reduces to (A1) if plants are of equal size. This can be seen by writing $s_r = n_r / n$, noting that $H = 1 / n = \sum_r n_r / n^2$, and likewise for the t -subscripted terms. However, if the spillovers extend over q periods then (A2) generalises to:

$$\hat{p}_q^{na} = \frac{\sum_{r=1}^{r=R} \left\{ s_r^2 - \sum_{t=q+1}^{t=T} \left[2 s_{rt} \sum_{\tau=1}^{\tau=q} s_{rt-\tau} + s_{rt} (s_{rt} - 1) \right] \right\} - H}{1 - \sum_{t=q+1}^{t=T} \left[2 s_t \sum_{\tau=1}^{\tau=q} s_{t-\tau} + s_t (s_t - 1) \right] - H}. \quad (\text{A3})$$

Following the same reasoning as above, this reduces to (16).

Appendix B: Industry Classification and Frequency Estimates

Industry		Number of investments	\hat{p}	\hat{p}_1^{na}	\hat{p}_2^{na}	\hat{p}_3^{na}	\hat{p}_5^{na}
3	Food, Beverages and Tobacco (15 and 16)	360	0.115	0.115	0.115	0.115	0.116
4	Textiles and Textile Products (17 and 18)	170	0.145	0.143	0.141	0.140	0.138
5	Leather and Leather Products (19)	26	0.194	0.184	0.183	0.189	0.195
6	Wood and Wood Products (20)	94	0.135	0.134	0.133	0.133	0.135
7	Pulp, Paper and Paper Products (21)	201	0.169	0.169	0.168	0.167	0.169
8	Publishing and Printing (22)	131	0.213	0.211	0.212	0.214	0.213
9	Coke, Refined Petroleum Products (23)	25	0.154	0.150	0.149	0.152	0.153
10	Chemicals (24, excl. 24.4)	473	0.133	0.132	0.131	0.131	0.130
11	Pharmaceuticals (24.4)	231	0.162	0.161	0.160	0.159	0.158
12	Rubber and Plastic Products (25)	418	0.123	0.123	0.123	0.123	0.123
13	Mineral Products (26)	155	0.134	0.134	0.134	0.134	0.132
14	Basic Metals (27)	193	0.147	0.146	0.145	0.145	0.144
15	Metal Products (28)	380	0.145	0.145	0.145	0.145	0.145
16	Machinery (29)	797	0.129	0.129	0.129	0.129	0.129
17	Office Machinery (30)	131	0.158	0.158	0.158	0.158	0.157
18	Electrical Machinery (31)	399	0.126	0.126	0.126	0.126	0.127
19	Electronic Components (32.1)	493	0.216	0.216	0.215	0.213	0.210
20	TV and Radio (32.2 and 32.3)	271	0.133	0.132	0.131	0.132	0.131
21	Medical and Optical Instruments (33)	339	0.126	0.126	0.126	0.126	0.125
22	Motor Vehicles (34)	779	0.176	0.175	0.175	0.175	0.175
23	Other Transport (35)	189	0.126	0.124	0.125	0.125	0.124
24	Furniture and Leisure Goods (36)	217	0.148	0.146	0.144	0.144	0.145
Mean		294	0.150	0.149	0.148	0.149	0.149

Notes: Industries defined according to NACE (rev. 1), with the 2 or 3-digit code given in parentheses. Total number of projects over 1985-05. \hat{p} is calculated using (7) and \hat{p}_q^{na} using (16) for $q = 1, 2, 3$ and 5.

Appendix C: Decomposition of Index over Longer Time Horizons

Manufacturing Industry	$\hat{\gamma}$	$\hat{\gamma}_q^s$				$\hat{\gamma}_q^{na}$			
		3 yr	5 yr	7 yr	10 yr	3 yr	5 yr	7 yr	10 yr
Food, Beverages and Tobacco	-0.67	0.32	-0.16	-1.00	-1.41	-0.77	-0.60	-0.08	0.40
Textiles and Textile Products	2.72	6.43	5.87	5.68	0.76	0.75	0.05	-0.61	2.15
Leather and Leather Products	8.29	12.77	5.71	4.11	6.52	4.38	5.69	5.87	3.39
Wood and Wood Products	1.56	1.58	0.38	-0.39	-1.32	1.08	1.39	1.79	2.56
Pulp, Paper and Paper Products	5.44	1.48	0.22	-0.77	-3.15	4.98	5.34	5.89	7.80
Publishing and Printing	10.43	-0.31	0.99	-0.66	-0.50	10.52	9.97	10.82	10.80
Coke, Refined Petroleum Products	3.75	0.17	0.91	-2.94	-12.46	2.11	1.75	3.89	11.52
Chemicals	1.36	2.10	1.96	0.58	-0.29	0.72	0.47	1.01	1.58
Pharmaceuticals	4.66	3.63	4.29	4.46	2.00	3.55	2.70	2.04	3.15
Rubber and Plastic Products	0.28	0.62	-0.02	-0.53	-0.54	0.09	0.29	0.59	0.68
Mineral Products	1.54	0.55	1.99	1.65	0.69	1.37	0.64	0.57	1.02
Basic Metals	2.93	2.81	2.37	1.39	1.27	2.07	1.85	2.11	1.98
Metal Products	2.73	0.08	0.07	0.11	-1.55	2.71	2.70	2.67	3.89
Machinery	0.92	0.58	0.04	-0.25	-0.68	0.74	0.90	1.06	1.43
Office Machinery	4.23	2.24	1.82	1.52	1.40	3.54	3.40	3.34	3.18
Electrical Machinery	0.62	0.50	-0.19	-0.43	-0.51	0.47	0.71	0.87	1.00
Electronic Components	10.84	3.47	4.56	5.24	2.80	9.77	8.76	7.76	8.73
TV and Radio	1.36	1.15	1.56	0.07	-0.38	1.01	0.65	1.32	1.64
Medical and Optical Instruments	0.59	0.81	0.93	0.91	0.47	0.35	0.17	0.06	0.24
Motor Vehicles	6.23	0.54	0.69	0.45	0.26	6.06	5.91	5.96	6.03
Other Transport	0.61	3.02	3.02	4.15	-4.42	-0.32	-0.77	-1.83	3.93
Furniture and Leisure Goods	3.15	6.89	4.72	3.13	1.57	1.03	1.00	1.31	1.97
Mean	3.27	2.34	1.90	1.20	-0.43	2.56	2.41	2.56	3.59

Notes: Decomposition of geographic concentration index for spillovers of up to $q = 10$ years. Calculated on the same basis as Table 2, where indices for 3 and 5 years are reproduced from this table. Indices expressed as %.

Appendix D: Estimates of Spillovers using Alternative Weightings

Manufacturing Industry	New weighting in (19)				Old weighting in (18)			
	1 yr	2 yr	3 yr	5 yr	1 yr	2 yr	3 yr	5 yr
Food, Beverages and Tobacco	0.37	0.23	0.32	-0.17	0.39	0.24	0.32	-0.16
Textiles and Textile Products	7.02	6.40	6.69	6.49	6.90	6.20	6.43	5.87
Leather and Leather Products	29.74	23.04	13.66	6.40	31.76	22.42	12.77	5.71
Wood and Wood Products	0.84	1.60	1.39	0.47	1.05	1.90	1.58	0.38
Pulp, Paper and Paper Products	0.83	1.33	1.51	0.23	0.85	1.32	1.48	0.22
Publishing and Printing	3.81	1.45	-0.28	0.93	4.37	1.58	-0.31	0.99
Coke, Refined Petroleum Products	1.61	2.26	0.12	0.65	2.83	3.06	0.17	0.91
Chemicals	2.37	2.43	2.32	2.27	2.31	2.30	2.10	1.96
Pharmaceuticals	3.22	3.31	3.31	3.99	3.74	3.70	3.63	4.29
Rubber and Plastic Products	0.76	0.82	0.62	-0.03	0.78	0.84	0.62	-0.02
Mineral Products	0.07	0.22	0.61	2.55	0.07	0.22	0.55	1.99
Basic Metals	3.85	3.54	2.74	2.62	4.15	3.74	2.81	2.37
Metal Products	0.98	0.69	0.08	0.08	0.98	0.68	0.08	0.07
Machinery	0.71	0.58	0.64	0.05	0.70	0.54	0.58	0.04
Office Machinery	1.84	1.45	1.91	1.59	2.26	1.77	2.24	1.82
Electrical Machinery	1.63	0.91	0.48	-0.18	1.76	0.97	0.50	-0.19
Electronic Components	1.59	1.80	3.90	5.84	1.70	1.86	3.47	4.56
TV and Radio	1.92	1.75	1.24	2.00	2.16	1.74	1.15	1.56
Medical and Optical Instruments	0.91	0.54	0.85	1.01	0.94	0.54	0.81	0.93
Motor Vehicles	0.31	0.30	0.52	0.68	0.32	0.32	0.54	0.69
Other Transport	7.69	4.85	2.65	2.67	8.98	5.55	3.02	3.02
Furniture and Leisure Goods	7.46	6.88	6.10	4.64	8.99	8.14	6.89	4.72
Mean	3.62	3.02	2.34	2.04	4.00	3.17	2.34	1.90

Notes: Estimates of $\hat{\gamma}_q^s$ for $q = 1, 2, 3$ and 5 using (19) and (18). Indices expressed as %.

References

- Alegria, R. (2009), 'The Location of Multinational Firms in the UK: Sectoral and Functional Agglomeration', *Spatial Economics Research Centre Annual Conference*, London School of Economics, May.
- Barrios, S. and Strobl, E. (2004), 'Industry Mobility and Geographic Concentration in the European Union', *Economics Letters*, 82, 71-5.
- Barrios, S., Bertinelli, L., Strobl, E. and Teixeira, A-C., (2005), 'The Dynamics of Agglomeration: Evidence from Ireland and Portugal', *Journal of Urban Economics*, 57, 1, 170-88.
- Blomström, M. and Kokko, A. (1998), 'Multinational Corporations and Spillovers', *Journal of Economic Surveys*, 12, 247-77.
- Brülhart, M. and Traeger, R. (2004), 'An Account of Geographic Concentration Patterns in Europe', *Regional Science and Urban Economics*, 35, 597-624.
- Clarke, R. (1985), *Industrial Economics*, Blackwell, Oxford.
- Combes, P. P. and Overman, H. G. (2004), 'The Spatial Distribution of Economic Activities in the European Union', in J. V. Henderson and J-F. Thisse (eds), *Handbook of Regional and Urban Economics*, Volume 4, Elsevier, North Holland, Amsterdam.
- Dimitropoulou, D., Burke, S. and McCann, P. (2006), 'The Determinants of the Location of Foreign Direct Investment in UK Regions', *Regional Science Association International: British and Irish Section Conference*, Jersey, August.
- Döring, T. and Schnellenbach, J. (2006), 'What Do We Know about Geographical Knowledge Spillovers and Regional Growth? A Survey of the Literature', *Regional Studies*, 40.3, 375-95.

- Dumais, G., Ellison, G. and Glaeser, E. L. (2002), 'Geographic Concentration as a Dynamic Process', *Review of Economics and Statistics*, 84.2, 193-204.
- Duranton, G. and Puga, D. (2004), 'Micro-Foundations of Urban Agglomeration Economies', in J. V. Henderson and J-F. Thisse (eds), *Handbook of Regional and Urban Economics*, Volume 4, Elsevier, North Holland, Amsterdam.
- Ellison, G. and Glaeser, E. L. (1994), 'Geographic Concentration in US Manufacturing Industries: A Dartboard Approach', *NBER Working Paper no. 4840*, Cambridge MA.
- Ellison, G. and Glaeser, E. L. (1997), 'Geographic Concentration in US Manufacturing Industries: A Dartboard Approach', *Journal of Political Economy*, 105.5., 889-927.
- Ellison, G. and Glaeser, E. L. (1999), 'Evolution of the Geographic Concentration of Industry', *American Economic Review*, 89, 2, 311-16.
- Glaeser, E. L., Kallal, H. D., Scheinkman, J. A. and Shleifer, A. (1992), 'Growth in Cities', *Journal of Political Economy*, 100, 1126-52.
- Henderson, J. V., Kuncoro, A. and Turner, A. (1995), 'Industrial Development in Cities', *Journal of Political Economy*, 103, 1067-85.
- Henderson, J. V. (1997), 'Externalities and Industrial Development', *Journal of Urban Economics*, 42, 449-70.
- Henderson, J. V. (2003), 'Marshall's Scale Economies', *Journal of Urban Economics*, 53, 1-28.
- Holmes, T. J. and Stevens, J. J. (2002), 'Geographic Concentration and Establishment Scale', *Review of Economics and Statistics*, 84.4, 682-90.
- Jones, J. and Wren, C. (2010), 'Assessing the Regional Impact of Grants on FDI Location: Evidence from UK Regional Policy, 1985-05', *Journal of Regional Science*, forthcoming.
- Kim, S. (1999), 'Regions, Resources, and Economic Geography: Sources of US Regional Comparative Advantage, 1880-1987', *Regional Science and Urban Economics*, 29, 1-32
- Krugman, P. (1991), *Geography and Trade*, MIT Press, Cambridge, MA.

- Lamorgese, and Ottaviano, (2002), Space, Factors and Spillovers, *mimeo*.
- Maurel, F. and Sédillot, B. (1999), 'A Measure of Geographic Concentration in French Manufacturing Industries', *Regional Science and Urban Economics*, 29, 575-604.
- Mood, A. M., Graybill, F. A. and Boes, D. C. (1974), *Introduction to the Theory of Statistics*, McGraw-Hill, Singapore, Third Edition.
- Neary, J. P. (2001), 'Of Hype and Hyperbolas: Introducing the New Economic Geography', *Journal of Economic Literature*, 39, 536-61.
- Quigley, J. (1998), 'Urban Diversity and Economic Growth', *Journal of Economic Perspectives*, 12, 127-38.
- Rosenthal, S. and Strange, W. (2001), 'The Determinants of Agglomeration', *Journal of Urban Economics*, 50, 191-229.
- Rosenthal, S. and Strange, W. (2003), 'Geography, Industrial Organization, and Agglomeration', *Review of Economics and Statistics*, 85, 377-93.
- Rosenthal, S. and Strange, W. (2004), 'Evidence on the Nature and Sources of Agglomeration Economies', in J. V. Henderson and J-F. Thisse (eds), *Handbook of Regional and Urban Economics*, Volume 4, Elsevier, North Holland, Amsterdam.
- Scitovsky, T. (1952), 'Two Concepts of External Economies', *Economic Journal*, 62, 54-67.
- Wren, C. and Jones, J. (2009), 'Re-investment and the Survival of Foreign-Owned Plants', *Regional Science and Urban Economics*, 39, 214-23.