
La localización de nuevos establecimientos industriales: estudios previos y evidencia para los municipios catalanes

40

Este artículo presenta una revisión de los estudios empíricos sobre localización industrial y una aplicación ilustrativa de los resultados que proporcionan. Se describen los dos modelos más utilizados: modelos de elección discreta (en los que el análisis recae sobre el agente decisor) y de recuento (focalizado en el territorio), siendo este último el que se toma como referencia en este estudio. Pero la ausencia de datos apropiados en España solo permite realizar inferencias basadas en estos últimos. Las estimaciones para los municipios catalanes indican que factores institucionales como la capitalidad de provincia son claramente relevantes, mientras que las economías de aglomeración y localización apenas lo son.

Artikulu horrek industriaren lokalizazioari buruzko berrazterketa bat agertzen du baita emandako emaitzen aplikazio adierazgarri bat ere. Gehien erabil diren ereduak deskribatzen dira: Hautaketa diskretuko ereduak (non garrantzia eragile erabakitzaileari dagokion) eta zenbatzekoak (lurraldeari dagokio garrantzia). Hain zuzen ere azken honetan hartzen da erreferentziatzat. Kataluniako zenbatespenek adierazten dute faktoreek, probintziaren hiriburutasunak, adibidez, garrantzi nabaria dutela lokalizatzeko eta metatzeko ekonomiek ordea kasik bat ere ez duten artean.

This paper reviews the empirical studies on industrial location and presents an illustration of the results they provide. Usual specifications in the literature are discrete choice models (in which the emphasis lies on the decision agent) and count data models (in which the emphasis lies on the territory). However, the lack of appropriate data in Spain only allows to make inferences from the latter. Estimates from the Catalan municipalities indicate that institutional characteristics such as being capital of a province are clearly relevant whereas agglomeration and location economies barely are.

ÍNDICE

1. Introducción
 2. Revisión de los estudios realizados
 3. Nuevos establecimientos industriales en Cataluña
 4. Conclusiones
- Referencias bibliográficas

N.º de clasificación JEL: C25, R10, R30

Palabras clave: ciudades, localización industrial

1. INTRODUCCIÓN

La entrada de empresas en un mercado es sin duda un fenómeno de gran relevancia económica, tanto cuantitativa como cualitativamente (Geroski, 1995). Los datos del Directorio Central de Empresas del INE, por ejemplo, indican que en España se crearon 374.094 empresas en el año 2005 (compárese con las 3.064.129 existentes, según estas mismas fuentes). Por otra parte, diversas teorías y la evidencia empírica procedente de distintos marcos institucionales demuestran que estos agentes juegan un papel fundamental en cuestiones tan importantes como la generación de empleo y la introducción de innovaciones (Audretsch, 1995). No resulta entonces sorprendente que su estudio constituya una

de las áreas más activas de la disciplina «Economía y Organización Industrial» (Caves, 1998; Segarra, 2002).

Un aspecto crítico en la decisión de crear una empresa es precisamente dónde hacerlo. «¿Cuál puede ser una buena localización para mi negocio?» es una pregunta a la que empresas y empresarios suelen dedicar profunda reflexión, conscientes de que una respuesta adecuada puede determinar el éxito o el fracaso de una nueva actividad productiva. Las administraciones públicas tratan en consecuencia de orientar estas decisiones y aumentar la probabilidad de éxito a través de iniciativas como las «incubadoras de empresas», la mejora de infraestructuras (suelo industrial, vías de comunicación, etc.) o las ayudas fiscales y financieras (Owens y Sarte, 2002). La localización de empresas y establecimientos de nueva creación supone por lo tanto una continua (re)asignación de sustanciales recursos públicos y privados. Sin embargo, nuestro conocimiento sobre qué determina que una empresa o establecimiento se localice en un área geográfica de-

* Este trabajo se ha beneficiado de la financiación pública ofrecida en el marco de las CICYT SEJ2004-07824/ECON y SEJ2004-05860/ECON del Ministerio de Educación y Ciencia así como del apoyo del CIDEM (Generalitat de Catalunya). Los errores que puedan aparecer son solo responsabilidad de los autores.

terminada, en detrimento de otras potenciales localizaciones, es todavía relativamente limitado (Hayter, 1997).¹

Este artículo pretende, en primer lugar, ofrecer una síntesis de este conocimiento a partir de la revisión de los principales estudios empíricos sobre localización industrial.² Para facilitar la exposición, se ha optado por agruparlos en torno a dos grandes aproximaciones metodológicas: la basada en modelos de elección discreta y la basada en modelos para datos de recuento.³ Los modelos de elección discreta permiten analizar la localización desde la óptica de la empresa que entra en el mercado, de manera que las investigaciones que los emplean suelen prestar particular atención al efecto que determinadas características individuales (dimensión y sector, por ejemplo) tienen en la decisión de entrada. Por su parte, los modelos de recuento permiten analizar la localización de nuevas empresas desde la óptica del espacio geográfico escogido (municipio, provincia, región, etc.), revelando así qué características del mismo (población, tejido industrial, infraestructuras,

etc.) afectan a la probabilidad de ser escogido para iniciar una actividad productiva.

En segundo lugar, este artículo pretende ofrecer una evidencia empírica de los determinantes de la localización industrial en España. Desgraciadamente, un estudio basado en un modelo de elección discreta resulta difícil de llevar a cabo en el caso español debido a la falta de datos apropiados. La fuente estadística oficial que proporciona información relativa a la localización de nuevas actividades productivas es el Registro de Establecimientos Industriales (REI), el cual tan solo ofrece datos individuales relativos al número de trabajadores, la inversión realizada y la potencia eléctrica instalada. Investigaciones previas han tratado de evitar los efectos de esta restricción combinando información relativa al establecimiento y al territorio (Arauzo y Manjón, 2004). No obstante, es evidente que el reducido número de variables explicativas hace que las inferencias así obtenidas deban ser analizadas con cierta cautela.

Estas limitaciones explican que en este trabajo se haya preferido explotar la dimensión territorial y circunscribir el análisis a la determinación de qué entornos territoriales estimulan de forma más clara la localización de nuevas actividades. Este planteamiento nos va a permitir analizar qué factores atribuibles bien a la iniciativa pública (dotación de infraestructuras, calificación del capital humano, etc.), bien a la iniciativa privada (economías de aglomeración, especialización o diversificación productiva), o a la mera disposición geográfica del territorio (distancia a capitales, a municipios más poblados, etc.) afectan al número esperado de establecimientos localizados en un determinado territorio durante cierto período de tiempo («la *ratio* de ocurrencia» de una distribución de Poisson). En concreto, el estudio explota la información relativa a la creación de nuevos establecimientos industriales en los nove-

¹ Quizás el hecho de que el estudio sistemático de las decisiones empresariales de localización no se inicie hasta finales de la década de los setenta (Carlton, 1979; Johnson y Cathcart, 1979) pueda explicar la falta de consenso académico sobre los determinantes de la localización empresarial.

² Hemos limitado nuestro interés a la industria porque los servicios constituyen un grupo mucho más heterogéneo y, comparativamente, menos estudiado (véase Holl, 2004a para un análisis conjunto de la industria y los servicios).

³ Véase, por ejemplo, Green (2000) para una introducción a estos modelos. Es evidente que existen estudios en este ámbito que emplean otras especificaciones econométricas, pero el interés en ese caso radica en el stock de empresas (o de trabajadores) más que en la propia decisión de entrada (véase, por ejemplo, Glaeser *et al.*, 1992 y Henderson *et al.*, 1995). En consecuencia, se ha optado por no incluir estos trabajos en nuestra particular revisión de la literatura dedicada a este tema.

cientos cuarenta y seis municipios catalanes durante el período 2001-2004. Los datos relativos a los establecimientos se obtuvieron del «Registre d'Establiments Industrials de Catalunya» (REIC), mientras que los relativos a los municipios proceden del «Institut d'Estadística de Catalunya» y de la base de datos construida por Trullén y Boix (2004).⁴

2. REVISIÓN DE LOS ESTUDIOS REALIZADOS

Las decisiones que toman las empresas al escoger el emplazamiento en el que realizarán sus actividades pueden analizarse desde dos perspectivas: la del agente que entra en el mercado y la del territorio donde se localiza ese agente. Como veremos inmediatamente, esta distinción es un tanto artificial, pero permite poner de manifiesto que, según la perspectiva que se adopte, la metodología econométrica que se requiere es diferente, como también lo es la información estadística a utilizar. A continuación se discuten estas diferencias y se revisan, por orden cronológico, los trabajos más destacados de cada perspectiva.

2.1. Estudios que emplean modelos de elección discreta

El cuadro n.º 1 contiene un sumario con las principales características de los estu-

dios que emplean modelos de elección discreta, incluyendo un detalle del ámbito territorial, el período, la desagregación sectorial, los factores determinantes de la localización y la especificación econométrica. Respecto a esta última, el punto de partida en estos estudios es la modelización de la probabilidad de que una empresa escoja un determinado territorio como localización para un nuevo establecimiento (su sede social si la empresa solo dispone de un establecimiento) de entre un conjunto de territorios que pueden constituir una localización alternativa. Esta probabilidad puede analizarse empleando diferentes especificaciones, si bien prácticamente todas derivan, de una u otra manera, del modelo de elección discreta propuesto originalmente por McFadden (1974).

No obstante, en la práctica estas especificaciones difieren en el tipo de datos que requieren para su estimación (y en la interpretación de los coeficientes del modelo). Por un lado, en un modelo *logit* condicional esta probabilidad es una función no lineal de las características del territorio que potencialmente van a afectar a los beneficios esperados de la elección. Esto significa que en este tipo de especificaciones las variables a tener en cuenta corresponden a medidas referidas a las localizaciones (no a los agentes).⁵ Por otro lado, en un modelo *logit* multinomial tan solo se tendrían en cuenta las características del agente entrante (y no las de la localización escogida). Finalmente,

⁴ El estudio realizado en este trabajo podría extenderse fácilmente al resto de municipios españoles, pero entonces no tendríamos de la riqueza informativa que proporciona esta base de datos para los municipios catalanes. La falta de datos adecuados para realizar este tipo de estudios en España no se limita únicamente a la variable dependiente, sino, desgraciadamente, también a las explicativas. Este es un problema crónico al que se enfrentan los investigadores interesados en el caso español y que, entre otras cuestiones, complica la obtención de evidencia empírica sólida y dificulta la realización de comparativas internacionales.

⁵ Esta caracterización se mantiene en lo esencial en modelos en los que las elecciones se efectúan en diferentes niveles o ramas, como en el caso del *logit* anidado. En este tipo de modelos, el agente escoge en primer lugar las localizaciones correspondientes a un nivel de agregación geográfica (por ejemplo, una provincia) y posteriormente escoge entre las localizaciones de un nivel de agregación geográfica inferior, pertenecientes a la localización escogida en el primer nivel (por ejemplo, los municipios pertenecientes a la provincia escogida en primer lugar).

Cuadro n.º 1

Estudios que emplean modelos de elección discreta

Estudio	Ámbito territorial	Período	Desagregación sectorial	Factores determinantes de la localización	Especificación econométrica
Arauzo y Manjón (2004)	Comarcas y provincias (Cataluña)	1987-1996	Sector industrial (agregado)	Dimensión del establecimiento e inversión realizada, economías y deseconomías de urbanización y densidad poblacional	Logit condicional y logit multinomial
Bartik (1985)	Estados (EE.UU.)	1972-1978	Sector industrial (agregado)	Tasa de filiación sindical, niveles impositivos y provisión de servicios públicos	Logit condicional y logit anidado
Baudewyns <i>et al.</i> (2000)	Comunas de Bruselas y arrondissements de Valonia (Bélgica)	1981-1991 y 1990-1994	Sectores industriales y de servicios	Infraestructuras de transporte y economías de aglomeración	Logit condicional
Baudewyns (1999)	Comunas de Bruselas (Bélgica)	1981-1991	Sectores industriales y de servicios	Infraestructuras de transporte y economías de aglomeración	Logit multinomial
Carlton (1979, 1983)	SMSA (EE.UU.)	1967-1971	3 sectores industriales	Dimensión de la empresa, coste de la energía y niveles de ocupación. (Poca incidencia del nivel impositivo y de los programas públicos de apoyo a las nuevas empresas)	Logit condicional
Coughlin <i>et al.</i> (1991)	Estados (EE.UU.)	1981-1983	Sector industrial (agregado)	Renta per cápita, economías de aglomeración, tasa de paro, sindicación, infraestructuras de transporte y promoción económica	Logit condicional
Friedman <i>et al.</i> (1992)	Estados (EE.UU.)	1977-1988	Sector industrial (agregado)	Dimensión del mercado, infraestructuras de transporte, promoción económica y nivel impositivo	Logit condicional

.../...

Cuadro n.º 1 (continuación)
Estudios que emplean modelos de elección discreta

Estudio	Ámbito territorial	Período	Desagregación sectorial	Factores determinantes de la localización	Especificación econométrica
Guimarães <i>et al.</i> (2000)	Concelhos (Portugal)	1985-1992	Sector industrial (agregado)	Economías de aglomeración	Logit condicional
Hansen (1987)	Estado de São Paulo (Brasil)	1977-1979	Sector industrial (agregado)	Economías de urbanización, economías de localización, distancia al centro del estado	Logit anidado
Head <i>et al.</i> (1995)	Estados (EE.UU.)	1980-1989	Sector industrial (4 dígitos)	Economías de localización y economías de aglomeración en general, vinculaciones de tipo input-output	Logit condicional
Smith y Florida (1994)	Condados (EE.UU.)	-	Sectores industriales vinculados a la industria del automóvil	Vinculaciones de tipo input-output. Economías de localización en las manufacturas, concentraciones poblacionales, mano de obra calificada, infraestructuras de transporte	Logit condicional
Woodward (1992)	Condados y estados (EE.UU.)	1980-1989	Sector industrial (agregado)	Diferencias entre estados (tamaño de mercado y sindicación) y condados (capital humano, economías de aglomeración y tasa de paro)	Logit condicional

Fuente: Elaboración propia.

también existen especificaciones que combinan ambos tipos de datos, como por ejemplo el denominado modelo logit mixto o el logit condicional con productos cruzados entre características de las empresas y variables ficticias (*dummies*) territoriales para representar la clasificación territorial de los datos.

Los modelos logit condicional fueron aplicados por primera vez al estudio de la localización industrial por Carlton (1979, 1983). A partir de los resultados obtenidos por McFadden (1974), estos trabajos muestran que la maximización de una función de beneficios lineal en el vector de explicativas x_k , y en un término de perturbación con distribución i.i.d. (idéntica e independiente distribución) de tipo Weibull, es coherente con que la probabilidad de que la empresa i escoja la localización j para su nueva actividad productiva venga dada por la siguiente expresión:

$$Pr(j) = \frac{\exp \sum_K \beta_K \ln \chi_K(j)}{\sum_{m=1}^M \exp \sum_K \beta_K \ln \chi_K(m)} \quad (1),$$

siendo $m = 1, \dots, M$ el total de localizaciones posibles y β un vector de parámetros a estimar. Posteriormente, Bartik (1985) extendió la propuesta básica de D.W. Carlton en dos direcciones. En primer lugar, al emplear un modelo logit anidado relajó el supuesto de independencia de las alternativas irrelevantes (véase también Hansen, 1987).⁶

⁶ Este supuesto implica la sustitución proporcional entre alternativas (la *ratio* entre las probabilidades asociadas a dos alternativas es independiente de las probabilidades asociadas a las otras alternativas). Los modelos logit condicional y multinomial, por ejemplo, requieren de este supuesto para su construcción, mientras el logit anidado (o cualquier otro modelo «de valor extremo generalizado») y el logit mixto no imponen este supuesto.

Esto le permitió considerar la existencia de una correlación ρ entre los términos de perturbación asociados a los estados pertenecientes a una misma región a partir de la siguiente expresión de la probabilidad de que se escoja el estado s si ha escogido la región r :

$$Pr(s|r) = \frac{\exp[(1/\rho) \beta' \chi_K]}{\sum_{j \in r} \exp(\beta' \chi_j + D'R_j)} \quad (2),$$

siendo R un vector de variables *dummies* regionales y D un vector de efectos región constantes.

La segunda contribución del trabajo de Bartik (1985) consiste en resolver la imposibilidad práctica de estimar un modelo con todas las alternativas a las que realmente se enfrentan las empresas. La solución a esta restricción computacional pasa por agrupar las alternativas en grupos homogéneos, de forma que lo que se acaba estimando es la probabilidad de escoger uno de esos grupos (h) de entre todos los grupos (g) existentes:

$$Pr(h) = \frac{\exp[\beta' \bar{\chi}_h + (1-w) \ln N_h]}{\sum_g \exp[\beta' \bar{\chi}_g + (1-w) \ln N_g]} \quad (3),$$

siendo $\bar{\chi}_g$ el vector de características medias de todas las alternativas en el interior del grupo, w la correlación del término de perturbación entre todas las alternativas en el interior de un grupo, N_h el número de alternativas en el interior del grupo y N_g el número de grupos. En particular, los grupos escogidos en este trabajo para relacionar las alternativas iniciales a nivel local son los estados estadounidenses.

A partir de estas contribuciones seminales, las aplicaciones posteriores se centran,

especialmente durante la década de los noventa, en el caso particular que constituye la inversión directa extranjera (IDE). Así, por ejemplo, Coughlin *et al.* (1991), Friedman *et al.* (1992), Woodward (1992), Smith y Florida (1994) y Head *et al.* (1995) analizan el caso estadounidense empleando modelos logit condicional. Más recientemente y para Portugal, Guimarães *et al.* (2000) y Figueiredo *et al.* (2002) presentan aplicaciones análogas pero con la ventaja que supone disponer de un reducido nivel de agregación territorial (los *concelhos*). Tal y como argumentan estos autores, el trabajar con ámbitos territoriales más extensos y heterogéneos debería incidir negativamente en las propiedades estadísticas de los estimadores de máxima verosimilitud empleados.⁷

A la vista de esta evolución en la aplicación de estos modelos, la conclusión que cabe extraer es que la utilización de sencillos modelos de elección discreta para el estudio de la localización industrial es un recurso metodológico prácticamente agotado. Por el contrario, el uso de modelos más sofisticados (como, por ejemplo, el logit mixto) con métodos de estimación alternativos (por ejemplo, el método generalizado de momentos) basados en la simulación de probabilidades está todavía poco explorado (Train 2003). Por lo tanto, es de esperar que esta sea una vía por la que discurrirán futuras investigaciones en esta área.

2.2. Estudios que emplean modelos para datos de recuento

El cuadro n.º 2 contiene un sumario de los estudios sobre localización industrial

que emplean modelos para datos de recuento. Este tipo de modelos, basados en la distribución de Poisson y distribuciones relacionadas (por ejemplo, la distribución binomial negativa), permiten analizar los determinantes del número esperado de nuevas empresas o establecimientos creados en un determinado emplazamiento (por ejemplo, un municipio) por unidad de tiempo (por ejemplo, un año), por lo que constituyen el recurso econométrico natural cuando lo que se pretende es analizar la localización desde la óptica del territorio escogido. Además, presentan el atractivo de que el modelo de decisión implícito en su formulación es análogo al de los modelos de elección discreta.

Así, asumiendo que el criterio decisor se basa en la maximización de beneficios (esperados) y suponiendo que los mismos contienen un término estocástico i.i.d., la probabilidad de escoger el emplazamiento j por parte de la empresa i puede expresarse en términos de una variable aleatoria discreta que recoja el resultado de esta elección. Esta variable es la base sobre la que se construyen los modelos de elección discreta (Guimarães *et al.*, 2003). No obstante, combinando estas decisiones individuales dicha probabilidad podría pasar fácilmente a referirse al número de entradas efectuadas en un determinado territorio y período de tiempo, una variable aleatoria cuya distribución sería del tipo Poisson (o una distribución análoga).

Matemáticamente, si y_j es una realización de esta variable aleatoria gobernada por una distribución de Poisson de parámetro μ (la *ratio* de ocurrencia del evento de interés), entonces la función de densidad, dado un vector de variables explicativas x_k , vendrá dada por:

⁷ Sobre este punto véase también Arauzo y Manjón (2004).

Cuadro n.º 2

Estudios que emplean modelos de recuento

Estudio	Ámbito territorial	Período	Desagregación sectorial	Factores determinantes de la localización	Especificación econométrica
Arauzo (2005)	Municipios (Cataluña)	1987-1996	5 sectores industriales	Economías y deseconomías de urbanización, diversidad industrial, intensidad de commuting, porcentaje de ocupación industrial, porcentaje de ocupación en servicios y distancia a las capitales de comarca y principales ciudades	Poisson
Arauzo y Manjón (2004)	Municipios y comarcas (Cataluña)	1987-1996	Sector industrial (agregado)	Economías y deseconomías de urbanización, diversidad industrial, densidad poblacional, capital humano y economías de localización	Poisson
Basile (2004)	Provincias (Italia)	1998-1999	Sector industrial (agregado)	Dummies regionales, economías de aglomeración, demanda, asimetrías informativas, infraestructuras, costes laborales unitarios y tasa de paro	Poisson, binomial negativa con efectos aleatorios, Poisson alterado en el cero
Cie_Ilik (2005)	Provincias (Polonia)	1993-1998	Conjunto de la economía	Dummies estatales, PIB, nivel salarial, capital humano, tasa de paro, infraestructuras viarias, ferroviarias y portuarias	Binomial negativa
Coughlin y Segev (2000)	Condados (California, EE.UU.)	1989-1994	Sector industrial (agregado)	Dummies regionales y de dimensión poblacional, capital humano, nivel salarial, nivel impositivo, infraestructuras viarias, población de color	Binomial negativa
Holl (2004a)	Municipios (Portugal)	1986-1997	12 sectores industriales 9 sectores de servicios y construcción	Población, infraestructuras viarias, diversidad sectorial, especialización sectorial, nivel salarial y capital humano	Binomial negativa con efectos fijos

.../...

Cuadro n.º 2 (continuación)
Estudios que emplean modelos de recuento

Estudio	Ámbito territorial	Período	Desagregación sectorial	Factores determinantes de la localización	Especificación econométrica
Holl (2004b)	Municipios (España)	1980-1994	10 sectores industriales	Infraestructuras viarias, accesibilidad a los proveedores, accesibilidad a la demanda intrarregional e interregional	Poisson condicional con efectos Fijos
Holl (2004c)	Municipios (Portugal)	1986-1997	Sector industrial (agregado)	Población, infraestructuras viarias, diversidad sectorial, nivel salarial y capital humano	Poisson con efectos fijos, y binomial negativa con efectos fijos
List (2001)	Condados (California, EE.UU.)	1983-1992	Sector industrial (agregado) y sectores contaminantes y no contaminantes	Economías de aglomeración, dimensión del mercado, accesibilidad y superficie del condado	Poisson y Poisson alterado en el cero
List y McHome (2000)	Condados (Nueva York, EE.UU.)	1980-1990	7 sectores industriales	Regulación medioambiental, nivel salarial y nivel impositivo	Poisson condicional para datos de panel
Wu (1999)	Área metropolitana de Guangzhou (China)	1981-1991	Conjunto de la economía	Distancia al distrito central de negocios, accesos a la población, mercado de trabajo y existencia de hoteles de alto standing	Poisson y binomial negativa

Fuente: Elaboración propia.

$$P(y_j = 0, 1, 2, \dots | x_k) = \frac{e^{-\mu_j} \mu_j^{y_j}}{y_j!} \quad (4),$$

siendo la esperanza condicional de la variable $E[y_j | x_k] = \mu_j = e^{\beta' x_k} = Var[y_j | x_k]$.

El uso de modelos de recuento permite subsanar dos importantes limitaciones de los modelos de elección discreta. En primer lugar, las localizaciones no escogidas tienen una contribución nula a la función de verosimilitud. Esto significa que el hecho de que un territorio no reciba ninguna entrada es una información irrelevante para los modelos de elección discreta. En cambio, en los modelos de recuento es posible analizar de manera explícita (e incluso en cierta manera independiente) los casos en los que la variable de interés toma valor nulo (obsérvese que $y_j = 0$ corresponde a territorios en los que no se ha localizado ninguna entrada). En segundo lugar, los modelos de elección discreta presentan dificultades computacionales cuando el número de alternativas es elevado y el proceso de cálculo del estimador emplea métodos numéricos para hallar el máximo de la función de verosimilitud. Estas dificultades pueden convertirse en insalvables cuando se manejan datos geográficos desagregados (por ejemplo, en el ámbito municipal), mientras que en los modelos de recuento el número de territorios analizados (= tamaño muestral = número de alternativas en los modelos de elección discreta) no suele suponer un obstáculo para la estimación.⁸

Estas características seguramente han incidido en que las contribuciones más recientes en esta área empleen mayoritariamente este tipo de especificaciones. Así, entre dichos trabajos podemos destacar los que se centran en la inversión directa extranjera, como los de List (2001) y Coughlin y Segev (2000), para los condados (*counties*) de California, Cieřlik (2005) para las provincias polacas, Basile (2004) para las provincias italianas o Wu (1999) para el caso de la ciudad china de Guangzhou. Entre los trabajos que no tienen en cuenta la procedencia de la empresa entrante tenemos los de Arauzo (2005) y Arauzo y Manjón (2004) para los municipios catalanes, Holl (2004b) para los municipios españoles, Holl (2004a y 2004c) para los municipios portugueses o List y McHome (2000) para los condados (*counties*) del estado de New Cork (EE.UU.).

La conclusión que cabe extraer de la utilización de este tipo de modelos es que se ha pasado, progresivamente, de las estimaciones conceptualmente más sencillas, basadas en modelos de Poisson, a otras más realistas (y, de hecho, más acordes con las características de la realidad objeto de estudio) donde son más habituales los modelos binomiales negativos o inflados en ceros. Asimismo, la creciente disponibilidad de paneles de datos ha permitido el empleo de estimadores con mejores propiedades estadísticas, subsanando de este modo algunas de las deficiencias asociadas a la explotación de secciones cruzadas (mejor control de la heterogeneidad inobservable, mayor eficiencia).

3. NUEVOS ESTABLECIMIENTOS INDUSTRIALES EN CATALUÑA

3.1. Análisis descriptivo: tamaño, sector y entorno urbano

El fenómeno de la localización puede analizarse descriptivamente desde diver-

⁸ En el caso de los municipios catalanes, por ejemplo, 946 alternativas es una cifra imposible de gestionar para los modelos (logit multinomial, logit condicional, logit anidado, etc.) y *softwares* (Gauss, Stata, LIMDEP, etc.) más comunes. Una solución, propuesta por Guimarães *et. al.* (2003), es recurrir a un modelo de Poisson para, bajo determinadas condiciones, obtener estimaciones idénticas a las de un modelo logit condicional (véase también Arauzo y Manjón 2004).

sos puntos de vista. No obstante, los datos proporcionados por el REIC tan solo permiten explorar dos dimensiones básicas: tamaño (número de trabajadores) y sector (dos dígitos de la CNAE-93). Respecto al primero, el Panel A del cuadro n.º 3 muestra que la entrada anual de nuevos establecimientos industriales en Cataluña presenta una cierta estabilidad durante el período 2001-2004. En cambio, la creación de ocupación asociada a estas actividades productivas es muy variable: de los 5.120 puestos de trabajo generados en el año 2001 a los 18.892 del año 2003. Esta variabilidad provoca que la dimensión media de los nuevos establecimientos industriales se mueva en estos

años entre los cinco y los dieciséis trabajadores.

Sectorialmente, las cifras del cuadro n.º 4 indican que los nuevos establecimientos replican en lo sustancial la distribución del tejido industrial catalán. En particular, destaca el peso de la «Fabricación de productos metálicos, excepto maquinaria y equipo»: alrededor de uno de cada cuatro nuevos establecimientos industriales pertenece a este sector. Estos establecimientos, además, están entre los mayores generadores de empleo (en torno a un 20% del total). De hecho, si les sumamos los pertenecientes a la «Industria de la madera y del corcho, excepto muebles; cestería y espartería» y a la «Industria de la construcción de

Cuadro n.º 3
Entrada de establecimientos industriales y creación de empleo en Cataluña 2001-2004

Panel A: Datos Anuales.

Año	Entrantes	Entrantes	Empleos por establecimiento
2001	915	5.120	5,59
2002	1.117	10.011	8,96
2003	1.153	18.892	16,38
2004	1.088	11.509	10,57
Total	4.273	45.532	10,65

Panel B: Estadísticos descriptivos (total período 2001-2004).

Medida	Desviación típica	Mínimo	Máximo
1,12	4,14	0	70

Fuente: Elaboración propia a partir del REIC.

maquinaria y equipo mecánico», obtenemos casi la mitad de los empleos generados por el conjunto de los establecimientos industriales durante el período 2001-2004.

De la misma manera, dado que en el REIC se indica el municipio de destino de los nuevos establecimientos, estas dimensiones básicas (tamaño y sector) pueden cruzarse con otras variables que definan el tipo de entorno urbano, como por ejemplo la dimensión del municipio. Este ejercicio nos permite observar la estrecha relación existente entre la dimensión del municipio receptor y la del nuevo establecimiento.⁹ El Panel A del cuadro n.º 5 muestra como los establecimientos de menor dimensión tienden a localizarse en los municipios más pe-

queños, a excepción de la ciudad de Barcelona. Sin embargo, el grueso de la ocupación generada en esos municipios no corresponde a estos establecimientos de menor dimensión (Panel B del cuadro n.º 5). Por el contrario, son los establecimientos de más de noventa y nueve trabajado-

⁹ También parece existir cierta correlación entre la dimensión del municipio y la especialización sectorial de los nuevos entrantes. Así, por ejemplo, los municipios más pequeños parecen ser la localización preferida para establecimientos de sectores con un nivel tecnológico bajo (de acuerdo con la clasificación que al respecto realiza la Organización para la Cooperación y el Desarrollo Económicos, OCDE), mientras que los municipios más grandes, especialmente Barcelona, son los que absorben la mayoría de los establecimientos de sectores tecnológicamente avanzados.

Cuadro n.º 4

**Entrada de establecimientos industriales en Cataluña por sector de actividad
2001-2004**

CNAE-93	Establecimientos entrantes	Empleos	Empleos por establecimiento
10- Extracción y aglomeración de antracita, hulla, lignito y turba	1	1	1,0
14- Extracción de minerales no metálicos ni energéticos	13	49	3,8
15- Industria de productos alimenticios y bebidas	306	1.989	6,5
16- Industria del tabaco	1	3	3,0
17- Industria textil	256	2.571	10,0
18- Industria de la confección y de la peletería	284	2.106	7,4
19- Preparación, curtido y acabado del cuero; fabricación de artículos de marroquinería y viaje; artículos de guarnicionería, talabartería y zapatería	13	87	6,7
20- Industria de la madera y del corcho, excepto muebles; cestería y espartería	287	9.436	32,9

.../...

Cuadro n.º 4 (continuación)

**Entrada de establecimientos industriales en Cataluña por sector de actividad
2001-2004**

CNAE-93	Establecimientos entrantes	Empleos	Empleos por establecimiento
21- Industria del papel	82	743	9,1
22- Edición, artes gráficas y reproducción de soportes grabado	232	2.072	8,9
23- Coquerías, refino de petróleo y tratamiento de combustibles nucleares	1	4	4,0
24- Industria química	138	1.665	12,1
25- Fabricación de productos de caucho y materias plásticas	180	2.112	11,7
26- Fabricación de otros productos minerales no metálicos	159	1.312	8,3
27- Metalurgia	33	323	9,8
28- Fabricación de productos metálicos, excepto maquinaria y equipo	1.192	9.028	7,6
29- Industria de la construcción de maquinaria y equipo mecánico	450	4.250	9,4
30- Fabricación de máquinas de oficina y equipos informáticos	1	20	20,0
31- Fabricación de maquinaria y material eléctrico	125	2.270	18,2
32- Fabricación de material electrónico; fabricación de equipos y aparatos de radio, televisión y comunicaciones	44	739	16,8
33- Fabricación de equipos e instrumentos médico-quirúrgicos, de precisión, óptica y relojería	55	983	17,9
34- Fabricación de vehículos de motor, remolques y semirremolques	94	1.554	16,5
35- Fabricación de otro material de transporte	35	309	8,8
36- Fabricación de muebles; otras industrias manufactureras	291	1.906	6,5
TOTAL	4.273	45.532	10,7

Fuente: Elaboración propia a partir del REIC.

Cuadro n.º 5

Entrada de establecimientos industriales y creación de empleo en Cataluña por dimensión del establecimiento y del municipio receptor 2001-2004

Panel A: Número de entrantes		Dimensión del municipio (habitantes)					TOTAL
		Menos de 2.000	De 2.000 a 10.000	De 10.001 a 50.000	De 50.001 a 100.000	De 100.001 a 1.000.000	
Dimensión del establecimiento (trabajadores)							
Menos de 10	341	862	917	320	591	206	3.237
Entre 10 y 49	87	238	249	93	180	45	892
Entre 50 y 99	4	17	15	5	9	4	54
Más de 99	2	6	5	4	6	1	24
Sin datos	12	15	18	7	12	2	66
TOTAL	434	1.123	1.186	422	786	256	4.273

Panel B: Creación de empleo		Dimensión del municipio (habitantes)					TOTAL
		Menos de 2.000	De 2.000 a 10.000	De 10.001 a 50.000	De 50.001 a 100.000	De 100.001 a 1.000.000	
Dimensión del establecimiento (trabajadores)							
Menos de 10	1.317	3.584	3.881	1.344	2.507	726	13.359
Entre 10 y 49	1.736	4.316	4.592	1.680	3.172	882	16.378
Entre 50 y 99	221	1.061	1.070	307	629	271	3.559
Más de 99	7.900	1.208	929	1.364	694	141	12.236
Sin datos	0	0	0	0	0	0	0
TOTAL	11.174	10.169	10.472	4.695	7.002	2.020	45.532

Fuente: Elaboración propia a partir del REIC.

res los que generan unas dos terceras partes de esos empleos. Curiosamente, en el resto de municipios la situación es la inversa, siendo los establecimientos de menos de 50 trabajadores los que cargan con el grueso de la contratación.

3.2. Modelización econométrica

Los descriptivos presentados ofrecen sin duda algunos resultados interesantes, pero conviene subrayar que carecen de capacidad explicativa. Explorar la existencia de relaciones de causalidad entre las decisiones de localización y las características de los territorios en las que se materializan requiere de algún tipo de análisis econométrico. En particular, si nuestro interés se limita a analizar las «respuestas medias» entonces nuestro objetivo se concreta en la estimación de la esperanza del número de establecimientos industriales creados en un municipio y período determinado, condicionada por un conjunto de factores explicativos. En definitiva, nuestro objetivo es estimar $E[y_j | x_k]$ para los municipios catalanes en el período 2001-2004.¹⁰

Especificaciones

La propuesta inicial para la estimación de esta esperanza condicional podría ser un modelo de Poisson como el descrito en la sección anterior. No obstante, esta especificación no tiene en cuenta la existencia de factores no observables (historia, reputación, redes de contactos informales, etc.) en el ámbito municipal que pueden incidir

en la relación entre la dependiente y las explicativas. Una forma de incorporar al modelo componentes estocásticos inobservables (e independientes de x) es recurrir al denominado Poisson mixto o generalizado. Denotando con la letra v estos componentes y asumiendo una vez más una forma multiplicativa y una linealidad en el vector de variables explicativas, la especificación resultante sería:

$$E(y | x, v) = \mu(x, \beta) v = \exp(x\beta) v \quad (5).$$

La introducción de heterogeneidad inobservable en el modelo de Poisson tiene dos consecuencias interrelacionadas. En primer lugar, el modelo mixto presenta «sobredispersión», es decir, la variancia condicional de la variable dependiente es mayor que su esperanza condicional. En segundo lugar, la proporción de ceros que genera la correspondiente distribución mixta de y es más grande que la que genera la distribución de Poisson. Como se puede apreciar en Panel B del cuadro n.º 3, las entradas en los municipios catalanes parecen ajustarse bien a ambas características, lo que plantea dudas sobre el supuesto de que los datos hayan sido generados por una distribución de Poisson y apunta la necesidad de considerar distribuciones alternativas como base para nuestras especificaciones.¹¹

Una de las más utilizadas es la distribución mixta (continua) Poisson-Gamma, también conocida como distribución binomial negativa. Si suponemos que $\exp(v)$ sigue

¹⁰ El detalle técnico de los elementos estadísticos y econométricos discutidos en esta sección puede hallarse en Cameron y Trivedi (1998).

¹¹ Dado que si el modelo tiene constante la esperanza condicional no dista mucho de la media muestral y puesto que la variancia condicional suele ser inferior a la no condicional, los descriptivos del cuadro n.º 3 apuntan claramente a la existencia de sobredispersión en los datos. Además, es fácil comprobar que la frecuencia de ceros en nuestra variable dependiente (en torno al 70% de la muestra) es mucho más elevada que la esperada si los datos hubieran sido efectivamente generados por una distribución de Poisson.

una distribución Gamma con esperanza unitaria y varianza α , se puede demostrar que la variable y tendrá como esperanza condicional β (igual que antes) y como varianza $\mu + \alpha\mu^2$. Por lo tanto, una manera de contrastar estadísticamente la existencia de sobredispersión es a partir de la hipótesis nula de que $\alpha = 0$. Este contraste se puede llevar a cabo utilizando, por ejemplo, la *ratio* de verosimilitud como estadístico de prueba. Este estadístico lo denotaremos como «LR test α » en los cuadros de resultados.

Una modelización alternativa, que también permite recoger la sobredispersión y el exceso de ceros, es la basada en modelos mixtos (finitos) construidos sobre distribuciones discretas. Este es el caso de los denominados modelos inflados de cero, en los que las observaciones que toman el valor cero se suponen generadas por una distribución diferente que la generadora de las observaciones estrictamente positivas. Distribuciones habitualmente utilizadas para modelizar los ceros son la logística (modelo logit) y la normal (modelo probit). Respecto a las observaciones positivas, la distribución habitual es la Poisson, pero también se puede recurrir a la binomial negativa para introducir heterogeneidad inobservable en esta parte del modelo.¹²

El problema práctico al que nos enfrentamos es que no existe una forma simple de distinguir entre estos dos mecanismos

generadores de sobredispersión y exceso de ceros.¹³ Por un lado, el modelo binomial negativo (BINNEG) controla por heterogeneidad inobservable pero supone que el proceso generador es igual para los ceros y las observaciones positivas. Por otro, el modelo de Poisson inflado (POISINF) diferencia estos procesos generadores, pero no controla por heterogeneidad inobservable, algo que sí hace el modelo binomial negativo «inflado». No obstante, la interpretación de los parámetros en modelos «inflados», especialmente en lo que respecta a las variables que aparecen en las dos partes del modelo, no siempre resulta fácil. Además, incluir dos mecanismos generadores en la especificación supone emplear un modelo poco eficiente que da lugar a resultados altamente indeterminados.¹⁴

El estudio de las medidas de ajuste (denotado por «Pseudo-R²» en los cuadros de resultados), los contrastes de significación conjunta («LR test») y los logaritmos de la función de verosimilitud («Log L») pueden ayudar a discriminar entre especificaciones alternativas. No obstante, a menudo sus resultados no son concluyentes, como tampoco lo son los del procedimiento de contraste propuesto por Vuong (1989). Este contraste para modelos no anidados (Vuong Test) permite en principio discriminar entre el modelo de Poisson y el Binomial Negativo y sus respectivos análogos inflados. Desgraciadamente, el contraste

¹² El estimador de máxima verosimilitud del modelo de Poisson es coherente en tanto en cuanto la esperanza condicional de la variable dependiente esté correctamente especificada. Esta coherencia puede verse por lo tanto afectada por la presencia de heterogeneidad inobservable o la existencia de diferentes distribuciones para ceros y observaciones positivas. La sobredispersión, en cambio, no afecta a la coherencia del estimador, pero sí a su eficiencia (los errores estándares son en general más grandes que los obtenidos empleando la binomial negativa o un modelo inflado).

¹³ De hecho, todavía podría existir una tercera causa explicativa de estos efectos. Uno de los supuestos implícitos en el uso del proceso de Poisson es que los eventos (la creación de un establecimiento industrial) son independientes. Si este supuesto no se cumple (situación plausible en el contexto de la localización), el proceso mostraría sobredispersión.

¹⁴ De hecho, nuestros intentos de estimar el modelo binomial negativo inflado no lograron alcanzar la requerida convergencia del correspondiente algoritmo de maximización.

presenta áreas de indeterminación en las que no es posible concluir si se rechaza o no la hipótesis nula.

Una última cuestión que complica la modelización econométrica de los determinantes de la localización es la disponibilidad de datos en forma de panel, es decir, de observaciones repetidas de los mismos individuos en momentos temporales equidistantes. La explotación de esta estructura de los datos puede suponer una mejora en la eficiencia de nuestras estimaciones y nos permitirá introducir heterogeneidad inobservable en el modelo de una forma alternativa a la que plantea la binomial negativa (por lo que, en cierta manera, nos permitirá evaluar la posibilidad de un error de especificación en la distribución de ν). No obstante, para el corto período de tiempo analizado en este trabajo, el supuesto de que la *ratio* de ocurrencia es constante no resulta del todo inadmisibles. Bajo este supuesto, nuestra variable dependiente corresponderá a las entradas totales del período 2001-2004 y nuestras explicativas podrían estar calculadas como medias del período. Por otra parte, como la mayoría de estas variables están referidas al año 2001, otra posibilidad es restringir la estimación a las entradas de ese año. En todo caso, ambas aproximaciones consideran los datos como una sección cruzada y pueden explorarse empleando las especificaciones descritas.

No obstante, si consideramos que los datos tienen una estructura de panel, necesitamos adaptar nuestros modelos a la presencia de efectos individuales.¹⁵ Asumiendo que los regresores son exógenos y que los efectos individuales se incorporan al modelo en forma multiplicativa, se plantea la elección básica entre un estimador de «efectos fijos» (basado en la existencia de un estadístico suficiente para los efectos individuales) y un estimador de «efectos va-

riables» (que sigue una distribución Gamma). Ambos estimadores son coherentes bajo el supuesto de independencia entre regresores y efectos individuales (aunque la coherencia del estimador de efectos fijos no depende de este supuesto), pero el de efectos aleatorios es más eficiente (si bien incoherente si no existe independencia). Cabe entonces recurrir a un *test* de Hausman para contrastar la hipótesis nula de independencia y decidir en consecuencia qué estimador es el más adecuado para nuestros modelos. Este estadístico aparece como «Hausman Test» en los cuadros de resultados.¹⁶

La elección de las variables explicativas

En principio, una vez escogida la especificación y el método de estimación más adecuados, la teoría económica debería proporcionar los elementos definitorios del vector de variables explicativas. Desgraciadamente, no existe una teoría generalmente aceptada del fenómeno de la localización industrial y sus determinantes. Así, tras revi-

¹⁵ En concreto, los resultados presentados en este trabajo se obtuvieron empleando un modelo de Poisson. Esto significa imponer equidispersión, pero el uso de la binomial negativa para corregir esta restricción supone controlar por heterogeneidad inobservable en una estructura de datos que ya permite tener en cuenta este aspecto. Existen ejemplos en la literatura en los que la binomial negativa resultó ser la distribución más adecuada (Hausman *et al.*, 1984), pero en nuestro caso, igual que ocurría al emplear los datos en forma de sección cruzada, los algoritmos no lograron alcanzar la convergencia.

¹⁶ Independientemente de los valores que tome el estadístico, existen otros aspectos a valorar. Por ejemplo, el hecho de que si se estima el modelo por efectos fijos, los coeficientes de las variables sin variación temporal no estén identificados. Dado que muchas de las variables explicativas en este estudio son de este tipo hemos creído interesante presentar los resultados para los dos estimadores propuestos (efectos fijos y aleatorios).

sar la literatura pertinente, Hayter (1997) concluye que se pueden distinguir tres aproximaciones básicas: la neoclásica, la conductista y la institucional.¹⁷ La denominada aproximación neoclásica plantea un marco de decisión optimizadora, en el que las pautas de localización están gobernadas por la maximización/minimización de beneficios/costes. La mayoría de los trabajos que hemos comentado en las páginas precedentes se encuadran dentro de esta aproximación. Por su parte, la aproximación conductista cuestiona los supuestos implícitos de información completa y certeza sobre el entorno de la aproximación neoclásica¹⁸ e introduce la posibilidad de que factores no económicos, ligados al perfil personal del emprendedor, puedan incidir en las decisiones de localización (Figueiredo *et al.*, 2002). Finalmente, para la aproximación institucional los planteamientos de la aproximación conductista (y, hasta cierto punto, también los de la neoclásica) son excesivamente reduccionistas. Sin tratar de minimizar la importancia que tienen las características de la empresa o el empresario como sujetos decisores, a la hora de analizar las pautas de localización, la aproximación institucional defiende la necesidad de considerar el marco institucional

¹⁷ Obsérvese que esta clasificación excluye otras aportaciones teóricas de gran interés, como pueden ser la teoría del ciclo del producto y la teoría de la incubadora (Duranton y Puga, 2001). Sin embargo, para los propósitos de este trabajo dichas teorías resultan comparativamente menos relevantes.

¹⁸ Un planteamiento que subyace tras la aproximación conductista es que el análisis de localizaciones potencialmente óptimas es creciente con la dimensión de la empresa, mientras que las empresas de menores dimensiones se localizan preferentemente cerca del lugar de residencia del empresario (véase Michelacci y Silva, 2005, para el análisis del caso español y Arauzo y Manjón, 2004, para una discusión más extensa sobre este particular). Desgraciadamente, la base de datos que utilizamos para recoger las entradas no contiene información sobre la figura del empresario.

en el que se desarrollan (clientes, proveedores, sindicatos, sistemas regionales de innovación, presión fiscal, administraciones públicas, otras empresas, etc.). Ejemplos ilustrativos de esta aproximación se hallan en los trabajos de Charney (1983), Bartik (1985), Coughlin y Segev (2000) y Gabe y Bell (2004), entre otros.

Es evidente que estas aproximaciones están estrechamente relacionadas, pero también que difieren críticamente en su concepción de cuáles son los principales determinantes de la localización. Así, para la aproximación neoclásica, empresas y empresarios toman las decisiones ponderando qué características del territorio pueden afectar a sus beneficios netos esperados. Estas características incluyen, entre otras, los costes de transporte y de personal, así como la existencia de economías externas. Por su parte, para la aproximación conductista, las características personales del empresario (edad, experiencia previa, etc.) y su limitada capacidad para conocer el entorno (conocerá mejor aquellas localizaciones más cercanas a su lugar de residencia y considerará más localizaciones como potenciales cuanto mayor sea el tamaño de la empresa y, por lo tanto, mayores los recursos que pueda dedicar a obtener información) constituyen factores determinantes de la elección de un territorio. Finalmente, para la aproximación institucional, cuestiones como el nivel de salarios, el grado de afiliación a los sindicatos o las regulaciones impositivas deberían necesariamente formar parte de nuestro vector de variables explicativas.

Desde el punto de vista del trabajo empírico, lo ideal sería poder incorporar todas estas propuestas (aunque solo fuera para evitar posibles errores de especificación asociados a la omisión de variables relevantes). Pero obviamente este es un objeti-

Cuadro n.º 6

Variables explicativas (definición, fuentes y años disponibles)

Variables	Definición (Código identificativo)	Fuente estadística	Años disponibles
Aproximación neoclásica			
Densidad de población	Densidad de habitantes (DENS)	Trullén y Boix (2004)	2001-2004
Economías de urbanización	Densidad de ocupación (URBA)	Trullén y Boix (2004)	2001
Deseconomías de urbanización	Economías de urbanización al cuadrado (DESURBA)	IDESCAT	2001
Diversidad industrial	Inversa del índice de Hirschmann-Herfindahl (DIV)	Trullén y Boix (2004)	2001-2004
Economías de localización	Trabajadores en el sector "i" (17 sectores) sobre superficie del municipio (ECOLOC1, ..., ECOLOC17)	IIDESCAT	2001
Commuting	Ratio de movilidad (COMMOV)	Trullén y Boix (2004)	2001
	Tiempo de desplazamiento (COMTEMPS)	Trullén y Boix (2004)	2001
Capital humano	Porcentaje de residentes ocupados en ciencia y tecnología según la Clasificación Nacional de Ocupaciones (CHRHCT)	Trullén y Boix (2004)	2001
	Residentes con nivel de educación terciaria (CHUNIV)	Trullén y Boix (2004)	2001
	Media de años de educación de la población mayor de 25 años (CHEDU)	Trullén y Boix (2004)	2001
Ocupación industrial	Ocupación en sectores industriales sobre ocupación total (OCUIND)	IDESCAT	2001
Ocupación servicios	Ocupación terciaria sobre ocupación total (OCUSERV)	Trullén y Boix (2004)	2001-2004
Emprendedores	Ocupados declarados empresarios agrícolas y no agrícolas (con y sin trabajadores) sobre el total de ocupados (ENTREP)	IDESCAT	2001
Infraestructuras de transporte	Tiempo medio de desplazamiento por carretera a las 4 capitales provinciales (INFRACAP)	Trullén y Boix (2004)	2001
	Tiempo medio de desplazamiento por carretera al aeropuerto más cercano (INFRAAERO)	Trullén y Boix (2004)	2001
	Tiempo medio de desplazamiento por carretera al puerto de mercancías más cercano (INFRAPORT)	Trullén y Boix (2004)	2001
	Existencia de estación de ferrocarril con parada (INFRAFERR)	Trullén y Boix (2004)	2004

.../...

Cuadro n.º 6 (continuación)
Variables explicativas (definición, fuentes y años disponibles)

Variables	Definición (Código identificativo)	Fuente estadística	Años disponibles
Aproximación Institucional			
Identificador de provincia	<i>Dummy</i> de provincia (PROV1 es Barcelona, PROV2 es Girona y PROV3 es Lleida. Tarragona es la categoría residual)	IDESCAT	
Identificador de capital comarca	<i>Dummy</i> de capital de comarca (CAPCOM)	IDESCAT	
Identificador de población del litoral	<i>Dummy</i> de municipio del litoral (LITORAL)	IDESCAT	
Área metropolitana	<i>Dummies</i> identificadoras de las áreas metropolitanas de Barcelona (METBCN), Girona (METGIR), Lleida (METLLE), Manresa (METMAN) y Tarragona (METTGN)	Trullén y Boix (2004)	2001

Nota: La relación de los sectores utilizados es la siguiente: 1.- Agricultura y ganadería. 2.- Pesca. 3.- Industrias extractivas. 4.- Industrias manufactureras. 5.- Electricidad, gas y agua. 6.- Construcción. 7.- Comercio y reparación. 8.- Hostelería. 9.- Transporte y comunicaciones. 10.- Intermediación financiera. 11.- Inmobiliarias, alquileres y servicios empresariales. 12.- Administración Pública, defensa y SS. 13.- Educación. 14.- Sanidad y servicios sociales. 15.- Otros servicios. 16.- Personal domestico. 17.- Organismos extraterritoriales.

vo que raramente puede alcanzarse. En primer lugar, porque a menudo resulta difícil asignar una variable a una única aproximación (una misma medida puede estar recogiendo factores asociados a más de una aproximación). En segundo lugar, porque no siempre es posible disponer de variables representativas de las tres aproximaciones.

Estas dificultades tienen un reflejo directo en nuestra base de datos. Tal y como se puede apreciar en el cuadro n.º 6, no se dispone de variables relativas a la aproximación conductista. Además, la aproximación institucional tan solo ha podido ser recogida en forma de variables *dummies* relativas a la organización territorial de Ca-

taluña. En definitiva, el núcleo principal de nuestras variables explicativas lo constituyen medidas relativas a la aproximación neoclásica.¹⁹ No obstante, en su favor hay que destacar la riqueza informativa que

¹⁹ Este «desequilibrio» entre aproximaciones teóricas en los regresores podría introducir sesgos en las estimaciones, si bien es difícil evaluar su magnitud en modelos no lineales como los empleados en este estudio y en ausencia de estudios previos que ofrezcan una indicación de la importancia relativa de cada aproximación en las decisiones de localización. Además, conviene tener en cuenta que nuestra unidad estadística de referencia es el territorio y no la empresa, por lo que los sesgos que puedan originarse por la omisión de factores ligados a la aproximación conductista deberían ser poco importantes.

proporciona esta base de datos²⁰ respecto a factores como las economías de localización, el capital humano, las infraestructuras de transporte y las áreas metropolitanas.²¹

Resultados

El cuadro n.º 7 presenta las estimaciones de los modelos BINNEG (binomial negativa) y POISINF (Poisson inflado sobre la base de un logit para los ceros) para las dos estructuras de datos de sección cruzada descritas previamente (referidas al año 2001 y al período 2001-2004). Como era de esperar, los contrastes realizados (LR Test α , Vuong Test) indican la necesidad de recoger la sobredispersión de la variable dependiente, pero es difícil concluir cuál es la manera más apropiada de hacerlo. Los contrastes de significación conjunta y el logaritmo de la función de verosimilitud parecen recomendar el empleo de un Poisson inflado. No obstante, estos estadísticos no son directamente comparables entre modelos, por lo que podría ser arriesgado rechazar completamente la validez de los re-

sultados obtenidos empleando la binomial negativa.

Independientemente de cuál sea la especificación econométrica escogida, las variables relacionadas con la aproximación institucional son mayoritariamente estadísticamente significativas. En particular, destacan los diferenciales efectos positivos de la provincia de Barcelona, la capitalidad de comarca y la proximidad a la costa mediterránea. Por lo tanto, en Cataluña parece coexistir un proceso de concentración intermetropolitano, que beneficiaría al área de Barcelona, con un proceso de desconcentración intrametropolitano, que beneficiaría a localidades de menor tamaño. No obstante, de este último solo se beneficiarían aquellas con una mínima dimensión (como la que proporciona, por ejemplo, la capitalidad comarcal), así como las situadas en zonas próximas al litoral. De la misma forma, los municipios con mayor densidad de puestos de trabajo tienen comparativamente peor atractivo como localizaciones potenciales, tal y como indican los coeficientes relativos a la ocupación, a las economías de urbanización y a las economías de localización de los diferentes sectores manufactureros analizados.²²

En cambio, los efectos no son tan claros en las variables relativas a la aproximación neoclásica, quizás a excepción de la ocupación en el sector servicios y las infraestructuras portuarias. Destaca especialmente el caso de la variable densidad de población, ya que a pesar de que presenta el signo (negativo) esperado en la mayoría de las especificaciones, solo en una de ellas es significativo. No obstante, este resultado se ajusta al obtenido para el mismo

²⁰ Existen otros aspectos que pueden resultar relevantes para comprender las diferencias locales en términos de localización de empresas. Uno de ellos, por ejemplo, lo constituyen las diferencias locales en la disponibilidad de financiación ajena. Sin embargo, dado que se trabaja con la totalidad de municipios de Cataluña, no resulta posible disponer de variables que puedan medir este fenómeno en el conjunto de dichos municipios.

²¹ Con objeto de poner de manifiesto más claramente el efecto de cada uno de estos factores, los cuadros de resultados incluyen un contraste de tipo Wald de significación conjunta de las variables relativas a cada uno de ellos. Así, «Wald Test E.L.» alude a la significación conjunta de las variables relativas a las economías de localización, «Wald Test CH» a las relativas al capital humano, «Wald Test Infra» a las relativas a las infraestructuras de transporte y «Wald Test Met» a las relativas a las áreas metropolitanas.

²² En todo caso, los resultados para las economías de localización muestran una elevada variabilidad sectorial.

Cuadro n.º 7

Modelos econométricos (datos de sección cruzada)

Variable	BINNEG (2001-2004)	POISINF (2001-2004)	BINNEG (2001)	POISINF (2001)
DENS	-0,0013 (0,0030)	0,0001 (0,0009)	-0,0011 (0,0028)	-0,0037 (0,0016)**
URBA	-0,0001 (0,0069)	-0,0056 (0,0021)***	0,0001 (0,0061)	0,0040 (0,0037)
DESURBA	0,0000 (0,0000)	0,0000 (0,0000)***	0,0000 (0,0000)*	0,0000 (0,0000)
DIV	0,0826 (0,0898)	-0,1597 (0,0267)***	0,1523 (0,0636)**	-0,0652 (0,0328)**
ECOLOC1	-0,0072 (0,0406)	-0,0430 (0,0117)***	-0,0230 (0,0338)	-0,0044 (0,0207)
ECOLOC2	-0,0515 (0,0876)	-0,1428 (0,0409)***	0,0134 (0,0425)	-0,0295 (0,0232)
ECOLOC3	-0,1893 (0,2103)	-0,1005 (0,0737)	-0,7317 (0,8754)	-3,4848 (0,6817)***
ECOLOC4	0,0015 (0,0026)	0,0031 (0,0006)***	0,0019 (0,0020)	-0,0026 (0,0016)
ECOLOC5	-0,3931 (0,1665)**	-0,8583 (0,0611)***	-0,4733 (0,1823)***	-0,7674 (0,1444)***
ECOLOC6	0,0482 (0,0252)*	0,0780 (0,0075)***	0,0460 (0,0218)**	0,0531 (0,0121)***
ECOLOC7	0,0090 (0,0190)	0,0234 (0,0053)***	0,0188 (0,0137)	0,0377 (0,0080)***
ECOLOC8	0,0155 (0,0265)	0,0464 (0,0071)***	-0,0275 (0,0271)	-0,0258 (0,0171)
ECOLOC9	-0,0190 (0,0195)	-0,0391 (0,0054)***	-0,0333 (0,0214)	0,0640 (0,0245)***
ECOLOC10	0,0353 (0,0608)	-0,0297 (0,0156)*	-0,0430 (0,0655)	-0,0779 (0,0481)
ECOLOC11	0,0219 (0,0286)	0,0680 (0,0097)***	0,0308 (0,0233)	-0,0398 (0,0198)**
ECOLOC12	-0,0429 (0,0330)	-0,0208 (0,0094)**	0,0333 (0,0294)	0,0888 (0,0266)***
ECOLOC13	0,0139 (0,0361)	0,0054 (0,0107)	-0,0231 (0,0225)	-0,0215 (0,0151)
ECOLOC14	-0,0039 (0,0169)	-0,0499 (0,0049)***	-0,0154 (0,0163)	-0,0272 (0,0126)**
ECOLOC15	0,0044 (0,0176)	0,0128 (0,0040)***	-0,0011 (0,0187)	0,1206 (0,0579)**

.../...

Cuadro n.º 7 (continuación)

Modelos econométricos (datos de sección cruzada)

Variable	BINNEG (2001-2004)	POISINF (2001-2004)	BINNEG (2001)	POISINF (2001)
ECOLOC16	-0,1462 (0,0782)*	-0,2274 (0,0252)***	-0,0608 (0,0643)	-0,1410 (0,0562)**
ECOLOC17	-0,5530 (5,8221)	-8,3716 (1,6743)***	-5,0723 (5,4686)	7,4937 (5,4199)
OCUIND	-0,0384 (1,6979)	1,0452 (0,4698)**	-1,1473 (1,1712)	1,8917 (0,5871)***
OCUSERV	1,0132 (0,5372)*	0,7459 (0,1495)***	0,5307 (0,3728)	0,7075 (0,1910)***
ENTREP	0,7668 (1,6696)	0,6164 (0,4690)	-0,9304 (1,1700)	-0,2597 (0,7540)
COMMOV	-0,0137 (0,0111)	-0,0095 (0,0074)	-0,0091 (0,0202)	-0,0356 (0,0176)**
COMTEMPS	0,0000 (0,0000)	0,0000 (0,0000)***	0,0000 (0,0000)	-0,0000 (0,0000)
CHRHCT	-0,3724 (4,0312)	3,7857 (0,9785)***	-0,1040 (2,5872)	1,4093 (1,4167)
CHUNIV	-0,9818 (4,1137)	-3,4854 (1,0900)***	0,2908 (2,5865)	5,0055 (1,3569)***
CHEDU	0,0591 (0,1756)	-0,0578 (0,0521)	-0,0521 (0,1163)	-0,3405 (0,0720)***
INFRACAP	0,0161 (0,0184)	-0,0019 (0,0048)	-0,0252 (0,0124)**	-0,0182 (0,0076)**
INFRAAERO	-0,0040 (0,0069)	-0,0082 (0,0021)***	-0,0080 (0,0052)	-0,0141 (0,0027)***
INFRAPORT	-0,0090 (0,0159)	0,0114 (0,0041)***	0,0249 (0,0109)**	0,0284 (0,0066)***
INFRAFERR	-0,1648 (0,4686)	-0,7559 (0,1267)***	0,2450 (0,3254)	-0,0711 (0,1582)
PROV1	2,2338 (0,1860)***	1,5369 (0,0694)***	1,8567 (0,2368)***	1,1375 (0,1456)***
PROV2	0,3324 (0,1952)*	0,0331 (0,0863)	-0,0848 (0,2744)	-0,1442 (0,2151)
PROV3	-0,1336 (0,2032)	-0,2826 (0,0973)***	-0,7757 (0,3100)**	-0,9511 (0,2995)***
CAPCOM	2,4908 (0,2812)***	1,6942 (0,0448)***	2,6507 (0,3278)***	1,3952 (0,1132)***
LITORAL	0,7565 (0,2262)***	0,4035 (0,0451)***	0,6551 (0,2928)**	0,3619 (0,1200)***

.../...

Cuadro n.º 7 (continuación)

Modelos econométricos (datos de sección cruzada)

Variable	BINNEG (2001-2004)	POISINF (2001-2004)	BINNEG (2001)	POISINF (2001)
METBCN	0,1659 (0,4319)	0,0853 (0,1187)	0,2869 (0,2709)	0,0265 (0,1487)
METGIR	0,1004 (0,6018)	-0,9151 (0,1997)***	-0,2246 (0,3787)	-1,0515 (0,2094)***
METLLE	0,1992 (0,5101)	0,3005 (0,1411)**	-0,2428 (0,3966)	-1,1524 (0,3649)***
METMAN	-0,3442 (0,8444)	-1,1424 (0,2703)***	-0,7827 (0,5887)	0,0207 (0,3949)
METTGN	-0,1085 (0,6041)	-0,1477 (0,1615)	0,2974 (0,3800)	-0,0319 (0,2153)
Observaciones	946	946	899	899
Log L	-1781,75	-3129,91	-836,91	-870,87
LR Test	382,65***	4795,03***	261,07***	866,66***
Pseudo-R ²	0,10		0,13	
LR Test α	4616,16***		724,48***	
Vuong Test		7,18***		5,68***
Wald Test E.L.	24,25	575,22***	14,76	185,86***
Wald Test CH	0,16	20,45***	0,25	31,54***
Wald Test Infra.	2,01	59,52***	6,12	32,75***
Wald Test Met.	0,68	49,43***	4,79	37,72***

Fuente: Elaboración propia.

NOTA: ***, ** y * indican coeficientes estadísticamente significativos para un nivel de confianza de, respectivamente, 1%, 5% y 10%. Entre paréntesis los errores estándar. Las definiciones de las variables explicativas se hallan en el cuadro n.º 6.

ámbito territorial por Arauzo (2005). Por su parte, la incidencia parcialmente negativa de la educación universitaria y de la media de años de escolarización puede resultar, en principio, sorprendente. Sin embargo, una vez más se trata de un resultado recurrente en trabajos previos (Arauzo, 2002, 2005) y para el que pueden plantearse tres explicaciones alternativas o complementarias: a) la estructura industrial catalana está formada básicamente por actividades de nivel tecnológico medio y bajo, para las cuales el capital humano no es un «input lo-

cacional»; b) la no coincidencia espacial entre lugar de residencia y lugar de trabajo; c) los coeficientes asociados a estas variables podrían estar reflejando los efectos negativos del nivel salarial sobre las decisiones de localización (Bartik, 1985).²³

²³ De hecho, la evidencia empírica sobre la aportación del capital humano a la localización industrial es contradictoria, de modo que hay trabajos en los que la incidencia del capital humano es positiva (véase, por ejemplo, Woodward, 1992 y Coughlin *et. al.*, 2000) y trabajos en los que es negativa o no significativa (véase, por ejemplo, Bartik, 1985).

El resto de determinantes considerados no son estadísticamente significativos o presentan signos contradictorios. Por otro lado, los contrastes de significación conjunta de las economías de localización, el

capital humano, las infraestructuras de transporte y las áreas metropolitanas siempre proporcionan elevados estadísticos de prueba en la especificación binomial negativa.

Cuadro n.º 8

Modelos econométricos (datos de panel)

Variable	EF	EA	Variable	EF	EA
DENS	0,0010 (0,0006)*	0,0008 (0,0005)	ECOLOC12	-0,0071 (0,0064)	-0,0093 (0,0062)
URBA	-0,0025 (0,0013)*	-0,0022 (0,0012)*	ECOLOC13	0,0014 (0,0069)	-0,0027 (0,0064)
DESURBA	0,0000 (0,0000)**	0,0000 (0,0000)**	ECOLOC14	-0,0048 (0,0029)*	-0,0035 (0,0028)
DIV	-0,0007 (0,0146)	0,0009 (0,0143)	ECOLOC15	0,0006 (0,0025)	0,0012 (0,0024)
ECOLOC1	-0,0085 (0,0074)	-0,0100 (0,0072)	ECOLOC16	-0,0066 (0,0133)	-0,0030 (0,0122)
ECOLOC2	0,0041 (0,0159)	0,0034 (0,0155)	ECOLOC17	-0,7202 (0,9684)	-0,4188 (0,9106)
ECOLOC3	-0,0881 (0,0499)*	-0,0902 (0,0494)*	OCUIND	0,2878 (0,2605)	0,3378 (0,2556)
ECOLOC4	0,0003 (0,0004)	0,0003 (0,0004)	OCUSERV	0,0198 (0,0921)	0,0598 (0,0901)
ECOLOC5	-0,0411 (0,0262)	-0,0579 (0,0257)**	ENTREP	-0,1063 (0,2715)	-0,0677 (0,2649)
ECOLOC6	0,0078 (0,0044)*	0,0101 (0,0042)**	COMMOB	-0,0039 (0,0063)	-0,0041 (0,0057)
ECOLOC7	-0,0032 (0,0033)	-0,0009 (0,0028)	COMTEMPS	-0,0000 (0,0000)***	-0,0000 (0,0000)**
ECOLOC8	0,0021 (0,0037)	0,0008 (0,0036)	CHRHCT	-0,9824 (0,5826)*	-1,0305 (0,5734)*
ECOLOC9	-0,0005 (0,0036)	-0,0016 (0,0034)	CHUNIV	0,9357 (0,6240)	0,9166 (0,6106)
ECOLOC10	0,0300 (0,0087)***	0,0278 (0,0083)***	CHEDU	0,0284 (0,0266)	0,0280 (0,0260)
ECOLOC11	0,0013 (0,0054)	0,0010 (0,0052)	INFRACAP	-0,0009 (0,0036)	0,0005 (0,0034)

.../...

Cuadro n.º 8 (continuación)

Modelos econométricos (datos de sección cruzada)

Variable	EF	EA	Variable	EF	EA
INFRAAERO	-0,0020 (0,0020)	-0,0030 (0,0018)*	METGIR	0,0292 (0,0819)	0,0231 (0,0805)
INFRAPORT	0,0030 (0,0034)	0,0026 (0,0031)	METLLE	0,1909 (0,1135)*	0,1540 (0,1077)
INFRAFERR	0,0457 (0,0729)	0,0417 (0,0709)	METMAN	-0,1037 (0,1398)	-0,1307 (0,1367)
PROV1	—	2,2476 (0,1768)***	METTGN	-0,0581 (0,0989)	-0,0556 (0,0967)
PROV2	—	0,3317 (0,1942)*	Observaciones	1723	3656
PROV3	—	-0,1969 (0,2018)	Log L	-1516,28	-3.290,66
CAPCOM	—	2,4427 (0,2749)***	LR Test ^a	44,27	416,11***
LITORAL	—	0,7207 (0,2243)***	LR Test α	5.252,18***	
METBCN	0,0030 (0,0655)	-0,0026 (0,0640)	Hausman Test	4,95	
			Wald Test E.L.	25,85	28,06**
			Wald Test CH	4,83	4,99
			Wald Test Infra.	1,59	3,19
			Wald Test Met.	4,01	3,55

Fuente: Elaboración propia.

NOTA: ***, ** y * indican coeficientes estadísticamente significativos para un nivel de confianza de, respectivamente, 1%, 5% y 10%. Entre paréntesis los errores estándar. EF indica efectos fijos y EA indica efectos aleatorios. Las definiciones de las variables explicativas se hallan en el cuadro n.º 6. “WaldTest” es un test de significación conjunta del modelo (asintóticamente equivalente al “LRTest” presentado en el cuadro n.º 7 para los modelos estimados con datos de sección cruzada).

«Wald Test» es un test de significación conjunta del modelo (asintóticamente equivalente al «LRTest» presentado en el cuadro n.º 7 para los modelos estimados con datos de sección cruzada).

En general, los resultados no varían sustancialmente al emplear una u otra estructura de datos de sección cruzada, pero sí lo hacen al controlar por heterogeneidad inobservable en modelos para datos de panel. En concreto, una comparativa de los cuadros n.º 7 y 8 muestra que las principales discrepancias se observan en los resultados relativos a la aproximación neoclási-

ca. Un ejemplo muy claro puede observarse para las variables que miden la incidencia de las infraestructuras de transporte, relevantes en las estimaciones realizadas con datos de sección cruzada mostradas en el cuadro n.º 7 (especialmente para el caso de las infraestructuras portuarias), pero no en las estimaciones con modelos para datos de panel del cuadro n.º 8. No obstante,

de entre las variables que pretenden recoger factores asociados a esta aproximación solo son estadísticamente significativas las relativas al capital humano (de acuerdo con el contraste tipo Wald) y a las economías de urbanización y desurbanización (aunque con signo contrario al esperado). Por su parte, los resultados relativos a las variables vinculadas a la aproximación institucional coinciden ampliamente con los obtenidos a partir de los datos de sección cruzada, excepto para el caso de las *dummies* identificadoras de áreas metropolitanas, que carecen en este caso de significación estadística. Por último, queremos mencionar que las conclusiones apenas varían si se emplea un estimador de efectos fijos (EF) o un estimador de efectos aleatorios (EA).

4. CONCLUSIONES

Este trabajo presenta una revisión de los estudios empíricos sobre localización industrial en la que se hace especial hincapié en las metodologías econométricas que emplean, básicamente, modelos de elección discreta y modelos para datos de recuento. Esta revisión nos ha permitido, por un lado, constatar las dificultades existentes en España para disponer de datos adecuados para la realización de este tipo de estudios. Por otro lado, tras revisar las características básicas de los modelos empleados en la literatura económica, parece probable que futuras investigaciones en esta área recurran a sofisticados modelos

de elección discreta o a explorar las complejidades del tratamiento estadístico de los números enteros no negativos.

Además, a modo de ilustración de los resultados que pueden extraerse en este tipo de estudios, este trabajo proporciona una aplicación de los modelos de recuento a la localización de establecimientos industriales en los municipios catalanes durante el período 2001-2004. En las estimaciones obtenidas destaca el poco peso de los determinantes «clásicos» de la localización (densidad, urbanización, etc.) y, simultáneamente, la importancia de factores de carácter institucional (capitalidad) y geográfico (proximidad al litoral). Estos resultados parecen coherentes con el empleo de diferentes especificaciones (modelo de Poisson inflado, distribución binomial negativa), diferentes estructuras de datos (sección cruzada, panel) y con la existencia de heterogeneidad inobservable entre los municipios considerados.

Desde el punto de vista de la política económica, los resultados muestran una heterogeneidad sectorial y territorial que desaconseja el uso de medidas uniformes y transversales. En este sentido, las políticas centradas en el fomento de la entrada de establecimientos industriales deberían tener en cuenta, como mínimo, tres ámbitos fundamentales: la dimensión de los establecimientos, el sector de actividad en el que operan y el municipio en el cual se localizan. En caso contrario, se estarían obviando aspectos clave en las decisiones de localización de las empresas.

REFERENCIAS BIBLIOGRÁFICAS

- ARAUZO, J.M. (2002): «Pautes de localización industrial: Estructura productiva i capital humà als municipis Catalans», *Revista Econòmica de Catalunya* 45: 43-55.
- ARAUZO, J.M. y MANJÓN, M. (2004): «Firm Size and Geographical Aggregation: An Empirical Appraisal in Industrial Location», *Small Business Economics* 22: 299-312.
- ARAUZO, J.M. (2005): «Determinants of Industrial Location. An Application for Catalan Municipalities», *Papers in Regional Science* 84 (1): 105-120.
- AUDRETSCH, D. (1995): *Innovation and Industry Evolution*, The MIT Press.
- BARTIK, T.J. (1985): «Business Location Decisions in the U.S.: Estimates of the Effects of Unionization, Taxes, and Other Characteristics of States», *Journal of Business and Economic Statistics* 3: 14-22.
- BASILE, R. (2004): «Acquisition versus greenfield investment: the location of foreign manufacturers in Italy», *Regional Science and Urban Economics* 34: 3-25.
- BAUDEWYNS, D. (1999): «La localisation intra-urbaine des firmes: une estimation logit multinomiale», *Revue d'Économie Régionale et Urbaine* 5: 915-930.
- BAUDEWYNS, D. ; SEKKAT, K. y M. BEN-AYAD (2000) : «Infrastructure publique et localisation des entreprises à Bruxelles et en Wallonie», en M. Beine y F. Docquier (eds.), *Convergence des Régions: cas des régions belges*, De Boeck (p. 280-303) : Bruselas.
- CAMERON, A.C. y TRIVEDI, P.K (1998): *Regression Analysis of Count Data*, Cambridge University Press.
- CARLTON, D. (1979): «Why new firms locate where they do: An econometric model», en W. Wheaton (ed.) *Interregional Movements and Regional Growth*. The Urban Institute, Washington.
- CARLTON, D. (1983): «The location and employment choices of new firms: An econometric model with discrete and continuous endogenous variables», *Review of Economics and Statistics* 65: 440-449.
- CAVES, R.E. (1998): «Industrial Organization and New Findings on the Turnover and Mobility of Firms», *Journal of Economic Literature* 36 (4): 1947-1982.
- CHARNEY, A. (1983): «Intraurban Manufacturing Location Decisions and Local Tax Differentials», *Journal of Urban Economics* 14: 184-205.
- CIEŹLIK, A. (2005): «Location of foreign firms and national border effects: the case of Poland», *Tijdschrift voor Economische en Sociale Geografie* 96(3): 287-297.
- COUGHLIN, C.C. y SEGEV, E. (2000): «Location determinants of new foreign-owned manufacturing plants», *Journal of Regional Science* 40: 323-351.
- COUGHLIN, C.C.; TERZA, J.V. y ARROMDEE, V. (1991): «State characteristics and the location of foreign direct investment within the United States», *The Review of Economics and Statistics* 73: 675-683.
- DURANTON, G. and PUGA, D. (2001): «Nursery cities: Urban diversity, process innovation and the life cycle of products», *American Economic Review* 91: 1454-1477.
- FIGUEIREDO, O.; GUIMARÃES, P. y Woodward, D. (2002): «Home-field advantage: location decisions of Portuguese entrepreneurs», *Journal of Urban Economics* 52: 341-361.
- FRIEDMAN, J.; GERLOWSKI, D.A. y SILBERMAN, J. (1992): «What attracts foreign multinational corporations? Evidence from branch plant location in the United States» *Journal of Regional Science* 32: 403-418.
- GABE, T. y BELL, K.P. (2004): «Tradeoffs between local taxes and government spending as determinants of business location», *Journal of Regional Science* 44(1): 21-41.
- GEROSKI, P. (1995): «What do we know about entry?», *International Journal of Industrial Organization* 13 (4): 421-440.
- GLAESER, E. L.; KALLAL, H.; SCHEINKMAN, J. A. y SHLEIFER, A. (1992): «Growth in Cities», *Journal of Political Economy* 100: 1126-1152.
- GREEN, W.H. (2000): *Econometric Analysis*, Prentice Hall International.
- GUIMARÃES, P., FIGUEIREDO, O. y WOODWARD, D. (2000): «Agglomeration and the Location of Foreign Direct Investment in Portugal», *Journal of Urban Economics* 47: 115-135.
- GUIMARÃES, P.; FIGUEIREDO, O. and WOODWARD, D. (2003): «A Tractable Approach to the Firm Location Decision Problem», *Review of Economics and Statistics* 85(1): 201-204.
- HANSEN, E.R. (1987): «Industrial Location Choice in São Paulo, Brazil: A Nested Logit Model», *Regional Science and Urban Economics* 17: 89-108.
- HAUSMAN, J.A.; HALL, B.H. y GRILICHES, Z. (1984): «Econometric Models for Count Data with an Ap-

- plication to the Patents - R&D Relationship», *Econometrica* 52: 909-938.
- HAYTER, R. (1997): *The dynamics of industrial location. The factory, the firm and the production system*, New York: Wiley.
- HEAD, K; RIES, J. y SWENSON, D. (1995): «Agglomeration benefits and location choice: Evidence from Japanese manufacturing investments in the United States», *Journal of International Economics* 38 (3/4): 223-247.
- HENDERSON, V.; KUNCORO, A. y TURNER, M. (1995): «Industrial development in cities», *Journal of Political Economy* 103 (5): 1067-1090.
- HOLL, A. (2004a): «Transport infrastructure, agglomeration economies, and firm birth. Empirical evidence from Portugal», *Journal of Regional Science* 44 (4): 693-712.
- HOLL, A. (2004b): «Manufacturing location and impacts of road transport infrastructure: Empirical evidence from Spain», *Regional Science and Urban Economics* 34 (3): 341-363.
- HOLL, A. (2004c): «Start-ups and relocations: Manufacturing plant location in Portugal», *Papers in Regional Science* 83 (4): 649-668.
- JOHNSON, P. S. y CATHCART, D. G. (1979): «New Manufacturing Firms and Regional Development: Some Evidence from the Northern Region», *Regional Studies* 13: 269-280.
- LIST, J.A. (2001): «US county-level determinants of inbound FDI: evidence from a two-step modified count data model», *International Journal of Industrial Organization* 19: 953-973.
- LIST, J.A. y McHONE, W.W. (2000): «Measuring the effects of air quality regulations on «dirty» firm births: Evidence from the neo and mature-regulatory periods», *Papers in Regional Science* 79: 177-190.
- McFADDEN, D. (1974): «Analysis of qualitative choice behaviour», en P. Zarembka (ed.) *Frontiers in Econometrics*, Academic Press.
- MICHELACCI, C. y SILVA, O. (2005): «Why So Many Local Entrepreneurs?», Working Papers 2005-06 CEMFI.
- OWENS, R.E. y SARTE, P.-D. (2002): «Analyzing firm location decisions: is public intervention justified?», *Journal of Public Economics* 86: 223-242.
- SEGARRA, A.; J.M. ARAUZO; N. GRAS; M. MANJÓN; F. MAÑÉ; M. TERUEL y B. THEILEN (2002): *La creación y la supervivencia de las empresas industriales*, Editorial Cívitas: Madrid.
- SMITH, D.F. y FLORIDA, R. (1994): «Agglomeration and Industrial Location: An Econometric Analysis of Japanese-Affiliated Manufacturing Establishments in Automotive-Related Industries», *Journal of Urban Economics* 36 (1): 23-41.
- TRAIN, K.E. (2003): *Discrete Choice Methods with Simulation*, Cambridge University Press.
- TRULLÉN, J. y BOIX, R. (2004): *Indicadors 2005*, Diputació de Barcelona y Universitat Autònoma de Barcelona.
- Vuong, Q. (1989): «Likelihood ratio tests for model selection and non-nested hypotheses», *Econometrica* 57: 307-333.
- WOODWARD, D. (1992): «Locational determinants of Japanese manufacturing start-ups in the United States», *Southern Economic Journal* 58: 690-708.
- Wu, F. (1999): «Intrametropolitan FDI Firm Location in Guangzhou, China: A Poisson and Negative Binomial Analysis», *The Annals of Regional Science* 33: 535-555.