

Martínez-Martín, Jaime

Article

General equilibrium long-run determinants for Spanish FDI: A spatial panel data approach

SERIEs - Journal of the Spanish Economic Association

Provided in Cooperation with:

Spanish Economic Association

Suggested Citation: Martínez-Martín, Jaime (2011) : General equilibrium long-run determinants for Spanish FDI: A spatial panel data approach, SERIEs - Journal of the Spanish Economic Association, ISSN 1869-4195, Springer, Heidelberg, Vol. 2, Iss. 3, pp. 305-333, <https://doi.org/10.1007/s13209-011-0058-3>

This Version is available at:

<https://hdl.handle.net/10419/77750>

Standard-Nutzungsbedingungen:

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

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

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

Terms of use:

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

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

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



<http://creativecommons.org/licenses/by/2.0/>

General equilibrium long-run determinants for Spanish FDI: a spatial panel data approach

Jaime Martínez-Martín

Received: 12 April 2010 / Accepted: 2 April 2011 / Published online: 27 April 2011
© The Author(s) 2011. This article is published with open access at SpringerLink.com

Abstract While general equilibrium theories of trade stress the role of third-country effects, little work has been done in the empirical foreign direct investment (FDI) literature to test such spatial linkages. This paper aims to provide further insights into long-run determinants of Spanish FDI by considering not only bilateral but also spatially weighted third-country determinants. The few studies carried out so far have focused on FDI flows in a limited number of countries. However, Spanish FDI outflows have risen dramatically since 1995 and today account for a substantial part of global FDI. Therefore, we estimate recently developed spatial panel data models by maximum likelihood (ML) procedures for Spanish outflows (1993–2004) to top-50 host countries. After controlling for unobservable effects, we find that spatial interdependence matters and provide evidence consistent with new economic geography theories of agglomeration, mainly due to complex (vertical) FDI motivations. Spatial error models estimations also provide illuminating results regarding the transmission mechanism of shocks.

Keywords Foreign direct investment · Spatial econometrics · Panel data

JEL Classification F21 · F23 · C31 · C33

1 Introduction

According to [Blonigen \(2005\)](#) “there is an increasing recognition that understanding the forces of economic globalization requires looking first at foreign direct investment

J. Martínez-Martín (✉)
AQR-IREA Research Group, Universitat de Barcelona and BBVA Research,
Torre IV, Av. Diagonal 690, 08034 Barcelona, Spain
e-mail: Jaime.martinezmartin@ub.edu

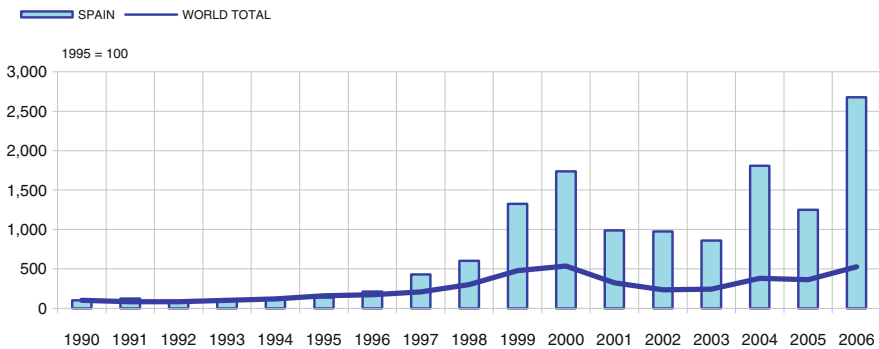


Fig. 1 Spanish and world outward foreign direct investment transactions. Source: UNCTAD

(FDI) by multinational corporations (MNCs): that is, when a firm based in one country locates or acquires production facilities in other countries”.¹ Indeed, while real world GDP grew at an annual rate of 2.5% and real world exports grew by 5.6% annually from 1986 through 2005, UNCTAD data show that real world FDI inflows grew by 17.7% over the same period.

FDI has grown at a remarkable rate since 1980. This surge has occurred worldwide, but it has been particularly dramatic in Spain. Spain’s outward FDI flows have recently outpaced world FDI transactions, especially in the second half of the nineties when Spanish firms began to internationalize.² Initially a net importer, Spain’s outflows have steadily increased and become more active, eventually making the country a current net capital exporter. According to UNCTAD figures, Spain’s cumulative investment abroad scarcely represented 3% of its GDP in the early nineties, but by 2006 outward FDI stock had risen to 41% of GDP. Thus, the relative weight of Spanish investment in world FDI rose to approximately 6% on average in 2001–2006 (see Fig. 1).

The positive qualities associated with this sort of international capital flow, that is, its relative stability (by reducing vulnerability to specific conditions of a domestic or foreign market), its potential for spurring productivity and diffusing technology, and the fact that it permits the spatial fragmentation of production processes, have meant that increasing attention has been devoted to the effects and determinants of FDI.³

Much of the literature is based on analyses using partial equilibrium models of individual firm-level FDI decisions. Researchers looking at world FDI patterns have generally used variations of a gravity framework to model FDI, specifying parent- and host-country GDPs along with distance as its core determinants. These models appear able to describe FDI patterns statistically, but while [Anderson and van Wincoop](#)

¹ According to the IMF’s definition, FDI is the acquisition of 10% or more of the assets of a foreign firm. It is often defined as an investment involving a long-term relationship and reflecting a lasting interest and control of a resident entity in one economy (foreign direct investor or parent enterprise) in a firm resident in an economy other than that of the foreign direct investor (FDI enterprise or affiliate enterprise of a foreign affiliate). It implies that the investor exerts a significant degree of influence on the management of the enterprise resident in the other economy.

² See [Gordo et al. \(2008\)](#).

³ See for instance [Romer \(1993\)](#), [Rappaport \(2000\)](#) and [Rodríguez-Clare \(1996\)](#).

(2003) have solidified an appropriate gravity specification as theoretically valid for trade patterns, it is not clear that this is true for FDI patterns.

The bulk of the theory on the creation of multinational enterprises (MNE) stems from the general equilibrium models of Markusen (1984) and Helpman (1984) which use two-country frameworks. Since then, richer general equilibrium models have been developed that allow for more complex forms of imperfect competition (e.g. Markusen 2002; Helpman et al. 2004). Nevertheless, most FDI models maintain the simple two-country, two-factor framework.

With the recent development of spatial econometrics models, the theoretical literature has recognized that the complex motivations for FDI probably require modelling in a multilateral context, a context in which a MNE considers home, host, and third country characteristics when choosing firm activities. Hence, while general equilibrium theories of trade stress the role of third-country effects, little work has been done in the empirical FDI literature to test such spatial linkages.

In this paper, though, recently developed spatial panel data models proposed by Elhorst (2003, 2009) are estimated by maximum likelihood (ML) for Spanish outflows for 1993–2004 to top-50 host countries. By observing the spatial distribution pattern of Spanish FDI outflows during this period (see Fig. 2), spatial interdependence claims to be tested. The model approach is similar to Blonigen et al. (2007) and Garretsen and Peeters (2008) since spatial interactions are captured under a multilateral framework. By estimating spatial lag models (the former) and, additionally, spatial error models (the latter) both these studies find significant evidence for spatial FDI interdependences (i.e. US and Dutch outflows) after controlling for fixed effects.

In spite of recent trends in international applied research, no empirical approach has focused on Spanish FDI outflows from a multilateral point of view for the time being.

Thus, the null hypotheses under analysis in this work are twofold.

First, the influence of spatial interactions on Spanish FDI outflows is tested by estimating recently developed spatial models with unobservable effects. Second, no structure is imposed to isolate one particular multilateral effect (i.e. horizontal or vertical specialization motivations, among others); rather, the net effects of such forces are estimated. In this regard, the correlation signs of the spatial autoregression and spatial error models may provide evidence for or against alternative theories for FDI motivations.

The main empirical results prove the relevance of spatial interdependence in Spanish FDI outflows after controlling for unobservable effects. Additionally, we provide evidence consistent with new economic geography (NEG) theories of agglomeration, mainly due to complex (vertical) FDI motivations. And last but not least, spatial error model estimations show traces of the transmission mechanism of shocks. However, as in the related literature, the results turn out to be sensitive to sample selection.

The remainder of the paper proceeds as follows. Section 2 reviews the related literature, emphasizing approaches that consider third-country effects on FDI location decisions; Sect. 3 discusses the empirical model, while Sect. 4 describes the data and provides a brief overview of Spanish FDI geographical patterns; Sect. 5 highlights the main results, and Sect. 6 concludes.

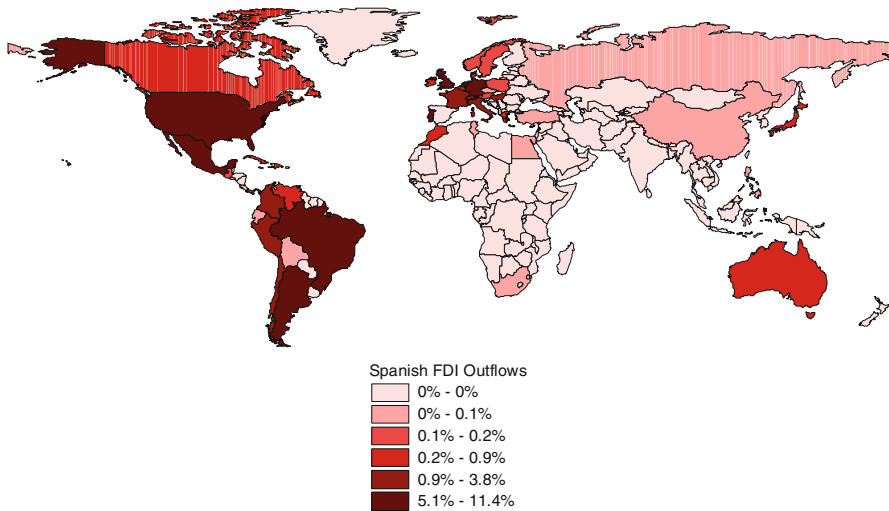


Fig. 2 Spanish foreign direct investment outflows (1993–2004). Source: Registro de Inversiones Extranjeras (RIE)

2 Theoretical background

Studies of FDI flows are lagging some way behind the parallel trade literature. More specifically, research literature dealing with spatial interactions on FDI theory is quite recent. Therefore, only a few number of recent papers have applied spatial econometric techniques to allow for the interdependence of FDI activity (the dependent variable) across host countries. [Coughlin and Segev \(2000\)](#) estimated that FDI in neighbouring provinces increases FDI in a Chinese province and consider this to be evidence of agglomeration externalities. [Baltagi et al. \(2007\)](#) develop a model of MNE activity in a multi-country world that predicts how a variety of neighbouring country characteristics (GDP, trade costs, endowments, etc.) should affect FDI in a focus country depending on MNE motivations (horizontal, vertical, export-platform, etc.). However, the properties of the spatial lag parameter included in their specification are quite a long way from the elasticity, for two main reasons: first, due to the misinterpretation of the spatial lag parameter, and second, a spreading spatial effect on exogenous variables, as [Anselin \(2003\)](#) points out.

[Baltagi et al. \(2007\)](#) and [Blonigen et al. \(2007\)](#) are two innovative studies in this topic. The first analyzes US outward FDI stock in country–industry pairs (1989–1999) and the second focuses on FDI from the US to 20 OECD countries (1980–2000). Baltagi’s model specification is more general since it includes country–industry-pair effects and also spatially weighted exogenous variables. Errors are spatially correlated when the spatial lag coefficient is not null. However, recall that OLS estimation is still consistent but is inefficient. To solve the endogeneity problem, [Blonigen et al. \(2007\)](#) apply a ML method following [Elhorst \(2003\)](#) while [Baltagi et al. \(2007\)](#) uses a fixed and random effects 2SLS estimator (using the second and third order spatial lags of the exogenous regressors as instruments). In both studies, the estimations exhibit a

significant spatial dependence, which is negative in [Blonigen et al. \(2007\)](#) and positive in [Baltagi et al. \(2007\)](#). In addition, spatial correlation of errors is only detected in [Baltagi et al. \(2007\)](#) by using the moments approach method ([Kapoor et al. 2007](#)).

More recently, [Garretsen and Peeters \(2008\)](#) provide similar evidence for spatial dependence of Dutch FDI outflows (1984–2004) benchmarking Blonigen's model. Additionally, they estimate a spatial error model, and find significant results. The recent estimates by [Hall and Petroulas \(2008\)](#) are similar to Baltagi's model in that they use spatially weighted exogenous variables and test for spatial autocorrelation with a wider sample selection of host and origin countries. An earlier but less general model allowing for spatial autocorrelation was estimated by [Abreu \(2005\)](#). By means of a bi-parametric spatial panel data estimator, she tested the impact of tax policy on FDI in attracting FDI. Her findings point out that taxes are an effective policy tool, but only in small countries (i.e. countries with small markets).

Unlike recent trends in international applied research, no empirical approach has focused on Spanish FDI outflows from a multilateral point of view. Some empirical papers on FDI determinants use discrete choice data for Spanish firms under a bilateral trade model framework: [Canals and Noguer \(2006\)](#) prove that distance discourages FDI, while size and sharing a language encourage it; these findings are confirmed by recent studies by [Gordo and Tello \(2008\)](#) and [Barrios and Benito-Ostolaza \(2009\)](#). However, the potential interdependence of FDI across potential host countries is not taken into account. The current paper is the first attempt to do so.

2.1 Spatial FDI interactions

By definition, an econometric bilateral model of FDI does not take into account the specificities of the neighbouring host countries.⁴ Hence, in order to control for the correlation between outward FDI to one country and outward FDI to its neighbours, recently developed spatial panel data model estimation methods have emerged as useful tools, providing consistent and efficient estimations in order to capture third-country effects which would otherwise lead to misspecification errors. Accordingly, by estimating spatial effect models, no structure is imposed to isolate one particular multilateral effect, such as agglomeration or vertical specialization motivation; rather, the net effect of such forces is estimated. A primary benefit of this procedure compared to discrete location choice models is that the correlation signs of spatial autoregression and spatial error models may provide evidence for or against alternative theories for FDI motivations.

Based on [Blonigen et al. \(2007\)](#) and following previous theoretical works, four multinational firm strategies may be linked to the spatial lag coefficient combined with expected signs of (surrounding) market potential variables. Since a mixture of these motivations may occur, no testing for the existence of one over the other applies, and we will focus on identifying net effects.

The first strategy *Pure Horizontal FDI* relies on high trade costs between countries whereby a firm from host country x can serve foreign markets i and j by setting up

⁴ See [Blanchard et al. \(2008\)](#).

Table 1 Summary of hypothesized spatial lag coefficient and market potential effect for various forms of FDI

FDI motivation	Sign of spatial lag	Sign of market potential variable
Pure horizontal	0	0
Export platform	–	+
Pure vertical	–	0
Vertical specialization with agglomeration	+	0/+

Source: [Blonigen et al. \(2007\)](#)

production in i and j or simply by exports. Undoubtedly the chance to avoid trade costs becomes an incentive for horizontal (or “market seeking”) FDI as opposed to exporting, however high setting up production costs in such countries may likely discourage FDI. This may be reflected by the statistically non-significance of the spatial lag, since firm’s decision from country x on whether or not to engage in FDI in country i has nothing to do with its decision to do so in country j .⁵ Likewise, with regard to FDI decision of an x firm to start its production in country i , the market size of other countries j has no bearing with pure horizontal FDI. It will be translated to Table 1 into the form of 0-entry for the market potential variable.

The expected sign for spatial lag coefficient regarding the second strategy *Pure Vertical FDI*, would be negative. One might expect for such FDI strategy from home country x to host country i , that market size of countries $j (\neq i)$, and thus market potential, not to be relevant. As FDI theory states, vertical FDI is mainly driven by factor costs differences across countries and argues that the home country firm may try to set up (some parts of) production in the country with lowest factor costs. Eventually, this implies vertical FDI from country x to country j to be at the expense of vertical FDI from x to host country i .

The third strategy *Export Platform FDI* under consideration is based on the notion that trade costs between potential host countries i and j are lower than between home country x with respect to i and j . Therefore, those firms from country x may decide to engage in FDI by setting up production in host country i (avoiding trade costs between the countries x and i) and thereby using FDI as a platform to export from i to market in country j .⁶ Under such a setting, market potential variable is expected to yield a positive impact on FDI since larger (and closer) markets in countries $j (\neq i)$ make country i a more attractive location for FDI. However, a negative sign may be expected for the spatial lag coefficient since serving the combined markets i and j is more efficient from a single FDI location and setting up a plant in another country is costly (production takes place under increasing returns to scale). In other words, *ceteris paribus*, an increase in export platforms motivated FDI from home country x to third country j would yield less FDI from x to i . It may be even clearer when

⁵ Under such framework, it is worthy to mention that spatial lag is extremely useful to test whether FDI from x into host i is affected by the FDI going from x to j taking the distance between i and j into account.

⁶ An extended explanation of FDI export-platform model is made in [Ekholm et al. \(2003\)](#).

distance between i and j is relatively small. As a result, a negative sign for the spatial lag coefficient is expected.

Finally, the more controversial and the last but not least strategy is called *Vertical Specialization with Agglomeration* (also known as *Complex FDI*). The underlying assumption relies on the fact that whenever country x invests around, it implies that some part of such production takes place in host countries i and j , and multinational firms “slice up the value chain” of their production process by seeking out (low cost) suppliers in multiple (closer) countries.⁷ In this sense, if closer countries i and j share similar supply (network) features, MNEs may try to set up production in i (whenever they find it profitable) given that they already produce from (nearby) country j as well. Thus, FDI from home country x to third country j may be seen as complement for FDI from home country x to host country i even more so if i and j are neighbouring countries. One might expect geographical clustering of such FDI flows for supply reasons. However, whether market potential matters is far than clear. In case market potential captures agglomeration effects, one might expect a positive sign; whereas, in case it captures demand or market-size reasons, one might expect a 0-sign as summarized in Table 1.

It is worthy to mention that such underlying multinational firm strategies to engage FDI (summarized in Table 1) do not cover all spatial FDI interactions. For instance, what if a financial crisis in country x generates a contagion effect in neighbouring countries (and the contagion effect is decreasing with distance)? In that case, FDI may drop not only in country x , but also in neighbouring countries. In that case, even FDI would be spatially related, it would not have anything to do with causes specifically related to FDI. Thus, such economies would be simply spatial interacting and any reached conclusions under such a setting must be taken carefully.

In addition, the alternative scenarios mentioned above may imply different relationships between FDI locations. In the export platform models, plant-level fixed costs create more incentive to have a single plant in one country and less incentive to expand into nearby countries. Of course, these savings must be balanced against trade costs that increase with distance, implying that the degree of substitution is a decreasing function of distance. Agglomeration economies with respect to other Spanish investments, on the other hand, suggest that proximity to other FDI increases the incentive to invest in nearby countries.⁸

This provides us with new information regarding the impact of agglomeration and substitution effects, as well as estimates that are more comparable to the bulk of the FDI literature which considers the level of FDI activity. Furthermore, we consider distance effects that extend beyond bordering locations, something that Head et al. (1995) do not do.

Finally, it is worthy to mention that previous findings evidenced the fact that spatial effects by and largely cease to be significant when country fixed effects are included

⁷ See Garretsen and Peeters (2008) for further details.

⁸ Agglomeration externalities may occur between any firms, but what matters in this context would be such externalities between Spanish investment and Spanish firms in neighbouring countries. See Blomström and Kokko (1998), for example, for a general discussion of how agglomeration economies may arise in the context of FDI.

(i.e. suggesting that spatial effects, if at all relevant, are mainly *cross-sectional*).⁹ In this sense, one of our aims is to provide empirical evidence whether spatial effects are no longer relevant when controlling for country-fixed effects. Similarly, our data set only contains aggregate annual outwards FDI and thereby is the summation of all FDI decisions undertaken by firms in a given year neglecting that these various FDI decisions may be the result from rather different motives.

3 Empirical model

To start with, the empirical modelling uses the Gravity Model approach as a point of reference, which is arguably the most widely used empirical specification of FDI. It has been modified based on the recent literature to include variables measuring host skill endowments and (surrounding) market potential.

In particular, variables are measured in logs:

$$FDI = \alpha + \beta_1 \text{ Host Variables} + \beta_2 \text{ Market Potential} + \rho WFDI + \varepsilon$$

where $\varepsilon \sim N(0, \sigma^2 I)$ and where FDI is a vector ($N \times 1$), with row j equal to log FDI from Spain—the parent country—to host country j . The variable $WFDI$ is the spatial weighted FDI whose coefficient ρ measures the intensity of FDI interdependences. *Host Variables* are defined as a matrix of k exogenous variables, ε white noise disturbance, and N number of observations. A traditional measure of *Market Potential* and an alternative *Surrounding Market Potential* (wider than the traditional approach) are sequentially included as well, since recent studies find such characteristics significant in explaining the observed variation in FDI. Finally, ρ becomes the autoregressive parameter which reflects the intensity of interdependences across sample observations. Given the skewness of our FDI data sample, the model is specified in log-linear form. This model leads to better-behaved residuals.¹⁰

We also estimate a spatial error model in order to understand the mechanism of transmission of shocks in terms of FDI flows.

$$FDI = \alpha + \beta_1 \text{ Host Variables} + \beta_2 \text{ Market Potential} + \varepsilon$$

where $\varepsilon = \lambda W\varepsilon + \mu$. Thus, a shock affecting the Spanish FDI outflows to host country i would have an impact on the Spanish FDI outflows to host country j . The impact magnitude (proxied by the spatial error coefficient, λ) depends on the distance between the host countries i, j measured by the weighting matrix. Furthermore, the related literature suggests that a spatial error model may be relevant.¹¹

⁹ See Blonigen et al. (2007).

¹⁰ Further technical details on recently developed spatial panel data estimation methods by ML are provided in the Appendix.

¹¹ See Coughlin and Segev (2000), Abreu (2005), Baltagi et al. (2007) and Garretsen and Peeters (2008).

3.1 Host variables

The main reason for including host variables is to capture the standard gravity-model variables for the host countries (GDP, population, distance from parent to host country, sharing a common language and trade/investment friction variables), as well as a measure of host skilled-labour endowments. Based on previous results in the literature, the priors reveal that the higher the *host GDP*, the higher the FDI. Holding GDP constant, increasing a country's *population* reduces its per capita GDP and thus FDI as well. Populations are therefore included to control for the known tendency for FDI to move between wealthy markets. Negative coefficients on population are to be expected. However, an agglomeration effect would lead to upward pressure on this parameter, and, the coefficient result would therefore be ambiguous. With regard to *trade costs*, if FDI is undertaken to exploit vertical linkages, then higher host trade costs reduce its value. Alternatively, if FDI is primarily horizontal and intended to replace Spanish exports, then higher host trade costs should induce tariff-jumping FDI. Thus, the effect of trade costs becomes ambiguous. As in the traditional gravity model, *distance* between the home country (Spain) and host countries is included, which may capture both higher management costs (which reduce FDI) and higher trade costs (with an ambiguous effect). Finally, a country dummy trying to capture cultural links such as sharing a common language is considered. However, their time-invariant condition will restrict their estimation results after controlling for fixed effects.

3.2 Market potential variable

Following [Blonigen et al. \(2007\)](#), the surrounding market potential variable for a country j is defined as the sum of (inverse) distance-weighted real GDPs of all other $k \neq j$ countries in the world by year. Since a great deal of work has focused on the robustness of the market potential weighting scheme, the MP weight matrix used is driven by 224 countries and not only by sample ($n = 50$) countries. This approach is similar to [Head and Mayer \(2004\)](#)'s measure of the market potential of neighbouring regions. Thus, for the construction of this variable we do not employ exactly the same set of weights as those used for the spatial lag term of *WFDI* (whereby spatial interdependences are measured), but the functional form on distance is the same. Note that there is little theory to guide the choice of weights. The empirical analysis section explores the robustness of our results to various weighting schemes and market potential variables.

By extension, a "traditional" market potential variable is introduced as an alternative, based on the sum of host and weighted GDPs of other host countries. The host region's GDP is added as a proxy to capture the "traditional" market potential effects, even though lower identification power may arise. The data do not clearly reject a common coefficient on host GDP but the focus will be on surrounding market potential in order to better identify the various forms of FDI.¹² The spatial lag and

¹² Even though host GDP may become an important factor for both pure horizontal and export-platform Spanish outflows motivations, the market potential surrounding a host region should have an impact only on export-platform MNE activities.

the traditional and surrounding market potential coefficients' signs and significance will provide evidence on different FDI motivations. Hence, the expected sign for the (surrounding) market potential variable is not clear a priori and will depend mainly on the motivations of Spanish MNEs to invest abroad.

3.3 Spatially dependent FDI

In general, it comes as no surprise that regional science theories focus on spatial interactions since “everything is related to everything else, but near things are more related than distant things”.¹³ So, there are several reasons why we might be interested in spatial effects in fitting data with a spatial model (see [Anselin 1988](#)). First, a spatial autocorrelation or “spatial error” model places additional structure on the unobserved determinants of FDI which would otherwise be captured by the traditional error term.¹⁴ Second, and of particular interest in testing the theories of FDI offered above, the estimation of a spatial autoregressive or “spatial lag” model accounts directly for relationships between dependent variables that are believed to be related in some spatial way. As such, these methods allow the data to reveal patterns of substitution or complementarity, as well as the strength of any such patterns, through the estimated spatial lag coefficient.

Spatial dependence is multidirectional (a region may be affected not only by other proximate region but also by some other neighbouring regions, and this region may influence the others as well). In other words, the main problem the spatial context faces is the border effect, which arises as a consequence of the meaning of spatial dependence, not only limited to sample regions but also extended to spatial units for which information is not available ([Griffith 1985](#)). As [Florax \(1992\)](#) points out, there is no single, commonly accepted solution to this problem.

The solution of the multidirectionality in the spatial context comes from the definition of the weight spatial matrix, W :

$$W_y = \begin{bmatrix} 0 & w_y(d_{i,j}) & w_y(d_{i,k}) \\ w_y(d_{j,i}) & 0 & w_y(d_{j,k}) \\ w_y(d_{k,i}) & w_y(d_{k,j}) & 0 \end{bmatrix}$$

W is a quadratic and non-stochastic matrix whose elements w_{ij} reflect the intensity of the existing interdependence between each pair of regions i and j .

To define the mentioned weights, recall the inexistence of a unanimously accepted W , although these weights must be non-negative and finite ([Anselin 1980](#)). Despite this, researchers usually resort to the first order physical contiguity concept, used initially by [Moran \(1948\)](#) and [Geary \(1954\)](#), where w_{ij} is equal to 1 if regions i and j are physically adjacent and 0 otherwise. Although the contiguity matrix is often used due

¹³ Tobler's first law of geography (1970).

¹⁴ Spatially correlated errors are analogous to clustering error terms where the OLS assumption of independence between all errors is relaxed. Instead, here we assume that while the errors are independent across groups they need not be independent within groups.

to its simplicity, it has some serious limitations which impose an excessive number of restrictions.

In our case, then, we apply a simple inverse distance function to define the weights of the spatially dependent variables (i.e. *WFDI* and *Market Potential*), where the shortest distance within the sample is assigned a weight of unity, following [Blonigen et al. \(2007\)](#), [Garretsen and Peeters \(2008\)](#), and [Hall and Petroulas \(2008\)](#). Moreover, a common practice of spatial models is to row-standardize the weight matrix.¹⁵

4 Data

For the various spatial panel data models estimated, we use a panel of annual data on Spanish FDI activity in the countries where Spain made high investments in the 1993–2004 period (i.e. top-50 host countries). Our focus will be limited to 50 host countries (top-50 FDI receivers) for at least two reasons.

First, focusing on these sample countries will probably limit vertical specialization as a primary motivation of FDI, allowing us to better disentangle the factors behind any spatial interdependence. In this sense, it may simplify the comparisons between results on FDI motivations. As demonstrated by [Markusen and Maskus \(2001\)](#) and [Blonigen et al. \(2003\)](#), these sub-samples, which include the large majority of FDI activity provide robustness. The cost of this, however, is that it assumes that the excluded countries exert no influence on FDI patterns within the remaining data. For the European countries, this might raise special concern, due to the increased openness of the Central and Eastern European countries during the nineties. Nevertheless, the Spanish FDI towards these specific countries during the whole period was negligible, as shown in the following subsection. Second, limiting to 50 countries is due to estimations convergence constraints since we are dealing with ML spatial panel data models. However, note that these countries account for the lion's share of Spanish FDI throughout the world. In particular, for the years in our sample, these countries hosted an average of 95% of Spanish outbound FDI.

We restrict ourselves to outbound data from a common parent country, Spain. Existing FDI theory provides obvious reasons to expect that a parent country's FDI in host markets is interdependent, but little attention has been paid to the interdependence of FDI decisions by parent countries in a common host country (although if one considers competition in goods or host-country factor markets, there could well be such a link). This may be the reason why [Markusen and Maskus \(2001\)](#) and [Blonigen et al. \(2003\)](#) find that the determinants of FDI activity for US inbound and outbound data yield very different estimates.

The dependent variable FDI is measured by the aggregated gross effective outflows in millions of euros as reported in the foreign investment register (RIE) of the Ministry of Industry, Tourism and Trade. Conversion using a price index of gross fixed capital formation (Penn World Tables, PWT 6.2) is computed. The cumulated sum of gross effective outflows is taken as a proxy of the outward stock of capital.

¹⁵ In unreported results, several tests of alternative weighting schemes were also applied. In general, these tests yielded broadly similar results and are available upon request.

Table 2 Descriptive statistics

Variable	Mean	SD	Min	Max
Real FDI (in thousands)	2,661,806	6,850,836	0	40,200,000
Host population (thousands)	80,201	218,523	393	1,294,846
Host trade costs	0.019	0.011	0.003	0.063
Host GDP (in millions)	708,000	1,510,000	13,200	10,700,000
Host distance from Spain in km	5,488	3,876	501	17,593
Host skills	37.44	19.90	4.30	89.90
Common language	0.30	0.46	0	1
Trend (1993 = 1)	6.50	3.45	1	12
Trend ²	54.17	46.14	1	144
Traditional market potential	199,000	84,600	85,000	685,000
Surrounding market potential (in millions)	667,000	1,440,000	4,015	10,300,000

The reason comes from the long-run modelling specification. In other words, using an FDI aggregate measure instead of FDI flows relies on the assumption that MNE investment strategies are known to be long run decisions, potentially non-fully captured by temporal flows.¹⁶

The set of explanatory variables included in the spatial interactive model are the following: host country real domestic product (GDP) in current prices and population data come from the PWT, which reports these data for 1950 through 2004. The host trade-cost measure is the inverse of the openness measure, which itself is equal to exports plus imports divided by GDP. To control for distance, we followed the literature by using great circle distances between capital cities, measured in kilometres, which are drawn from the CEPii database. Host country skills are measured as the gross enrolment rate (tertiary education) provided by the World Development Indicators from the World Bank. Linear interpolation was used for several years. Finally, host (surrounding) market potential variable is measured as the inverse-distance-weighted real GDP of other host countries ($n = 224$). By extension, a “traditional” market potential variable is introduced as an alternative based on the sum of inverse-distance-weighted host GDP and weighted GDPs of other host countries.

The final sample spans from 1993 to 2004 for 50 countries. See data appendix for data specific definitions and sources, and Table 2 for summary descriptive statistics of the variables.

4.1 Spanish FDI spatial distribution

In order to investigate spatial interdependences of Spanish FDI outflows several spatial statistics techniques have been applied to the data. FDI geographical patterns have evolved since 1993: the clustering process detected in the second half of the nineties

¹⁶ Bajo and Montero (1999).

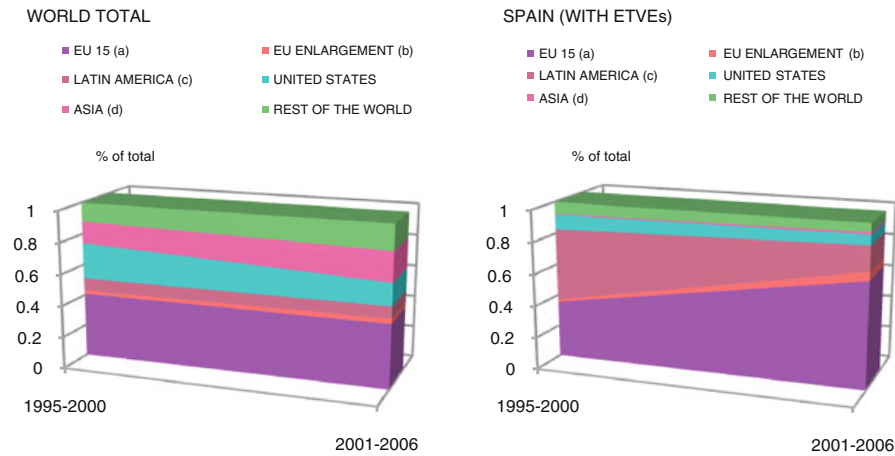


Fig. 3 Spanish and world outward foreign direct investment transactions. Breakdown by geographical area. Source: Bank of Spain and UNCTAD. *a* The EU 15 includes the euro area (excluding slovenia), the United Kingdom, Sweden and Denmark. *b* Candidates for EU enlargement in 2004 (Cyprus, Czech Republic, Estonia, Hungary, Lithuania, Latvia, Malta, Poland, Slovenia and Slovakia) plus Rumania and Bulgaria. *c* Includes South America, Central America (excluding Belize), Mexico, Cuba, Haiti and the Dominican Republic. *d* Turkey is not included

became more diversified as time went by, as explained in Sect. 3.3. Similarly, the Spanish spatial distribution pattern of FDI outflows behaves quite differently to FDI world transactions (see Fig. 3).

In this regard, the second half of the nineties was marked by a process of internationalization of Spanish firms in Latin American markets. The driving engines of the outbound flows were based on the deregulation of several sectors,¹⁷ the privatization process of state-owned companies and, probably, the access to expanding markets. Undoubtedly, cultural similarities played a key role as well.¹⁸

Quantitatively speaking, while 45% of total Spanish outward FDI targeted Latin American markets, EU15 countries attracted nearly 40% during this period. This *centre-periphery* pattern was completed by less significant flows to host countries such as the US, in contrast to worldwide trends (see Fig. 3).¹⁹

In recent years, closer regions have emerged as the main targets for Spanish FDI outflows, accounting for 85% of the total amount of OECD members in 2006. Broadly speaking, investment in European markets was revitalized, in particular, in the United Kingdom. Spanish firms became less oriented to emergent economies, such as Asian and Eastern European countries, than to worldwide flows.²⁰

¹⁷ For further details on Latin American investments, see López-Duarte and García-Canal (2002).

¹⁸ As empirically proved by Barrios and Benito-Ostolaza (2009).

¹⁹ For a detailed breakdown at the firm level see Guillén (2004) and Santiso-Guimaras (2007).

²⁰ A broader explanation for eastern European countries is provided by Turrión and Velázquez (2004).

4.2 Exploratory spatial data analysis

The use of econometric techniques under an exploratory spatial data analysis framework was also considered, in an attempt to perform a more in-depth study of the geographical distribution of Spanish FDI activity. It is quite useful to test the likely presence of spatial interdependences across regions in terms of FDI performance. Specifically, we compute the *Moran I test* to see if a random distribution exists or if, on the contrary, closer countries tend to show similar FDI patterns. Spatial dependence, or spatial autocorrelation, is said to exist when the values observed at one location (for instance, in one country) depend on the values observed in its neighbouring locations. Although various statistics have been proposed for verifying the existence of spatial autocorrelation in a specific variable, one of the most widely used is the *Moran I test* (Moran 1948), which is computed as follows:

$$I = \frac{N}{S} \frac{\sum_i \sum_h w_{ih} z_i z_h}{\sum_i z_i^2}$$

where N is the number of observations, w_{ih} is the element of the spatial weights matrix W that expresses the potential interaction between two regions i and h , S is the sum of all the weights (all the elements in the weights matrix) and z_i represents the normalized value of a variable x being analyzed in region i .

A significant and positive value for this statistic indicates a trend for similar values of the variable to *cluster* in space (positive spatial dependence). On the other hand, when the test is significant and negative, the trend is for dissimilar values to cluster in neighbouring locations (negative spatial dependence). The latter case might illustrate a situation where the strength of centripetal forces within the region is such that it prevents the diffusion of FDI activities to its neighbours. Non-significance of the *Moran I test* implies the acceptance of the null hypothesis, that is, the non-existence of spatial autocorrelation, indicating the prevalence of a random distribution of the variable throughout space.

Once we have obtained the indices from 1993 to 2004 in order to study the evolution of the concentration pattern of Spanish FDI activity around the world, we are able to provide the expected significant interdependence results on clustering FDI outflows during the early nineties, losing track from then on due to dispersion motivations (see Table 8 in Appendix).

5 Results

5.1 Full sample panel data results

Table 3 presents our initial results with column (1), showing the OLS results of our pooled model without the variables that may capture potential spatial patterns in the data: that is, spatial lag or (surrounding) market potential variables. Columns (2) and (3) present OLS estimates that sequentially add the market potential variables starting from not including the spatial lag. One reason for sequentially adding in the spatial lag and the surrounding market potential variable is to be able to examine the potential

Table 3 Spatial analysis of Spanish outflows FDI—panel data model

	Full sample					
	Pooled (1)	Pooled (2)	Pooled (3)	RE (4)	FE (5)	2WFE (6)
Ln (host population)	−1.6186 [0.1835]***	−1.5143 [0.1866]***	−1.7909 [0.1812]***	−0.9760 [0.3821]***	4.8507 [2.3030]**	4.9735 [2.2982]**
Ln (host GDP)	2.5821 [0.1933]***	2.4121 [0.2024]***	3.6666 [0.2702]***	2.4830 [0.6565]***	−5.2214 [3.3230]	−4.8570 [3.3757]
Ln (host trade cost)	−0.6493 [0.2502]***	−0.3583 [0.2715]	−0.5167 [0.2447]**	−0.1889 [0.4062]	−0.0290 [0.4827]	0.0655 [0.4856]
Ln (distance from Spain in km)*	−0.8701 [0.1395]***	−0.8675 [0.1387]***	−0.6992 [0.1391]***	−0.9080 [0.3987]**	—	—
Ln (host skills)	−0.0305 [0.0066]***	−0.0302 [0.0066]***	−0.0331 [0.0065]***	−0.0111 [0.0093]	0.0001 [0.0113]	0.0012 [0.0113]
Common language	3.8242 [0.3077]***	3.8313 [0.3060]***	3.6495 [0.3011]***	3.1973 [0.8107]***	—	—
Trend (1993 = 1)	0.8528 [0.1147]***	0.8159 [0.1149]***	0.8701 [0.1117]***	0.9746 [0.0705]***	0.9473 [0.0851]***	0.9177 [0.1091]***
Trend ²	−0.0250 [0.0086]***	−0.0230 [0.0086]***	−0.0258 [0.0083]***	−0.0325 [0.0051]***	−0.0327 [0.0051]***	−0.0287 [0.0073]***
Traditional market potential	—	0.9341 [0.3487]***	—	—	—	—
Surrounding market potential	—	—	−1.0499 [0.1877]***	−0.8247 [0.5313]*	5.9262 [3.5309]*	5.4489 [3.5887]
Constant	−21.2309 [3.0479]***	−35.5192 [6.1356]***	−20.9986 [2.9668]***	−8.6572 [6.1744]	−54.2341 [27.2138]***	−53.1642 [27.2030]***
Country dummies	No	No	No	No	Yes	Yes
Observations	600	600	600	600	600	600
Adj <i>R</i> ²	0.5660	0.5709	0.5889	0.7463	0.7524	0.7606

Standard error in parentheses

* Significant at 10%, ** significant at 5%, *** significant at 1%

power explanation of the spatially dependent variables and the omitted variable bias. If country-*k* GDP correlates with FDI in country *k* and also with FDI in country *j*, then including country-*k* FDI in the prediction of *j*'s FDI (e.g. through ρ *W* FDI) while not directly including country-*k*'s GDP leaves the estimation of ρ prone to bias. Of course, including market potential without ρ *W* FDI would also yield biased estimates of the effect of market potential (e.g. [Head and Mayer 2004](#)). Column (4) shows the results for the random effects model, column (5) for the fixed effects model and column (6) for the two-way fixed effects model.²¹

²¹ Common language variable has been omitted in Tables 4 and 5 since once included, convergence by ML became harder to achieve and the statistical gain was hardly noticeable.

Several interesting observations emerge from Table 3. Since we are mainly concerned with the relevance of space for Spanish FDI, the market potential variable is of particular interest.

The traditional market potential variable sequentially added may be rejected by the data in favour of including separate terms for host-country GDP and the surrounding market potential (since $\text{Adj-}R^2$ falls from 0.59 to 0.57). Moreover, the unexpected negative sign of the parameter is inconsistent with all of the MNE motivations discussed above. Likewise, the introduction of unobservable effects with country dummies (i.e. controlling for time-invariant unobserved effects specific to each country) substantially reduces the statistical significance of the surrounding market potential. We will provide alternative hypotheses for this result below.

Once the country specific dummy variables are introduced, some standard gravity model variables such as host GDP and host skills become statistically insignificant. However, in line with the previous empirical literature, the host population prevails (with an elasticity of 4.8, given the variables are in logs). A very significant non-linear time trend captures the home country time series variation in FDI. To sum up, in agreement with the literature on Spanish FDI determinants, distance retains a negative sign across models and sharing a common language boosts FDI (Barrios and Benito-Ostolaza 2009).

Both OLS and RE estimations had to be rejected in favour of the Fixed Effects model after obtaining Hausman's specification test and F-tests for Random Effects. In order to control for the events which might affect all the countries in the sample (i.e. international financial constraints, recessions, etc.), temporal dummies have been estimated by two-way fixed effects. Nevertheless, no substantial statistical benefit emerges.

The results therefore suggest that Spanish FDI outflows are affected not only by a host country's large market potential, but also by (inverse) distance-weighted relatively large GDP levels of surrounding countries.

5.2 Spatial (panel) lag model results

Table 4 shows the estimation results allowing for a spatial lag model. Pooled, Random Effects and Fixed Effects models are presented by adding market potential variables sequentially.

In agreement with Blonigen et al. (2007) and Garretsen and Peeters (2008), the spatial lag coefficient decreases dramatically once fixed effects are included: from 0.50 to 0.14 in the former, from 0.44 to 0.07 in the latter, and from 0.81 to 0.31 in our case. Intuitively, these results are due to the fact that spatial autocorrelation may be captured by country dummies as well. However, the spatial lag parameter (ρ) is still clearly significant. Therefore, one might infer the following interpretation of the spatial lag coefficient (column 2): there is approximately an 8% increase in Spanish FDI into a host country x for a 10% increase in the distance-weighted FDI going into surrounding markets (average distance 5,500 km approx.). Nonetheless, in contrast to the findings of Blonigen et al. (2007), our spatial effects for Spanish FDI outflows do not appear to be completely cross-sectional.

In order to tackle this question and to analyze FDI theories and Spanish motivations to invest abroad (outlined in Table 1), column (6) shows the most representative

Table 4 Spatial lag specification of Spanish outflows FDI

	Full sample					
	SAR (1)	SAR (2)	SAR RE (3)	SAR RE (4)	SAR FE (5)	SAR FE (6)
Ln (host population)	−0.0284 [0.1403]	−0.0513 [0.1547]	0.2005 [0.2679]	0.1983 [0.2834]	6.3208 [2.3409]***	6.3239 [2.3432]***
Ln (host GDP)	0.2775 [0.0811]***	0.2549 [0.1041]**	0.1087 [0.0934]	0.1061 [0.1398]	0.0727 [0.0972]	0.0766 [0.1563]
Ln (host trade cost)	0.6501 [0.2685]**	0.6241 [0.2790]**	0.3744 [0.3491]	0.3698 [0.3933]	0.1941 [0.3707]	0.2013 [0.4330]
Ln (distance from Spain in km)*	−0.4133 [0.1244]***	−0.4069 [0.1258]***	−0.2467 [0.1992]	−0.2463 [0.1998]	—	—
Ln (host skills)	−0.0027 [0.0061]	−0.0035 [0.0065]	0.0133 [0.0089]	0.0133 [0.0090]	0.0224 [0.0100]***	0.0224 [0.0100]***
Trend (1993 = 1)	0.2018 [0.1432]	0.2041 [0.1433]	0.6693 [0.1327]***	0.6690 [0.1329]***	0.7534 [0.1457]***	0.7538 [0.1465]***
Trend ²	−0.0048 [0.0105]	−0.0050 [0.0105]	−0.0256 [0.0073]***	−0.0256 [0.0074]***	−0.0304 [0.0074]***	−0.0304 [0.0074]***
Surrounding market potential	—	0.0598 [0.1713]	—	0.0070 [0.2822]	—	−0.0106 [0.3329]
Constant	2.2716 [1.8291]	1.6639 [2.5409]	2.2163 [2.9951]	2.1334 [4.4255]	−58.0186 [29.8336]***	−57.8958 [29.8295]***
Spatially weighted FDI (ρ)	0.8130 [0.0342]***	0.8120 [0.03445]***	0.4650 [0.0816]***	0.4650 [0.0816]***	0.3190 [0.0954]***	0.3190 [0.0954]***
Country dummies	No	No	No	No	Yes	Yes
Observations	600	600	600	600	600	600
Adj R^2	0.5468	0.5468	0.8341	0.8342	0.8487	0.8487

Standard error in parentheses

* Significant at 10%, ** significant at 5%, *** significant at 1%

spatial lag specification. The combination of a statistically zero market potential coefficient and a positive spatial lag coefficient may suggest the existence of agglomeration effects across country borders amongst Spanish investing firms (consistent with complex-vertical motivations for MNE activity). However, although one might expect geographical clusterings of Spanish FDI mainly for supply reasons, market potential sign might be capturing demand or market-size reasons. As previously stated, it is open to debate. In any case, complex (vertical) FDI models with agglomeration economies may prevail, as the evolution of the Spanish FDI outflow data suggest.

In line with the findings of [Baltagi et al. \(2007\)](#) and [Blonigen et al. \(2007\)](#), if spatial interactions are stable over time, country dummies may be capturing spatial effects as

well. Interestingly, when fixed effects are included the surrounding market potential variable becomes statistically zero, rather than the more puzzling negative sign.²²

Though in the empirical FDI literature it is hard to find similar evidence of third-country effects that are stable over time captured by country dummies, we find an analogy in the international trade literature. [Feenstra \(2002\)](#) found that third-country interdependence in gravity model estimation (pointed out by [Anderson and van Wincoop 2003](#)) may be well captured under a panel with a country-level fixed effects framework.

As far as the population variable is concerned, its positive coefficient and its undoubted statistical importance remain clear even though the reason for this result is unfortunately hard to grasp. Population growth would seem to discourage FDI due to the wealth effect, mostly reflected in GDP per capita, as mentioned in Sect. 3.1. Nonetheless, what if those investments were focused on services? Greater population might mean higher returns for those Spanish investors who decide to enter these markets, even if the GDP per capita of potential consumers is lower. Agglomeration theories, then, would exert upward pressure on this parameter.

In any case, by observing the statistical evolution of Host GDP variable one might (carefully) conclude that Spanish FDI abroad does not seem to be related to the expansion in markets with very high growth rates.

With respect to the strength of the spatial lag relationship, our estimates reveal that, on average, FDI invested in the average country in our sample is positively associated with proximity-weighted FDI in other countries. The data support the notion that spatial autoregression does not vary to a large extent across time ([Garretsen and Peeters 2008](#)).

By comparing the spatial fixed effects lag model with market potential variables (column 6) to the FE model (column 5, Table 3), we observe that the standard determinants results are robust to the inclusion of a spatial lag.

Previous studies include a common observation referring to the sensitivity of results to the selection of the host countries. By re-estimating the basic spatial lag and spatial error models for different subsamples (European, Latin American and OECD countries) broadly similar conclusions are obtained. However, it is especially interesting that spatial linkages do not vary in terms of relevance. To sum up, the sensitivity analysis conducted with different weighting matrices yields very similar results and adds robustness to our estimations.

5.3 Spatial (panel) error model results

So far, we have focused our attention on the spatial lag model since its results may provide evidence consistent with several FDI theory motivations. However, apart from the channels already identified by FDI theory, other transmission mechanisms of

²² A hypothesis for a negative sign of the surrounding market potential parameter is the negative competitive impact of firms in these neighbouring markets. This may happen when companies in surrounding countries have greater competitive advantages for serving the host market than Spanish firms.

Table 5 Spatial error specification of Spanish outflows FDI

	Full sample					
	SEM (1)	SEM (2)	SEM RE (3)	SEM RE (4)	SEM FE (5)	SEM FE (6)
Ln (host population)	0.1981 [0.1402]	0.1300 [0.1484]	−0.0150 [0.2726]	0.0059 [0.2923]	6.3755 [2.5224]***	6.3970 [2.5262]***
Ln (host GDP)	0.2106 [0.0800]***	0.1138 [0.1057]	0.1236 [0.0946]	0.1454 [0.1442]	0.0713 [0.0974]	0.0839 [0.1577]
Ln (host trade cost)	0.1187 [0.2636]	−0.0160 [0.2799]	0.4574 [0.3533]	0.4957 [0.4010]	0.2080 [0.3703]	0.2309 [0.4337]
Ln (distance from Spain in km)*	−0.8553 [0.1707]***	−0.8572 [0.1705]***	−0.2004 [0.2087]	−0.2023 [0.2088]	—	—
Ln (host skills)	0.0070 [0.0063]	0.0048 [0.0065]	0.0114 [0.0092]	0.0116 [0.0093]	0.0212 [0.0101]**	0.0212 [0.0101]**
Trend (1993 = 1)	1.1768 [0.9093]	1.1820 [0.9202]	1.2586 [0.1324]***	1.2604 [0.1324]***	1.1534 [0.1188]***	1.1546 [0.1193]***
Trend ²	−0.0428 [0.0681]	−0.0438 [0.0689]	−0.0468 [0.0099]***	−0.0468 [0.0099]***	−0.0450 [0.0084]***	−0.0450 [0.0084]***
Surrounding market potential		0.2324 [0.1669]		−0.0582 [0.2928]		−0.0340 [0.3344]
Constant	7.9732 [3.2447]***	5.6508 [3.6733]	6.7999 [3.0990]**	7.4388 [4.4555]**	−56.7085 [29.8623]***	−56.4305 [29.8482]***
Spatial autocorrelation (λ)	0.8550 [0.02722]***	0.8570 [0.0268]***	0.3728 [0.0952]***	0.3707 [0.0954]***	0.2900 [0.0987]***	0.2889 [0.0988]***
Country dummies	No	No	No	No	Yes	Yes
Observations	600	600	600	600	600	600
Adj R^2	0.3763	0.3703	0.833	0.8329	0.8452	0.8452

Standard error in parentheses

* Significant at 10%, ** significant at 5%, *** significant at 1%

shocks may arise.²³ By estimating a model which allows for spatial autocorrelation in the error term, the significance of the autocorrelation coefficient (λ)²⁴ would provide consistent evidence on whether or not a shock in the Spanish FDI to the host country j ($\neq i$) may have an impact on Spanish FDI to host country i , where the magnitude of the impact will depend on the weighted distance (W) between the two countries i and j . Regarding the question of whether it makes sense or not to include the spatial autocorrelation model, the LM test for the spatial error specification (against the pooled

²³ See Coughlin and Segev (2000), Abreu (2005), Baltagi et al. (2007) and Garretsen and Peeters (2008) who defend the relevance of the spatial error model.

²⁴ Coming from $\varepsilon_t = \lambda W\varepsilon_t + \mu$ as specified in Sect. 3.

model without spatial effects) failed to reject the null (spatial error term inclusion) hypothesis.

Table 5 provides the estimation results for the spatial error model by sequentially adding the market potential variable. Since we are mainly concerned with the relevance of spatial links for Spanish FDI outflows, the spatial autocorrelation variable takes on particular interest.

Even though the spatial autocorrelation coefficient is far from insignificant and positive, the spatial error model does not provide sufficient evidence to test the substitution or complementarity of FDI across countries. Nonetheless, according to this result, it comes as no surprise that shocks to Spanish FDI outflows to third country j may influence the Spanish FDI outflows to host country i . Regardless of the model (and subsamples) estimated, the coefficient is invariably statistically significant.

Once the country dummies are included, host skills and host population emerge as significant variables while the spatial autocorrelation coefficient (even though significant) declines from 0.85 to 0.29. This may suggest that apart from the channels already identified by FDI theory, some other transmission channels of shocks may arise.

Following the empirical results shown in Table 5 one may conclude that Spanish FDI outflows would be reliably reflected by spatial autocorrelation patterns. However, the significance of the spatial error coefficient may sometimes be driven by a mis-specification of the underlying model in terms of omitted variables. As a result, the spatial error specification might be considered a “catch-all” for omitted spatially autocorrelated regressors.²⁵ Hence the preference for the spatial lag model.

6 Conclusions

In this essay, spatial econometric techniques are used to analyze the patterns of Spanish FDI outflows from 1993 to 2004. The vast majority of previous empirical work has examined bilateral data while ignoring the potential interdependence in FDI across locations. A few recent exceptions, however, have used multilateral approaches in this context, i.e. Coughlin and Segev (2000), Abreu (2005), Baltagi et al. (2007), Blonigen et al. (2007), Garretsen and Peeters (2008) and Hall and Petroulas (2008).

The related research has so far focused on FDI flows of a few countries. However, Spanish FDI outflows have risen dramatically since 1995 and today account for a substantial part of global FDI.²⁶ Previously a net importer of FDI, the Spanish economy as a whole is now a net exporter.

The hypotheses under analysis in this study question both the relevance of space in Spanish MNEs’ investments abroad and the nature of the firms’ motivations. We estimate several Spatial Panel Data models recently developed by Elhorst (2003, 2009) using ML for Spanish outward FDI into top-50 host countries. Secondly, no structure is imposed to isolate one particular multilateral effect (i.e. horizontal, vertical,

²⁵ See Fingleton and López-Bazo (2006).

²⁶ The relative weight of Spanish investment in world FDI rose to approximately 6% on average between 2001 and 2006.

export-platform, etc.); rather, the net effects of these forces are estimated and related to specific FDI theoretical motivations.

The paper's approach is similar to [Blonigen et al. \(2007\)](#) and [Garretsen and Peeters \(2008\)](#) since, first and foremost, a spatial lag model is estimated in order to test whether spatial effects remain relevant after controlling for fixed effects. As a result, we found evidenced significant omitted variable bias for the spatially dependent variables, namely, surrounding market potential measure and spatial lag. This result has an important bearing on previous works on Spanish FDI determinants, since this is the first attempt to include spatial effects.

Consistent with our priors and based on an exploratory spatial data analysis, the data support the notion that spatial linkages for Spanish FDI outflows do not vary to a large extent across (sample) time ([Garretsen and Peeters 2008](#)). Additionally, spatial error model estimations suggest that apart from channels already identified by FDI theory (based on spatial lag and market potential signs and significance) some other transmission channels of shocks across Spanish FDI outflows may arise.

Finally, after controlling for unobservable effects, we conclude that spatial linkages matter for Spanish outbound FDI. We also find evidence consistent with NEG theories of agglomeration, mainly due to complex (vertical) FDI motivations. Thus, Spanish multinational firms decisions have contributed to generate geographical clusters of FDI abroad based on supply factors: conducting vertical fragmentation processes by seeking out (low cost) suppliers in multiple (closer) countries. However, these results may be data-driven and disaggregated sub-samples across sectors must be studied in greater depth. The results also suggest the difficulty of disentangling the channels through which third-country effects affect Spanish outbound FDI and how these channels may vary across space and time. Introducing the variables such as corporate taxation and foreign capital restrictions would be a useful extension to the literature.

Acknowledgments I am extremely grateful to Rosina Moreno, Ferdinand Paraguas, Charles Plagin and J.Paul Elhorst for their replies to my inquiries, and the seminar participants at the 2007 European Regional Science Association Summer Institute at the University of Bratislava for their comments and suggestions. I owe my gratitude as well to two anonymous referees for valuable comments that substantially improved the analysis of this paper. Special thanks to Enrique López-Bazo for his invaluable advice. All errors or omissions are my sole responsibility.

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution and reproduction in any medium, provided the original author(s) and source are credited.

Appendix: Estimation of spatial panel models

The estimation of panel data models that include spatially lagged dependent variables and/or spatially correlated error terms follows as a direct extension of the theory developed for the single cross-section.²⁷ In the former, we must deal with the endogeneity problem of the spatial lag, while in the latter, we must account for the non-spherical nature of the error variance covariance matrix.

²⁷ See [Anselin et al. \(2008\)](#).

Even though a moments approach method is also suggested in the literature by Kapoor et al. (2007), our approach is focused on the ML principle. Our estimations are based on a model with a parameterized form for spatial dependence, specified as a spatial autoregressive process. In practice, ML estimation consists of applying a non-linear optimization to the log-likelihood function, which yields a consistent estimator from the numerical solution to the first order conditions. Thus, asymptotic inference is based on asymptotic normality, with the asymptotic variance matrix derived from the information matrix. As usual, the second order partial derivatives of the log-likelihood are required. However, computing a Jacobian determinant in single cross-section becomes a problem for the implementation of ML estimates, so the classic solution in panel data models is to decompose the Jacobian in terms of the eigenvalues of the spatial weights matrix even though this is computationally costly.

The estimation procedure has focused on controlling heterogeneity sequentially within a panel (with and without spatial effects). From a pooled model where the intercept is common for all cross-section units:

$$Y_{it} = \alpha + \beta_1 X_{lit} + \varepsilon_{it}$$

to a random effects model, controlling for the “individual” performance of each unit.

$$Y_{it} = \alpha + \beta_1 X_{lit} + \mu_i + \varepsilon_{it}$$

The random effects model allows the assumption whereby each cross-section unit has a different intercept. In other words, instead of considering α as fixed, it is taken as a random variable with an average value α and a standard deviation μ_i .

Another way of modelling the “individual” feature of every region involves the use of the fixed-effects model. This model does not allow for different random values across regions, but considers them as constant or fixed by estimating each single intercept. A fixed-effects model has also been estimated:

$$Y_{it} = v_i + \beta_1 X_{lit} + \varepsilon_{it}$$

where v_i is a vector of binary dummy variables for each region.

In both cases, (robust) LM tests for Random Effects and F tests for Fixed Effects were driven as well. The RE model may be tested against the FE model using Hausman’s specification test (with and without spatial effects). Since some common events may have affected all sample countries during the period 1993–2004, the model estimation by two-way fixed effects would reduce biased results.²⁸

However, we focus our attention mainly on the spatial (panel) lag FE model and the spatial (panel) error FE model. Like Hausman’s specification test, the SAR model may be tested against the SEM model using LM tests. For further details on Hausman’s specification’s test and Random Effects SAR and SEM models (see Elhorst 2009) (see Table 6).

²⁸ The two-way fixed effects model estimated is: $Y_{it} = v_i + \eta_t + \beta_1 X_{lit} + \varepsilon_{it}$, where η_t represents a vector of dummy variables for each year.

Table 6 Data sources and definitions

Variable	Description	Source	Website
Real FDI	Aggregated annual FDI in millions of € converted with a price index of gross fixed capital formation (PWT)	Spanish Ministry of Industry, Tourism and Trade (DataInVex)	http://datainvex.comercio.es/
Host GDP	Real GDP in current prices in millions of dollars	Penn World Tables (PWT)	
Host population	Population in millions	Penn World Tables (PWT)	http://pwt.econ.upenn.edu/php_site/pwt62/pwt62_form.php
Trade costs	Inverse of the openness measure	Penn World Tables (PWT)	
Distance	Great circle distances between capital cities, measured in kilometers	Centre D'Etudes Prospectives et D'Informations	http://www.cepii.fr/anglaisgraph/bdd/distances.htm
Host skills	Gross enrollment rate (tertiary education). Linear interpolation 1993–1994.	Internationales(CEPii) World Development Indicators (World Bank)	http://web.worldbank.org/
Traditional market potential	Distance-weighted host and weighted gross domestic product of other host countries in the sample ($n = 224$) (host + weighted GDPs)	PWT/CEPii	http://www.cepii.fr/anglaisgraph/bdd/distances.htm
Surrounding market potential	Distance-weighted gross domestic product (weighted GDPs)	PWT/CEPii	http://www.cepii.fr/anglaisgraph/bdd/distances.htm

A.1 Spatial (panel) lag model estimation

As mentioned above, a spatial lag model includes a spatially lagged dependent variable on the regression specification (Anselin 1988).

$$y = \rho(I_T \otimes W_N)y + X\beta + \varepsilon$$

where the ρ parameter is the spatial autoregressive coefficient and measures the intensity of interdependences across regions.

In a cross-section setting, a spatial lag model is typically considered as the formal specification for the equilibrium outcome of a spatial or social interaction process, in which the value of the dependent variable for one agent is jointly determined with that of the neighbouring agents.

At first sight, the extension of the spatial lag model to a panel data context would presume that the equilibrium process at hand is stable over time (constant ρ and constant W). Nevertheless, the inclusion of the time dimension allows much more flexible specifications.

Let us consider the pooled spatial lag model given in the above equation. Assuming a Gaussian distribution for the error term, with $\varepsilon_t \sim N(0, \sigma_\varepsilon^2 I_{NT})$, the log-likelihood (ignoring the constants) follows as:

$$L = \ln |I_T \otimes (I_N - \rho W_N)| - \frac{NT}{2} \ln \sigma_\varepsilon^2 - \frac{1}{2\sigma_\varepsilon^2} \varepsilon' \varepsilon$$

where $\varepsilon = y - \rho(I_T \otimes W_N)y - X\beta$, and $|I_T \otimes (I_N - \rho W_N)|$ as the Jacobian determinant of the spatial transformation.

Given the block diagonal structure of the Jacobian, the log-likelihood further simplifies to:

$$L = T \ln |I_N - \rho W_N| - \frac{NT}{2} \ln \sigma_\varepsilon^2 - \frac{1}{2\sigma_\varepsilon^2} \varepsilon' \varepsilon$$

boiling down to a repetition of the standard cross-sectional model in T cross-sections. Generalizing this model slightly, we assume $\varepsilon_t \sim N(0, \Sigma)$ to allow for more complex error covariance structures. Thus, the log-likelihood remains essentially the same, except for the new error covariance term:

$$L = T \ln |I_N - \rho W_N| - \frac{1}{2} \ln |\Sigma| - \frac{1}{2} \varepsilon' \Sigma^{-1} \varepsilon$$

In our case, the estimation may be simplified by first calculating the eigenvalues of W , ω_i , as

$$\log |I - \rho W| = \sum_{i=1}^n \log(1 - \rho \omega_i)$$

The standard formula for calculating the spatial fixed effects is based on [Elhorst and Freret \(2007\)](#) and [Elhorst \(2009\)](#), who proposed a ML estimator derivation with which to address the endogeneity problem.²⁹ The log-likelihood function of the model is as follows:

$$\begin{aligned} \text{Log} L = & \frac{NT}{2} \log(2\pi\sigma^2) + T \log |\mathbf{I}_N - \delta \mathbf{W}| \\ & - \frac{1}{2\sigma^2} \sum_{i=1}^N \sum_{t=1}^T \left(y_{it} - \delta \sum_{j=1}^N w_{ij} y_{jt} - x_{it} \beta - \mu_i \right)^2 \end{aligned}$$

Once the partial derivatives with respect to μ_i are taken, and solving μ_i , the standard formula for calculating fixed effects is obtained:

$$\mu_i = \frac{1}{T} \sum_{t=1}^T \left(y_{it} - \delta \sum_{j=1}^N w_{ij} y_{jt} - x_{it} \beta \right), \quad i = 1, \dots, N$$

By substituting the solution of μ_i in the log-likelihood function, the resulting function with respect to β , δ and σ^2 is:

$$\begin{aligned} \text{Log} L = & \frac{NT}{2} \log(2\pi\sigma^2) + T \log |\mathbf{I}_N - \delta \mathbf{W}| \\ & - \frac{1}{2\sigma^2} \sum_{i=1}^N \sum_{t=1}^T \left(y_{it}^* - \delta \left[\sum_{j=1}^N w_{ij} y_{jt} \right] * -x_{it}^* \beta \right)^2 \end{aligned}$$

where the asterisk denotes the demeaning procedure (i.e. $y_{it}^* = y_{it} - \frac{1}{T} \sum_{t=1}^T y_{it}$ and $x_{it}^* = x_{it} - \frac{1}{T} \sum_{t=1}^T x_{it}$). For further technical details on the asymptotic properties of the estimator (similar to a generalized least squares estimator of a linear regression model) or the asymptotic variance matrix of the parameters (see [Elhorst and Freret 2007](#); [Elhorst 2009](#)) (see Table 7).

A.2 Spatial (panel) error model estimation

In contrast to the spatial lag model, a spatial error specification does not require a theoretical model for spatial interaction, but, instead, is a special case of a non-spherical error covariance matrix. As proved by [Anselin et al. \(2008\)](#), an unconstrained error covariance matrix at time t , $E[\varepsilon_{it}, \varepsilon_{jt}]$, $\forall i \neq j$ contains $N \frac{(N-1)}{2}$ parameters.

²⁹ According to [Anselin et al. \(2006\)](#) an endogeneity problem arises from $\sum_i w_{ij} y_{it}$. This yields to fail the standard regression assumption properties (i.e. $E[(\sum_i w_{ij} y_{it}) \varepsilon_{it}] = 0$) such that this simultaneity must be accounted for.

The log-likelihood function for the spatial error model considered here follows directly as a special case of the standard result for ML estimation with non-spherical error covariance (Magnus 1978). With $\varepsilon_t \sim N(0, \Sigma)$ as the error vector, the expression for the log-likelihood is (ignoring the constant terms):

$$L = -\frac{1}{2} \ln |\Sigma| - \frac{1}{2} \varepsilon' \Sigma^{-1} \varepsilon$$

In the pooled model with SAR error terms, the relevant determinant and the inverse matrix are:

$$\left| I_T \otimes (B'_N B_N)^{-1} \right| = |B_N|^{-2T}$$

and:

$$\Sigma_{NT}^{-1} = \frac{1}{\sigma_\mu^2} \left| I_T \otimes (B'_N B_N) \right|$$

The corresponding log-likelihood function is then:

$$L = -\frac{NT}{2} \ln \sigma_\mu^2 + T \ln |B_N| - \frac{1}{2\sigma_\mu^2} \varepsilon' [I_T \otimes (B'_N B_N)] \varepsilon$$

where $\varepsilon = y - X\beta$. The estimates for the regression coefficient β are the result of a spatial feasible generalized least squares (FGLS) using a consistent estimator:

$$\hat{\beta} = [X' (I_T \otimes B'_N B_N) X]^{-1} X' (I_T \otimes B'_N B_N) y$$

Exploiting the block diagonal nature of $B'_N B_N$ would be equivalent to a regression of the stacked spatially filtered dependent variables, $(I_N - \theta W_N)X_t$, as a direct generalization of the single cross-section case.

Table 7 Countries included in the analysis

Argentina	Chile	Greece	Mexico	Russia
Australia	China	Guatemala	Morocco	South Africa
Austria	Denmark	Hungary	Netherlands	Sweden
Belgium	Dominican Republic	India	Nicaragua	Switzerland
Bolivia	Ecuador	Ireland	Norway	Tunisia
Brazil	Egypt	Israel	Panama	Turkey
Canada	El Salvador	Italy	Peru	United Kingdom
Colombia	Finland	Japan	Philippines	United States
Cuba	France	Jordan	Poland	Uruguay
Czech Republic	Germany	Luxembourg	Portugal	Venezuela

Table 8 Spatial global autocorrelation test (I Moran)^a

Weight matrix	Physic contiguity (Queen 8 neighbours)	Distance (Euclidean)	Distance (Arc)	K-nearest 20 neighbours
Total	No Rho	No Rho	Rho	Rho
1993	No Rho	No Rho	No Rho	Rho
1994	Rho	Rho	Rho	No Rho
1995	No Rho	No Rho	No Rho	Rho
1996	Rho	Rho	Rho	Rho
1997	Rho	Rho	Rho	Rho
1998	Rho	Rho	Rho	Rho
1999	Rho	Rho	Rho	Rho
2000	No Rho	No Rho	Rho	Rho
2001	No Rho	No Rho	No Rho	Rho
2002	No Rho	No Rho	No Rho	Rho
2003	No Rho	No Rho	No Rho	Rho
2004	No Rho	No Rho	No Rho	Rho

^aHo: no spatial autocorrelation

Hence, in our special case the log-likelihood function whereby the spatial specific effects are fixed is:

$$\begin{aligned}
 \text{Log}L = & \frac{NT}{2} \log \left(2\pi\sigma^2 \right) + T \log |I_N - \rho W| \\
 & - \frac{1}{2\sigma^2} \sum_{i=1}^N \sum_{j=1}^T \left\{ y_{it}^* - \rho \left[\sum_{j=1}^N w_{ij} y_{jt} \right]^* - \left(x_{it}^* - \rho \left[\sum_{j=1}^N w_{ij} y_{jt} \right]^* \right) \beta \right\}^2
 \end{aligned}$$

By solving the first-order maximizing conditions the concentrated log-likelihood function form of ρ is:

$$\text{Log}L = \frac{NT}{2} \log [e(\rho)'e(\rho)] + T \log |I_N - \rho W| ;$$

where $e(\rho) = Y^* - \rho(I_T \otimes W)Y^* - [X^* - \rho(I_T \otimes W)X^*]\beta$.

Finally, the ML estimator of ρ , given β and σ^2 , is computationally straightforward to obtain once the previous function is maximized with respect to ρ (see Table 8).

References

- Abreu M (2005) Spatial determinants of foreign direct investment. Paper presented at the annual meeting of the Latin American and Caribbean Economic Association (LACEA), Paris, 27–29 October 2005
- Anderson JE, van Wincoop E (2003) Gravity with gravitas: a solution to the border puzzle. *Am Econ Rev* 93(1):170–192

- Anselin L (1980) Estimation methods for spatial autoregressive structures. Regional science dissertation and monograph series No 8, Cornell University, Ithaca, NY
- Anselin L (1988) Spatial econometrics: methods and models. Kluwer, Boston
- Anselin L, Bera AK, Florax R, Yoon MJ (1996) Simple diagnostic tests for spatial dependence. *Regional science and urban economics*, Vol. 26, Issue 1. Elsevier, pp 77–104
- Anselin L (2003) Spatial externalities, spatial multipliers, and spatial econometrics. *Int Reg Sci Rev* 26:153
- Anselin L, Le Gallo J, Jayet H (2008) Spatial panel econometrics. In: Matyas L, Sevestre P (eds) *The econometrics of panel data, fundamentals and recent developments in theory and practice*, 3rd edn. Kluwer, Dordrecht
- Bajo O, Montero M (1999) Foreign direct investment and trade: a causality analysis. Working Paper 9902, Universidad Pública de Navarra
- Baltagi BH, Egger P, Pfaffermayr M (2007) Estimating models of complex FDI: are there third-country effects?. *J Econom* 140:260–281
- Barrios S, Benito-Ostolaza JM (2009) The location decisions of Multinationals and the cultural link: evidence from Spanish direct investment abroad. Working Paper 0804, Universidad Pública de Navarra
- Blanchard P, Gaigné C, Mathieu C (2008) Foreign direct investment. the lessons from panel data. In: Mathyas L, Sevestre P (eds) *The econometrics of panel data: fundamentals and recent developments in theory and practice*. Springer, Berlin
- Blomström M, Kokko A (1998) Multinational corporations and spillovers. *J Econ Surv* 12:247–277
- Blonigen BA (2005) A review of the empirical literature on FDI determinants. NBER Working Paper No. 11299
- Blonigen BA, Davies RB, Head K (2003) Estimating the knowledge-capital model of the multinational enterprise: comment. *Am Econ Rev* 87(4):520–544
- Blonigen BA, Davies RB, Waddell GR, Naughton HT (2007) FDI in space: spatial autoregressive relationships in foreign direct investment. *Eur Econ Rev* 51:1300–1325
- Canals C, Noguer M (2006) The determinants of cross-border investment: a value-chain analysis. La Caixa Working Paper Series No 05/2006
- Coughlin C, Segev E (2000) Foreign direct investment in China: a spatial econometric study. *World Econ* 23(1):1–23
- Ekholm K, Forslid R, Markusen JR (2003) Export-platform foreign direct investment. NBER Working Paper No 9517
- Elhorst JP (2003) Specification and estimation of spatial panel data models. *Int Reg Sci Rev* 26(3):244–268
- Elhorst JP, Freret S (2007) Yardstick competition among local governments: French evidence using a two-regimes spatial panel data model. European and North American RSAI meetings paper. Paris, November, 2007
- Elhorst JP (2009) Spatial panel data models. In: Fischer MM, Getis A (eds) *Handbook of applied spatial analysis*, chap C.2. Springer, Berlin
- Feenstra RC (2002) Border effects and the gravity equation: consistent methods for estimation. *Scott J Political Econ* 49(5):491–506
- Fingleton B, López-Bazo E (2006) Empirical growth models with spatial effects. *Pap Reg Sci* 85(2): 177–198
- Florax R (1992) The university: a regional booster? Economic impacts of academic knowledge infrastructure. PhD dissertation, Wageningen University
- Garretsen H, Peeters J (2008) FDI and the relevance of spatial linkages: do third country effects matter for Dutch FDI? CESifo Working Paper No. 2191
- Geary R (1954) The contiguity ratio and statistical mapping. *Incorp Stat* 5:115–145
- Griffith DA (1985) An evaluation of correction techniques for boundary effects in spatial statistical analysis: contemporary methods. *Geogr Anal* 17:81–88
- Gordo E, Martín C, Tello P (2008) La Internacionalización de las Empresas Españolas a través de la Inversión Extranjera Directa. *Economic Bulletin*, Bank of Spain, No. 1, pp 91–104
- Gordo E, Tello P (2008) Determinantes Microeconómicos de la Decisión de Localización de la Inversión Directa en el Exterior de las Empresas Españolas. *Economic Bulletin*, Bank of Spain, No. 9, pp 61–74
- Guillén M (2004) La internacionalización de las empresas Españolas. *Información Comercial Española (ICE)* 812:211–224
- Hall SG, Petroulas P (2008) Spatial interdependence of FDI locations: a lessening of the tyranny of distance? Working Paper 67, Bank of Greece

- Head K, Ries J, Swenson D (1995) Agglomeration benefits and location choice: evidence from Japanese manufacturing investments in the United States. *J Int Econ* 38(3–4):223–247
- Head K, Mayer T (2004) Market potential and the location of Japanese investment in the European Union. *Rev Econ Stat* 86(4):959–972
- Helpman E (1984) A simple theory of international trade with multinational corporations. *J Political Econ* 92(3):451–471
- Helpman E, Melitz MJ, Yeaple SR (2004) Export versus FDI with heterogeneous firms. *Am Econ Rev* 94(1):300–316
- Kapoor M, Kelejian HH, Prucha IR (2007) Panel data models with spatially correlated error components. *J Econom* 140(1):97–130
- López-Duarte C, García-Canal E (2002) La inversión directa de las empresas Españolas en Latinoamérica. *Revista Asturiana de Economía* 23:27–45
- Magnus JR (1978) Maximum likelihood estimation of the GLS model with unknown parameters in the disturbance covariance matrix. *J Econom* 7:281–312
- Markusen JR (1984) Multinationals, multi-plant economies, and the gains from trade. *J Int Econ* 16(3–4): 205–226
- Markusen JR, Maskus KE (2001) General-equilibrium approaches to the multinational firm: a review of theory and evidence. NBER Working Papers 8334
- Markusen JR (2002) Multinational firms and the theory of international trade. MIT Press, Cambridge
- Moran P (1948) The interpretation of statistical maps. *J R Stat Soc B* 10:243–251
- Rappaport J (2000) How does openness to capital flows affect growth? Mimeo, Federal Reserve Bank of Kansas City
- Romer P (1993) Idea gaps and object gaps in economic development. *J Monet Econ* 32(3):543–573
- Rodríguez-Clare A (1996) Multinational, linkages, and economic development. *Am Econ Rev* 86:852–873
- Santiso-Guimaras J (2007) La internacionalización de las empresas Españolas: Hitos y retos. *Información Comercial Española (ICE)* 839:89–102
- Turrión J, Velázquez J (2004) Presencia empresarial de España en los países de la ampliación: ¿una oportunidad perdida?. *Información Comercial Española (ICE)* 818:165–184