

**DEFINICIÓN DE LA METODOLOGÍA DE DETECCIÓN E  
IDENTIFICACIÓN DE CLUSTERS INDUSTRIALES EN ESPAÑA**

Informe elaborado por el Grupo de Investigación  
**Análisis Cuantitativo Regional** de la Universidad  
de Barcelona, para la **Dirección General de Política  
de la Pequeña y Mediana Empresa**

---

<b>1. Introducción.....</b>	<b>1</b>
<b>2. Objetivos.....</b>	<b>6</b>
<b>3. Clusters industriales: definición, delimitación e interés.....</b>	<b>8</b>
3.1. Definición.....	8
3.2. Delimitación.....	12
3.3. Interés del concepto.....	15
<b>4. Métodos de detección e identificación de clusters.....</b>	<b>21</b>
<b>4.1. Método tradicional à la Porter.....</b>	<b>26</b>
4.1.1. <i>Elementos esenciales del método</i>	
4.1.2. <i>Ventajas e inconvenientes</i>	
4.1.3. <i>Criterio estadístico de detección de aglomeración</i>	
4.1.4. <i>Aplicaciones de la metodología</i>	
4.1.5. <i>Potencialidad de resultados obtenidos</i>	
<b>4.2. Método basado en la comparación de distribuciones.....</b>	<b>43</b>
4.2.1. <i>Elementos esenciales del método</i>	
4.2.2. <i>Ventajas e inconvenientes</i>	
4.2.3. <i>Criterio estadístico de detección de aglomeración</i>	
4.2.4. <i>Aplicaciones de la metodología</i>	
4.2.5. <i>Potencialidad de resultados obtenidos</i>	

<b>4.3. Método basado en la distribución de distancias.....</b>	<b>62</b>
4.3.1. <i>Elementos esenciales del método</i>	
4.3.2. <i>Ventajas e inconvenientes</i>	
4.3.3. <i>Criterio estadístico de detección de aglomeración</i>	
4.3.4. <i>Aplicaciones de la metodología</i>	
4.3.5. <i>Potencialidad de resultados obtenidos</i>	
<b>5. Valoración de la aplicación de los métodos propuestos al caso de la economía española.....</b>	<b>80</b>
<b>Bibliografía.....</b>	<b>89</b>

## 1. Introducción

Las economías, independientemente del ámbito geográfico que abarquen, muestran ciertos patrones de especialización, aprovechando su posición competitiva en uno o varios conjuntos interrelacionados de sectores, a los que se ha dado en denominar como clusters. En este sentido, existe un amplio consenso en que las empresas resultan más competitivas cuando se encuentran agrupadas espacialmente, explotando de esa forma ciertas ventajas productivas asociadas a la proximidad espacial con otras empresas y agentes que intervienen, directa o indirectamente, en la actividad productiva. Esta idea, que se remonta inicialmente a las aportaciones de Alfred Marshall, ha sido profusamente tratada en las últimas décadas tanto desde un punto de vista teórico como empírico en la literatura económica.

La investigación académica y las actuaciones de política económica en torno a los clusters han tenido como referencia básica la definición formulada por M. Porter de *clusters industriales* como “*las concentraciones geográficas de empresas interconectadas, proveedores especializados, proveedores de servicios, empresas en sectores próximos, e instituciones asociadas (como por ejemplo universidades, agencias gubernamentales, asociaciones empresariales, etc) en ámbitos particulares que compiten pero que también cooperan*”.

Una de las características básicas de la línea posiblemente más prolífica en el análisis de clusters industriales, aquella basada en el *análisis de caso*, es la asunción a priori de la existencia en un territorio determinado de uno o varios clusters que constituyen precisamente el objeto del estudio. Es decir, se suele asumir la presencia del cluster que se analiza sin contrastar previamente de forma rigurosa su existencia. Aun en el caso en que en ese tipo de estudios se hayan definido unos criterios para mostrar la presencia de los

clusters, el inconveniente es que dichos criterios suelen ser particulares a cada caso analizado, pudiendo verse afectados por diferencias en la escala geográfica y sectorial considerada. Son precisamente estas circunstancias las que han motivado recientemente la necesidad de establecer una metodología que permita identificar sistemática, robusta y objetivamente clusters industriales y conocer su ubicación precisa en el territorio, en lo que ha dado en denominarse como *mapeado de clusters* (*cluster mapping* en la nomenclatura anglosajona). De esta forma, han aparecido recientemente aportaciones que sugieren métodos que, apoyándose en la definición de cluster industrial, establecen criterios estadísticos para su detección e identificación. Estos métodos han sido ya aplicados a diversas economías habiendo permitido obtener los mapas de, por ejemplo, los Estados Unidos, Gran Bretaña, Suecia, Francia y Alemania.

La utilidad que tiene la disponibilidad de un mapa de clusters para la valoración de la situación existente y para el diseño de políticas, justifica el esfuerzo de su realización de forma sistemática y con una metodología común para el conjunto de la economía española y para cada uno de los niveles en que se puede dividir el conjunto del territorio español (por ejemplo CCAA y provincias).

No obstante, la existencia de diversas propuestas con aproximaciones metodológicas diferenciadas y la elevada demanda de información estadística detallada por parte de algunas de éstas, hacen aconsejable un estudio previo de la adecuación al caso concreto español de los diversos métodos existentes y de la posible necesidad de ajustes en la aplicación de estos. Adicionalmente, resulta conveniente valorar a priori i) la disponibilidad de información estadística de base, en su vertiente sectorial y territorial, ii) las magnitudes económicas de las que se requeriría información, dado que la misma podría condicionar la aplicación de algunos de los métodos disponibles, y iii) la complejidad que entrañaría la implementación de cada uno de los métodos disponibles, dado que esta circunstancia puede condicionar seriamente las posibilidades de actualización y extensión posterior del análisis a ámbitos territoriales y sectoriales concretos.

Las dificultades asociadas a la implementación de todas las propuestas existentes para la realización de un “*cluster mapping*” para la economía española, revisadas críticamente en este informe, no deberían llevar en modo alguno a eludir su realización. Esta afirmación se apoya en los beneficios que se obtendrían de la existencia de un mapa de este tipo para nuestra economía, entre los que destacamos:

- la identificación de puntos fuertes y débiles de competitividad de la economía española,
- la categorización del estado de desarrollo de los clusters existentes,
- la identificación de áreas de elevado potencial de crecimiento
- la identificación de áreas en las que existe margen para incentivar la colaboración entre instituciones y agentes productivos.

Las conclusiones y recomendaciones referidas a la aplicación de una metodología específica para el caso de la economía española, recogidas en este informe, se han obtenido tras realizar una exhaustiva revisión de la literatura existente, en la que se han procesado tanto trabajos estrictamente académicos como informes en los que se recoge la experiencia de la aplicación de algunas de las metodologías propuestas para otras economías. Esta tarea ha permitido identificar dos aproximaciones que incorporan explícitamente un método de detección e identificación de clusters industriales en el territorio. El primero se fundamenta en las aportaciones de Michael Porter y su equipo de colaboradores en el Institute for Strategy and Competitiveness en la Harvard Business School. Con este método, al que denominaremos “*à la Porter*”, esta institución ha elaborado un mapa exhaustivo de clusters para la economía norteamericana, habiendo intervenido también, directa o indirectamente, en la elaboración de mapas para otras economías. La segunda de las propuestas identificadas se configura en torno a los trabajos de Thomas Brenner, investigador del Max Planck Institute of Economics. Su propuesta se basa en la capacidad de ajustar la distribución observada de la actividad económica a través de una distribución teórica

asociada a una situación de aglomeración y a otra asociada a aleatoriedad. Por esta razón lo denominamos en el informe como “*método basado en la comparación de distribuciones*”.

Las diferencias entre ambos enfoques son sustanciales y conducen a pensar que los resultados obtenidos a partir de la aplicación de uno u otro método pueden diferir notablemente, lo que unido al coste asociado a la implementación de ambos, no hace más que justificar la necesidad de valorar a priori las ventajas e inconvenientes de su implementación para el caso de la economía española.

A pesar de no contener un procedimiento explícito que permita la detección e identificación inmediata de clusters, hemos incluido en el análisis efectuado una tercera aproximación, basada en la distancia que separa a las empresas localizadas en el territorio de una economía. Esta estrategia goza de una amplia difusión en el análisis de la localización en el ámbito de otras disciplinas (ecología, biología, epidemiología, etc) y ha sido adaptada para el caso del análisis de la localización de la actividad económica por Marcon y Puech (2003) y por Duranton y Overman (2005). De hecho, este último trabajo ha constituido la referencia fundamental que hemos utilizado para valorar las posibilidades de aplicación del método en él descrito para la realización de un “cluster mapping” para la economía española, al proponer una medida de (exceso de) localización que cumple con un conjunto de propiedades deseables que garantizan la obtención de resultados comparables entre sectores, economías e instantes de tiempo. Además sus resultados son robustos a las unidades geográficas de referencia y se enmarcan en un contexto estadístico que evita la introducción de consideraciones subjetivas. Pero después de considerar detalladamente estas aportaciones debemos concluir que no incorporan ningún instrumento que permita de forma directa la detección e identificación de clusters en el territorio. Aun así hemos optado por incluirlas en el informe dado su posible utilidad como instrumento para valorar la presencia de clusters en los sectores analizados y por la posibilidad de que en el futuro se desarrollen instrumentos que permitan extender los principios de los métodos basados en las distancias a la elaboración de mapas de clusters.

Después de recopiladas y sistematizadas las aportaciones fundamentales de las tres aportaciones metodológicas consideradas, se ha realizado un análisis crítico de las mismas con una doble perspectiva: por una parte teniendo en cuenta cuestiones generales acerca de su aplicabilidad, en la que se ha intentado valorar las posibilidades de mejora de los métodos e instrumentos sugeridos originariamente y, por otra, en referencia a su aplicación específica para la obtención de un mapa de los clusters existentes en la economía española, valorando los puntos fuertes y los débiles de cada una de las opciones a la vista de los condicionantes existentes. Este proceso se ha estructurado finalmente en el informe en:

- i. la descripción sintética de los elementos fundamentales de cada una de las metodologías,
- ii. la presentación de sus ventajas e inconvenientes más destacados,
- iii. la valoración de la existencia de un criterio estadístico que permita vaciar al máximo de subjetividad el proceso de detección e identificación de los clusters, de vital importancia teniendo en cuenta el interés en la implementación de un procedimiento objetivable,
- iv. la síntesis de las características básicas y los resultados más destacados en estudios previos que hayan utilizado la metodología analizada, y
- v. la potencialidad de los resultados que se obtendrían mediante su aplicación al caso de la economía española.



## 2. Objetivos

Como se ha indicado con anterioridad, la obtención de un mapa de clusters en España precisa de la aplicación de alguno de los métodos propuestos en la literatura y de la recopilación de información estadística detallada sobre diversas vertientes de la actividad productiva con elevado detalle tanto sectorial como territorial. Por ello, el objetivo básico de este estudio es el de determinar los puntos fuertes y las limitaciones de los distintos métodos existentes y valorar la viabilidad de su aplicación al caso de la economía española.

Este objetivo básico se puede detallar en otros específicos que a continuación se glosan:

- analizar los elementos básicos constituyentes del concepto de cluster industrial, con el objetivo primordial de valorar la adecuación de los métodos disponibles para su detección e identificación,
- elaborar un *estado del arte* acerca de las características fundamentales de los métodos existentes para la realización de un mapa de clusters,
- recopilar y sintetizar los resultados obtenidos en los *mapeados* realizados hasta el momento para otras economías
- anticipar las necesidades de información estadística de los distintos métodos y las posibilidades de su adecuación a los datos disponibles para la economía española
- determinar la potencialidad del output generado a través de cada uno de los métodos para valorar los resultados de las políticas existentes y contribuir al diseño de las futuras, y

- determinar la dependencia de la obtención de ulteriores análisis en relación al método de mapeado empleado, es decir, en qué medida los métodos de mapeado facilitarán, por ejemplo, la visualización geográfica y sectorial de clusters, la clasificación de éstos, la determinación de los factores que los favorecen y estimulan, sus efectos sobre la productividad y la generación de empleo en el territorio, y la detección de clusters “clave” en función de criterios pre-establecidos.

### **3. Clusters industriales: definición, delimitación e interés**

#### **3.1. Definición**

El concepto de cluster o agrupación industrial ha suscitado un elevado interés académico y político desde la última década del siglo pasado, a pesar de que algunos de los elementos fundamentales sobre los que se sustenta procedan de argumentos postulados con anterioridad (Marshall, 1890). Entre otras posibilidades para explicar la extensión y generalización del concepto se encuentra sin duda su potencial para describir y, al menos aparentemente, explicar la heterogeneidad en la distribución espacial de la actividad económica y la especialización y competitividad de las regiones, o más genéricamente territorios, en determinadas actividades productivas. Todo ello bajo la asunción del principio de rendimientos crecientes a escala vinculados a la localización, lo que permite obtener un mayor rendimiento de la concentración espacial de los factores productivos. A todo ello cabe añadir su supuesto atractivo para fundamentar políticas de desarrollo en regiones, ciudades y, en general, localidades con bajos niveles de desarrollo y con escasa densidad de actividad y, más recientemente, como un posible mecanismo con el que garantizar la *adhesión* de la actividad productiva al territorio ante la amenaza de la deslocalización de actividades hacia economías emergentes con menores costes productivos.

La forma a través de la cual los clusters industriales supuestamente permiten superar desventajas locacionales como las anteriores tiene que ver con la idea de que se alcanza un desarrollo industrial más potente si las empresas se agrupan creando una masa crítica suficiente que les permita alcanzar unas mayores pautas de crecimiento, aprovechar mejor las posibilidades de inversión, de colaboración formal e informal, y de compartir e intercambiar conocimiento. A pesar de que existe un cierto debate en cuanto a la verosimilitud de estos hechos, resulta indiscutible que la actividad productiva se distribuye

en el espacio de forma no aleatoria ni uniforme, siguiendo por el contrario patrones regulares y, en cierta medida, jerarquizados como resultado de la actuación simultánea de los mecanismos de rendimientos a escala asociados a la localización y a costes de transacción vinculados a la distancia (Krugman, 1993; Fujita y Krugman, 1995).

La heterogeneidad en la distribución espacial de la actividad está relacionada (puede ser entendida simultáneamente como causa y como efecto) con la especialización productiva de las unidades espaciales. De hecho, a pesar de que la especialización de una localidad en una actividad determinada puede haber tenido su origen en la disponibilidad de materias primas, en el acceso al mercado o en simples accidentes históricos, la persistencia de tal especialización se explica a través de mecanismos auto-sostenidos que tienen que ver con la localización de redes de proveedores de bienes intermedios, la existencia de un mercado de trabajo especializado, y externalidades tecnológicas producidas y aprovechadas por las empresas localizadas en ese territorio. A todo ello cabe añadir la generación de un ambiente y una cultura específica del tipo de actividades desarrolladas y que difícilmente son generables o trasladables a otros entornos.

Es la heterogeneidad en la distribución espacial de la actividad la que unida a la consecuente especialización sectorial de los territorios, y a las causas que las pueden estar provocando, la que conduce a la noción de cluster industrial. Pero una revisión de la literatura nos permite concluir que estos dos rasgos genéricos son los únicos que resultan comunes a la práctica totalidad de las numerosas contribuciones a la literatura académica y a aquellos estudios que han pretendido detectar y caracterizar clusters industriales en diversas economías. Como señalan Gordon y McCann (2000), es en esta ambigüedad del concepto en la que reside el principal problema de su tratamiento y uso. De hecho, ha sido práctica habitual adaptar la definición del fenómeno al interés concreto del estudio que en cada caso se estaba realizando, como se deduce de la recopilación de definiciones de cluster industrial aparecidas en la literatura en la Tabla 1 del artículo de Martin y Sunley (2003) y recogidas en el Cuadro 1.

**Cuadro 1. Recopilación de definiciones de cluster industrial.**

---

*“Un cluster es un grupo de empresas interconectadas e instituciones relacionadas en un determinado campo, que se encuentran próximas geográficamente, y que están vinculadas a través de elementos comunes y complementariedades.”* Porter (1998a)

*“El concepto más general de cluster sugiere menos estricto: una tendencia de las empresas dedicadas a actividades similares a colocalizarse, aunque sin tener una presencia particularmente importante en un área geográfica.”* Crouch y Farrell (2001)

*“Un cluster es utilizado de forma muy sencilla para representar concentraciones de empresas que son capaces de producir sinergias debido a su proximidad geográfica y a la existencia de interdependencias entre ellas, a pesar de que su peso en el empleo total no sea preponderante ni incluso relevante.”* Rosenfeld (1997)

*“Los clusters económicos no sólo tienen que ver con sectores e instituciones de apoyo, sino que tienen que ver con ellos tanto en cuanto sean más competitivos debido a las relaciones que establecen entre ellos.”* Feser (1998)

*“Los clusters son definidos aquí como un grupo de empresas pertenecientes a un sector localizadas en un mismo área geográfica.”* Swann y Prevezer (1996)

*“Un cluster es un amplio grupo de empresas de sectores similares en una localización determinada.”* Swann et al (1998)

---

*“Definimos un cluster innovador como un amplio número de empresas industriales o de servicios que tienen un alto nivel de colaboración, habitualmente a través de una supply chain, y operando bajo similares condiciones de mercado.”* Simmie y Sennett (1999)

*“Los clusters pueden ser caracterizados como redes de empresas muy interdependientes (que incluyen proveedores especializados) relacionadas entre sí a través de una cadena de producción de valor.”*

Roelandt y Den Hertag (1999)

*“El popular término cluster está fundamentalmente ligado a esta dimensión local o regional de las redes... La mayoría de las definiciones comparten el concepto de cluster como una red localizada de organizaciones especializadas, cuyos procesos productivos están estrechamente vinculados a través del intercambio de bienes, servicios y/o conocimiento.”*

Van der Berg et al (2001)

*“Un cluster regional es un cluster industrial in el que las empresas integrantes están próximas unas a otras.”* Enright (1996)

---

---

A todo ello cabría añadir la tendencia a utilizar términos como cluster, aglomeración, nuevo distrito industrial, milieu, y complejo para denotar supuestamente la misma realidad, a pesar de que resulten razonables las dudas sobre la similitud de la realidad a la que se les aplicaba en cada caso. En este sentido, Gordon y McCann (2000) sugieren tres tipos básicos de cluster industrial:

- el modelo clásico de aglomeración pura, basado en las ideas de Marshall y caracterizado por la ausencia de cooperación entre los agentes, la libre pertenencia y la importancia de la co-localización
- el modelo del complejo industrial, caracterizado por la existencia de relaciones estables e identificables entre empresas, que dan lugar a un club cerrado, y la necesidad de compartir una misma ubicación en el espacio
- el modelo de la red social, basado en la fuerza de las interacciones sociales entre empresas, en las que la confianza personal juega un papel preponderante y donde la pertenencia no es completamente libre aunque tampoco se trata de un club completamente cerrado. La localización espacial no es imprescindible, aunque sí favorece el establecimiento de los lazos de confianza que sustentan la red

En consecuencia, la ambigüedad y el carácter circunstancial que se desprende de las diversas definiciones tienen serias implicaciones para cualquier estudio que pretenda la detección objetiva de clusters y su posterior delimitación y análisis de características. Además, pensamos que toda esta confusión tiene una clara repercusión en cuanto a la eficacia de las medidas de política económica que pretendan estimular la creación o fortalecimiento de agrupaciones espaciales de empresas y otros agentes, lo suficientemente dinámicas como para favorecer el crecimiento económico y la generación de empleo. Esto es así dado que la metodología utilizada en cada caso, al basarse en la vaguedad del concepto puede haber conducido a resultados sesgados y, en consecuencia, haber dado lugar a conclusiones poco robustas.

### **3.2. Delimitación**

A pesar de la existencia de múltiples acepciones, existe amplio consenso en situar el origen del concepto o idea de clusters industriales, tal y como se ha venido utilizando en la última década, en la aportación de Michael Porter, quien los define como:

“Concentraciones geográficas de empresas interconectadas, proveedores especializados, servicios empresariales, compañías en sectores próximos e instituciones asociadas (como por ejemplo universidades, agencias gubernamentales, y cámaras de comercio) en actividades productivas específicas, que compiten pero también cooperan.” Porter (1998b, pág. 197)

Es pues la interrelación de los diversos agentes localizados en un espacio concreto la fuerza que estimula y favorece el desarrollo del cluster, a través de lo que Porter ha popularizado como el *diamante competitivo*. Y a su vez, el cluster no es más que la manifestación espacial de ese *diamante*. La generalidad de la definición permite aplicar el concepto a un amplio espectro de agrupaciones industriales, de niveles de relaciones *input-output*, de entornos institucionales y de ámbitos geográficos. Analizando la definición, se pueden aislar dos elementos fundamentales:

- las empresas del cluster están relacionadas bien formalmente a través de relaciones verticales u horizontales, bien de forma más difusa mediante el papel cohesionante ejercido por algún tipo de instituciones, dando lugar a un entramado o red empresarial
- el cluster está constituido por empresas relacionadas que además están próximas en el espacio, es decir hay co-localización empresarial dado que la decisión de localización espacial de una empresa depende en gran medida de las decisiones de localización que hayan tenido otras empresas con las que potencialmente puede establecer vínculos como los descritos anteriormente.

Así, vemos como tanto la dimensión sectorial como la geográfica juegan un papel fundamental en la definición, aunque el problema inmediato surge cuando se trata de delimitar ambas dimensiones (Martin y Sunley, 2003). No hay nada en la definición del concepto ni en los numerosos desarrollos en la literatura que nos ofrezca una guía fundamentada y basada en criterios objetivos para establecer los límites sectoriales y espaciales de un cluster industrial. Nada nos permite establecer a priori el detalle de la



desagregación sectorial a considerar. Ni tampoco el tipo e intensidad de los vínculos entre sectores y entre empresas necesario para considerarlas como integrantes de un cluster, o el grado de especialización que debería presentar el supuesto cluster para ser considerado como tal. La falta de precisión y la imposibilidad de generalización quedan patentes a partir de las propias palabras de Porter:

*“La delimitación de un cluster es frecuentemente una cuestión de matiz, implicando un proceso creativo que debe estar determinado por el conocimiento de los vínculos y complementariedades existentes entre empresas e instituciones”* Porter (1998a, pág 202)

El mismo grado de imprecisión aparece respecto a los límites espaciales o geográficos. Es obvio que la dimensión geográfica es consustancial al concepto tal y como se desprende de las propias palabras de Porter:

*“El proceso de agrupación, y el intenso intercambio entre empresas en el cluster, funciona mejor cuando éstas están geográficamente concentradas”* Porter (1990, pág. 157)

Pero es asimismo obvio que tampoco encontramos en este caso ninguna guía que nos permita establecer a priori si los límites geográficos de los clusters son más o menos amplios, ni cual es el ámbito espacial en el que deberían estar operando las fuerzas generadoras del cluster (vínculos empresariales, externalidades de conocimiento, externalidades pecuniarias, redes sociales, etc). En la definición tampoco existe ninguna referencia explícita y precisa al grado de densidad espacial de las actividades o de las interacciones entre agentes en un espacio delimitado. En consecuencia, resulta imposible establecer a priori un umbral a partir del cual concluir objetivamente acerca de la presencia de un cluster industrial. De hecho, tanto de las discusiones metodológicas como de las numerosas implementaciones prácticas puede deducirse que es posible encontrar clusters en cualquier nivel de desagregación geográfica. Así, no parece incompatible referirse a clusters de sectores y/o empresas a nivel nacional sin que sea necesario que estos se

encuentren agrupados espacialmente en el territorio de ese país, como a clusters de empresas intensamente relacionadas y localizadas en un área espacial muy delimitada. Ambos casos parecen merecer el mismo calificativo sin que resulte necesario explicitar las claras diferencias entre ambos fenómenos. La única justificación parece encontrarse en cual sea el objetivo y el ámbito del análisis que se está efectuando o cuáles son los intereses del destinatario último del estudio.

### 3.3. Interés del concepto

Esta imprecisión ha llevado a que, desde una perspectiva bondadosa, se haya calificado a la naturaleza del concepto de cluster como *amplia y metafórica* (DTI, 2001), mientras que desde ámbitos más críticos no se haya dudado en calificarlo como de *caótico* (Martin y Sunley, 2003). De hecho esta circunstancia ha llevado incluso a estos últimos autores a sugerir que la definición es intencionadamente *opaca y difusa*, y a que en muchos casos se haya recurrido a una redefinición específica adaptada a las características del análisis que se pretendía realizar. Asimismo, sugieren que en la escasa precisión del concepto reside una de las claves de su popularidad, dado que de esa forma se ha podido incluir en él a un amplio abanico de agrupaciones industriales, vínculos input-output, entornos institucionales, etc. Siempre ha resultado posible encontrar una acepción del concepto que permitiese calificar a un determinado fenómeno de agrupación de agentes y actividades productivas como un cluster industrial, seguramente porque de esa forma era más sencillo dotarlo de unas determinadas cualidades y características sin necesidad de probar su presencia. Así, los clusters industriales han llegado a tener “*el discreto encanto de los oscuros objetos del deseo*” (Steiner, 1998) y “*más que constituir un modelo o teoría rigurosamente contrastado y evaluado, la idea de cluster ha llegado a ser ampliamente aceptada en base a la fe como una forma de pensar válida y razonable acerca de la economía nacional y como un método mediante el cual descomponer la economía en agrupaciones geográfico-sectoriales diferenciadas, con el propósito de entender y promover la competitividad y la innovación*” (Martin y Sunley, 2003).

En contraposición a la corriente imperante, podríamos pensar que los clusters industriales son fenómenos excepcionales, una especie de caprichos de la naturaleza económica que como tales resultan difícilmente reproducibles. Esta visión puede ser más verosímil en un mundo como el actual en el que la globalización, el uso de las nuevas tecnologías y el flujo más fluido de factores, nos hace cuestionar la generalización del atractivo de una localización concreta. Es bajo estas circunstancias cuando la aproximación metodológica al uso en el análisis de los clusters industriales podría conducir a mayores errores en la extracción de conclusiones generales y en el diseño de políticas encaminadas a favorecer la formación y consolidación de clusters en otras ubicaciones. Esto es así ya que los resultados extraídos del análisis de uno o varios casos correspondientes a localidades y actividades muy concretas no podrían ser extrapolados a otras situaciones dado que estarían contaminados por un sesgo de selección (Markusen, 1996). O dicho con otras palabras, sería un error pensar que se puede crear un Silicon Valley en un número elevado de localidades. Gordon y McCann (2000) describen esta circunstancia:

*“Un tercer factor significativo parece ser la idea de la inadecuación de las aproximaciones tradicionales a la política espacial a la vista de los desarrollos macroeconómicos y estructurales de los años ochenta y noventa, que suscitó la necesidad de estrategias innovadoras para estimular el crecimiento endógeno en áreas poco desarrolladas y marginales. Esto ha suscitado varios intentos de extraer conclusiones genéricas a partir de un número muy limitado de casos exitosos...”* Gordon y McCann (2000 pág. 515)

Pero a pesar de ello, parece existir cierta fiebre por promover clusters de negocios en todo el mundo, dado que a la moda académica hay que sumar el interés político en este tipo de fenómeno, motivado por la asociación que se ha realizado entre el concepto de cluster y las ideas de capacidad innovadora, competitividad y motor de crecimiento y generación de empleo en el conjunto de la economía (OCDE, 1999; 2001). Siguiendo a Markusen (1996), la percepción favorable de los clusters, en sus diversas acepciones, se basa en la capacidad de estos para generar puestos de trabajo y para garantizar estabilidad y dinamismo a largo plazo. En concreto, se ha sugerido que:

- aseguran un crecimiento para el conjunto de la región en el que se encuentran por encima del nivel medio,
- impiden que la región sufra la pérdida de empleo y el cierre de empresas ocasionada por los momentos de crisis del ciclo económico
- proporcionan empleo de calidad, contrarrestando la tendencia al aumento de la desigualdad en la distribución de la renta, al prever la concentración excesiva de la riqueza
- fortalecen la representación y participación del factor trabajo en las decisiones empresariales
- incentivan la participación en la política regional

En cualquier caso, es importante destacar que existe suficiente unanimidad en cuanto a la ineficacia de las iniciativas orientadas a la creación o generación de un nuevo cluster en una localidad determinada (Schmitz y Nadvi, 1999), a la vez que se amplía el consenso respecto a la idea de que las actuaciones de los planificadores deben tener por objetivo facilitar un entorno vital que favorezca el despegue y consolidación de un cluster embrionario. Siguiendo a Scott (1998) y a la OCDE (1999), la actuación política debe encaminarse a garantizar la oferta de bienes públicos, no presentes debido a fallos de mercado, en el área geográfica donde se sitúa el cluster potencial. Martin y Sunley (2003) agrupan estos bienes en cuatro tipos:

- la presencia de redes cooperativas que requieren la organización de dialogo previo entre los agentes implicados,
- la existencia de marketing colectivo de la especialización industrial de la localidad,

- la presencia de servicios a las empresas en la localidad, de tipo financiero, de marketing, diseño, innovación, etc
- la detección de lagunas y debilidades en la cadena de valor y la generación de incentivos para facilitar la atracción de agentes productivos e inversores que cubran esos huecos.

Pero aún en ese caso, la guía de actuación dista de resultar nítida, entre otras razones por la propia indeterminación del concepto que se pretende potenciar. Si se desea cubrir la necesidad de bienes públicos necesarios para favorecer el despegue de clusters embrionarios, parece obvio que se debe poder detectar de forma relativamente precisa dónde se encuentran dichos clusters y cuál es la actividad que los caracteriza. Es decir, se precisa de información acerca del área geográfica y del sector de actividad en el que opera el supuesto cluster. Pero como se ha comentado anteriormente, no tenemos a nuestra disposición una clara delimitación de las unidades territoriales en las que deberían operar los clusters ni el detalle sectorial en el que deberíamos tratar de localizarlos. Esta circunstancia tiene repercusiones no únicamente en la detección de los supuestos clusters que merecerían atención específica, sino también para el propio diseño y ejecución de la atención dado que, como indican Enright y Ffowcs-Williams (2001), las políticas de apoyo a los clusters deberían ser realizadas por aquellos niveles de gobierno más próximos al ámbito geográfico en el que se circunscribe el cluster.

Tampoco resulta obvio que garantizar la oferta de bienes públicos en una localidad en la que pudiera encontrarse un cluster embrionario resulte siempre exitoso, a pesar de que se base en la experiencia extraída de clusters que son habitualmente tomados como modelo. Dicho de otra forma, cabe dudar de la generalización de las *best practices* o los *blueprints* a cualquier región o localidad, dado que difícilmente se reunirán todas las condiciones que condujeron al éxito de un caso concreto (Martin y Sunley, 2003). Sería interesante que para ello se cotejase la experiencia con intentos que no acabaron produciendo los resultados

deseados. Ese proceder se ajustaría en mayor medida a lo que cabría exigir a cualquier método con características científicas y permitiría valorar la efectividad y rendimiento social de las medidas de apoyo a los clusters. A esto cabe añadir las dudas que se pueden plantear en cuanto a la capacidad de los clusters para actuar realmente como motores de crecimiento regional y local. En primer lugar, se puede argumentar que mientras no cabe dudar de que las empresas compiten, y que en ese proceso competitivo se estimula la innovación, la aplicación de las mejores prácticas o la destrucción creativa, es menos inmediato asumir que las regiones o las localidades compitan y en qué medida esa competencia tiene efectos conjuntos positivos. Así, el establecimiento de un cluster en una región puede perjudicar a empresas de la misma actividad localizadas en otras regiones, de forma que el efecto neto sobre el conjunto de la economía resulte nulo o como mínimo mucho menor al que se deriva de la consideración aislada del cluster. Incluso la consolidación de un cluster podría tener efectos perversos sobre otras actividades localizadas en esa misma región pero que resulten ajenas al cluster. Por ejemplo, no se deberían olvidar los costes de congestión generados por un cluster en la localidad en la que se ubica y que acabarían afectando a todo su tejido productivo. Todo ello nos lleva a sugerir la necesidad de realizar una valoración del efecto neto de las políticas de apoyo a clusters industriales sobre el conjunto de la economía y a, en cualquier caso, suscribir la importancia de la coordinación supralocal/supraregional en la aplicación de tales medidas. Siguiendo a Markussen (1996) dicha coordinación debería basarse en:

- valorar cuántos clusters son sostenibles y adecuados para la economía en la que se esté interesado,
- desarrollar estrategias para estabilizar los clusters existentes y valorar las posibilidades reales de aquellos embrionarios en áreas deficitarias,
- impedir la utilización de recursos públicos para subsidiar la competencia entre regiones y localidades,

- acentuar las políticas de ámbito regional y local pero potenciando las alianzas entre regiones y localidades, dada la ausencia de correspondencia entre los límites administrativos y los correspondientes a los clusters industriales

Por último, y a modo de conclusión de esta sección, cabe indicar que creemos que en la indeterminación y poca precisión asociadas al concepto de cluster industrial puede residir la causa fundamental de la falta de sofisticación, rigurosidad y objetividad de los métodos de detección y cuantificación de clusters industriales. Pero también que la práctica al uso, consistente en el diseño de métodos *ad hoc* de detección, ha facilitado la proliferación de la citada indeterminación y, sobre todo, la ausencia de elementos que permitan garantizar la robustez de los resultados obtenidos en los diversos estudios. Por ello, creemos importante la valoración de los diferentes métodos disponibles, acentuando sus virtudes y puntos débiles, así como los requerimientos de información estadística de cada uno de ellos.

#### 4. Métodos de detección e identificación de clusters

En la línea de argumentación de la sección anterior, Rosenfeld (1997) manifestó que “*hay tantas definiciones de clusters como número de instituciones que utilizan el término*”. Podemos extender esta afirmación a las metodologías aplicadas en los estudios que han detectado e identificado clusters industriales, probablemente a causa de la propia generalidad del concepto y a la poca precisión en cuanto a los límites geográficos y sectoriales de los clusters. En cualquier caso, es posible agrupar los métodos propuestos y utilizados en la literatura para realizar el conocido como “*cluster mapping*”, es decir aquellos instrumentos que supuestamente permiten representar los clusters existentes en un área geográfica, incluyendo la información referida al volumen de negocio, los principales agentes implicados, y otros datos clave como el número de empresas, la cantidad de empleo, etc. En función del tipo de aproximación empleada para establecer el *mapping*, encontramos:

- a) **Métodos basados en el uso de herramientas analítico-estadísticas**, de distinto grado de sofisticación y cuya clave reside en la medición del grado de aglomeración económica existente en cada una de las áreas geográficas consideradas.

El instrumento cuantitativo más frecuente es el coeficiente de localización (CL), que aproxima el grado de ventaja competitiva de un sector en cada área geográfica respecto a la del conjunto de la economía analizada. Dadas las limitaciones que conlleva el uso aislado del CL, suele ser habitual acompañarlo de medidas complementarias. Así, en ocasiones se utiliza la matriz de crecimiento-participación, que está configurada, además de por el CL, por:



- una medida del tamaño absoluto del sector en la región, habitualmente aproximado por el empleo sectorial
- una medida del crecimiento del sector en la región, aproximada a través del crecimiento del empleo sectorial

Tanto en el caso de la utilización del CL de forma aislada, como combinado a través de la matriz de crecimiento-participación, esta aproximación está sujeta a las críticas de ser sensible a la clasificación sectorial utilizada, que además suele venir dada de forma exógena y rígida, y de no utilizar información acerca de la interdependencia entre sectores, que constituye precisamente un elemento esencial en la propia definición de cluster. Esta última circunstancia ha llevado a autores como Bergman y Feser (1999) a indicar que el CL sólo resulta útil si se complementa con una medida del grado de interdependencia entre sectores. Precisamente, la combinación de medidas del grado de concentración sectorial con medidas de la intensidad de relaciones intersectoriales es el elemento fundamental del método de “*cluster mapping*” propuesto y aplicado por M Porter en el contexto del Institute for Strategy and Competitiveness en la Harvard Business School.

En contraposición al método desarrollado por Porter y sus colaboradores, recientemente han aparecido propuestas alternativas también basadas en la aplicación de instrumentos analíticos y cuantitativos. Estas propuestas metodológicas serán descritas con mayor detalle en los próximos apartados de este informe, por lo que en este punto únicamente indicaremos que se basan en la utilización de información más detallada tanto a nivel de las unidades de observación como de las unidades geográficas consideradas y de la situación teórica de referencia en ausencia de aglomeración de la actividad.

En todo caso, y aún combinando información del grado de concentración y de interrelación sectorial, debe tenerse en cuenta que:

- este tipo de método no tiene en cuenta el ciclo vital de un cluster, es decir si éste se encuentra en gestación, en plena efervescencia o en declive. Esta circunstancia puede condicionar el proceso de detección de algunos clusters de especial relevancia,
- el método tiene validez limitada al no incorporar información de factores intangibles como, por ejemplo, relaciones basadas en la confianza, la presencia de capital social, de dinámica organizativa, etc.

b) **Métodos basados en el estudio de caso** de clusters individuales o de grupos de clusters, mediante el uso de técnicas cualitativas y de supuestos a priori sobre la existencia del (los) cluster(s) analizados.

Es utilizado por analistas que rechazan el supuesto de que los clusters sólo pueden ser identificados y cuantificados mediante la aplicación de técnicas estadísticas. El argumento básico que utilizan es que el crecimiento y la dinámica innovadora de un cluster únicamente pueden ser capturadas de forma correcta a través de la aplicación de técnicas cualitativas, por ejemplo mediante el recurso a entrevistas en profundidad con agentes representativos de los clusters. Un paradigma de tal aproximación lo constituye el estudio de Saxenian (1994) acerca de Silicon Valley en los noventa, a la que siguieron otros trabajos en la misma línea que trataron de comparar esa experiencia con la de otros supuestos clusters.

La conclusión fundamental que se desprende de los estudios que han utilizado este método cualitativo es que las únicas políticas efectivas parecen ser aquellas diseñadas para fortalecer clusters ya existentes, mientras que las tendentes a promover el nacimiento o formación de un cluster, a través de acciones *top-down*, aparecen como ineficaces.

El enfoque basado en el estudio de caso puede proporcionar conocimiento de la naturaleza y la dinámica de los clusters industriales presentes en un área geográfica

determinada, así como de las claves de su éxito. Pero difícilmente puede ser utilizada esta aproximación para la detección de tales clusters. O dicho de otro modo, el estudio de caso puede proporcionar un conocimiento profundo de un cluster o grupo de clusters previamente detectados mediante la aplicación de otros instrumentos. En caso contrario se puede acabar aceptando erróneamente su existencia. Adicionalmente, hay que tener en cuenta que no resultará adecuado comparar la evidencia obtenida de casos distintos si estos no se han derivado del mismo marco de análisis. Así, difícilmente se podrá dar respuesta a por qué unas experiencias han resultado exitosas mientras otras han fracasado, o por qué unas han resultado mejores que otras.

En consecuencia, podemos concluir que el método basado en el análisis de caso no constituye una aproximación metodológica adecuada para la realización de un “*cluster mapping*”.

- c) **Métodos centrados en el análisis de políticas públicas** y estrategias diseñadas para promover la gestión y/o fortalecimiento de un cluster o conjunto de clusters. Suelen corresponder a análisis efectuados por autoridades competentes en el desarrollo local y/o regional, e incorporan elementos tanto de a) como de b).

A partir de esta clasificación general de las aproximaciones utilizadas en la literatura para el análisis de los clusters industriales podemos concluir que cualquier estudio tendente a su detección y cuantificación en el conjunto del territorio español debería basarse en la aplicación de instrumentos cuantitativos. Debería también explicitar claramente qué aspectos concretos de la definición de cluster industrial adoptada se están recogiendo con los instrumentos utilizados, si estos se incorporan de forma parcial o total, y si se hace de forma directa o mediante algún tipo de aproximación. Ello permitirá valorar en su justa medida el mapa obtenido.

A continuación se describen en más detalle cada uno de los métodos propuestos en el ámbito de la aproximación cuantitativa, presentando de forma resumida los rasgos característicos de las técnicas empleadas, los requerimientos de información estadística, el ámbito territorial característico, las ventajas e inconvenientes de su aplicación y la potencialidad de los resultados que se podrían obtener.

## 4.1. Método à la Porter

Michael Porter y su equipo de colaboradores en el Institute for Strategy and Competitiveness en la Harvard Business School propusieron un método basado en la utilización de instrumentos cuantitativos para derivar la composición de economías regionales y locales, y los límites geográficos de los clusters de sectores interrelacionados. Dicho método pretende capturar los elementos básicos en la definición de cluster industrial del propio Porter, sirviéndose para ello de algunas medidas utilizadas con anterioridad para valorar el grado de concentración espacial de la actividad y la intensidad de las relaciones entre distintos sectores productivos. Al determinar el conjunto de sectores que se encuentran relacionados en uno o varios clusters concentrados geográficamente el método de Porter da lugar a un “*cluster mapping*” para el conjunto de la economía en que se esté interesado.

### 4.1.1. Elementos esenciales del método.

#### Definición de sectores

La implementación del método de Porter precisa de la clasificación de todos los sectores considerados en tres grandes tipos de agrupaciones con pautas de competencia espacial y patrones de localización geográfica muy dispares:

- a) **Sectores locales.** Son aquéllos distribuidos uniformemente entre regiones, de forma que su dimensión en cualquier región (medida por ejemplo a través del empleo) es proporcional a la población de la región. Por las características del producto de este tipo de actividades, las empresas de una región compiten sólo parcialmente con las localizadas en otras regiones. Aunque en la mayoría de casos corresponden a actividades de servicios y de construcción, también pueden considerarse en esta categoría algunas actividades manufactureras.

- b) **Sectores dependientes de recursos.** Son aquellos en los que la localización geográfica de la actividad está asociada a la presencia de determinados recursos naturales. Las empresas de estos sectores compiten con otras del mismo sector localizadas en otras regiones del país y con aquéllas de otros países.
- c) **Sectores comercializables.** Son aquéllos cuya producción tiene como mercado la propia región, el resto de regiones del país y potencialmente el resto de países, y cuya localización, además, no depende de la ubicación de determinados recursos naturales. La localización en una región específica se determina en función de un amplio abanico de factores de competitividad, de forma que la concentración geográfica de este tipo de sectores varía sustancialmente entre regiones.

### **Clasificación de sectores**

En la práctica, Porter sugiere:

- utilizar la distribución sectorial del empleo en un año determinado para clasificar los sectores considerados en los tres grupos anteriores, y
- emplear tres medidas de variación del empleo sectorial en el espacio para los distintos sectores:
  - i. el porcentaje que sobre el total de empleo nacional representa el empleo de aquellos estados con un  $CL \geq 1$ ,
  - ii. la media del CL para los cinco estados con un valor más elevado para ese coeficiente,
  - iii. el coeficiente de Gini del empleo.
- examinar el patrón de distribución regional del empleo para un amplio conjunto de sectores con el objetivo de determinar “puntos de corte” para cada una de las medidas

anteriores. En el caso concreto del estudio de Porter para los estados norteamericanos, estos se definieron como:<sup>1</sup>

- i. el 50% o más del total del empleo de un sector se encuentra en estados con  $CL \geq 1$ ,
- ii. la media del CL en los cinco estados con mayor valor para este coeficiente toma un valor mínimo de 2,
- iii. el coeficiente de Gini del empleo es igual o mayor a 0.3

Así, aquellos sectores en los que se cumplen simultáneamente los tres criterios son definidos como sectores comercializables. Por su parte, aquellos en los que no se cumple ninguno o únicamente uno de los criterios anteriores, son clasificados como locales.

En una segunda etapa se realiza un ajuste fino, una depuración de la clasificación obtenida como resultado de la aplicación de los criterios anteriores. Tal depuración se basa en elementos definidos de forma totalmente *ad hoc*:

- para los sectores en los que se cumplen dos de los tres criterios, se examina la distribución de su empleo y el tipo de actividad que realiza. En caso de que se considere que algunos de esos sectores comparten las características requeridas a un sector comercializable son clasificados como tales,
- para los sectores clasificados como comercializables, si se considera que el tipo de actividad para alguno de ellos se aleja de las características típicas de esa agrupación, y tras analizar detalladamente la distribución geográfica de su empleo, se reclasifican como locales,

---

<sup>1</sup> La arbitrariedad de este procedimiento es aceptada por el propio Porter al indicar que “a pesar de que los puntos de corte utilizados para realizar la clasificación es arbitraria, su modificación sólo ocasionaba pequeñas modificaciones en los resultados”.

- para los sectores clasificados como comercializables, pero cuyo tipo de actividad es altamente dependiente de la localización de recursos naturales, se reclasifican como sectores dependientes de recursos naturales

### **Agrupación de sectores**

Como resultado final de este proceso se obtienen los sectores comercializables, aquellos con marcadas pautas de concentración geográfica, y las regiones en las que cada uno de ellos presenta una mayor presencia. No obstante, con esa información no finaliza el proceso de *mapping*, dado que Porter reconoce que los sectores pueden no ser las unidades de análisis apropiadas, dada la más que probable existencia de externalidades entre sectores relacionados en el seno de los clusters. Y precisamente en este punto es donde se plantea una de las principales limitaciones del método aquí sintetizado: el que se carezca de un instrumental sólido que, junto a la ausencia de información estadística referida a la difusión de conocimiento y a externalidades pecuniarias, imposibilita determinar unívocamente la composición sectorial de los clusters para un amplio conjunto de economías regionales.

El reconocimiento de esta carencia obliga a la aplicación de un método indirecto, mediante el cálculo de la correlación de la distribución regional del empleo en el conjunto de los sectores comercializables. La idea es que elevadas correlaciones locacionales deben ser motivadas por las citadas externalidades, pudiendo ser utilizadas para definir los límites y composición de los clusters. Como ejemplo típico se puede señalar el de la asociación entre la localización geográfica del empleo en el sector de desarrollo de software y en el del hardware.

Las correlaciones estadísticamente significativas permiten identificar pares y después grupos de sectores que supuestamente presentan una estrecha relación. En la propuesta de Porter se rechaza el uso de algoritmos de agrupación basados en los coeficientes de correlación al no ser considerados como adecuados por excesivamente automatizados. En su lugar se aboga por una actuación pragmática en la que se comienza a través de un



pequeño grupo de sectores obviamente relacionados y a partir de éste se sigue la pista a otros sectores a partir de los patrones de correlaciones.

**Depuración de la agrupación de sectores**

Una vez llegado a este punto, las agrupaciones sectoriales obtenidas a través del análisis de correlaciones son depuradas de la posible presencia de correlaciones espurias entre algunos sectores. La presencia de correlaciones espurias tiene entre otros posibles motivos los siguientes:

- el empleo de una definición muy amplia de los sectores considerados,
- el que los datos no permitan diferenciar entre el empleo de las sedes y de las plantas de producción locales,
- el que sectores con una intensa presencia en regiones grandes puedan aparecer como altamente correlacionados incluso en ausencia de ningún tipo de relación entre ellos,
- el que sectores de escasa implantación en la economía analizada puedan no estar presentes o lo estén con niveles bajos de empleo en muchas regiones, dando lugar de esa forma a correlaciones significativas, y
- el que sectores fuertemente implantados en regiones grandes muestren, simplemente por esta circunstancia y sin que exista interrelación entre ellos, correlaciones elevadas.

La depuración se basa en el uso de información adicional acerca de los sectores considerados y de la información sobre los intercambios comerciales entre ellos, contenida en las tablas input-output. En concreto:

- se utiliza, en primer lugar, toda la información y conocimiento disponible de cada sector. Para ello se emplea la máxima desagregación sectorial posible y, por ejemplo, el listado de productos característicos de cada sector. Con esa información se concluye acerca de la probable presencia de externalidades “lógicas”,
- en segundo lugar, se comprueba la existencia de flujos comerciales entre sectores a través de la información contenida en las tablas input-output.

De esta forma, se elimina el vínculo entre cualquier par de sectores por espurio si entre ellos no se detecta una externalidad “lógica” ni flujos significativos de productos.

Adicionalmente, se incluyen asociaciones no detectadas a través del análisis de correlaciones si se detectan en esos casos flujos significativos de productos a través de la información revelada por las tablas input-output.

Finalmente, ante la presencia de solapamiento de sectores en más de un cluster, se procede a realizar una definición amplia y otra restrictiva de cluster. En la amplia no se impide que un sector pueda formar parte simultáneamente de más de un cluster, mientras que en la restrictiva un sector únicamente puede formar parte de uno de los clusters. Para ello se sigue el criterio de intensidad en la correlación locacional observada para ese sector, de forma que se asigna al cluster con el que presente mayor correlación. En este mismo sentido, es posible definir el núcleo del cluster como aquel conjunto de sectores que presentan correlaciones más intensas.

#### 4.1.2. *Ventajas e inconvenientes*

##### Ventajas

- Una de las principales ventajas de la utilización de un método à la Porter radica en la sencillez de su aplicación, entendiéndose como tal la facilidad en los cálculos a realizar (que se limitan a CL, porcentajes de empleo y correlaciones del empleo regional entre sectores) y la interpretación de estos y de los otros instrumentos a utilizar (como por ejemplo la información contenida en las tablas input-output). De esta forma, una vez se dispone de la información estadística de base, el tratamiento cuantitativo a realizar es simple, estando su volumen relacionado con el grado de detalle sectorial con el que se trabaje.
- Otra de las ventajas que se le deben asignar a este tipo de aproximación es la baja necesidad de información estadística, dado que para su aplicación únicamente se precisa de información acerca del volumen de empleo en los sectores definidos previamente y para el conjunto de regiones o localidades que se vayan a tomar como unidades geográficas. Aunque la utilización de otro tipo de información acerca de las características de los sectores y de los flujos comerciales entre ellos es de utilidad para la mejora en el proceso de agrupación sectorial, no constituye un elemento indispensable para la implementación, pudiendo ser sustituido por conocimiento no necesariamente basado en información estadística sistemática para todos los sectores y regiones.
- Asimismo, la difusión y reconocimiento de esta metodología entre, como mínimo, parte de la comunidad académica y de los responsables de las políticas industriales y de desarrollo, junto a su aplicación en diversos proyectos de “*cluster mapping*” realizados para diversas economías, facilita, en principio, la comparabilidad de los resultados obtenidos y la aceptación de los mismos.

- Finalmente, como otra ventaja inmediata se puede señalar la facilidad de su actualización en sucesivos periodos a medida que la información estadística de base se encuentre disponible. Esto permite además incorporar una característica innata en los clusters, como es la del seguimiento de su dinámica. Asimismo, es posible reproducir el análisis para otros niveles de desagregación geográfica y/o sectorial en el caso en que se considere oportuno y todo ello con un coste informativo y computacional limitado.

### **Inconvenientes**

- Entre los inconvenientes más inmediatos sin duda destaca el de la escasa robustez de los resultados que se obtienen a través de la aplicación de esta aproximación. El origen del problema de robustez se encuentra en la utilización de información agregada, tanto en lo referido a la vertiente sectorial como a la territorial. Así, los clusters detectados y la cuantía de su intensidad variarán necesariamente dependiendo del detalle sectorial utilizado, sin que a priori se pueda determinar si número e intensidad dependen directa o inversamente del nivel de desagregación sectorial adoptados. Estos también serán distintos para diferentes niveles de desagregación utilizados. Es decir, que los clusters detectados pueden diferir si el análisis se realiza considerando, por ejemplo, comunidades autónomas o provincias. Dada la incertidumbre acerca de las fronteras sectoriales y geográficas de un cluster típico, las implicaciones de los problemas de robustez de los resultados que se obtengan a partir de la aplicación de esta metodología resultan obvias.
- La determinación de los puntos de corte, necesaria para la clasificación de los sectores, es la misma para todos los sectores considerados. Esta circunstancia sólo sería razonable si todos ellos compartiesen la misma distribución espacial, lo que difícilmente se dará en la realidad. Imponer un mismo umbral puede ocasionar tanto la inclusión como la exclusión errónea de sectores en el grupo de los aglomerados.

- En relación con esta circunstancia, para el cálculo de las correlaciones de localización cobra especial relevancia la definición de las unidades geográficas o regiones. Para establecer esta definición se debe tener en cuenta i) la disponibilidad de datos con el detalle sectorial requerido para el nivel de desagregación espacial deseado, y ii) el hecho de que para un nivel de desagregación geográfica elevado el empleo en muchos de los sectores comercializables tenderá a ser escaso e incluso inexistente en muchas de las unidades espaciales. Esta circunstancia provocará que el coeficiente de correlación presente valores elevados entre muchos de los sectores, sin que ello sea debido a la presencia de relaciones intersectoriales.

Si a ello añadimos que la unidad geográfica relevante puede diferir entre clusters, y que para algunos de ellos los límites geográficos pueden no corresponder con los límites administrativos, se suele aconsejar realizar el estudio para varios niveles de desagregación territorial.

- Otra crítica obvia que se puede realizar a esta aproximación es la de basarse en exceso en criterios *ad hoc*. Todas las fases del procedimiento están caracterizadas por apoyarse en elementos subjetivos y difícilmente generalizables. De esta forma, en cada estudio se deben fijar atendiendo a la información contenida en los propios datos que se están utilizando y en el conocimiento de la economía analizada por parte del investigador o equipo de investigadores. Esta circunstancia ha quedado reflejada en la descripción de los elementos fundamentales de la metodología descritos en el apartado anterior, y que de forma breve se pueden sintetizar en i) la determinación de los puntos de corte en los coeficientes CL y de Gini, y en el proceso de depuración de la clasificación de los sectores en comercializables, locales y vinculados con recursos naturales, ii) la agrupación de sectores y su posterior depuración que, aún basándose en coeficientes de correlación entre sectores, no sigue ninguna pauta objetivable sino que se debe apoyar en conocimiento subjetivo de los sectores y regiones consideradas.
- Finalmente, y dado que para la detección de relaciones intersectoriales necesaria para realizar la agrupación de los sectores que integran cada uno de los clusters se emplea

una aproximación indirecta, en cualquier estudio debe tenerse presente que la economía objeto de análisis debe ser lo suficientemente grande como para que todos los sectores y clusters puedan estar potencialmente presentes en cada una de sus regiones, y que éstas sean numerosas, diversas e interdependientes. Aunque como se señala en Porter (2003) el caso de la economía americana no suscita dudas en este sentido, no se debería trasladar esta argumentación automáticamente al análisis en otros ámbitos.<sup>2</sup>

#### ***4.1.3. Criterio estadístico de detección de aglomeración***

En la aplicación de un método *à la Porter*, la determinación del nivel de aglomeración sectorial que deberíamos observar para concluir acerca de la posible pertenencia de un sector a un cluster es una cuestión abierta que, como hemos indicado anteriormente, se deja en manos del investigador. Es decir, el establecimiento de un punto de corte se basa en un criterio subjetivo y propio de cada caso, lo que provoca que pueda ser considerado como arbitrario.

Pero este proceder contrasta con la necesidad de identificar clusters de forma consistente y asumible por parte de cualquier persona interesada en ellos. Resultaría entonces conveniente definir un nivel crítico objetivable para el grado de aglomeración a partir del cual considerar la pertenencia de sectores a clusters. Tal nivel crítico debería ser factible empíricamente y estar basado en consideraciones teóricas.

Asumiendo estas premisas, O'Donoghue y Gleave (2004) propusieron el coeficiente de localización estandarizado como instrumento en el que basar la clasificación entre sectores aglomerados y no aglomerados, y consecuentemente la determinación de los sectores comercializables potencialmente pertenecientes a clusters industriales. Con ello recogen la sugerencia de Duranton y Overman (2005) al insistir en la importancia de establecer la

---

<sup>2</sup> En palabras de Porter, “esta aproximación no es factible en la mayoría, sino en todos, los demás países”.

significación estadística como un pre-requisito para la determinación de la composición de los clusters.

El coeficiente de localización es utilizado porque, en opinión de los citados autores, la manifestación esencial de todo cluster es su aglomeración en el espacio, y ésta es recogida por dicho coeficiente. Pero también señalan que las aglomeraciones deben constituir localizaciones excepcionales, lo que implica que deben presentar valores para el CL significativos estadísticamente para algunos de los sectores considerados. En consecuencia, proponen identificar localizaciones con concentraciones de actividad excepcionales a través de valores del CL estadísticamente significativos al 5%, o lo que es lo mismo, detectar observaciones claramente diferenciadas del resto en base a un criterio objetivo.

Los pasos a seguir en la implementación del procedimiento son:

- i) calcular el CL para todos los sectores al nivel de desagregación sectorial y regional determinado,
- ii) comprobar que el CL se distribuye como una normal (aplicando para ello el test de Kolmogorov-Smirnoff). Si se detecta fuerte asimetría, se debe transformar logarítmicamente el CL,
- iii) estandarizar el CL y seleccionar aquellos casos cuyo valor exceda, en valor absoluto, 1.96 veces la desviación estándar de los valores del CL

El requerimiento de la normalidad de la distribución del CL (o de su transformación logarítmica) supone una evidente limitación del método, así como el hecho de que no tenga en cuenta posibles diferencias en la distribución del tamaño empresarial entre las regiones consideradas. Es decir, el CL estandarizado no es capaz de discriminar entre un valor significativo causado por la existencia de una gran empresa en una región determinada, de otro motivado por la presencia de un entramado de numerosas pequeñas y medianas

empresas. La solución propuesta en este caso consiste en el cálculo de un coeficiente de localización ajustado:

$$CL_{\text{ajustado}} = \frac{e_{PYME,i} / e_{\text{Total}}}{E_{PYME,i} / E_{\text{Total}}}$$

donde  $e$  y  $E$  hacen referencia al empleo regional y nacional respectivamente, el subíndice PYME a las pequeñas y medianas empresas, el  $i$  a cualquiera de los sectores considerados, y Total al conjunto de la economía.

El  $CL_{\text{ajustado}}$  permite comprobar hasta qué punto la aglomeración es debida a la presencia de PYMEs y, en consecuencia, a la posible existencia de un cluster.

Otras propuestas para la aplicación de un criterio objetivable en la selección de los sectores que potencialmente pueden formar clusters industriales, dadas sus pautas de aglomeración geográficas, han abandonado el uso del CL por sus conocidas limitaciones. Por ejemplo, se ha argumentado que el valor máximo que puede alcanzar el CL depende del tamaño de las regiones analizadas (Isard et al, 1998). En este sentido, Karlsson et al (2005) sugieren basar la selección de sectores aglomerados en cada una de las regiones en los residuos atípicos (outliers) de la regresión entre el empleo sectorial en cada región y la población de la región por una parte, y en la regresión entre el número de plantas productivas en cada sector y región y la población regional por otra. Mediante dichas regresiones se sugiere detectar observaciones con concentraciones de empleo y número de plantas significativamente por encima de los medios. La idea es que un cluster debe ocasionar de forma simultánea una sobrerrepresentación del empleo y del número de plantas productivas de los sectores que lo forman en la región en la que se localiza.

El proceso se puede sintetizar en las siguientes etapas:



i) una vez determinado el detalle sectorial y territorial, se recoge la información referida al empleo (Empl) y al número de empresas (Plantas),

ii) con esa información se efectúan las siguientes regresiones:

$$\ln \text{Empl}_{ir} = \beta_0 + \beta_1 \ln \text{Pobl}_{ir} + \varepsilon_{ir}$$

$$\ln \text{Plantas}_{ir} = \alpha_0 + \alpha_1 \ln \text{Pobl}_{ir} + \nu_{ir}$$

donde Pobl denota la población como medida del tamaño de la región y los subíndices  $i$  y  $r$  hacen referencia al sector y región respectivamente.  $\varepsilon$  y  $\nu$  son términos de perturbación con las propiedades habituales,

iii) se retienen los residuos positivos de ambas regresiones que excedan el 10%. Las observaciones asociadas pertenecen a los sectores susceptibles de formar parte de un cluster en las regiones correspondientes.

#### **4.1.4. Aplicaciones de la metodología**

Se han realizado “*cluster mappings*” para diversas economías aplicando los elementos esenciales del método de Porter. De hecho, algunas de las experiencias existentes han contado de una u otra forma con la participación del Institute for Strategy and Competitiveness. En cualquier caso, todas ellas han debido ajustar los procedimientos descritos anteriormente al caso particular de la economía analizada, a la vez que han incorporado elementos adicionales en algunos de los pasos a seguir.

Quizás una de las experiencias más difundidas ha sido la realizada para el Reino Unido por parte de Trends Business Research para el Departamento de Comercio e Industria (DTI) de ese país, y que bajo el título “Business Clusters in the UK — A First Assessment” abordó la detección, identificación y cuantificación de clusters industriales en el conjunto del territorio mediante la adaptación del método propuesto por Porter.

Las etapas seguidas en ese trabajo, similares a las descritas anteriormente, se pueden sintetizar en:

- i) Identificación de los sectores sobrerrepresentados en cada región, en términos del empleo y que son “empleadores” significativos ( $CL > 1.25$  y porcentaje del empleo del sector en el empleo total de la región  $> 0.2$ ). Con esa información e incorporando “juicio e interpretación”, se clasifican los sectores para formar la base de los clusters.
- ii) Revisión de la estructura sectorial de cada región
  - a. examinar sectores con  $CL > 1.25$  pero con porcentaje del empleo regional  $< 0.2$ . Si se considera conveniente se incluyen en un cluster
  - b. examinar sectores con porcentaje del empleo regional  $> 0.2$  pero con  $CL < 1.25$ . Se sitúan en un cluster si no son sectores que sirven a demanda local. En algunos casos se valora la posibilidad de que constituyan el núcleo de un cluster
- iii) utilizando datos de empresas para detectar la actividad característica de grandes empresas en un conjunto de sectores, determinar la naturaleza específica de su producción. De esta forma se obtiene evidencia adicional de vínculos y relaciones potenciales. Para ello se toman los sectores con elevados valores del CL y que pueden llegar a formar parte de un cluster en una región determinada
- iv) agrupar los sectores aglomerados en clusters. Para ello se utiliza el conocimiento del concepto de cluster e información de cada uno de los sectores. En concreto, el conocimiento de:
  - a. los vínculos input-output y la evidencia disponible sobre la co-localización industrial,
  - b. las cadenas de valor,

- c. que sectores de una misma agrupación sectorial tienden a localizarse en una misma región,
  - d. la información procedente de otras fuentes de datos, que proporcionan información de vínculos entre sectores.
- v) analizar las unidades geográficas inferiores, dado que:
- a. pueden existir clusters sub-regionales significativos que no lleguen a ser visualizados cuando se realiza el análisis a nivel regional
  - b. la localización precisa de un cluster permite determinar si existe un centro geográfico en la región
  - c. un elevado CL puede deberse a la presencia de una única gran planta productiva, lo que en esencia no constituiría un cluster

Permite detectar concentraciones locales significativas para sectores específicos, y una vez identificados los clusters a nivel local compararlos con los obtenidos en el análisis regional.

- vi) los clusters detectados se clasifican atendiendo a:
- a. el estado de su desarrollo
  - b. su intensidad
  - c. la dinámica de su empleo
  - d. su importancia

- vii) detectar otros clusters, que resultan menos obvios. Para ello se hace uso del conocimiento de cada una de las regiones, junto a la utilización de la información referida a las empresas, que permite la identificación de clusters no basados en sectores. La idea es que el análisis formal puede no permitir la identificación de algunos clusters debido a que el detalle sectorial empleado no sea lo suficientemente detallado y/o que los clusters se encuentren en estado embrionario o sean de pequeña dimensión
- viii) se excluyen aquellos clusters identificados con la aplicación del método pero que corresponden a actividades peculiares, como las de defensa, educación o salud
- ix) como último paso se realiza la validación de los clusters detectados con representantes regionales y agentes e instituciones de índole diversa.

En cualquier caso, se insiste en que el análisis realizado debe ser interpretado como un proceso inicial de identificación de clusters que pueden existir en cada una de las regiones de la economía. En ese sentido, representa un punto de partida para un proceso de interacción con las estrategias de desarrollo. A modo de ejemplo, la detección de co-localización permite aventurar la posibilidad de existencia de un cluster potencial, de forma que determinadas actuaciones podrían mejorar las opciones de evolución hacia un cluster efectivo.

#### **4.1.5. Potencialidad de resultados obtenidos**

- La identificación de clusters industriales en el territorio siguiendo un procedimiento a *la Porter* permite la detección de agrupaciones de sectores co-localizados y a los que se atribuye un nivel mínimo de interacción. En consecuencia, mediante el empleo de información estadística correspondiente a un detalle sectorial y territorial elevado

sería posible la obtención de un mapa preciso de la localización de la actividad económica en la geografía española. Ese mapa permitiría ubicar en el territorio actividades económicas supuestamente interrelacionadas.

La riqueza del mapa dependería en gran medida del detalle sectorial y territorial que se pudiese llegar a emplear. Respecto al detalle sectorial, sería aconsejable el uso de la desagregación a 2, 3 y 4 dígitos de la clasificación CNAE (o equivalente), mientras que para el territorial el análisis a nivel de comunidades autónomas, provincias y municipios podría proporcionar evidencia más rica que en el caso de realizar el análisis únicamente tomando como referencia a las primeras.

- Además de la obtención de un mapa global, la combinación del método con información acerca de las características de determinadas actividades y sectores productivos debería permitir la identificación de clusters de especial interés, por su dinamismo y/o por resultar claves para la economía nacional o para alguno de los territorios.
- Haría posible la detección de regiones y/o localidades con implantación de clusters especialmente dinámicos o en clara recesión. En ambos casos la detección permitiría discernir el tipo de actividades o sectores implicados y el ámbito territorial afectado. A la vez, y en caso de que se considerase oportuna su realización, permitiría definir los límites sectoriales y territoriales de un análisis específico más detallado de cada uno de esos clusters.

A su vez, el análisis revelaría tanto aquellos sectores menos proclives a formar parte de clusters industriales como aquellas localizaciones menos tendentes a acogerlos. También en ese caso, el análisis permitiría delimitar sectores y territorios susceptibles de un estudio posterior más profundo, que tuviese por objeto, por ejemplo, determinar la potencialidad de los incentivos a la consolidación de clusters embrionarios o débiles en determinadas localidades.

## 4.2. Método basado en la comparación de distribuciones

En contraposición a la estrategia sugerida por Porter y colaboradores, recientes aportaciones han propuesto una metodología para identificar empíricamente clusters industriales en el territorio, basada en la comparación de la distribución espacial de cualquier sector analizado con otra que supuestamente caracterizaría una situación en la que operasen las fuerzas que causan la aglomeración espacial de la actividad.

El fundamento de esta estrategia lo encontramos en los trabajos de Ellison y Glaeser (1997 y 1999), quienes desarrollaron un estadístico para contrastar si los niveles de concentración observados son mayores a los que correspondería en una situación en la que la localización de las plantas productivas se determinase aleatoriamente. El estadístico se apoya en un modelo en el que tanto la existencia de externalidades de localización específicas para cada sector como ventajas locacionales naturales y factores aleatorios puros contribuyen a la concentración geográfica de la actividad productiva. El resultado es la obtención de un índice de concentración geográfica basado en un modelo, y que presenta un par de propiedades deseables: i) toma valor nulo en caso de que la distribución existente esté tan concentrada como lo estaría si fuese determinada por factores aleatorios (como si se tirasen dardos en un mapa para elegir las localizaciones de las empresas), lo que implica que la distribución de referencia no es una en la que el empleo se encuentra uniformemente distribuido en el espacio, y ii) el índice permite la comparación entre sectores, países e instantes del tiempo.

No obstante, el índice de concentración geográfica de Ellison y Glaeser no resulta útil para la realización de un “*cluster mapping*” dado que únicamente permite valorar el grado de concentración geográfica de un sector, pero no determinar en qué regiones o localidades se producen las concentraciones significativas de actividad.

Thomas Brenner (2003, 2005) utiliza la misma idea referida a la comparación de la distribución de la actividad de un sector en el conjunto de localidades o regiones con la derivada de un modelo teórico en ausencia y en presencia de economías de aglomeración,

para dar un paso adelante y sugerir un método de *mapping* basado en la comparación de distribuciones. Los elementos fundamentales del mismo serán expuestos a continuación, aunque antes, y debido a las similitudes entre la aproximación sugerida por este autor y las medidas de concentración geográfica a la Ellison y Glaeser, vamos a describir brevemente estas últimas y a mostrar como únicamente podrían ser utilizadas como un instrumento de apoyo en los análisis de identificación y cuantificación de clusters industriales.

### **Medidas de concentración geográfica basadas en distribuciones**

Ellison y Glaeser (1997) formulan un modelo simple en el que “*la concentración geográfica de un sector es el resultado de una secuencia de decisiones de localización maximizadoras del beneficio realizadas por las plantas productivas*”. Los mecanismos que ocasionan la aglomeración espacial de las plantas productivas tienen que ver con las ventajas naturales de algunas localizaciones y con la presencia de externalidades específicas para el sector, mientras que cuestiones idiosincrásicas de cada planta son las que impiden una concentración espacial absoluta.

Supongamos que tenemos K plantas productivas en el sector i, que deben seleccionar su localización secuencialmente del conjunto de las R regiones con el objetivo de maximizar sus beneficios. El beneficio percibido por una planta k cuando se localiza en la región r está determinado por:

$$\ln \pi_{ikr} = \ln \pi_{ir} + g_r(v_1, v_2, \dots, v_{k-1}) + \eta_{ir} + \varepsilon_{ikr}$$

donde  $\pi_{ir}$  es el beneficio esperado asociado a la localización específica en la región r, y la función  $g_r$  refleja el efecto sobre los beneficios de la región r causado por el hecho de que las plantas 1, 2, ..., k-1 han seleccionado previamente localizarse en  $v_1, v_2, \dots, v_{k-1}$ .  $\eta_{ir}$  es un componente aleatorio inobservable asociado a la localización en la región r y  $\varepsilon_{ikr}$  es una perturbación inobservable específica de la planta k.

Como se ha indicado antes, en el modelo de Ellison y Glaeser, las plantas de un sector pueden agruparse espacialmente debido a que: i) prefieran localizarse en regiones con menores costes de producción observables, ii) prefieran localizarse en regiones con

menores costes de producción inobservables, y iii) a la existencia de externalidades localizadas geográficamente.

Asumiendo determinadas características para la distribución de los efectos de las ventajas naturales sobre los beneficios y sobre el mecanismo a través del que las externalidades entre plantas afectan a sus beneficios, el índice de concentración geográfica para un sector dado:

$$\gamma \equiv \frac{G - \left(1 - \sum_r x_r^2\right)H}{\left(1 - \sum_r x_r^2\right)(1-H)} = \frac{G / \left(1 - \sum_r x_r^2\right) - H}{(1-H)}$$

captura consistentemente la magnitud de las fuerzas de aglomeración del modelo. G es una medida de concentración geográfica del sector definida como  $G \equiv \sum_r (s_r - x_r)^2$  donde  $s_r$  es el porcentaje del empleo del sector en la región r y  $x_r$  el porcentaje del empleo total en la región r.  $H = \sum_{j=1}^N z_j^2$  es el índice de Herfindahl de la distribución del tamaño relativo de las plantas (z) en el sector.

El índice tomará valor nulo en ausencia de fuerzas de aglomeración, mientras que no se encuentra acotado superiormente. Ellison y Glaeser (1997) sugieren considerar a los sectores como altamente concentrados cuando  $\gamma > 0.05$ , y como escasamente concentrados a aquéllos para los que  $\gamma < 0.02$ .

Apoyándose en el mismo modelo de localización, los citados autores sugirieron un índice de co-aglomeración. Este índice pretende cuantificar en qué medida un conjunto de sectores pertenecientes a un grupo tienden a localizarse conjuntamente en el espacio. Se define como:

$$\gamma^c \equiv \frac{\left[ G / \left(1 - \sum_r x_r^2\right) \right] - H - \sum_i \hat{\gamma}_i \omega_i^2 (1 - H_i)}{1 - \sum_i \omega_i^2}$$



donde  $G$  es la medida de concentración geográfica para el conjunto de sectores del grupo,  $H = \sum_i \omega_i^2 H_i$  es el índice de Herfindahl para el grupo, con  $\omega_i$  el porcentaje del empleo del sector  $i$ -ésimo. Finalmente,  $\hat{\gamma}_i$  es el valor del índice de concentración para el sector  $i$ .

Un valor nulo para el índice de co-aglomeración indica que no hay más aglomeración de las plantas en el grupo de sectores que el atribuible a la tendencia de éstas a localizarse cerca de otras plantas del mismo sector y donde exista una elevada concentración del empleo total.

De esta sintética exposición de los índices propuestos por Ellison y Glaeser, se desprende su utilidad para detectar el grado de concentración territorial de un conjunto de sectores y para valorar patrones de co-localización de algunos sectores. Pero al ser medidas globales para el conjunto de la economía, estos índices no permiten realizar ningún juicio acerca de los territorios en los que se encuentran las concentraciones. Por otra parte, requieren de la definición a priori de los sectores y territorios objeto de análisis, junto a la utilización de información referida al empleo en cada sector y región, y al empleo en las plantas de cada sector para el conjunto de la economía. Por tanto, y a pesar de su extensa difusión en los últimos años (véase por ejemplo Maurel y Sédillot, 1999), estos índices no permiten realizar un “*cluster mapping*”, aunque sí sugeriríamos su empleo previo junto a otras estrategias para determinar en base a criterios robustos qué sectores se encuentran concentrados geográficamente y son, por tanto, susceptibles de formar parte de clusters industriales.

#### **4.2.1. Elementos esenciales del método.**

En este apartado describimos los puntos básicos del método propuesto por T Brenner para realizar un “*cluster mapping*”. Como hemos indicado anteriormente, el método sugerido por este autor se basa en la comparación de la distribución de la actividad productiva en el territorio para cualquier sector con la distribución que teóricamente correspondería por una parte a una situación sin fuerzas que causasen aglomeración y, por otra, a una en la que estuviesen presentes clusters industriales como resultado de la presencia de las ciudades

fuerzas. El elemento diferencial fundamental que permite la identificación de los clusters en este caso es la posibilidad de estimar un umbral en el volumen de actividad en cada uno de los territorios considerados. De esa forma, tendremos indicios de que un sector puede formar parte de un cluster en una región cuando en ella la presencia de actividad del sector supere dicho umbral. De hecho, la aplicación de un contraste estadístico formal permitirá comprobar la significación de la concentración del sector en la región.

La aplicación del método permite distinguir dos etapas:

- i) en primer lugar se enfrenta la distribución empírica asociada a los datos de sectores y regiones con la que se deduce de las predicciones teóricas. El resultado nos permite determinar aquellos sectores que muestran indicios de localización en aglomeraciones industriales,
- ii) a continuación, y utilizando la información de la estimación sobre el umbral de actividad proporcionado por el propio método, se identifican las regiones en las que se encuentran localizadas las aglomeraciones en cada sector.

### **Supuestos iniciales**

- Para la derivación de las distribuciones teóricas de referencia, en el caso de presencia y ausencia de clusters, se modeliza estocásticamente la localización de las plantas productivas y la capacidad de atracción de las regiones. En ambos casos se reconoce la imposibilidad de considerar todos los elementos relevantes en el modelo, dada la ausencia de información estadística sobre los mismos (por ejemplo los vínculos con cada localidad de los empresarios o la influencia de ciertos intangibles como la cultura empresarial sobre el atractivo de las regiones). Ante esta circunstancia se opta por una solución radical al proponer utilizar únicamente el tamaño de las regiones

analizadas, mientras que todos los demás determinantes se modelizan de forma estocástica.

- Para determinar la forma de la distribución de la actividad productiva de un sector entre las distintas regiones se tiene en cuenta la distribución regional de algunos factores que se supone debe estar determinando la localización regional de las plantas productivas. De esta forma, la propuesta de Brenner se aleja de la especificación puramente aleatoria característica de la aproximación *dartboard* de Ellison y Glaeser al incorporar información sobre la distribución de los factores que pueden estar afectando tanto la decisión individual de localización de cada planta como el atractivo de cada región.<sup>3</sup>
- Se asume que los sectores con clusters locales se caracterizan por procesos locales de auto-sostenimiento, lo que provoca que las regiones en los que se emplazan presenten niveles de actividad económica mucho mayores y, en consecuencia, que se localicen en ellas muchas más plantas productivas.

#### **Determinación de la distribución regional de la actividad productiva (sin y con clusters)**

El primer paso en la aplicación del método consiste en la determinación de la distribución de las plantas productivas de un sector entre las distintas regiones, o lo que es lo mismo, establecer la función que caracteriza a la probabilidad de que  $f$  plantas de un sector  $i$  se localicen en una región  $r$  de tamaño  $s$ ,  $P(f|s)$ .

Para ello se selecciona un conjunto de factores que supuestamente influyen en las decisiones de localización de las empresas y en el atractivo locacional de las regiones, y se analizan las características de su distribución empírica. Para la aplicación que realiza a la economía alemana, Brenner selecciona el capital humano, aproximado a través del número de estudiantes por habitante, y las empresas de servicios (legales, marketing, consultoría, etc). De la inspección de la distribución de esas variables en los distritos administrativos alemanes, Brenner sugiere la combinación de una distribución Boltzmann y una

---

<sup>3</sup> En ese caso la distribución que caracteriza el número de plantas localizadas en cada región es una binomial.

exponencial decreciente, a las que se añadiría una Binomial (Bn[.]) para incorporar también las predicciones de un modelo tipo *dartboard*.<sup>4</sup> De esa forma, la distribución de probabilidad que caracterizaría el caso de ausencia de clusters es:

$$P_n(f | s) = [1 - \xi_3(s) - \xi_5(s)] \cdot [1 - \xi_1(s)] \cdot \xi_1(s)^f + \xi_3(s) \cdot \frac{[1 - \xi_2(s)]^2}{\xi_2(s)} \cdot f \cdot \xi_2(s)^f + \xi_5(s) \cdot Bn[\bar{f}, f, \xi_4(s)]$$

donde  $\xi_1(s)$ ,  $\xi_2(s)$ ,  $\xi_3(s)$ ,  $\xi_4(s)$  y  $\xi_5(s)$  son parámetros y  $\bar{f}$  es el número total de plantas en el sector analizado. El subíndice n hace referencia a la distribución neutral, es decir a la que caracteriza una situación de ausencia de clusters.

El paso siguiente es determinar la distribución asociada al caso en el que existan clusters industriales. De nuevo, se asume la imposibilidad de modelizar adecuadamente todos los mecanismos que pueden dar lugar a las aglomeraciones o clusters, por lo que el método emplea una aproximación simple, basada en la idea de que las regiones que tengan un atractivo superior a un determinado umbral contendrán un cluster. Así, la existencia de un cluster en una región se traducirá en la presencia de un número significativamente mayor de plantas productivas. Por tanto, la distribución de plantas en presencia de clusters será similar a la asociada a  $P_n(f|s)$  pero conteniendo un número de regiones con un número de empresas proporcionalmente muy elevado. De tal forma que la distribución asociada a  $P_c(f|s)$ , es decir la probabilidad en el caso de presencia de clusters, presentará dos máximos, correspondiendo el segundo a aquellas regiones en las que se produce una sobreconcentración de plantas.

La forma en la que Brenner propone modelizar ese segundo máximo en la distribución es muy simple: asumir que la forma es similar a la del resto de la distribución y que se inicia a partir de un determinado número de plantas. Este segundo supuesto lleva a descartar la exponencial decreciente, mientras que el mejor ajuste de la Boltzmann frente a la Binomial

<sup>4</sup> No obstante, y dado que las distribuciones Binomial y Boltzmann son bastante parecidas, en algunas de sus aplicaciones Brenner sugiere combinar únicamente la exponencial decreciente y la Boltzmann.

le conduce a sugerir el uso de la primera distribución para capturar el efecto de los clusters sobre la distribución de las empresas en el territorio. En concreto, la distribución en presencia de clusters se representa mediante:

$$\begin{aligned}
 P_c(f | s) = & [1 - \xi_3(s) - \xi_5(s) - \xi_8(s)] \cdot [1 - \xi_1(s)] \cdot \xi_1(s)^f + \xi_3(s) \cdot \frac{[1 - \xi_2(s)]^2}{\xi_2(s)} \cdot f \cdot \xi_2(s)^f \\
 & + \xi_5(s) \cdot Bn[\bar{f}, f, \xi_4(s)] \\
 & + \xi_8(s) \cdot \frac{[1 - \xi_7(s)]^2}{\xi_7(s)} \cdot f \cdot \xi_7(s)^{f - \xi_6(s)} \quad \text{si } f \geq \xi_6(s) \\
 & 0 \quad \quad \quad \text{si } f < \xi_6(s)
 \end{aligned}$$

donde el último término del lado derecho de la igualdad es una distribución de Boltzmann que aparece únicamente en el caso en que el número de plantas del sector supere un umbral dado por  $\xi_6(s)$ . Por su parte,  $\xi_8(s)$  es un parámetro que recoge el porcentaje de regiones que son descritas por esa parte adicional de la distribución.

**Consideración del tamaño de las regiones**

Resulta razonable suponer que la distribución de las plantas productivas dependerá del tamaño de las regiones (s). Para recoger ese efecto se asume que el número medio esperado de plantas localizadas en una región aumenta linealmente con el tamaño de la región, tanto para el caso de la distribución de  $P_n(f|s)$  como de  $P_c(f|s)$ . En consecuencia, el número medio de plantas debe aumentar linealmente con s a lo largo de toda la distribución, mientras que la contribución relativa de cada una de sus componentes debe permanecer estable, es decir, independiente de s. Para ello se introducen los siguientes supuestos sobre los parámetros:

$$\xi_1(s) = \frac{\xi_1 \cdot s}{1 + \xi_1 \cdot s}$$

$$\xi_2(s) = \frac{\xi_2 \cdot s}{1 + \xi_2 \cdot s}$$

$$\xi_3(s) = \xi_3$$

$$\xi_4(s) = \xi_4 \cdot s$$

$$\xi_5(s) = \xi_5$$

$$\xi_6(s) = \xi_6 \cdot s$$

$$\xi_7(s) = \frac{\xi_7 \cdot s}{2 + \xi_7 \cdot s}$$

$$\xi_8(s) = \xi_8$$

Dado que  $\xi_3$ ,  $\xi_5$  y  $\xi_8$  determinan el peso de cada una de las tres distribuciones, se suponen específicos de cada sector pero independientes del tamaño de la región. No sucede lo mismo con el resto de parámetros.

Por último, para la definición del tamaño de cada región, Brenner sugiere utilizar el porcentaje del empleo en la región respecto al del total del país.

### **Restricciones y estimación de parámetros**

El último término de  $P_c(f|s)$  es el fundamental para la detección e identificación de los clusters. Como se ha indicado, es una distribución tipo Boltzmann desplazada que aparecerá sólo en el caso en que el número de plantas exceda el umbral dado por  $\xi_6 \cdot s$ . La contribución de ese término a la distribución total dependerá además de  $\xi_8$ . En la implementación del procedimiento, se restringe el rango de valores que pueden tomar tanto  $\xi_6$  como  $\xi_8$ . Las razones esgrimidas por Brenner para la imposición de las restricciones sobre esos parámetros son las siguientes:

- dado que para cualquier sector que presente aglomeración los clusters deben constituir una excepción, sólo unas pocas regiones deben presentarlos. O dicho de otra forma, no se concluirá que el sector se encuentra aglomerado en clusters

industriales si éstos se encontrasen en todas las regiones. En consecuencia, el porcentaje de regiones que potencialmente pueden contener un cluster de un sector debe ser pequeño. En Brenner (2005) ese umbral se establece en un 5%, mientras que en Brenner (2003) se fija en un 10%, o lo que es lo mismo, se impone que  $\xi_8 \leq 0.05$  en el primer caso y  $\xi_8 \leq 0.1$  en el segundo,

- sólo las regiones con un número de plantas que claramente exceda un umbral determinado serán candidatas a presentar un cluster industrial en el sector considerado. El umbral viene dado por el parámetro  $\xi_6$ , y en otras aproximaciones se ha fijado como el número medio de plantas en el sector, de forma que todas las regiones que presentasen un número de plantas superior a la media nacional eran consideradas como regiones con un cluster. En otras ocasiones se ha impuesto un criterio más estricto, considerando por ejemplo un valor umbral de 3. En su lugar, Brenner propone un criterio algo más complejo, basado en la idea de que en presencia de un cluster, el último término de  $P_c(f|s)$  debe estar en cierta medida *alejado* del resto de la distribución. Es decir, que el grado de solapamiento de las dos partes de la distribución debe ser pequeño, siendo éste debido a las regiones que contienen un mayor número de plantas causados por la primera parte. La restricción introducida garantiza que  $\xi_8$  excede al menos en 5/6 el porcentaje de regiones con un número de plantas superior al umbral dado por  $\xi_6$ 's:

$$\xi_8 = \frac{5}{6} \frac{n_{cl,n}}{N_r}$$

donde  $n_{cl,n}$  es el número de regiones que contienen un número de plantas debido a la primera parte de la distribución mayor a  $\xi_6$ 's, y  $N_r$  el número total de regiones.

Adicionalmente, se debe tener en cuenta que todos los parámetros deben ser positivos y que  $\xi_3 + \xi_5 + \xi_8 \leq 1$ .

La estimación de los parámetros desconocidos de  $P_c(f|s)$  para cada sector se obtiene a través del método de máxima verosimilitud. Para ello se maximiza el logaritmo de la función de verosimilitud de  $P_c(f|s)$  sujeta a las restricciones descritas anteriormente. La función de verosimilitud se define como:

$$L_c = -\ln \left[ \prod_{r=0}^{N_r} (P_c f(r) | s(r)) \right]$$

donde  $f(r)$  denota el número de plantas localizadas en la región  $r$  y  $s(r)$  es el tamaño de esa región. La solución del problema de maximización con restricciones se puede conseguir a través de la aplicación de un método numérico, por ejemplo de los implementados en el modulo MAXLIK de Gauss u otro similar de algún software disponible.

Siguiendo el mismo procedimiento se pueden estimar los parámetros de la distribución en ausencia de clusters, de  $P_n(f|s)$ . En este caso la función de verosimilitud se define como:

$$L_n = -\ln \left[ \prod_{r=0}^{N_r} (P_n f(r) | s(r)) \right]$$

De esta forma, la presencia de aglomeraciones en clusters en un sector provocará que se alcance un mejor ajuste a través de  $P_c(f|s)$  que de  $P_n(f|s)$ , mientras que si no hay una diferencia sustancial entre ambas, es decir que no se mejora el ajuste al añadir la componente diferencial de  $P_c(f|s)$ , se debería concluir en contra de la presencia de clusters en el sector considerado. Recuérdese que esta circunstancia se producirá cuando  $f < \xi_6 \cdot s$ .

### **Identificación de los clusters industriales**

Una vez estimados los parámetros de la distribución de  $P_c(f|s)$  es posible realizar la identificación de los clusters en los sectores en los que se haya detectado el fenómeno de la aglomeración. Las regiones que contengan un cluster serán descritas por una distribución tipo Boltzmann con un valor mínimo del parámetro  $\xi_6(s)$ . Es decir que  $\xi_6(s)$  es el umbral de plantas productivas que dan lugar a un cluster. Dado que  $\xi_6(s) = \xi_6 \cdot s$  y que  $s$  varía entre



regiones, el umbral de plantas será distinto para cada región siempre y cuando éstas difieran en su tamaño.<sup>5</sup>

Por tanto, se considera que existe un cluster en todas aquellas regiones en las que el número de empresas en el sector analizado supera el valor umbral de  $\xi_6(s)$ .

#### 4.2.2. *Ventajas e inconvenientes*

##### Ventajas

- Entre las principales ventajas de esta metodología de detección e identificación de clusters destaca el que se base en un criterio estadístico, fijado a priori, objetivo y reproducible. Esta circunstancia permite realizar el “*cluster mapping*” con un mínimo de imposición de criterios propios de los investigadores que lo ejecutan, a la vez que facilita al analista y a cualquier usuario del mismo una interpretación y juicio inmediato del método empleado y de los resultados obtenidos.
- El método resulta atractivo intuitivamente al sugerir la comparación del ajuste alcanzado por un modelo teórico en ausencia de aglomeración con otro en el que están presentes clusters industriales. De la selección del modelo que mejor capture la estructura subyacente en los datos disponibles (la distribución espacial de la actividad del sector analizado) se infiere la presencia o ausencia de clusters y, a su vez, se obtienen los elementos necesarios para su identificación.
- El criterio para la detección e identificación de los clusters es específico de cada sector y región, lo que introduce un grado elevado de flexibilidad al considerar las peculiares características de la distribución de la actividad en los sectores y en el territorio.

---

<sup>5</sup> En el caso de fijar el mismo valor de  $s$  para todas las regiones (Brenner, 2003, lo fija como  $s=1/N_r$ ), el umbral no varía entre regiones.

- Considera simultáneamente información referida al número de plantas del sector analizado instaladas en cada región e información del volumen de empleados del sector en la región, para tener en cuenta la dimensión de ésta. Al determinar el umbral que posibilita la detección e identificación de los clusters en función del número de empresas se evita el riesgo de considerar un cluster formado por un pequeño número de grandes empresas.

**Inconvenientes**

- La detección e identificación de clusters únicamente tiene en cuenta la presencia de un número anormalmente elevado de empresas del sector analizado en un conjunto limitado de regiones. En consecuencia no se incorpora información relativa a uno de los elementos básicos de la definición de cluster industrial, como es la existencia de interacción entre agentes productivos ocasionada por externalidades de tipo tecnológico o pecuniario. En el caso de que la interacción se refiera a la producida entre plantas productivas de un mismo sector se asume que la co-localización es precisamente su consecuencia más inmediata, aunque en ningún caso se realiza un contraste directo de este supuesto, lo que no excluye que las aglomeraciones sean causadas por otras circunstancias (por ejemplo asociadas a condicionantes históricos o a la dependencia de recursos naturales). En el caso de las posibles interacciones entre plantas de distintos sectores, éstas se ignoran completamente en el proceso de detección e identificación. De esta forma, si en una o varias regiones se detecta un cluster en diversos sectores no es posible, con la información proporcionada por el método, discernir si todos o un subconjunto de ellos constituyen un cluster multisectorial o si, por el contrario, no existe interacción significativa entre ellos.
- Las distribuciones teóricas asociadas a las situaciones en ausencia y en presencia de clusters se derivan a partir de las distribuciones geográficas que, al menos supuestamente, caracterizan a los factores determinantes de la localización de las

empresas y del propio atractivo locacional de cada una de las regiones. Pero no se proporciona una guía explícita acerca de cuáles pueden ser esos factores. Aún recurriendo a la literatura, el grado de discrecionalidad por parte del investigador en la determinación de los mismos será elevado. A ello habrá que sumar que la elección estará marcada por la disponibilidad de información estadística sobre esos factores al nivel de desagregación territorial fijado en el análisis. En todo caso, no existen a priori garantías de que las distribuciones sugeridas y empleadas por Brenner para el caso alemán sean las más adecuadas para capturar la distribución de la actividad productiva en la economía española. Cabría en consecuencia realizar un análisis específico en ese sentido.

- El procedimiento impone algunos criterios de forma *ad hoc*. Por ejemplo, i) se asume que el número de empresas incrementa linealmente con el tamaño de la región a lo largo de toda la distribución, ii) el tamaño de las regiones se aproxima a través del empleo, lo que puede estar penalizando a regiones pequeñas pero altamente productivas, iii) se fija el porcentaje máximo de regiones que pueden presentar clusters en el 5% (Brenner, 2005) ó el 10% (Brenner, 2003), o iv) se determina la influencia de la distribución de Boltzmann que caracteriza los clusters en base a una compleja función con elementos aparentemente discrecionales.
- En el caso de sectores con poca implantación en el conjunto de la economía, es decir con un número escaso de plantas productivas, es posible detectar erróneamente la presencia de clusters debido a que un número reducido de plantas en unas pocas regiones (especialmente si estas son pequeñas) pueden ser interpretadas por el procedimiento como aglomeraciones en comparación con la ausencia de actividad en las restantes regiones. Este problema se planteará con mayor probabilidad en el caso de trabajar con un detalle sectorial y territorial elevado.
- La detección e identificación de los clusters precisa de la estimación de los parámetros que caracterizan la distribución en presencia de estos. La obtención de esa estimación no es inmediata, requiriendo la aplicación de algoritmos numéricos de

optimización con restricciones. En consecuencia la implementación del método es relativamente compleja y costosa en términos computacionales, precisando de software que permita la aplicación de los citados algoritmos o, directamente, la estimación por el método de máxima verosimilitud. Adicionalmente, se debe tener presente que la combinación de diversas distribuciones puede dificultar la obtención de las estimaciones y que incluso se podrían presentar problemas de convergencia en el algoritmo o de pobre identificación de los parámetros. En este sentido, en los trabajos de Brenner no se realiza una discusión detallada acerca del proceso de estimación (tipo de algoritmo recomendado, valores iniciales, propiedades de la estimación, etc).

- Las características del proceso de estimación de los parámetros y del contraste de la hipótesis de aglomeración (basados en la máxima verosimilitud) exigen la utilización de un número de regiones relativamente elevado. Esta circunstancia es señalada en los trabajos de Brenner como un requerimiento importante para garantizar la fiabilidad de los resultados obtenidos. La repercusión de este hecho para cualquier aplicación a realizar en la economía española es inmediata: no se debería implementar para el caso de las Comunidades Autónomas y sólo con prudencia para el de las provincias. Parece pues más adecuado utilizar unidades territoriales inferiores, como los municipios, que garanticen la disposición de, como mínimo, algunos cientos de observaciones. Además, debe tenerse presente que, por diversos motivos ya discutidos en otras secciones, no existen garantías de que los resultados sean robustos al nivel de desagregación territorial seleccionado.
- Para la implementación del proceso se precisa disponer de información estadística referida tanto al empleo como al número de plantas productivas de cada uno de los sectores analizados y en cada una de las unidades territoriales consideradas. En consecuencia, podemos considerar que el método impone un requerimiento informativo medio.

- La escasez de aplicaciones existentes, al menos por el momento, dificulta la comparación con otras experiencias y con los resultados obtenidos en las mismas. A ello cabe añadir el que la difusión de la metodología en el contexto académico sea, al menos por el momento y hasta donde nosotros conocemos, limitada, sin que haya sido sometida a un riguroso proceso de validación y discusión.

#### **4.2.3. Criterio estadístico de detección de aglomeración**

En la estrategia propuesta por T Brenner, la detección de aglomeraciones o clusters en un sector se apoya en el resultado de un contraste estadístico basado en la comparación de la verosimilitud alcanzada por el modelo que incorpora la presencia de clusters frente a aquel que no lo hace. Así, se concluye a favor de la existencia de clusters cuando el ajuste a los datos alcanzado por el modelo basado en  $P_c(f|s)$  supera significativamente el que se obtiene a través de  $P_n(f|s)$ , o lo que es lo mismo cuando  $\hat{L}_c > \hat{L}_n$ . En caso contrario,  $\hat{L}_c = \hat{L}_n$  por lo que la distribución tipo Boltzmann asociada al término propio de los clusters no se ve refrendada por los datos, indicando que la presencia de clusters es poco verosímil.

Tras la estimación de los parámetros de los modelos con y sin clusters para un sector dado, la comprobación de si la distribución con clusters describe de forma más fiel la realidad se realiza a través de un test de razón de verosimilitud:

$$\lambda = 2 \left[ \ln \hat{L}_c - \ln \hat{L}_n \right]$$

Bajo la hipótesis nula de ausencia de clusters, es decir de igualdad de las distribuciones de  $P_c(f|s)$  y  $P_n(f|s)$ ,  $\lambda$  se distribuye como  $\chi^2$  con 3 grados de libertad —la diferencia entre el número de parámetros de  $P_c(f|s)$  y de  $P_n(f|s)$ . El rechazo de la hipótesis nula lleva a asumir la existencia de clusters industriales para el sector analizado.

Adicionalmente, se sugiere comprobar si las distribuciones teóricas en ausencia y en presencia de clusters describen los datos empíricos adecuadamente. Para ello se emplea un contraste de Kolmogorov-Smirnov, que compara la función de distribución acumulada de

las distribuciones teórica y empírica. No habrá distancia significativa entre ambas si las distribuciones teórica y empírica son idénticas, mientras que si no lo son se observarán entre ambas desviaciones significativas. Se trata pues de comparar tanto la función de distribución acumulada teórica correspondiente a  $P_c(f|s)$  y a  $P_n(f|s)$  con la distribución empírica asociada a los datos del número de empresas del sector analizado en cada una de las regiones.

#### 4.2.4. Aplicaciones de la metodología

Hasta donde conocemos, el método únicamente ha sido aplicado por Brenner para el caso de la economía alemana. En Brenner (2005) se aplica al caso de los 23 sectores manufactureros de la clasificación estándar a 2 dígitos en Alemania, mientras que en Brenner (2003 y 2004) el análisis se amplía a los 104 sectores manufactureros de la clasificación a 3 dígitos. En ambos casos el ámbito territorial es el de los 441 distritos administrativos existentes en Alemania a 30 de junio de 2001.

En ambos estudios las líneas fundamentales de la estrategia adoptada son similares aunque hay diferencias en algunos aspectos. De entre estos destaca que para la desagregación a tres dígitos el análisis se efectúa tanto en términos absolutos como en términos relativos. Para ello, en el primer caso no se tiene en cuenta la dimensión de las regiones, asignando un mismo valor de  $s$  para todas ellas igual a  $1/441$ . Por el contrario, en el segundo se relativiza el número de empresas localizadas en cada región a su dimensión, aproximada a través del porcentaje del empleo de la región en el total de la economía alemana, siguiendo la pauta descrita anteriormente. Adicionalmente, en ese mismo estudio se simplifican  $P_c(f|s)$  y  $P_n(f|s)$ , excluyendo la componente asociada a la distribución binomial en ambos casos, de forma que las funciones empleadas son:

$$P_c(f | s) = [1 - \xi_3(s) - \xi_8(s)] \cdot [1 - \xi_1(s)] \cdot \xi_1(s)^f + \xi_3(s) \cdot \frac{[1 - \xi_2(s)]^2}{\xi_2(s)} \cdot f \cdot \xi_2(s)^f$$

$$+ \xi_8(s) \cdot \frac{[1 - \xi_7(s)]^2}{\xi_7(s)} \cdot f \cdot \xi_7(s)^{f - \xi_6(s)} \quad \text{si } f \geq \xi_6(s)$$

$$0 \quad \text{si } f < \xi_6(s)$$

$$P_n(f|s) = [1 - \xi_3(s)] \cdot [1 - \xi_1(s)] \cdot \xi_1(s)^f + \xi_3(s) \cdot \frac{[1 - \xi_2(s)]^2}{\xi_2(s)} \cdot f \cdot \xi_2(s)^f$$

Las razones de esta simplificación no son explicitadas claramente, indicándose únicamente que la distribución tipo Boltzmann parece aproximar mejor que la Binomial la distribución empírica de un mayor número de sectores. No obstante, no descartamos que tal decisión pueda también deberse a cuestiones computacionales dado que la versión simplificada de las distribuciones se utiliza en la aplicación en la que se emplean un mayor número de sectores y donde el número de firmas por sector en cada región debe ser notablemente más reducido.

En cualquier caso, los resultados para la economía alemana sugieren que la contribución de las distintas componentes de las distribuciones teóricas varían entre sectores, aunque con un cierto predominio de la tipo Boltzmann. En cuanto a la detección de clusters, se obtienen en 4 de los 23 sectores a dos dígitos y en 54 de los 104 a tres dígitos, lo que indica que el detalle sectorial empleado tiene un efecto importante sobre la detección de aglomeraciones industriales. Por otra parte, la aplicación del contraste de Kolmogorov-Smirnov revela que para la mayoría de sectores no se puede rechazar la igualdad de las distribuciones teóricas, tanto la neutral como la asociada a la presencia de clusters, y la empírica.

#### **4.2.5. Potencialidad de resultados obtenidos**

- La identificación de clusters industriales en el territorio siguiendo el procedimiento basado en la comparación de distribuciones permite determinar la existencia de aglomeraciones de plantas productivas en cada uno de los sectores considerados. El método asume que tales aglomeraciones son causadas por la presencia de externalidades con una dimensión espacial que hacen más rentable a las empresas co-localizarse en el territorio, dando lugar a clusters industriales. No obstante, para garantizar que esto es realmente así se debería complementar el análisis efectuado con

este método con otro que incorporase de forma directa o indirecta información acerca de las interacciones entre las plantas productivas. En cualquier caso, el uso de información estadística referida a un detalle sectorial y territorial elevado permitiría la obtención de un mapa preciso de las aglomeraciones existentes en el territorio español para cada uno de los sectores analizados.

- La riqueza del mapa dependería en gran medida del detalle sectorial y territorial que se pudiese llegar a emplear. Respecto al detalle sectorial, sería interesante obtener un mapa tanto con un detalle sectorial elevado como otro correspondiente a agrupaciones sectoriales más genéricas. Por su parte, la necesidad de disponer de un número de observaciones suficientemente elevado para garantizar las propiedades asintóticas del proceso de estimación e inferencia propios de esta metodología, desaconsejan su aplicación para el caso de las Comunidades Autónomas e incluso de las provincias. De esta forma el ámbito territorial más adecuado podría ser el de los municipios o el de alguna agrupación natural de éstos.
  
- La utilización de información territorial detallada permitiría la replica del análisis para cada una de las Comunidades Autónomas de forma aislada. Este ejercicio permitiría obtener mapas detallados de la localización de aglomeraciones a nivel regional, que podrían proporcionar evidencia complementaria a la obtenida para el conjunto del territorio español. Por ejemplo, permitiría la detección de concentraciones de actividad que, aunque modestas a nivel global, podrían llegar a tener su relevancia para algunas regiones.



### 4.3. Método basado en la distribución de distancias

Los dos métodos expuestos anteriormente utilizan unidades geográficas como países, regiones, condados, etc., como elementos en los que se mide el volumen de actividad económica en cada sector, que sirve de base para constatar la presencia de clusters. Es decir, realizan la detección e identificación de clusters para un nivel de agregación espacial fijado a priori. La agregación espacial conlleva una obvia ventaja computacional pero con el inconveniente de despreciar mucha información relativa a la localización precisa de las plantas productivas. Además lleva asociados problemas entre los que destacan:

- la existencia de un escaso número de niveles de agregación espacial en cada país,
- la dificultad de comparar los resultados obtenidos para distintos niveles de agregación espacial,
- el hecho de que las unidades espaciales se definan según criterios administrativos y no económicos,
- el que las unidades espaciales consideradas difieran en cuanto a su tamaño, población, dimensión del mercado, etc.,
- el problema de la unidad espacial modificable (el conocido entre los geógrafos como MAUP), que en el caso concreto que nos ocupa implica que la agregación de plantas en las unidades espaciales consideradas puede provocar la existencia de correlación espuria entre las variables de interés,
- la simetría en el tratamiento de las unidades espaciales consideradas, lo que supone que una región vecina próxima sea considerada igual que otra localizada a mil kilómetros. Esta circunstancia supone un problema especialmente grave cuando la localización de un sector sobrepasa las fronteras administrativas de las regiones.

Para tratar de superar los problemas asociados a la agregación en las medidas y contrastes de localización, Durantón y Overman (2005) señalan que cualquiera de ellas para ser realmente adecuada debe evitar la utilización de unidades espaciales agregadas sugiriendo, en su lugar, realizar el análisis para el espacio continuo. Es decir, utilizar la distancia entre las plantas productivas para valorar los patrones de localización espacial de la actividad.

La aplicación del método basado en la distribución de las distancias permite pues la obtención de una medida de localización, y de un contraste del supuesto de aleatoriedad en la distribución frente al de aglomeración, con interesantes propiedades y que supera en gran medida los inconvenientes señalados para los otros procedimientos propuestos en la literatura. Pero no constituye en sí mismo un método que permita directamente la detección e identificación de clusters industriales. Sin embargo, lo recogemos en este estudio dado que creemos que se debería valorar su utilización para contrastar de forma más robusta la existencia de aglomeración en cada uno de los sectores considerados, y la posibilidad de usar la información que proporciona para, combinándola con otros elementos, poder realizar un proceso de “*cluster mapping*” alternativo a los anteriormente expuestos.

#### ***4.3.1. Elementos esenciales del método.***

##### **Supuestos iniciales**

Los citados autores consideran crucial, tanto desde un punto de vista teórico como de política económica, conocer i) la escala espacial de los clusters, ii) si es una estructura de pequeñas y medianas empresas o un pequeño grupo de grandes empresas el causante de una elevada concentración, y iii) el alcance sectorial de la localización. Para aportar conocimiento sobre estas cuestiones clave, Durantón y Overman proponen un nuevo contraste basado en una medida de localización, con fundamentos de la estadística espacial. En su opinión, cualquier medida de localización debe tener las siguientes características:

- controlar por la tendencia global a la aglomeración de la industria,

- controlar la concentración sectorial,
- ser insesgada respecto a la escala y la agregación espacial, y
- proporcionar el nivel de significación de los resultados, es decir que se base en la aplicación de un enfoque probabilístico.

Los índices propuestos en Ellison y Glaeser (1997) y Maurel y Sédillot (1999) cumplen con 1) y 2), pero no con 3) y 4) dado que agregan las plantas productivas en regiones definidas a un determinado nivel de agregación. Por el contrario, el índice propuesto por Duranton y Overman cumple con todas las propiedades, al considerar la distribución de las distancias entre pares de plantas en un sector y compararlo con aquéllas de sectores hipotéticos con el mismo número de plantas que se encuentran aleatoriamente distribuidas, condicionadas a la distribución del conjunto de la industria. Así, el método de “*cluster mapping*” que es posible implementar en función de la sugerencia de los citados autores se basa en la comparación de la distribución real de las distancias entre plantas productivas de un determinado sector con la asociada a un modelo de localización lo más simple posible: el de aleatoriedad condicionada a la distribución de la actividad en el conjunto de la industria.

### **Metodología**

La implementación del método basado en la distribución de las distancias entre plantas productivas requiere el seguimiento de las siguientes fases:

- 1) Selección de las plantas productivas relevantes. Se debe determinar a priori cuál es el universo de plantas de cada sector que deben ser consideradas en el análisis. La opción más inmediata es la utilización de la información de todas ellas, aunque la inclusión de un elevado número de plantas de dimensión muy pequeña (por ejemplo menos de 10 trabajadores) puede distorsionar los resultados, dado que las decisiones de localización de éstas pueden ser muy diversas y distintas de las correspondientes a

empresas de mayor dimensión. Para ilustrar esta circunstancia, Durantón y Overman discuten el caso del sector de la construcción naval, que en el Reino Unido presenta un gran número de plantas muy pequeñas en las zonas no costeras mientras que todas las grandes empresas se encuentran localizadas en la costa. Concluyen que es muy probable que, aunque clasificadas en el mismo sector, unas y otras realicen actividades muy distintas. Por ello plantean dos alternativas, una consistente en establecer un umbral de tamaño a partir del cual incluir a las plantas en el análisis y otra basada en la ponderación de la contribución de cada planta según el tamaño de la misma. Estas opciones dan lugar a resultados distintos dependiendo de la distribución de la dimensión de las plantas productivas en el espacio.

2) Cálculo de la densidad de las distancias (densidades-K) para todos los pares de empresas en cada sector. Para ello:

- i. se calcula la distancia euclídea,  $D_{ij}$ , entre todo par de plantas del sector analizado. Para un sector con  $n$  plantas se calcularán un total de  $n(n-1)/2$  distancias
- ii. se estima la densidad de las distancias bilaterales para toda distancia  $d$  a través del método kernel (ver, por ejemplo, Silverman, 1986):

$$\hat{K}(d) = \frac{1}{n(n-1)h} \sum_{i=1}^{n-1} \sum_{j=i+1}^n f\left(\frac{d - d_{ij}}{h}\right)$$

donde  $h$  es el ancho de banda y  $f$  la función kernel.

En el caso de emplear una ponderación por la dimensión de las plantas, por ejemplo a través del número de empleados,  $e$ , la densidad-K se define como:

$$\hat{K}^{\text{emp}}(d) = \frac{1}{h \sum_{i=1}^{n-1} \sum_{j=i+1}^n e(i)e(j)} \sum_{i=1}^{n-1} \sum_{j=i+1}^n e(i)e(j) f\left(\frac{d - d_{ij}}{h}\right)$$

La densidad-K ha sido frecuentemente utilizada en otros ámbitos (geografía, biología, etc) para el análisis de la distribución espacial de individuos de interés, aunque en este caso Duranton y Overman sugieren alisar las  $K(d)$  a través de la estimación de funciones de densidad mediante la aplicación del método kernel con el objeto de minimizar los errores que se pudieran cometer en el cómputo de las distancias.

- 3) Construcción de distribuciones contrafactuales. Estas sirven de referencia para poder discernir entre situaciones hipotéticas de aleatoriedad y de aglomeración. Para ello se consideran sectores virtuales con el mismo número de plantas, que se encuentran aleatoriamente situadas en una de las localizaciones realmente ocupadas por una planta de cualquiera de los sectores que componen la industria. Sería posible obtener distribuciones contrafactuales bajo supuestos alternativos, pero se prefieren los citados dado que se pretende controlar por la tendencia global de la industria a concentrarse en el espacio, así como por el hecho de que existan lugares en el territorio en los que no es posible la localización de plantas productivas (tanto por razones naturales, i.e. existencia de un lago, como por reglamentaciones específicas). Por lo tanto en la obtención de las distribuciones virtuales de referencia se asume que el espacio de todas las localizaciones realmente existentes por todas las plantas de la industria, independientemente del sector al que pertenezcan, constituyen el conjunto de todas las posibles localizaciones para cualquier planta de un sector analizado.

Siguiendo esa premisa se generan distribuciones contrafactuales mediante un muestreo de la población del conjunto de localizaciones existentes. Para cada sector en el caso del Reino Unido, Duranton y Overman obtienen 1000 simulaciones. Cada una de estas simulaciones da lugar a la localización de las  $n$  plantas del sector en  $n$  localizaciones del conjunto de todas las posibles. A continuación se calcula  $K(d)$  alisada para cada simulación.

- 4) Comparación de la distribución real de distancias con la distribución de distancias teóricamente asociada a una situación de aleatoriedad en la localización de las plantas productivas (únicamente condicionada a la aglomeración observada en el conjunto de la industria). Las  $K(d)$  correspondientes a las distribuciones contrafactuales simuladas bajo el supuesto de aleatoriedad permiten determinar valores críticos empíricos con los que comparar a las  $K(d)$  asociadas a la distribución real. La comparación se puede realizar tanto para el conjunto de distancias en el intervalo que se considere relevante como a nivel global. Esto permite determinar si el sector analizado presenta globalmente un exceso de localización respecto al conjunto de la industria y para qué distancia se cumple tal circunstancia.

#### **Identificación de los clusters industriales**

Como se ha indicado anteriormente, en el trabajo de Duranton y Overman no se plantea la detección e identificación de clusters industriales sino únicamente el diseño de un contraste de aglomeración espacial. Sin embargo, creemos que la aplicación de su método permitiría, como mínimo, comprobar de forma robusta la presencia de clusters en cada uno de los sectores analizados. Y además sería interesante investigar la potencialidad de los resultados que se obtienen con la aplicación de esa metodología para la realización de un “*cluster mapping*”. Únicamente a modo de ejemplo, dado que no constituiría un método riguroso ni objetivo, sugerimos combinar la información que nos indica que un sector presenta un exceso de localización en términos globales y para qué rango de distancias lo presenta, con la simple visualización de las plantas productivas de ese sector en un mapa. De esa forma se podría fácilmente determinar las aglomeraciones de plantas más significativas del sector analizado. De hecho este proceder es seguido en el trabajo de Duranton y Overman para ilustrar la existencia de algunos clusters en sectores seleccionados para el caso del Reino Unido. Opciones más sofisticadas implicarían la aplicación de algoritmos que agrupasen plantas a partir de la información de la distancia que hay entre ellas, como por ejemplo los empleados en los habituales análisis multivariantes de agrupación de individuos.

**Relación con la función K de Ripley**

Marcon y Puech (2003) señalan que la densidad-K definida en el trabajo de Duranton y Overman es muy parecida a una medida de concentración espacial basada en el análisis de la distribución de puntos ampliamente utilizada en disciplinas como la ecología y la epidemiología. Esta medida se conoce como función K de Ripley (1976, 1977). La idea básica de esta medida es muy simple. Si se toman las plantas localizadas en un espacio determinado, éste presentará igual densidad de plantas en el caso de que cualquiera de las localizaciones que lo configuran resulte igualmente atractiva a todas ellas. Es posible entonces definir un marco de referencia calificado como de aleatoriedad espacial total, en el que las plantas se localizan en cualquier lugar con igual probabilidad y lo hacen de forma independiente a la localización de las otras plantas. Por el contrario, si la localización de una planta depende de la localización de otras, entonces cuando tomemos a una de ellas encontraremos un número mayor de plantas en su proximidad que el asociado a una situación de aleatoriedad. En consecuencia, podemos extraer conclusiones sobre el nivel de aglomeración de un sector a partir del número medio de plantas *vecinas* para cada distancia.

La función K de Ripley permite describir la distribución espacial de un conjunto de puntos. Denotando la densidad media de los puntos mediante  $\lambda$ , el número esperado de puntos en una situación de aleatoriedad en la distribución espacial en un círculo de radio  $r$  es  $\lambda\pi r^2$ . La función  $K(r)$  se define como el número medio de vecinos dividido por  $\lambda$ , es decir  $K(r) = \pi r^2$ . Este constituye el valor de referencia de la situación de aleatoriedad, con el que se compararán los valores de la función K de Ripley obtenidos, en nuestro caso, para el número medio de plantas vecinas en una distancia determinada.

Para evitar la comparación con  $\pi r^2$ , Besag (1977) normaliza la función para obtener como referencia el valor nulo, definiendo así la función  $L(r) = \sqrt{\frac{K(r)}{\pi}} - r$ . De esta forma,  $L(r) > 0$  nos indica que la distribución observada está aglomerada para una distancia de radio  $r$

mientras que si  $L(r) < 0$  la evidencia apunta a dispersión en la distribución para esa distancia.<sup>6</sup>

En consecuencia, y siguiendo un procedimiento similar al expuesto anteriormente para el caso del método de Duranton y Overman, se trata de calcular la función L para un amplio rango de distancias y comparar los valores obtenidos con los asociados a la hipótesis nula de una distribución aleatoria. Pero a pesar de ello, la función L no tiene en cuenta la heterogeneidad innata en la distribución de la actividad económica, es decir el hecho de que no todas las localizaciones del territorio resulten igualmente atractivas para las plantas del sector analizado, dado que la distribución de referencia que considera es la de aleatoriedad pura. Para incorporar la heterogeneidad espacial, Marcon y Puech sugieren definir una función D(r) mediante la que se compara la función K(r) observada para el sector analizado como una función K(r) asociada a una distribución de control: la correspondiente a las plantas de todos los demás sectores. De esta forma la función D(r) se define como la diferencia entre ambas funciones K. En consecuencia, la función D(r) permite detectar desviaciones del patrón de localización de las plantas de un sector respecto del existente para el conjunto de plantas, excluidas las del propio sector analizado.

A pesar de que las funciones R y L tienen un sólido bagaje en otros ámbitos científicos, Marcon y Puech (2003) reconocen que estas medidas precisan de adaptaciones para alcanzar todas las propiedades que cumple la densidad-K propuesta en Duranton y Overman.

---

<sup>6</sup> El cálculo de L requiere la definición de un círculo de radio r en torno a la localización de cada una de las plantas productivas. Esto ocasiona problemas en las inmediaciones de las fronteras del espacio analizado, dando lugar al calificado como *edge effect*. El sesgo inducido por este efecto se puede corregir utilizando únicamente la parte del área del círculo incluida en el espacio considerado.



#### 4.3.2. *Ventajas e inconvenientes*

##### Ventajas

- La medida de localización propuesta en Duranton y Overman, y el contraste basado en ésta, cumple con cuatro propiedades deseables: i) garantiza la comparación entre sectores, ii) controla por el nivel de concentración de la industria en su conjunto, iii) es insesgada respecto a la escala y la agregación espacial, y iv) proporciona información acerca del nivel de significación estadística de los resultados.
- En consecuencia, la identificación de exceso de localización (aglomeración) o de dispersión en un sector se basa en la utilización de un criterio estadístico, fijado a priori, objetivo y reproducible. Aunque se debe recordar que ello no posibilita la detección e identificación directa e inmediata del espacio geográfico en el que, en su caso, se sitúa la aglomeración, y por tanto la realización del “*cluster mapping*” a través de un procedimiento totalmente objetivo.
- Se tienen en consideración las características específicas de cada sector a la hora de computar los criterios estadísticos en los que se apoya el método. En consecuencia, los intervalos y bandas de confianza locales y globales son específicos para cada sector.
- Permite controlar por diferencias en la distribución del tamaño entre plantas productivas de forma que, además de tener en cuenta la localización precisa de cada planta, incorpora información del empleo que se ubica en cada una de ellas. Asimismo, permite distinguir las pautas de localización de plantas de distinta dimensión.
- En comparación con otros métodos basados en funciones de distancias, no requiere de la corrección por efectos frontera (*edge effects*).

- Dado que la propuesta de Duranton y Puga apareció publicada en un medio de reconocido prestigio entre la profesión, podemos señalar que el método goza de reconocimiento académico amplio.

### **Inconvenientes**

- El mayor inconveniente de este procedimiento es que no contiene ningún método explícito de detección e identificación de clusters industriales, lo que impide su aplicación directa para la realización de un “*cluster mapping*”. En consecuencia, únicamente es posible realizar esta operación de manera subjetiva a partir de la localización en el territorio de las plantas productivas pertenecientes a los sectores para los que se concluya a favor de la existencia de aglomeración.
- Aún en el caso de plantear la extensión del método para realizar la detección e identificación de clusters, no se incorporaría en el análisis uno de los elementos constituyentes de la definición de cluster industrial: el de la presencia de externalidades entre sectores que, al estar acotadas en el espacio, provocan la localización de plantas de distintos sectores en un mismo territorio.
- El método plantea una elevada necesidad de información estadística, al precisar de información de la localización espacial precisa de todas las plantas productivas de la economía analizada. Además, en el caso en el que se desee tener en cuenta la distribución de la dimensión empresarial se debe disponer de información de, por ejemplo, el número de empleados en cada una de las plantas productivas.
- La implementación de este método es relativamente compleja e intensiva en cálculo al requerir el cómputo de las distancias entre todos los pares de empresas de cada uno de los sectores considerados y de las distancias en el caso de cada una de las distribuciones simuladas.

- Las distribuciones contrafactuales que permiten contrastar el supuesto de aleatoriedad en la distribución espacial de la actividad frente a los de aglomeración o dispersión se obtienen a partir de la asunción de un modelo muy sencillo, que simplemente reproduce la localización espacial de las plantas en el conjunto de la industria, sin que se incorporen supuestos más complejos o evidencia adicional acerca de las pautas de localización de la actividad o de los factores que pueden estar ejerciendo de atractores o repulsores para actividades específicas. Además, la misma distribución hipotética sirve como referencia para todos los sectores, lo que evidencia que no se tienen en cuenta posibles particularidades de cada uno de ellos.
- En la construcción de las distribuciones contrafactuales todas las empresas son consideradas del mismo tamaño y, por tanto, se asume que su localización es intercambiable. Este supuesto es difícilmente verosímil en el caso en que la distribución del tamaño empresarial sea muy dispersa, dado que una empresa muy grande no podrá ocupar el lugar de una empresa muy pequeña y tampoco es racional imaginar que una pequeña ocupará todo el espacio de una muy grande.
- La determinación del umbral de la distancia máxima considerada, necesario para la implementación del método, se realiza de forma *ad hoc*.
- Dado que tanto la función de densidad real como la virtual son normalizadas para que sumen la unidad, desviaciones positivas de la primera en algunas distancias se deben compensar necesariamente con desviaciones negativas para otras distancias.
- El método puede tener problemas en detectar aglomeración en sectores en los que la actividad se concentre preferentemente en corredores (a lo largo de un río, en la costa, etc.) donde la distancia física sea mayor. Precisamente, la consideración de la distancia física como único factor de proximidad entre plantas constituye otra de las debilidades del método, dado que en algunos territorios puede existir divergencia

entre la distancia en términos físicos y la existente, por ejemplo, en términos de tiempo de desplazamiento.

- El método cuenta todavía con escasa difusión, posiblemente debido a su relativa novedad, lo que impide disponer de aplicaciones del mismo y de análisis críticos. La casi nula disponibilidad de aplicaciones dificulta también la posibilidad de comparar los resultados parciales y globales que se pudieran obtener de su aplicación a la economía española con los de economías de nuestro entorno.
- Aunque no es estrictamente necesario para su implementación, las características del método aconsejan que se disponga de un número mínimo de plantas en cada uno de los sectores analizados para garantizar la robustez de los resultados. De hecho no conocemos, y creemos que sería interesante conocer, la sensibilidad del índice y del contraste de localización al número de plantas.

#### ***4.3.3. Criterio estadístico de detección de aglomeración***

Como se ha indicado anteriormente, el contraste de localización presentado en esta sección se basa en la comparación de la distribución observada de las distancias entre pares de plantas del sector analizado y las correspondientes a distribuciones contrafactuales generadas para una situación de aleatoriedad únicamente condicionada a la tendencia a la aglomeración observada en la industria en su conjunto. Dado que se persigue basar la conclusión en un criterio de tipo estadístico, se construyen intervalos de confianza para las  $K(d)$  que permiten determinar la significación de la densidad tanto de forma local como global.

#### **Intervalos de confianza locales**

Para la construcción de los intervalos de confianza, el primer paso consiste en determinar el rango de distancias que se considera relevante para el análisis. Duranton y Overman sugieren utilizar como umbral la mediana de la distribución de las distancias observadas

entre todos los pares de plantas del conjunto de la industria, lo que para el caso del Reino Unido corresponde a 180km. De esa forma el rango de distancias analizadas es aquél en el intervalo  $[0, 180]$ . Una vez determinado el rango de distancias, para cada kilómetro en ese intervalo se ordenan las distribuciones simuladas en orden ascendente, seleccionándose el percentil 5 y el 95 para obtener los intervalos de confianza correspondientes al 5% inferior y al 5% superior, respectivamente. Denotándolos mediante  $K_5(d)$  y  $K_{95}(d)$ , si para un sector se obtiene que  $K(d) > K_{95}(d)$  se debe concluir que el sector presenta localización a una distancia  $d$ . Por el contrario, si  $K(d) < K_5(d)$  el sector presentará dispersión a la distancia  $d$ . Un gráfico de la densidad-K en el intervalo de distancias  $[0, 180]$ , junto a los valores asociados a cada distancia de  $K_5(d)$  y  $K_{95}(d)$  permite visualizar fácilmente el patrón de localización del sector analizado.

### **Bandas de confianza globales**

Además de obtener información acerca de las pautas de localización para cada valor de la distancia, resulta interesante extraer una conclusión acerca del patrón de localización global de los sectores analizados. La cuestión es que incluso ante una situación real de aleatoriedad es posible obtener para algún nivel de distancia en el rango considerado un valor de la densidad-K por encima (debajo) del intervalo local superior (inferior), dado que en cada caso estamos asumiendo una probabilidad de error del 5%. Es decir, que en cada 100 distancias esperaríamos obtener 5 para las que erróneamente concluiríamos que se produce un exceso de localización. La solución sugerida por Duranton y Overman consiste en seleccionar aquellos niveles de confianza locales inferior y superior para los que, cuando se consideran todas las distancias en  $[0, 180]$ , coincidan únicamente el 5% de las densidades-K simuladas.<sup>7</sup> Así, denotando por  $K^{\text{sup}}(d)$  la banda de confianza superior, si  $K(d) > K^{\text{sup}}(d)$  para al menos un  $d \in [0, 180]$ , la conclusión es que el sector presenta localización global. Por su parte, la conclusión será que el sector muestra dispersión global cuando  $K(d) > K^{\text{inf}}(d)$  para al menos un  $d \in [0, 180]$  y el sector no presenta localización (donde  $K^{\text{inf}}(d)$  es la banda de confianza global inferior al 5%). Por lo tanto, la localización global supone que la densidad-K se encuentra por encima de la banda de confianza superior

---

<sup>7</sup> Como indican los autores, la existencia de autocorrelación positiva entre distancias impide la aplicación del método tradicional de Bonferroni para la corrección del nivel de significación en la aplicación de contrastes secuenciales.

para al menos un nivel de distancia, mientras que la dispersión global requiere que la densidad-K se encuentre por debajo de la banda inferior para al menos una distancia y que en ninguna sobrepase la banda superior.

#### ***4.3.4. Aplicaciones de la metodología***

Las únicas aplicaciones de las medidas de localización basadas en distancias son las realizadas para el Reino Unido por Durantón y Overman (2005) y por Fratesi (2005), y para el área de París por Marcon y Puech (2003).

En el caso de los dos primeros trabajos, los autores explotan la exhaustiva información contenida en la base de datos ARD, producida por el instituto nacional de estadística británico y que contiene información acerca del número de empleados, el sector de actividad (según la clasificación SIC) y el código postal para todos los establecimientos del Reino Unido. Durantón y Overman limitan su análisis a los establecimientos productivos del sector industrial, considerando agrupaciones sectoriales de 2 a 5 dígitos SIC. La información de los códigos postales es combinada con la referida a las coordenadas espaciales de todos los códigos postales en el Reino Unido contenida en la base de datos Code-Point del Ordnance Survey. Esto permite a los autores obtener información bastante precisa de la localización geográfica de todas las plantas productivas del sector industrial en el Reino Unido.

Una vez obtenida la localización espacial de las plantas, Durantón y Overman implementan su procedimiento para la totalidad de las plantas, diferenciando por tamaño al considerar aquellas con 10 o más trabajadores y con menos de 10 trabajadores, y para la totalidad de las plantas pero ponderando por el número de empleados de cada una de ellas, de forma que en este último caso los resultados que obtienen deben ser interpretados en términos de localización del empleo.

Los resultados obtenidos permiten a los autores obtener las siguientes conclusiones acerca de los patrones de localización de la actividad productiva de los sectores industriales en el Reino Unido:

- el 52% de los sectores están localizados (presentan un exceso de localización, es decir están aglomerados), mientras que un 24% están dispersos, todo ello a un 5% de nivel de confianza. El resto no se desvían significativamente de la aleatoriedad,
- la localización se produce básicamente entre 0 y 50 km de distancia
- el grado de localización es muy distinto entre sectores
- los sectores que pertenecen a una misma rama de actividad suelen presentar patrones de localización similares
- al diferenciar por el tamaño de las plantas se obtienen patrones de localización muy diversos dependiendo de los sectores. En algunos sectores son las plantas pequeñas las más aglomeradas mientras que en otros la localización es mayor en las grandes.
- no hay diferencias sustanciales en los patrones de localización entre los sectores a 4 y 5 dígitos, pero sí a 3 dígitos. Para este último caso, la localización se produce tanto en distancias cortas (0 a 50 km) que pueden ser asociadas al ámbito local, como a escala regional, con distancias entre 80 y 140 km.

Por su parte, Fratesi (2005) analiza en detalle la localización de dos sectores de la clasificación SIC a 3 dígitos para el Reino Unido: el SIC 244 correspondiente al sector Farmacéutico y el SIC 334 de Óptica y Fotografía. Para ambos considera todos los subsectores de la clasificación a 5 dígitos. El interés del análisis específico de estos sectores radica, según el propio autor, en que se corresponden con actividades industriales avanzadas para las que *“los aspectos organizativos deberían jugar un papel más importante que las ventajas comparativas geográficas o las razones históricas”*. Para ello

utiliza índices de localización tradicionales como el de Gini, índices del tipo Ellison y Glaeser y las modificaciones de Maurel y Sedillot (1999) y Devereux et al (1999), y medidas basadas en la distribución de distancias, siguiendo la aproximación de Duranton y Puga.

Los resultados del citado trabajo confirman que se obtienen distintas conclusiones acerca del grado de aglomeración al utilizar las medidas basadas en la densidad de las distancias, aunque se debe tener presente que los resultados obtenidos mediante los índices de concentración y los basados en distancias no son directamente comparables dado que los primeros se refieren a la concentración del empleo mientras el segundo se basa en el índice no ponderado según el volumen de empleo en cada planta, por lo que hace referencia a la localización de las plantas productivas. Fratesi considera que *“las distintas medidas revelan aspectos complementarios de los patrones de localización, por lo que es razonable la utilización simultánea de todas las medidas en los trabajos empíricos”*. En cualquier caso, los resultados obtenidos a través de las densidades-K revelan diferencias sustanciales en el grado de localización entre los subsectores de los dos sectores considerados y como, en los casos en los que existe exceso de localización, ésta se produce para distancias no mayores a 50 km.

Finalmente, cabe señalar que Marcon y Puech (2003) analizan los patrones de localización en los sectores industriales del área de París y del conjunto de Francia, utilizando la función L de Besag, la cual, como indicamos anteriormente, es una modificación de la función K de Ripley. Para ello consideran las plantas de los 14 sectores industriales (5739 en el caso de París y 25186 para el del conjunto de Francia), concluyendo que los patrones de localización en el área urbana de París difieren de los del conjunto de Francia. Así, en el caso parisino detectan 5 sectores significativamente concentrados y 9 con dispersión significativa, mientras que para el conjunto del territorio francés son sólo 4 los sectores donde la concentración es estadísticamente significativa, y también 4 los que muestran dispersión significativa. Los restantes 6 muestran dispersión o concentración para algún rango de distancia. Además, en ambos ámbitos territoriales los patrones de concentración



geográfica difieren entre sectores, variando por ejemplo la distancia en la que aparecen los picos de mayor concentración.

#### **4.3.5. Potencialidad de resultados obtenidos**

- Como se ha reconocido al principio de esta sección, las aportaciones existentes en la literatura referidas al método basado en las distancias no contienen un procedimiento explícito de detección e identificación de clusters industriales, por lo que su utilización para el caso de la economía española no permitiría en sí misma la obtención de un mapa de los clusters existentes. No obstante, creemos interesante plantear, al menos con vistas al futuro, la utilización de medidas y tests basados en las distancias para la detección robusta de aglomeración en los sectores analizados. Creemos que es el método que garantizaría una mayor fiabilidad y riqueza de resultados, al controlar por la tendencia global a la aglomeración de la industria y por la concentración sectorial, por ser insesgada respecto a la escala y la agregación espacial, y por proporcionar el nivel de significación de los resultados. Dado que la identificación de los sectores que presentan aglomeración de la actividad es un paso previo en cualquier método de detección e identificación de clusters, creemos interesante poder incorporar las medidas propuestas por Duranton y Overman a cualquiera de los procedimientos de “*cluster mapping*”, bien sustituyendo o complementando el uso de medidas basadas en la agregación territorial.
- Las medidas basadas en distancias no consideran de forma directa la interacción entre las empresas por lo que no permiten extraer conclusiones acerca de la existencia de externalidades acotadas espacialmente, pero sí permiten comprobar la tendencia de las plantas productivas de distintos sectores a co-localizarse en el espacio, mediante una sencilla modificación de la densidad-K.<sup>8</sup>

---

<sup>8</sup> Véase el apartado 6 de Duranton y Overman (2005).

- El método basado en las distancias precisa de las coordenadas espaciales de la localización de todas las plantas, lo que puede resultar en un grave inconveniente para el caso de la economía española. No obstante, aún en el caso de no disponer de esa información georeferenciada, sería posible obtener las medidas de localización en el caso de disponer, por ejemplo, del municipio en el que se localizan las plantas. En ese caso, se debería realizar una imputación de una localización espacial precisa a cada una de las plantas de todos los municipios, incurriendo en un error que está en relación al área del menor de los municipios.
- Sería posible contrastar la existencia de aglomeración y de dispersión tomando como referencia tanto el conjunto de la economía española como territorios específicos de la misma. Así se podría realizar el análisis para cada una de las CCAA, simplemente replicando el procedimiento con las plantas localizadas en cada una de las regiones.
- Sería posible obtener resultados referidos tanto a los patrones de localización de las plantas productivas como del empleo, en el caso en el que fuese posible disponer, no únicamente de la localización de cada una de las plantas sino también de su volumen de empleo.

## **5. Valoración de la aplicación de los métodos propuestos al caso de la economía española**

Las características que consideramos más relevantes para realizar una valoración de la posibilidad de aplicación de cada uno de los procedimientos de detección e identificación de clusters industriales revisados en la sección anterior, son sintetizadas en las Tablas 1 a 3. En ellas se recopilan, para cada uno de los métodos, las principales ventajas e inconvenientes, el criterio en el que se basa la decisión de calificar a un sector analizado como aglomerado, la delimitación del ámbito geográfico en el que sería posible realizar la aplicación para nuestro país, los requerimientos de información estadística, y la posibilidad de comparación de los resultados que se obtendrían con los obtenidos en estudios previos.

Cabe señalar en primer lugar que, del exhaustivo análisis de los tres métodos considerados en este informe, se desprende la existencia de limitaciones e inconvenientes para cualquiera de ellos. Creemos importante explicitar estas limitaciones, especialmente para el método o métodos que sean finalmente utilizados para que, de esta forma, se contextualicen adecuadamente los resultados obtenidos y para que no se extiendan en exceso las implicaciones de política económica que se puedan derivar de ellos.

El estudio de las propuestas que proponen métodos basados en una función de las distancias, y en especial el de Duranton y Overman (2005), nos lleva a descartar esa aproximación como procedimiento fundamental de “*cluster mapping*”. Aunque creemos factible que en el futuro se desarrolle un instrumento objetivo y cuantificable que permita la detección e identificación de clusters industriales a partir de la información obtenida a través de las distancias bilaterales entre plantas, por el momento tal procedimiento está pendiente de desarrollo. Aunque la detección e identificación se podría apoyar en la inspección visual de las plantas productivas en un mapa del territorio español para aquellos

## Valoración de la aplicación de los métodos propuestos al caso de la economía española

sectores en los que se haya detectado aglomeración, no creemos que tal proceder cumpla con los imprescindibles requisitos de objetividad y reproducibilidad.

Sí creemos interesante que, en el caso de poder llegar a disponer de la información georeferenciada para todas las plantas productivas existentes en la economía española, se utilicen las medidas y contrastes basados en las funciones de distancia como paso previo para discriminar entre los sectores que presentan un exceso de localización, es decir aglomeración, de aquéllos que se encuentran aleatoriamente distribuidos o dispersos. El que este tipo de medidas presenten propiedades deseables ausentes en otros de los índices de localización existentes (muy especialmente su robustez a la desagregación territorial) nos lleva a sugerir su aplicación de forma complementaria a la implementación de cualquier otro método.

En este sentido, se debe tener presente la posibilidad de implementar esas medidas aún en el caso de no disponer de forma precisa de las coordenadas de todas las plantas. Sería posible hacerlo, con un nivel de error asumible, en caso de disponer del municipio en el que se localizan las empresas, dado que se podría simular una localización geográfica concreta en los límites espaciales de cada municipio para cada una de las empresas en ellos ubicadas.

Respecto a la aplicación de un método *à la Porter* para la realización de un “*cluster mapping*” en la economía española, no cabe duda de que esta opción tendría la clara ventaja de la sencillez de su implementación, tanto en lo referido al instrumental empleado como respecto a la información estadística requerida. Además es el único de los métodos revisados para el que se dispone de un número relativamente elevado de aplicaciones previas, lo que facilitaría la comparación con los resultados que se obtendrían y la utilización de alguna de las variaciones que se han ido realizando a la propuesta original.

Pero como se ha señalado en la sección 4.1, este tipo de metodología presenta serias limitaciones que nos hacen dudar de la robustez de los resultados que se alcanzarían en el caso en el que fuese aplicada para la economía española. En primer lugar, todos los indicios nos llevan a plantear que los resultados podrían diferir notablemente dependiendo del

## Valoración de la aplicación de los métodos propuestos al caso de la economía española

detalle sectorial y territorial empleados. Dado que no disponemos de elementos consistentes que nos hagan decantarnos por la aplicación para las CCAA, las provincias o los municipios, lo razonable sería reproducir los resultados para cada uno de esos ámbitos territoriales y analizar en detalle los resultados, confrontando los clusters detectados en cada uno de esos niveles territoriales. Pero aún en ese caso, nos cabría la duda de si la economía española es lo suficientemente amplia y variada como para que no se llegase a incurrir en el problema de representatividad reconocido por el propio Porter. Recuérdese que este autor llega a dudar de la fiabilidad de la aplicación de su método para una economía de dimensiones más modestas que la estadounidense.

En segundo lugar, otro factor que nos hace plantear dudas sobre la aplicación de esta metodología es la necesidad de adoptar muchas decisiones discrecionales, que deberán basarse exclusivamente en el criterio del equipo de investigadores que realice el estudio. Aun en el caso en que se fundamenten en un conocimiento profundo de la actividad productiva de la economía española, de todos sus sectores y territorios, la subjetividad afectará a los resultados obtenidos, siendo difícil que, por las propias características del método, se puedan explicitar nítidamente los criterios adoptados.

Finalmente, queremos insistir en las limitaciones del coeficiente de localización geográfico (CL) para valorar el grado de aglomeración de cada uno de los sectores considerados. Esta medida de localización no cumple con la mayoría de propiedades deseables para ese tipo de medidas, a pesar de que recientemente se haya sugerido una sencilla adaptación que permite valorar estadísticamente su significación bajo supuestos bastante generales. Por ello, en el caso en que se utilizase total o parcialmente un método *à la Porter*, recomendaríamos sustituir, o en todo caso complementar, el CL por otro índice de localización que cumpla con más propiedades deseables (por ejemplo el de Ellison y Glaeser o modificaciones de éste).

Ya hemos señalado que la falta de información acerca de la desagregación territorial más adecuada y la conocida sensibilidad de los resultados a la misma, nos hace sugerir la detección de clusters en los tres niveles territoriales habituales para la economía española.

Una circunstancia similar se plantea para el caso del detalle sectorial más adecuado. Dada la posibilidad de disponer de información para la clasificación CNAE a 2, 3 y 4 dígitos, creemos interesante que se realizase un análisis de sensibilidad de los resultados a esos niveles de detalle sectorial. No obstante, debe advertirse que aventuramos problemas en el caso en que el análisis se realice al mayor nivel de desagregación sectorial y territorial, dado que el elevado número de municipios con escasa o nula representación de actividad en muchos de los sectores conduciría a resultados espurios en cualquier índice de localización empleado.

En cualquier caso, la implementación de un método *à la Porter* precisaría de la recopilación del número de ocupados y del número de empresas en cada uno de los sectores considerados y para cada uno de los niveles territoriales para los que se desea realizar el análisis.

Por su parte, la aplicación de un método basado en la comparación de distribuciones, aún adoleciendo de algunos inconvenientes, plantea algunos rasgos interesantes que lo hacen especialmente atractivo para la realización de un “*cluster mapping*” en la economía española. En primer lugar, y como se ha indicado anteriormente, resulta un procedimiento objetivo y reproducible siempre que se expliciten claramente los no excesivos elementos que quedan a la discreción del equipo investigador. Además, es un método que incorpora ciertos criterios económicos que permiten realizar una interpretación más sólida de los resultados obtenidos, a la vez que incorpora algunas de las características propias de los sectores y regiones analizadas a la hora de fijar los umbrales que permiten delimitar la extensión de los clusters.

En todo caso, acompañamos la recomendación de la aplicación de este método con una solicitud de prudencia respecto a la interpretación de los resultados obtenidos y la derivación de conclusiones a partir de los mismos. Esto es así dado que, como hemos detallado en el apartado 4.2, el método es susceptible de problemas de robustez a la desagregación territorial y sectorial empleada. Por ello, de nuevo en este caso, recomendamos la implementación a todos los niveles posibles y la confrontación de los

## Valoración de la aplicación de los métodos propuestos al caso de la economía española

clusters detectados en cada uno de los casos. Adicionalmente, se debe tener presente que el método no incorporará uno de los elementos que parecen ser fundamentales en la definición de cluster industrial, el de la interrelación entre empresas de un mismo y distintos sectores debida a la existencia de externalidades acotadas espacialmente. Por ello, podría ser interesante realizar un análisis posterior de co-localización, aunque sea mediante la utilización de información cualitativa o procedente de otras fuentes, en línea con la agrupación de sectores realizada en el método de Porter (descrito en el apartado 4.1.1).

Recomendamos también que se proporcione toda la información que permita valorar el proceso de estimación de los parámetros de las funciones en presencia y en ausencia de aglomeración, y de los contrastes que permiten concluir acerca de la presencia de cluster en los sectores analizados. En concreto, se debería demandar información sobre el algoritmo numérico empleado, las condiciones de convergencia establecidas y si éstas se han alcanzado satisfactoriamente en cada caso. Asimismo, se debería realizar una valoración de la adecuación para cada sector de las funciones teóricas empleadas.

En cuanto al detalle territorial en el que se debería realizar el análisis, ya hemos señalado que ante la ausencia de elementos que nos permitan decantarnos claramente por uno de ellos y ante los indicios de sensibilidad de los resultados a distintos niveles de agregación, nuestra propuesta es la de realizar el análisis para todos los ámbitos posibles. En este caso, no obstante, debemos descartar el análisis para las CCAA dado que las características asintóticas de los procedimientos empleados desaconsejan su aplicación para un número reducido de regiones. Incluso los resultados para las provincias deberían ser tomados con precaución, especialmente para los sectores en los que los contrastes no resulten contundentes. De esta forma, teniendo en cuenta las características del método, el ámbito municipal y el de agrupaciones naturales de municipios (en por ejemplo comarcas) podrían, a priori, conducir a los resultados más adecuados.

Respecto al detalle sectorial, de nuevo en este caso planteamos un análisis a distintos niveles de desagregación dada la esperada sensibilidad de los resultados al detalle sectorial empleado. Así, sugerimos realizar el análisis para la clasificación CNAE a 2, 3 y 4 dígitos,

## Valoración de la aplicación de los métodos propuestos al caso de la economía española

a pesar de que se debe tener presente el problema que pudiera aparecer en el caso de la desagregación más elevada, por la ausencia de actividad en muchos de los municipios al emplear esas unidades territoriales. Esta misma circunstancia es la que nos lleva a desaconsejar a priori la utilización de la desagregación a 5 dígitos.

En cualquier caso, el nivel de desagregación empleado puede estar condicionado por la necesaria disponibilidad de información acerca del número de plantas productivas y del número de ocupados en todos los sectores y territorios considerados.

En resumen, teniendo en cuenta las características de cada uno de los procedimientos analizados, su potencialidad para proporcionar información acerca de la localización precisa de los clusters industriales en el territorio español, y las esperadas limitaciones de información estadística nuestra recomendación es:

- realizar la detección e identificación de clusters industriales en España mediante la implementación del método basado en la comparación de distribuciones espaciales propuesto por T Brenner,
- realizar el proceso para todos los niveles sectoriales y territoriales posibles con el objetivo de valorar la sensibilidad de los resultados a los distintos niveles de agregación,
- incorporar información acerca de la co-localización de sectores en los clusters detectados, siguiendo por ejemplo el procedimiento empleado en la propuesta de Porter,
- interpretar los resultados con la cautela necesaria basada en el conocimiento de las limitaciones de los métodos empleados.



*Tabla 1. Principales características de un método à la Porter.*

Ventajas	Inconvenientes	Criterio	Ámbito Geográfico	Datos	Comparabilidad
- <b>Sencillez</b> de cálculo e interpretativa	-Basado en <b>información agregada</b> (problema de robustez)	- <b>Adhoc</b> (basado en distribución <b>CL</b> y opinión expertos)	- Debe ser delimitado <b>a priori</b>	- <b>Empleo</b> desagregado sectorial y espacialmente	- <b>Posible</b> con experiencias en: <b>US, UK, nórdicos, etc.</b>
- <b>Baja</b> necesidad informativa	- <b>No invariante</b> a ámbito geográfico (MAUP) y definición sectorial	- Control expost y redefinición	- <b>CCAA, Provincias, Municipios</b>	- <b>Número empresas</b> desagregado sectorial y espacialmente	- <b>Dificultad</b> por uso de <b>criterios adhoc.</b>
- <b>Comparabilidad</b> con otras experiencias	-Alto contenido de <b>criterios adhoc</b> (difícil reproducción)	- Propuesta reciente de <b>criterio estadístico</b> basado en <b>estandarización de CL</b>			
- <b>Reconocimiento</b> académico y social	- <b>No</b> incorpora <b>particularidades de sectores</b>				
- Fácil <b>reproducibilidad</b> y <b>extensión</b> (geográfica y temporal)	- <b>Imposibilidad discriminación</b> clusters vs agrupaciones				
	- <b>Medida de aglomeración</b> utilizada <b>no</b> tiene en cuenta <b>tendencia a localización</b>				
	- <b>Economía</b> debe ser <b>suficientemente grande</b>				

*Tabla 2. Principales características de un método basado en la comparación de distribuciones.*

Ventajas	Inconvenientes	Criterio	Ámbito Geográfico	Datos	Comparabilidad
- <b>Objetividad y reproducibilidad</b> (basado en criterios estadísticos)	- Basado en <b>información agregada</b> (problema de robustez) - <b>No invariante</b> a ámbito geográfico (MAUP) y definición sectorial	- <b>Estadístico</b> , basado en comparación con <b>distribuciones teóricas</b>	- Debe ser delimitado a <b>priori</b>	- <b>Empleo</b> desagregado sectorial y espacialmente	- <b>Limitada</b> por escasez de aplicaciones.
- <b>Rigurosidad y actualidad</b> - Incorporación de <b>criterios económicos</b>	- Impone <b>algunos criterios adhoc</b> - <b>No interrelación sectorial</b> - <b>Imposibilidad discriminación</b> clusters vs agrupaciones	- Posibilidad de <b>control expost y redefinición</b> de umbral para detección de clusters	- <b>Provincias</b> (con reservas), <b>Municipios o agrupación de estos</b> (recomendable)	- <b>Número empresas</b> desagregado sectorial y espacialmente	- <b>Alemania.</b>
- Incorpora <b>particularidades de sector y región</b>	- <b>Complejidad cálculo</b> - <b>Problema de detección</b> en caso de <b>sectores con escasa implantación</b>		- <b>Desaconsejada su aplicación para CCAA</b> dado el escaso número de unidades territoriales para ese ámbito		
- Combina información de <b>número de plantas y de empleo</b>	- Necesidad informativa <b>media</b> - Requiere <b>número elevado de unidades territoriales</b>				
- <b>Control</b> por tendencia a <b>aglomeración</b>	- Asunción de <b>hipótesis económicas</b> - <b>Limitada comparabilidad</b> con otras experiencias  - <b>Reproducibilidad y extensión costosas</b> (geográfica y temporal)				

*Tabla 3. Principales características de un método basado en una función de las distancias bilaterales entre empresas.*

<b>Ventajas</b>	<b>Inconvenientes</b>	<b>Criterio</b>	<b>Ámbito Geográfico</b>	<b>Datos</b>	<b>Comparabilidad</b>
- <b>Objetividad y reproducibilidad</b> (basado en criterios estadísticos)	- <b>No método explícito y objetivo de detección e identificación de clusters</b>	- <b>Estadístico</b> , basado en comparación con <b>distribuciones empíricas</b>	- No delimitado, basado en localización espacial precisa	- <b>Coordenadas espaciales de empresas</b>	- <b>Limitada</b> por escasez de aplicaciones.
- <b>Reconocimiento académico, rigurosidad y actualidad</b>	- <b>No interrelación sectorial</b>	- Posibilidad de <b>detección de aglomeración global y para rangos de distancias</b>	- <b>Posibilidad de aplicar en caso de información agregada para municipios</b>	- <b>Clasificación sectorial de las empresas</b>	- <b>UK, Francia.</b>
- Incorporación de <b>criterios económicos</b>	- <b>Imposibilidad discriminación</b> clusters vs agrupaciones			- <b>Empleo</b> en cada una de las empresas	
- Incorporación de <b>particularidades de sector</b>	- <b>Complejidad cálculo</b>				
- Permite <b>controlar</b> por <b>diferencias</b> en la <b>distribución del tamaño empresarial</b>	- Necesidad informativa <b>elevada</b>				
- <b>Control</b> por tendencia a <b>aglomeración</b>	- Asunción de <b>hipótesis económicas</b>				
- <b>Invariante a escala geográfica, evita MAUP</b>	- <b>Modelo</b> localización de referencia <b>muy simple</b>				
- <b>Escasa imposición de criterios adhoc</b>	- <b>Limitada comparabilidad</b> con otras experiencias				
- <b>Amplia aplicación en otras disciplinas</b>	- <b>Reproducibilidad y extensión costosas</b> (geográfica y temporal)				
	- <b>Limitado al concepto de distancia física</b>				

**BIBLIOGRAFIA**

- Bergman E y Feser E (1999) *Industrial and Regional Clusters: Concepts and Comparative Applications*. Web Book in Regional Science, Regional Research Institute, West Virginia University
- Besag J (1977) Comments on Ripley's paper, *Journal of the Royal Statistical Society B* 39: 193-195
- Brenner T (2003) An identification of local industrial clusters in Germany, *Papers on Economics & Evolution* # 0304, Max Plack Institute, Jena
- Brenner T (2004) *Local industrial clusters: existence, emergence and evolution*. Routledge, Londres
- Brenner T (2005) A methodology to identify local industrial clusters and its application to Germany, Paper presented at the Spatial Econometrics Workshop, Kiel
- Crouch C y Farrell H (2001) Great Britain: Falling through the holes in the network concept. En C Crouch, P Le Galés, C Trogilia y H Voelzkow (eds) *Local production system in Europe: Rise or Demise*, Oxford University Press, Oxford
- Devereux M, Griffith R y Simpson H (2004) The geographic distribution of production activity in the UK, *Regional Science and Urban Economics* 34: 533-564
- DTI (2001) Business clusters in the UK — A first assessment, Informe elaborado por Trends Business Research (<http://www.dti.gov.uk/clusters/map/>)
- Duranton G y Overman H (2005) Testing for localization using micro-geographic data, *Review of Economic Studies* 72: 1077-1106
- Ellison G y Glaeser E (1997) Geographic concentration in US manufacturing industries: A dashboard approach, *Journal of Political Economy* 105: 889-927
- Ellison G y Glaeser E (1999) The geographic concentration of industry: Does natural advantage explain agglomeration? *American Economic Review Papers and Proceedings* 89: 311-316

- Enright M (1996) Regional clusters and economic development: A research agenda. En U Staber, N Schaefer y B Sharma (eds) *Business networks: Prospects for regional development*, Walter de Gruyter, Berlín
- Enright M y Ffowcs-Williams I (2001) Local partnership, clusters and SME globalisation, *OCDE Working Paper*
- Feser E (1998) Old and new theories of industrial clusters. En M Steiner (ed) *Clusters and regional specialisation: On geography, technology and networks*, Pion Ltd, Londres
- Fratesi U (2005) Measuring and explaining localisation: Evidence from two British sectors, Mimeo, Univ Luigi Bocconi
- Fujita M y Krugman P (1995) When is the economy monocentric? Von Thünen and Chamberlin unified, *Regional Science and Urban Economics* 25: 505-528
- Gordon I y McCann P (2000) Industrial clusters: Complexes, agglomerations and/or social networks?, *Urban Studies* 37: 513-532
- Isard W, Azis I, Drennan M, Miller R, Saltzman S y Thorbecke E (1998) *Methods of Interregional and Regional Analysis*. Aldershot, Ashgate
- Karlsson C, Mellander C, Paulsson T (2005) A spatial ICT clusters in Sweden — An empirical method to identify necessary conditions for existence. In Johansson B, Karlsson C y Stough R *Entrepreneurship and dynamics in a knowledge economy*, Routledge, Londres
- Krugman P (1993) First nature, second nature and metropolitan location, *Journal of Regional Science* 34: 129-144
- Marcon E y Puech F (2003) Evaluating the geographic concentration of industries using distance-based methods, *Journal of Economic Geography* 3: 409-428
- Markusen A (1996) Sticky places in slippery space: A Typology of Industrial Districts, *Economic Geography*, 72: 293-313
- Marshall A (1890) *Principles of economics*. Macmillan, Londres
- Martin R y Sunley P (2003) Deconstructing clusters: Chaotic concept or policy panacea?, *Journal of Economic Geography* 3: 5-35
- Maurel F y Sédillot B (1999) A measure of the geographic concentration in French manufacturing industries, *Regional Science and Urban Economics* 29: 575-604
- OCDE (1999) *Boosting innovation: The cluster approach*. OCDE, París

- OCDE (2001) *Innovative clusters: Drivers of national innovation systems*. OCDE, París
- O'Donoghue D y Gleave B (2004) A note on methods for measuring industrial agglomeration, *Regional Studies* 38: 419-427
- Porter M (1990) *The competitive advantage of nations*. Macmillan, London
- Porter M (1998a) *On competition*. Harvard Business School Press
- Porter M (1998b) Clusters and the new economics of competition, *Harvard Business Review*, Nov/Dec
- Porter M (2003) The economic performance of regions, *Regional Studies* 37: 549-578
- Ripley B (1976) The second-order analysis of stationary point processes, *Journal of Applied Probability* 13: 255-266
- Ripley B (1977) Modelling spatial patterns, *Journal of the Royal Statistical Society B* 39: 172-212
- Roelandt T y Den Hertog P (1999) Cluster analysis and cluster-based policy making in OECD countries: An introduction to the theme. En *Boosting innovation: The cluster approach*. OCDE, París
- Rosenfeld S (1997) Bringing business clusters into the mainstream of economic development, *European Planning Studies* 5: 3-21
- Saxenian A (1994) *Regional networks: Industrial adaptation in Silicon Valley and Route 128*. Harvard University Press, Cambridge
- Schmitz H y Nadvi K (1999) Clustering and industrialization: Introduction, *World Development* 27: 1503-1514
- Scott A (1998) *Regions and the world economy*. Oxford University Press, Oxford
- Silverman B (1986) *Density estimation for statistics and data analysis*. Chapman and Hall, Nueva York
- Simmie J y Sennett J (1999) Innovation in the London metropolitan region. En D Hart, J Simmie, P Wood y J Sennet (Eds) *Innovative clusters and competitive cities in the UK and Europe*. Oxford Brookes School of Planning
- Steiner M (1998) *Clusters and regional specialization: On geography, technology and networks*. Pion Ltd, Londres
- Swann G y Prevezer M (1996) A comparison of the dynamics of industrial clustering in computing and biotechnology, *Research Policy* 25: 139-157

- Swann G, Prevezer M y Stout D (1998) *The dynamics of industrial clustering: International comparisons in computing and biotechnology*, Oxford University Press, Oxford
- Van der Berg L, Braun E y Van Winden W (2001) Growth clusters in European cities: an integral approach, *Urban Studies* 38: 186-206