

NBER WORKING PAPER SERIES

FDI IN SPACE: SPATIAL AUTOREGRESSIVE RELATIONSHIPS  
IN FOREIGN DIRECT INVESTMENT

Bruce A. Blonigen  
Ronald B. Davies  
Glen R. Waddell  
Helen T. Naughton

Working Paper 10939  
<http://www.nber.org/papers/w10939>

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
December 2004

We thank Peter Egger, Eric Strobl, participants at the Fall 2004 Mid-West International Economics Group meeting and participants at the DIW/GEP Workshop on FDI and International Outsourcing. Any errors or omissions are the responsibilities of the authors. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

© 2004 by Bruce A. Blonigen, Ronald B. Davies, Glen R. Waddell, Helen T. Naughton. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

FDI in Space: Spatial Autoregressive Relationships in Foreign Direct Investment  
Bruce A. Blonigen, Ronald B. Davies, Glen R. Waddell, Helen T. Naughton  
NBER Working Paper No. 10939  
December 2004  
JEL No. F21, F23

**ABSTRACT**

Theoretical models of foreign direct investment (FDI) have only recently begun to model the role of third countries, and the empirical FDI literature has almost exclusively examined bilateral FDI data without recognizing the potential interdependence between FDI decisions to alternative host countries. This paper uses spatial econometric techniques to examine the spatial correlation between FDI to alternative (neighboring) regions. The sign of such correlations can provide evidence for or against alternative theories for FDI motivations. Using data on OECD countries from 1980-2000, we find evidence consistent with export platform FDI in Europe.

Bruce A. Blonigen  
Department of Economics  
1285 University of Oregon  
Eugene, OR 97403-1285  
and NBER  
bruceb@uoregon.edu

Ronald B. Davies  
Department of Economics  
1285 University of Oregon  
Eugene, OR 97403-1285  
rdavies@uoregon.edu

Glen R. Waddell  
Department of Economics  
1285 University of Oregon  
Eugene, OR 97403-1285  
waddell@uoregon.edu

Helen T. Naughton  
Department of Economics  
1285 University of Oregon  
Eugene, OR 97403-1285  
htammela@darkwing.uoregon.edu

## 1. Introduction

Since 1980, worldwide foreign direct investment (FDI) has grown at a remarkable rate. According to Markusen (2002), in the latter half of the 1990s FDI flows grew by nearly 32 percent. When compared to the 1.5 percent increase in exports and the 0.6 percent increase in world GDP, it comes as no surprise that increasing attention has been devoted to the determinants of FDI. The bulk of the theory on the creation of multinational enterprises (MNEs) stems from the general equilibrium models of Markusen (1984) and Helpman (1984) which use two-country frameworks. Since these early models, the literature has progressed to richer general equilibrium models that allow for more complex forms of imperfect competition (e.g., Markusen, 2002; Helpman, Melitz, and Yeaple, 2004). Nevertheless, most FDI models maintain the simple two-country, two-factor framework.

With the advent of these richer models, there has also been recent recognition in the theoretical literature that the complex motivations for FDI likely require modeling in a multilateral context, a context in which an MNE considers home, host, and third country characteristics when choosing firm activities. If FDI into a particular country stems only from a market-access motive to substitute for trade flows (simple horizontal FDI), then a bilateral country framework is likely sufficient to explain patterns in FDI data. However, other motivations certainly exist. For example, an MNE may use a host country as a low-cost “export platform” to other nearby markets. In fact, evidence on US FDI in Ireland suggests that as much as 94 percent of production is intended for export, 76 percent of which is bound for the European Union (IDA, 2004). Recent work by Ekholm, Forslid, and Markusen (2003), Yeaple (2003), and Bergstrand and Egger (2004) formalize this export platform FDI. Alternatively, an MNE may set up its vertical chain of production across countries to exploit the comparative advantages of various locales. In either case, these possibilities each suggest that FDI decisions are multilateral in nature and are therefore not accounted for by a two-country FDI model. By extension,

observed FDI patterns are likely to be better captured by empirical specifications that take direct account of the interdependence implied.

Like much of the theoretical literature, an inherent assumption of almost all empirical work on FDI has been that MNE decisions are bilateral in nature with no interdependence between host destinations. Most empirical FDI analysis constructs bilateral panel data of FDI activity and specifies a list of gravity regressors that include the GDP of the host and home countries as well as distance between the home and host country. While recent work by Carr, Markusen and Maskus (2001) has introduced a modified gravity framework that is more grounded in theory, this work also relies on bilateral transaction data to estimate the efficacy of the theoretical predictions. The difficulty in this is that using bilateral panels to estimate FDI patterns is only consistent with an underlying theory where FDI into a particular host country is independent of FDI into alternative host countries. As the above models suggest, the decision of which location is best suited for either an export platform or a stage of a vertical chain is a multilateral, comparative decision that creates interdependence among the observed bilateral FDI observations.

Importantly, the above alternative scenarios can 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 substitutability is a decreasing function of distance. Agglomeration economies with respect to other US investments, on the other hand, suggest that proximity to other FDI increases the incentive to invest in nearby countries.<sup>1</sup>

---

<sup>1</sup> Agglomeration externalities may occur between any firms, but what matters in our context here would be such externalities between US investment and US firms in neighboring countries. See Blomström and Kokko (1998), for example, for a general discussion of how agglomeration economies may arise in the context of FDI.

There are a few notable exceptions in the empirical FDI literature that have modeled or otherwise examined the influence of third markets on FDI to a host country. Head and Mayer (2004) examine Japanese FDI patterns into regions of developed Europe with a particular focus on the measurement of market potential that extends beyond the host region. In particular, in addition to the standard inclusion of host GDP, they include the GDPs of adjacent regions in an empirical specification for Japanese plant location in Europe. They find that both regions with high GDPs and regions surrounded by large markets tend to attract more FDI. In contrast, Head, Ries, and Swenson (1995) focus on patterns of related producers in third markets in their examination of Japanese affiliate locations in US states. Their conditional logit empirical specification explicitly models interdependence of the location decisions across all possible locales. They find evidence of agglomeration effects between bordering states, primarily for the automobile industry. However, they do not include measures to control for the size of surrounding markets.

In this paper we take a more general look at the potential interdependence of FDI across possible host countries by estimating spatial autoregressive models of FDI, precisely suited for estimating the effect of FDI activity in proximate regions on FDI in another focus region. As a result, we are not imposing structure to isolate one particular multilateral effect, such as agglomeration or vertical specialization, but rather estimating the net effect of such forces. A primary benefit of this estimation procedure relative to conditional logit is that spatial autoregression allows us to consider the (log) level of FDI rather than the discrete location choice. This provides us both 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 Head, Ries, and Swenson (1995) do not do.

We estimate our spatial model of FDI using data on US outbound FDI to a panel of OECD countries from 1980 through 2000. We limit our focus to the OECD for at least two reasons. First, these countries account for the lion's share of FDI in the world. In particular, for the years in our sample, on average these countries hosted 91 percent of US outbound FDI. Second, focusing on the OECD is also likely to limit vertical specialization as a primary motivation of FDI allowing us to better disentangle the factors behind any spatial interdependence. In fact, Blonigen and Davies (2004) find substantial structural differences in the determinants of FDI in developed versus less-developed countries. As a result, pooling the rich and poor countries leads to significantly-biased point estimates.

Our results show that spatial relationships between potential host countries matter significantly. Initial estimates suggest that there is a significant positive correlation between FDI to a country and the FDI that neighboring countries receive. This would be consistent with either international agglomeration economies or third market demand effects. However, we find that this result is sensitive to the inclusion of variables capturing the market potential of a host (i.e., the size of proximate third country markets). If a region's FDI is positively correlated with its own market size, as most existing studies of FDI report, then omitting controls for market potential will tend to bias the estimated coefficient on the spatial relationship upwards. Consistent with our priors, we do find that the spatial estimates decline significantly after controlling for market potential. In fact, when we focus on Europe and include a measure of market potential, we find a strong negative correlation between FDI in proximate countries. This pattern of substitution between industrialized countries of Europe, after controlling for market demand in neighboring countries, provides perhaps the strongest evidence to date of export platform FDI.<sup>2</sup>

---

<sup>2</sup> To be clear, there are two necessary conditions for evidence of export platform FDI. The first is that FDI will be attracted to countries that are located near other valuable markets because these locations make good platforms for

Only two other papers have used spatial econometric techniques to examine FDI activity to our knowledge.<sup>3</sup> Coughlin and Segev (2000) consider US FDI across Chinese provinces and find a positive spatial lag coefficient, attributing this to agglomeration economies. However, they do not include controls for the market size of surrounding regions as we do, thus this result may be indicative of the importance of including market potential as a regressor, rather than of agglomeration externalities.<sup>4</sup> Baltagi, Egger and Pfaffermayr (2004) focus on US outbound FDI for seven manufacturing industries across both developed and less-developed destinations and find evidence for a negative spatial lag coefficient which is consistent with export platform FDI for the industries that comprise their sample. By way of contrast, in line with most existing work on FDI, we use country level data. Furthermore, we restrict ourselves to the developed countries in order to better compare the horizontal and export platform models.

An alternative to using FDI data to examine this issue is to consider the exports of affiliates directly. Unfortunately, the data made available by the BEA do not specify to what extent these flows are meant for immediate distribution to consumer markets or as intermediate inputs. The former would be more consistent with an export platform model, whereas the latter would be connected with a vertical specialization/fragmentation model, which would have quite different implications for the spatial lag as we discussed earlier. Thus, given the available data, the use of spatial econometrics using FDI data is a useful method of testing current models.

---

reaching these third countries. The second is that, in the presence of plant-level fixed costs, FDI in one country acts as a substitute for FDI into other countries since the MNE can more cheaply export from a single, well-situated subsidiary. Thus, in order to test for export platform FDI, it is necessary to control for both spatial lags and market potential.

<sup>3</sup> In fact, we are aware of only one other paper applying spatial methods to trade issues. Keller and Shiue (forthcoming) analyze 18<sup>th</sup>-century trade patterns within China. Given the considerable interest in issues of trade and distance in the international economics (e.g., Anderson and van Wincoop, 2003; Redding and Venables, 2004) and the usefulness of spatial econometrics in their study, we hope that our paper fosters additional use of spatial techniques.

<sup>4</sup> In addition, they only consider FDI into a single country whereas we use a panel of twenty hosts. Furthermore, one can ask whether data originating in a centrally-planned economy such as China is representative of overall FDI patterns.

The remainder of the paper proceeds as follows. In the next section we discuss hypotheses concerning the implications of various models of multinational firm behavior for spatial relationships between FDI into various regions. Section 3 provides a brief overview of spatial econometric methods and discusses our data. Section 4 presents our estimates and highlights the importance of including both market potential and spatially-dependent variables. Section 5 concludes.

## **2. Hypotheses**

There are a variety of FDI motivations that have been illustrated in the literature, each with distinct implications for the spatial relationships we expect to observe through our estimates. The type of spatial relationship we focus on is what is referred to as spatial autoregression in the spatial econometrics literature. Analogous to a lagged dependent variable in time series analysis, in our context the estimated spatial lag coefficient characterizes the correlation between one region's FDI and geographically-proximate regions' FDIs. Before providing econometric details of our empirical specification and estimation procedure, we first outline the expected sign of this spatial lag coefficient for various models of FDI.

One of the most basic forms is horizontal FDI in which investment is motivated by market access and avoidance of trade frictions such as transport costs and import protection in the host country. In its simplest form, such a model would predict that there is no spatial relationship between markets, as the multinational firms make independent decisions for each destination market about the extent to which it wishes to serve that market through exports or affiliate sales. A sufficient condition for such a theoretical prediction is that the destination markets have sufficiently high trade protection against imports from other destination markets.

If trade protection between destination markets (or at least a group of destination markets) is low enough relative to trade frictions between the parent and destination countries, then export platform FDI is a plausible outcome. In this scenario, the multinational firm will

choose the most preferred destination market and use it as a platform to serve other markets through exports. This implies a negative spatial lag in FDI, as FDI to the platform comes at the expense of FDI to the other destination markets. In addition, the amount of FDI going into the export platform region will depend on the size of the proximate markets it will be serving through exports. Thus, if export platform FDI is occurring we would expect to find both a negative spatial lag and positive correlation between FDI and the market size of neighboring regions. This market potential effect is important to identify this form of FDI from the next form – vertical FDI.

The purest form of vertical FDI is a model where a multinational firm evaluates all potential destination markets to find the one that is the lowest cost provider of the activity it wishes to outsource. This clearly predicts a negative spatial lag coefficient, as the FDI going into the preferred region is at the expense of the other regions. However, the market potential of neighboring regions should be of no significance in this form of FDI. As we mentioned above, restricting our sample to only OECD countries makes it unlikely that we will find evidence for this type of FDI in our sample.

A more complicated variation of a vertical model is vertical specialization (or fragmentation) FDI, where multinational firms separate out a number of production activities, each of which may be in a separate geographic region (e.g. Davies, forthcoming). In this form of FDI and production, having suppliers (related or unrelated) in neighboring regions is likely to increase FDI to a particular market. More generally, there may be other cross-region forces that generate agglomeration incentives besides supplier networks. To the extent that these agglomerative forces are operating amongst US firms, we should expect to see a positive spatial lag coefficient in our estimates of US outbound FDI determinants. In these scenarios, market potential *per se* should not matter. However, the level of industrial production in neighboring countries should be correlated with increasing potential for vertical suppliers, including non-US

suppliers. Since industrial production and market potential measures will be highly correlated, our market potential variable likely proxies for both and we may therefore expect a positive coefficient on market potential if this model generally describes that data best. Thus, agglomeration externalities arising among US firms across regional borders would be evidenced by a positive spatial lag coefficient, whereas such agglomeration effects with non-US forces could potentially be captured by our market potential variable to the extent it is a close proxy for related industrial production in neighboring countries.

Table 1 summarizes our expected signs for various forms of FDI behavior. Of course, there may be a mixture of motivations behind the data we observe. Thus, we are only able to estimate net effects with our country-level data. To the extent that one form dominates the others, however, our results will identify that form.

### **3. Data and Empirical Methods**

In this section, we begin with an overview of spatial econometric techniques and then discuss our initial econometric specification for FDI and characterize the sample of countries on which we test the above models.

#### *3.1. Spatially-Dependent FDI*

In general, one would be interested in fitting data with a spatial model for one of two reasons.<sup>5</sup> First, a spatial autocorrelation or "spatial error" model places additional structure on the unobserved determinants of FDI that would otherwise be captured by the traditional error term.<sup>6</sup> 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

---

<sup>5</sup> See Anselin (1988) for detailed discussion.

<sup>6</sup> Spatially-correlated errors can be thought of as analogous to the better-known practice of clustering error terms where the econometrician is relaxing the OLS assumption of independence between all errors and assuming, instead, that while the errors are independent across groups they need not be independent within groups. If the researcher

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. For our purposes, the spatially-treated error structure is of secondary interest because although it may improve standard errors, it does not affect point estimates. In addition, it will not provide direct evidence of substitution or complementarity of FDI across countries and will therefore not inform theory, the objective and focus of this paper. In any case, we find little evidence of spatial errors in our data.<sup>7</sup>

### 3.2. A Modified Gravity Model

With theoretical ambiguity in the relationship across host country FDI, we therefore test for patterns in the data by employing a spatial lag model that allows the data to reveal whether and how the dependent variable in one observation relates to the dependent variables of other observations. For our empirical specification we begin with a “gravity” model, which is arguably the most widely used empirical specification of FDI (e.g., Eaton and Tamura, 1994; Brainard, 1997; Blonigen and Davies, 2004), and modify it based on the recent literature to include variables measuring skill endowments and the market potential of countries proximate to the host. In particular, ignoring time subscripts for notational purposes and where all non-discrete variables are measured in natural logs, our specification is:

$$[1] \quad FDI = \alpha_0 + \alpha_1 \text{Host Variables} + \alpha_2 \text{Market Potential} + \rho \cdot W \cdot FDI + \varepsilon,$$

where  $FDI$  is an  $n \times 1$  vector with row  $j$  equal to FDI from the US (the parent country) to host country  $j$ . While the standard specification would include characteristics of the parent country (e.g., real GDP, population and measures of trade costs, etc.), we discard such correlates since the parent country is always the US and these variables only have time-series variation. We

---

believes that “groups” are not so much defined by specifically observable characteristics but, rather, by “likeness” in a way that is best captured by geographic proximity, a spatial error model would correct for such relationships.  
<sup>7</sup> Moreover, the primary explanatory variables do not differ from OLS estimates in terms of either their point estimates or their significance. Therefore, we omit these results here and instead make them available on request.

instead capture such time-series variation in US FDI into our sample of countries by allowing for a quadratic trend in FDI. In unreported results, such home-variable controls are statistically insignificant and do not affect our results in any significant manner. We specify our model in log-linear form because, as documented by Blonigen and Davies (2004), such a model more likely leads to well-behaved residuals given the skewness of most FDI data samples. Such a log-linear model also allows for interactions of the underlying linear forms of the variables, as found in Carr, Markusen and Maskus (2001). We include host-skill variables and a measure of market potential as previous studies find such characteristics significant in explaining observed variation in FDI. Thus, to this point, our framework can be seen as a reduced form model that informally nests these previous specifications. The addition of  $\rho \cdot W \cdot FDI$  reflects the spatial autoregression term, discussed further below. Next we discuss the right-hand side variables of [1] in more detailed fashion.

“*Host Variables*” captures standard gravity-model variables for the host countries (GDP, population, distance from parent to host, and trade/investment friction variables), as well as a measure of skilled-labor endowments. Given the existing literature, our priors are that the higher is host GDP, the higher will be FDI. Holding GDP constant, increasing a country’s population reduces its per capita GDP and therefore FDI as well. Populations are therefore included to control for the known tendency for FDI to move between wealthy markets. We anticipate negative coefficients on population. With regard to trade costs, if FDI is undertaken to exploit vertical linkages, then higher host trade costs reduce the value to FDI. Alternatively, if FDI is primarily horizontal (as is likely in this particular sample) and intended to replace US exports, then higher host trade costs should induce tariff-jumping FDI. Thus, we remain agnostic on the effect of trade costs. Following Carr, Markusen, and Maskus (2001), we include information on skill endowments to proxy for the abundance of skilled laborers who are required for skilled-

labor intensive production by MNEs and expect that greater skill levels (particularly for the typically skill-deficient host) will be positively correlated with FDI. As a measure of investment risk we adopt a composite index that includes measures of political risk, financial risk, and other economic indicators. Our expectation is that higher risk is correlated with higher investment costs, implying lower FDI. As in the traditional gravity model, distance between the home and host is also included, which may proxy for both higher management costs (which reduce FDI) and higher trade costs (with an ambiguous effect).

The “*Market Potential*” variable for a country  $j$  is defined as the sum of distance-weighted GDPs of all other  $k \neq j$  countries in the sample, by year. This is similar to the Harris (1954) measure of market potential of neighboring regions which Head and Mayer (2004) find has the best explanatory power out of a number of market potential measures for their analysis of Japanese investment in the European Union. We use the same set of weights for construction of this variable as we will use for our construction of the spatial lag term which we discuss next. As we note in that discussion, there is little theory to guide the choice of weights. Thus, our empirical analysis section explores the robustness of our results to various weighting schemes.

Finally,  $\rho \cdot W \cdot FDI$  in Eq. [1] reflects the addition of the spatial lag term, where  $W$  is the spatial lag weighting matrix and where  $\rho$  is to be estimated and indicates the strength and sign of any spatial relationship.  $W$  itself is a matrix of dimension  $n \times n$ , block-diagonal with each block capturing a single year's observations. It is important to recognize that  $\rho \cdot W \cdot FDI$  captures the proximity of the observed host to other host countries;  $\rho \cdot W \cdot FDI$  should therefore not be confused with the standard gravity distance that measures the distance between the home and host countries.

For any year,  $y \in [1980, 2000]$ , the form of any  $W_y$  can be defined as:

$$[2] \quad W_y = \begin{bmatrix} 0 & w_y(d_{i,j}) & w_y(d_{i,k}) \\ w_y(d_{j,i}) & 0 & w_y(d_{j,k}) \\ w_y(d_{k,i}) & w_y(d_{k,j}) & 0 \end{bmatrix},$$

where  $w_y(d_{i,j})$  defines the functional form of the weights, declining in the distance,  $d_{i,j}$ , between any two host countries  $i$  and  $j$ . As distances are time-invariant, it will generally be the case that  $W_{1980} = W_{1981} = \dots = W_{2000}$ .<sup>8,9</sup> With our sample of FDI over years 1980 through 2000, the full weight matrix,  $W$ , is given by

$$[3] \quad W = \begin{bmatrix} W_{1980} & 0 & 0 \\ 0 & . & 0 \\ 0 & 0 & W_{2000} \end{bmatrix}.$$

In the construction of the weights themselves, the theoretical foundation for  $w_y(d_{i,j})$  is quite general and the particular functional form of any single element in  $W_y$  is therefore not prescribed. In our baseline results, we calculate weights using a simple inverse distance function where the shortest distance within the sample (the 173 kilometers separating Brussels and Amsterdam) gets a weight of unity and all other distances within the sample a weight that declines according to

$$[4] \quad w_y(d_{i,j}) = \frac{173}{d_{i,j}} \quad \forall i \neq j,$$

where  $d_{[i,j]}$  is the distance between hosts  $i$  and  $j$ , measured between capital cities. According to the above rule, a non-zero entry in the  $k$ th column of row  $j$  indicates that the  $k$ th observation will be used to adjust the prediction of the  $j$ th observation ( $j \neq k$ ).<sup>10</sup> The diagonal elements of  $W_y$  are set equal to zero in order that no observation of FDI predict itself. As is common, we use a row-

---

<sup>8</sup> The single exception to this in our sample is due to three missing observations in 1989. Thus,  $W_{1989}$  is of dimension 17x17 while all other years are of dimension 20x20.

<sup>9</sup> Since distance is time invariant, it is not surprising that in unreported results, after controlling for host-country fixed effects, we do not find a significant spatial lag. A similar result is found in Baltagi, Egger, and Pfaffermayr's (2004) fixed effects estimates.

standardized weighting matrix where  $W$  is normalized so that each row sums to unity. Multiplied by the vector of dependent variables, the spatially-weighted variable,  $W \cdot FDI$ , then has the simple interpretation of row-sums being a proximity-weighted average of FDI into alternative countries.

After reporting initial results below, we explore the effect of various weighting schemes on our estimates. Before continuing, however, it is important to note that the linear combination of the  $FDI$ 's appearing on the right-hand side of Eq. [1] is clearly endogenous and correlated with the error term. To see this point more formally, note that the random component of  $FDI_k$  is equal to the inner product of the  $k^{\text{th}}$  row of the matrix  $(I - \rho W)^{-1}$  and the vector of errors,  $\varepsilon$ . Each element of  $FDI$  thus depends on all of the error terms. As a result, each of the  $FDI_i$  on the right-hand side depends on the equation's error term. Thus, OLS estimates are inconsistent. We instead follow the literature by estimating the model in [1] using maximum likelihood (ML) methods, described in more detail in Appendix A.

### 3.3. Sample Data

In our estimation, we use a panel of annual data on US outbound FDI activity into the twenty OECD countries listed in Table 2 for the period 1980-2000. Our sample includes only those countries who were OECD members for the entire data period.<sup>11</sup> We restrict ourselves to the OECD countries in order to compare our results on export platform FDI to traditional horizontal FDI. As demonstrated by Markusen and Maskus (2002) and Blonigen, Davies, and Head (2003), this sample includes the large majority of US FDI activity, which in their results is consistent with the horizontal model.<sup>12</sup> We restrict ourselves to outbound data from a common

---

<sup>10</sup> For example, the distance between France and Germany will weight the US-outbound FDI to France in predicting the US-outbound FDI to Germany. Likewise, the distance between Great Britain and Germany will weight the US-outbound FDI to Great Britain in predicting the US-outbound FDI to Germany, and so on.

<sup>11</sup> After New Zealand joined OECD in May 29, 1973, no new members were admitted to OECD until Mexico in May 18, 1994.

<sup>12</sup> The cost to this, however, is that it assumes that excluded countries exert no influence on FDI patterns within the remaining data. For the European countries, one might be especially concerned about this due to the increased

parent country (the US) because existing FDI theory provides obvious reasons to expect that a parent country's FDI into host markets is interdependent, but little treatment of interdependence of FDI decisions by parent countries into a common host country (although if one considers competition in goods or host-country factor markets, there could well be such a link).<sup>13</sup> Second, as demonstrated by Markusen and Maskus (2001) and Blonigen, Davies, and Head (2003), when estimating the effects of variables such as the endowment of skilled labor on FDI, estimates relying on pooled outbound and inbound data can give very different estimates than separate regressions on each sub-sample.

Our measure of FDI is the real stock of FDI as reported by the Bureau of Economic Analysis. We use the FDI stock evaluated at historical cost and then convert it into real millions of dollars using the chain-type price index for gross domestic investment from the Economic Report of the President.<sup>14</sup> As these data are among the most widely used FDI data, it is straightforward to consider how spatial estimation adds to the results of earlier studies. Host country real gross domestic product (GDP) per capita and population data come from Penn World Tables (PWT), which reports such data for 1950 through 2000.<sup>15</sup> Our trade-cost measure is the inverse of the openness measure reported by the PWT, which itself is equal to exports plus imports divided by GDP.

Host country skill is measured by average years of schooling for those over age 25, reported every five years for 1960-2000.<sup>16</sup> Linear interpolation was used for other years. Host country investment costs are measured as the inverse of a composite index comprising operations

---

openness of the Central and Eastern European countries during the 1990s. However, as Markusen (2002) reports, these countries collectively receive less than four percent of annual world-FDI flows. Thus, we believe that this effect is likely small.

<sup>13</sup> This may be an explanation for why Markusen and Maskus (2001) and Blonigen, Davies, and Head (2003) find that the determinants of FDI activity for US inbound and outbound data yield very different estimates. We leave this issue for future research.

<sup>14</sup> The BEA's FDI data can be found at <http://www.bea.doc.gov/bea/di/di1usdbal.htm>. The price deflator can be found at <http://www.gpoaccess.gov/usbudget/fy05/sheets/b7.xls>.

<sup>15</sup> The PWT Version 6.1 data are available online at [http://pwt.econ.upenn.edu/php\\_site/pwt\\_index.php](http://pwt.econ.upenn.edu/php_site/pwt_index.php).

risk index, political risk index and remittance and repatriation factor index. These indices are developed by Business Environment Risk Intelligence S.A. and are available from 1980 to 2003.<sup>17</sup> Missing data from this source forces us to exclude Iceland and New Zealand. As such, our final sample spans from 1980 to 2000 for twenty OECD countries. Denmark, Finland and Portugal are also missing from our 1989 sample as FDI data were censored by the BEA due to confidentiality. To control for transport costs and other distance-related costs, we follow the literature in using great circle distances between capital cities, measured in 10,000s of kilometers.<sup>18</sup> Table 2 provides summary statistics of the variables in our model.

#### 4. Empirical Results

In this section, we present our initial results followed by some discussion of issues related to the specification of our baseline estimation.

##### 4.1. Initial Results

Table 3 present our initial results. Column (1) presents OLS results of our model without a spatial lag or market potential variable, and columns (2) and (3) present ML estimates that sequentially add the spatial lag and the market potential variables. One reason to sequentially add in the spatial lag and market potential variable is to examine the potential for omitted variable bias on the spatial lag from not including the market potential variable. If country- $k$  GDP correlates with FDI into country  $k$  and country- $k$  GDP also correlates with FDI into country  $j$ , then including country- $k$  FDI in the prediction of  $j$ 's FDI (e.g., through  $\rho \cdot W \cdot FDI$ ) while not directly including country- $k$ 's GDP leaves the estimation of  $\rho$  prone to bias. Of course, including market potential without  $\rho \cdot W \cdot FDI$  would also yield biased estimates of the effect of market potential (e.g., Head and Mayer, 2004).

---

<sup>16</sup> Acquired from Barro and Lee (2000), International Data on Educational Attainment.

<sup>17</sup> For more information see <http://www.beri.com>.

<sup>18</sup> With the exception of Belgium, these data were provided by Raymond Robertson at his website. Belgian distances were acquired from <http://www.indo.com>.

Turning to our results one sees that host country variables are generally significant in all specifications, with host GDP, host population, host skill, and host investment costs all matching their predicted coefficients. Consistent with nearly all FDI studies, distance between home and host is negatively correlated with FDI. Host trade costs are negative, the opposite of what the tariff-jumping horizontal theory predicts. Of particular interest is that the coefficients on these host variables are relatively stable as we add the spatial lag and market potential variables.

With respect to the estimated strength of the spatial lag relationship, our estimates reveal that, on average, FDI invested into the average country in our sample is positively associated with proximity-weighted FDI into other countries. The point estimates of our log specification also afford the simple interpretation that a 10% increase in FDI into the weighted-average proximate countries increases the FDI a destination country receives from the U.S. by approximately 5% (i.e., the estimated spatial lag coefficient suggests a proximity-weighted average elasticity of 0.5). This suggests the existence or dominance of agglomeration effects across country borders amongst US investing firms.

One would expect that the positive spatial lag would be accompanied by a positive sign on the market potential variable if such agglomeration externalities extended between US and non-US firms, provided the market potential variable proxies for industrial production in neighboring countries. This does not bear out in our data, as the results in Column (3) suggest a negative point estimate on the market potential variable. This negative sign on the market potential variable is a puzzle and motivates, in part, our sensitivity analysis with respect to various weighting schemes one might use for the spatial lag and market potential variable. Subsequently, we will also explore geographic subspaces of our data.

#### *4.2. Robustness*

As addressed in the construction of the initial assumptions on the weighting matrix,  $W$ , in Section 3.2, the theoretical foundation for  $w_y(d_{i,j})$  is quite general and the particular functional

form of any single element of  $W$  is therefore not prescribed. While our baseline results are founded on the reasonable assumption that weights decline according to a simple inverse distance function, we next explore the robustness of the reported results to alternative specifications of our weighting matrix. Equations [5] through [7] accomplish this, with results being reported in Table 4 for the following three alternatives to the weighting scheme of [4]:

$$[5] \quad w_y(d_{i,j}) = e^{-d_{i,j}/1000} \quad \forall i \neq j,$$

$$[6] \quad w_y(d_{i,j}) = \left(1 - \left(\frac{d_{i,j}}{11,155}\right)^2\right)^2 \quad \forall i \neq j \text{ if } d_{i,j} \leq 11,155; 0 \text{ otherwise,}$$

and,

$$[7] \quad w_y(d_{i,j}) = \left(1 - \left(\frac{d_{i,j}}{18,074}\right)^2\right)^2 \quad \forall i \neq j \text{ if } d_{i,j} \leq 18,074; 0 \text{ otherwise,}$$

where  $d_{[i,j]}$  remains the distance between hosts  $i$  and  $j$ , measured between capital cities. Like our baseline specification, [5] assigns positive weight to all other observations within each year but does so according to a different functional form. While assigning positive weight to all other host countries according to proximity is certainly a reasonable benchmark, we need not rule out that extreme distances are better modeled by zero weights, contributing nothing to the spatial lag. Within the sample of countries, 11,155 kilometers is the smallest maximum separation of any country-pair (i.e., the distance between Lisbon and Tokyo) and 18,074 kilometers is the maximum separation of any country-pair (i.e., the distance between Lisbon and Sydney). As such, Eq. [6] gives positive weight only to those pairings closer than 11,155. Eq. [7] adopts the same functional form as [6] but again allows all observations to receive some positive weight. As Table 4 reveals, the estimates of the spatial lag and market potential coefficients are fairly sensitive to the weighting matrix specification. This is particularly noted in Column (3) where

the spatial lag is negative and the magnitude of the market potential variable is sizably larger than in the other specifications.

#### 4.3. “Country” or “Continental” Agglomeration?

Given the extensive geographic separation of some countries from others within the sample of OECD we consider in our baseline specifications, a natural question to be addressed is whether the documented spatial dependence differs systematically for countries that are relatively remote within the sample of destination countries. In fact, considering the large area over which the OECD sample of countries lies, “continental agglomeration” may in fact be a more likely story and a more appropriate conclusion from the results reported thus far.

Within our sample of countries, Australia, Canada, and Japan are the most geographically remote from other potential host destinations. In addition, Australia and Japan lack easy access to the Atlantic Ocean, something our great circle distances do not account for. As an alternative, then, we exclude these three most-remote countries and re-estimate the empirical model for the sub-sample of European OECD countries. Table 5 presents results in similar format to Table 3, with OLS results presented in Column (1) followed by the sequential addition of the spatial lag term and the control for market potential.

Results from this sub-sample of European countries are quite different from those of our full sample of OECD countries. First, while the host variables are generally of the same sign as previous results, the effect of host GDP is less than half the size we estimated for the full sample, while the elasticity of FDI to distance is almost an order of magnitude larger. The sign of host population for the European sample is positive, where it was negative for the full sample. This highlights how sensitive determinants of FDI can be for various sub-samples of countries even before taking into account spatial considerations.

We also obtain quite different results in the European sample for our focus variables and reach strikingly different conclusions. Our full specification on European data, reported in

Column (3) of Table 5, provides strong evidence for platform FDI dominating as the primary form of FDI into Europe, with a negative spatial lag coefficient of almost -0.4 (i.e., a proximity-weighted average elasticity of -0.4) and a positive market potential elasticity above 1.3.

For this sub-sample, comparison of results with only a spatial lag (Column 2) to those including both a spatial lag and market potential variables (Column 3) suggests a substantial bias to the estimated spatial lag when the market potential variable is omitted. In fact, we estimate a positive spatial lag coefficient when the market potential variable is omitted, but a negative spatial lag coefficient when the market potential variable is included. These results suggest that there may be difficulties in interpreting the findings of Coughlin and Segev (2000) who estimated a spatial lag in a model of US FDI into China. There, they find positive spatial lags, which would indicate agglomeration. However, not including a measure of the influence of non-host markets raises concern that their result is due more to the potential found in nearby markets rather than to agglomerative-FDI patterns.

Importantly, we find that our results for the European sub-sample are robust to alternative specifications of the weighting matrix. Adopting the same alternative specifications of the forms in [5] through [7], Table 6 reports on the robustness of the results derived from the sample of European OECD countries.<sup>19</sup> In short, results are consistent in their pointing to FDI being strategically substituted across European countries and encouraged by proximate market potential, consistent with the predictions of export-platform motivated FDI.

## **5. Conclusion**

In this paper, we employ spatial econometric techniques to analyze patterns of US FDI into OECD countries. The vast majority of previous empirical FDI work has examined bilateral data while ignoring the potential for interdependence in FDI across regions. However, FDI

decisions likely involve such relationships across regions due to agglomeration economies, motivations to serve multiple markets from a foreign export platform, or a multilateral decision about the best location for a stage in a vertical chain of production. When we estimate the degree of spatial dependence between FDI into alternative countries using our entire OECD sample, we find evidence for agglomeration: FDI activity in proximate countries increases FDI in a country, *ceteris paribus*. However, this relationship disappears when we eliminate the remote countries of Australia, Canada and Japan and control for market potential in neighboring countries, where we find strong evidence that is consistent with an export-platform motivation for US FDI into Europe: a negative spatial lag coefficient and a positive coefficient on the market potential variable. The large differences in results for the European sample versus the full sample of OECD countries suggests that spatial relationships depend significantly on sub-samples of geographic space. Modeling this heterogeneity across different spatial dimensions would be a useful extension to the literature, as would be an examination of spatial relationships for FDI into less-developed countries.

---

<sup>19</sup> With the change in sample, we substitute the smallest maximum separation that was between with Lisbon and Tokyo with the distance between Ankara and Bern (2,183km) and substitute the maximum separation that was between Lisbon and Sydney with the distance between Ankara and Lisbon (3,585km).

**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	+	+

**Table 2: Descriptive Statistics**

Sample of OECD countries for the years 1980-2000. FDI is measured as the real stock of FDI as reported by the Bureau of Economic Analysis. Host real gross domestic product (GDP) per capita and population data come from Penn World Tables (PWT). Host trade costs are the inverse of the openness measure reported by the PWT, which itself is equal to exports plus imports divided by GDP. Host skill is measured by average years of schooling for those over age 25. Host investment costs are measured as the inverse of a composite index comprising operations risk index, political risk index and remittance and repatriation factor index, developed by Business Environment Risk Intelligence S.A.. Great circle distances between capital cities are used in all specifications, measured in 10,000s of kilometers. Market Potential is measured as the distance-weighted real gross domestic product of other host countries in the sample, with weights ascribed according to Eqs. [4]through [7], accordingly. For purposes of clarity, below we report market potential in millions of dollars.

Variable	Mean	Standard Deviation	Minimum	Maximum
FDI	19,206	29,676	149	233,128
Host GDP	15,690	5,922	2,473	32,057
Host Population	30,209	31,629	3,401	126,919
Host Trade Costs	0.022	0.012	0.005	0.070
Host Skill	8.328	1.978	2.800	11.860
Host Investment Costs	61.387	9.869	37.533	83.133
Host Distance from US in km	7,008	2,741	734	15,958
Market Potential (\$millions)	1,050.0	727.0	72.1	3,220.0

Sample countries included: Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Japan, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, Turkey, United Kingdom.

**Table 3: Spatial Analysis of US Outbound FDI.**

Sample of OECD countries for the years 1980-2000. Ln(FDI) is measured as the real stock of FDI as reported by the Bureau of Economic Analysis.

Independent Variable	OLS	Spatial ML	
	(1)	(2)	(3)
Ln(Host GDP)	1.703 (6.86)**	1.511 (6.35)**	1.711 (7.23)**
Ln(Host Population)	-0.491 (1.94)	-0.386 (1.60)	-0.509 (2.15)*
Ln(Host Trade Cost)	-1.258 (9.28)**	-0.945 (6.77)**	-1.491 (8.06)**
Ln(Host Skill)	1.177 (4.34)**	1.350 (5.20)**	1.026 (3.89)**
Ln(Host Investment Costs)	-1.499 (4.24)**	-1.389 (4.13)**	-1.400 (4.27)**
Ln(Host Distance from US in km)	-0.379 (4.52)**	-0.449 (5.57)**	-0.324 (3.88)**
Trend (1980 = 1)	-0.138 (4.37)**	-0.136 (4.54)**	-0.128 (4.38)**
Trend <sup>2</sup>	0.003 (2.56)*	0.002 (1.73)	0.001 (1.06)
Spatially weighted FDI <sup>a</sup> (i.e., $W \cdot FDI$ )		0.505 (5.75)**	0.552 (6.71)**
Market Potential <sup>a</sup> (i.e., weighted GDPs)			-0.314 (4.38)**
Constant	-28.618 (13.75)**	-28.312 (14.33)**	-27.512 (14.22)**
Observations	417	417	417
Adj R <sup>2</sup> / Log-Likelihood	.78	-489.77	-480.46

Standard errors in parentheses. \* significant at 10%; \*\* significant at 5%; \*\*\* significant at 1%. <sup>a</sup>Weights,  $W$ , are defined as  $w_y(d_{i,j}) = 173/d_{i,j} \quad \forall i \neq j$ .

**Table 4: Spatial Analysis of US Outbound FDI – Sensitivity Tests**

Sample of OECD countries for the years 1980-2000. Ln(FDI) is measured as the real stock of FDI as reported by the Bureau of Economic Analysis. The three alternative assumptions on the weighting matrix,  $W$ , are as follows (see text for details):

Column (1):  $w_y(d_{i,j}) = e^{-d_{i,j}/1000} \quad \forall i \neq j$  .

Column (2):  $w_y(d_{i,j}) = \left(1 - \left(\frac{d_{i,j}}{11,155}\right)^2\right)^2 \quad \forall i \neq j \text{ if } d_{i,j} \leq 11,155; 0 \text{ otherwise.}$

Column (3):  $w_y(d_{i,j}) = \left(1 - \left(\frac{d_{i,j}}{18,074}\right)^2\right)^2 \quad \forall i \neq j \text{ if } d_{i,j} \leq 18,074; 0 \text{ otherwise.}$

Independent Variable	Table 3(3)	Spatial ML		
		(1)	(2)	(3)
Ln(Host GDP)	1.711 (7.23)**	1.265 (5.93)**	1.762 (7.91)**	1.566 (7.50)**
Ln(Host Population)	-0.509 (2.15)*	-0.072 (0.33)	-0.536 (2.36)*	-0.451 (2.13)*
Ln(Host Trade Cost)	-1.491 (8.06)**	-1.338 (9.04)**	-1.883 (12.80)**	-1.773 (14.34)**
Ln(Host Skill)	1.026 (3.89)**	1.051 (4.41)**	0.509 (2.01)*	0.403 (1.71)
Ln(Host Investment Costs)	-1.400 (4.27)**	-1.766 (5.97)**	-1.480 (4.40)**	-1.680 (5.62)**
Ln(Host Distance from US in km)	-0.324 (3.88)**	-0.446 (6.23)**	-0.506 (6.60)**	-0.616 (8.42)**
Trend (1980 = 1)	-0.128 (4.38)**	-0.121 (4.67)**	-0.108 (3.66)**	0.030 (0.99)
Trend <sup>2</sup>	0.001 (1.06)	0.001 (0.99)	0.001 (1.12)	0.002 (1.50)
Spatially weighted FDI (i.e., $W \cdot FDI$ )	0.552 (6.71)**	0.691 (12.47)**	0.328 (3.16)**	-0.613 (2.91)**
Market Potential (i.e., weighted GDPs)	-0.314 (4.38)**	-0.078 (3.98)**	-0.534 (6.39)**	-1.592 (12.51)**
Constant	-27.512 (14.22)**	-29.191 (16.97)**	-20.219 (6.44)**	14.164 (3.34)**
Observations	417	417	417	417
Log-Likelihood	-480.46	-439.44	-462.10	-435.29

Standard errors in parentheses. \* significant at 10%; \*\* significant at 5%; \*\*\* significant at 1%.

**Table 5: Spatial Analysis of US Outbound FDI to European OECD.**

Sample of OECD countries for the years 1980-2000. Ln(FDI) is measured as the real stock of FDI as reported by the Bureau of Economic Analysis.

Independent Variable	Spatial ML		
	OLS (1)	(2)	(3)
Ln(Host GDP)	0.742 (3.40)**	0.704 (3.29)**	0.320 (1.57)
Ln(Host Population)	0.466 (2.11)*	0.463 (2.15)*	0.635 (3.17)**
Ln(Host Trade Cost)	-1.069 (7.24)**	-0.934 (5.98)**	-0.326 (1.97)*
Ln(Host Skill)	1.311 (5.37)**	1.349 (5.64)**	1.090 (4.87)**
Ln(Host Investment Costs)	-1.810 (5.39)**	-1.787 (5.44)**	-1.116 (3.54)**
Ln(Host Distance from US in km)	-4.683 (12.33)**	-4.443 (11.50)**	-4.364 (12.04)**
Trend (1980 = 1)	-0.076 (2.79)**	-0.078 (2.93)**	-0.127 (4.98)**
Trend <sup>2</sup>	0.003 (2.48)*	0.002 (2.02)*	0.005 (4.90)**
Spatially weighted FDI <sup>a</sup> (i.e., $W \cdot FDI$ )		0.236 (2.25)*	-0.377 (2.54)*
Market Potential <sup>a</sup> (i.e., weighted GDPs)			1.347 (7.58)**
Constant	16.949 (3.73)**	14.188 (3.08)**	2.451 (0.54)
Observations	354	354	354
Adj R <sup>2</sup> / Log-Likelihood	.85	-345.48	-317.90

Standard errors in parentheses. \* significant at 10%; \*\* significant at 5%; \*\*\* significant at 1%. <sup>a</sup> Weights,  $W$ , are defined as  $w_y(d_{i,j}) = 173/d_{i,j} \quad \forall i \neq j$ .

**Table 6: Spatial Analysis of US Outbound FDI to European OECD – Sensitivity Tests**

Sample of OECD countries for the years 1980-2000. Ln(FDI) is measured as the real stock of FDI as reported by the Bureau of Economic Analysis. The three alternative assumptions on the weighting matrix,  $W$ , are as follows (see text for details):

Column (1):  $w_y(d_{i,j}) = e^{-d_{i,j}/1000} \quad \forall i \neq j$  .

Column (2):  $w_y(d_{i,j}) = \left(1 - \left(\frac{d_{i,j}}{2,183}\right)^2\right)^2 \quad \forall i \neq j \text{ if } d_{i,j} \leq 2,183; 0 \text{ otherwise.}$

Column (3):  $w_y(d_{i,j}) = \left(1 - \left(\frac{d_{i,j}}{3,585}\right)^2\right)^2 \quad \forall i \neq j \text{ if } d_{i,j} \leq 3,585; 0 \text{ otherwise.}$

Independent Variable	Table	Spatial ML		
	5(3)	(1)	(2)	(3)
Ln(Host GDP)	0.320 (1.57)	-0.105 (0.50)	-0.108 (0.43)	-0.284 (1.14)
Ln(Host Population)	0.635 (3.17)**	1.058 (5.21)**	1.091 (4.51)**	1.325 (5.33)**
Ln(Host Trade Cost)	-0.326 (1.97)*	-0.361 (2.41)*	-0.551 (3.87)**	-0.498 (3.44)**
Ln(Host Skill)	1.090 (4.87)**	1.026 (4.63)**	0.861 (3.62)**	1.158 (5.07)**
Ln(Host Investment Costs)	-1.116 (3.54)**	-1.129 (3.75)**	-1.477 (4.94)**	-1.537 (5.18)**
Ln(Host Distance from US in km)	-4.364 (12.04)**	-4.348 (10.63)**	-4.733 (9.16)**	-5.039 (10.85)**
Trend (1980 = 1)	-0.127 (4.98)**	-0.121 (5.00)**	-0.077 (3.11)**	-0.145 (5.41)**
Trend <sup>2</sup>	0.005 (4.90)**	0.006 (5.37)**	0.005 (4.67)**	0.007 (6.28)**
Spatially weighted FDI (i.e., $W \cdot FDI$ )	-0.377 (2.54)*	-0.529 (3.35)**	-0.512 (4.30)**	-0.879 (4.76)**
Market Potential (i.e., weighted GDPs)	1.347 (7.58)**	1.711 (9.46)**	1.098 (8.91)**	2.212 (7.62)**
Constant	2.451 (0.54)	-0.927 (0.19)	12.642 (2.44)*	-5.852 (0.72)
Observations	354	354	354	354
Log-Likelihood	-317.90	-306.90	-312.41	-312.78

Standard errors in parentheses. \* significant at 10%; \*\* significant at 5%; \*\*\* significant at 1%.

Figure 1: Sample of OECD.

Sample countries included: Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Japan, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, Turkey, United Kingdom.

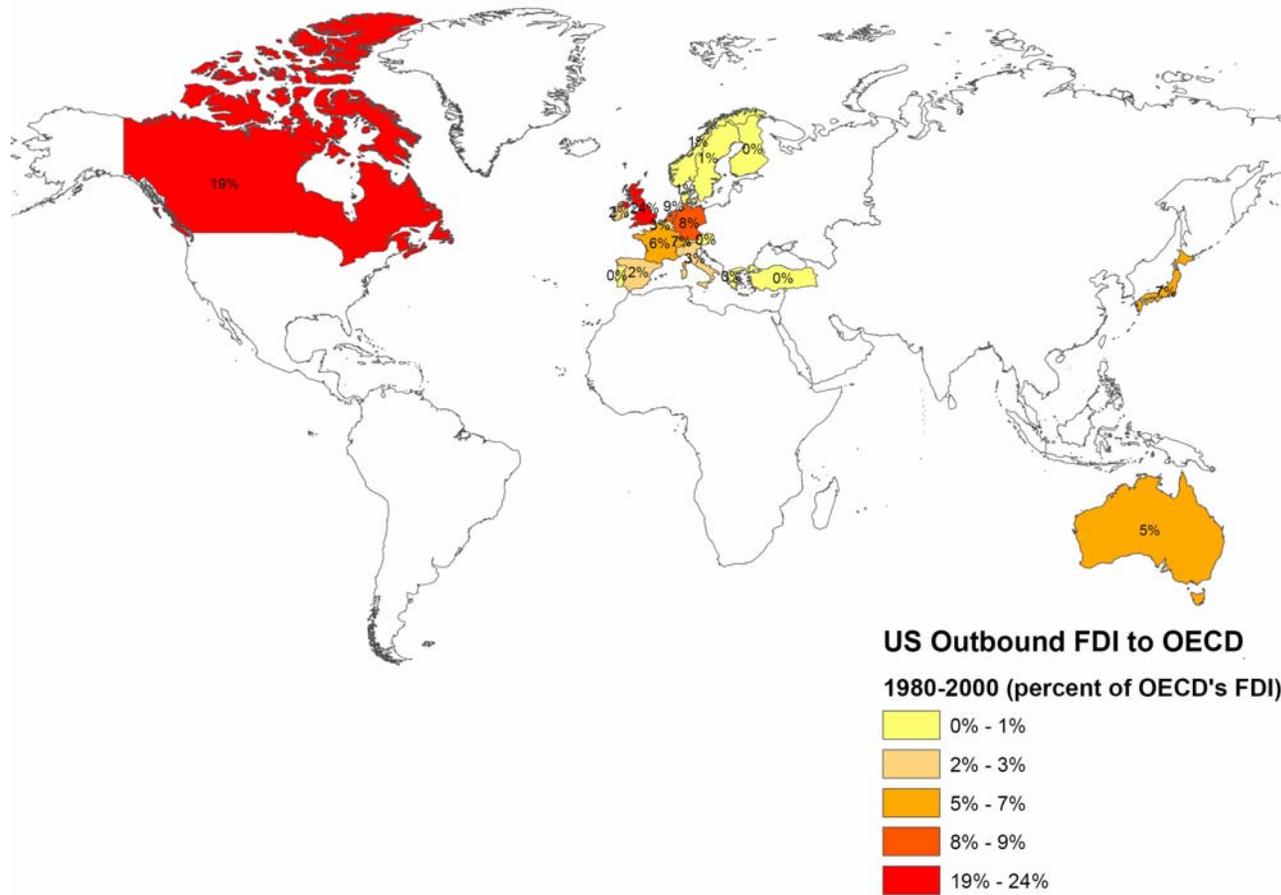
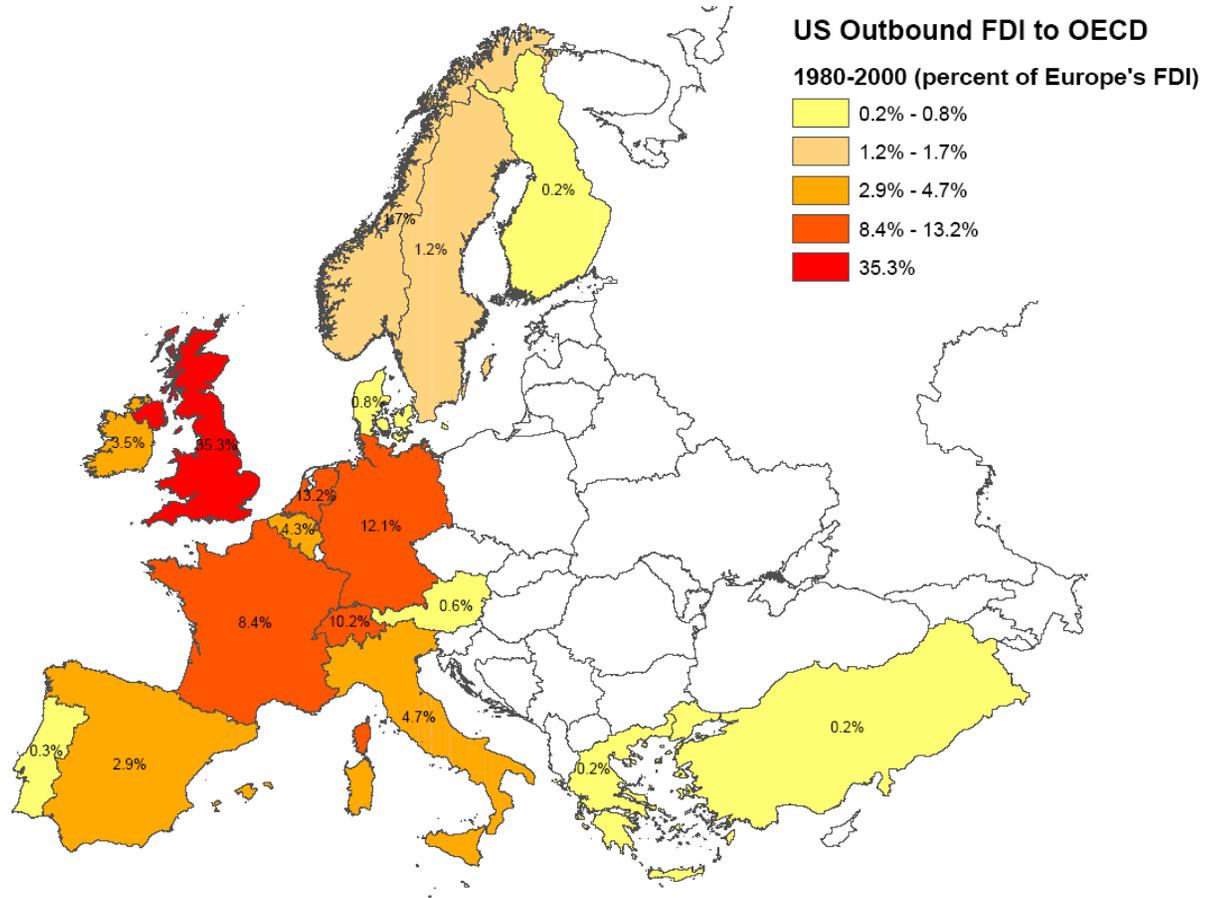


Figure 2: Sample of European OECD.

Sample countries included: Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, Turkey, United Kingdom.



## References

- Anderson, James E. and Eric van Wincoop. (2003) "Gravity with Gravitas: A Solution to the Border Puzzle," *American Economic Review* 93(1), 170-92.
- Anselin, Luc. (1988) *Spatial Econometrics: Methods and Models*, Kluwer Academic Publishers: Boston, MA.
- Baltagi, Badi H., Peter Egger, and Michael Pfaffermayr. (2004) "Estimating Models of Complex FDI: Are There Third-Country Effects?" Mimeo.
- Barro, Robert and J.W. Lee. (1996) "International Measures of Schooling Years and Schooling Quality." *American Economic Review: Papers and Proceedings* 86, 218-23.
- Bergstrand, Jeffrey H. and Peter Egger. (2004) "A Theoretical and Empirical Model of International Trade and Foreign Direct Investment with Outsourcing: Part I, Developed Countries." Mimeo.
- Blomström, Magnus and Ari Kokko. (1998) "Multinational Corporations and Spillovers," *Journal of Economic Surveys* 12, 247-77.
- Blonigen, Bruce A. and Ronald B. Davies. (2004) "The Effects of Bilateral Tax Treaties on US FDI Activity," *International Tax and Public Finance* 11(5), 601-22.
- Blonigen, Bruce A., Ronald B. Davies, and Keith Head. (2003) "Estimating the Knowledge-Capital Model of the Multinational Enterprise: Comment," *American Economic Review* 93(3): 980-94.
- Brainard, S. Lael. (1997) "An Empirical Assessment of the Proximity-Concentration Trade-off Between Multinational Sales and Trade," *American Economic Review* 87(4): 520-44.
- Carr, David L., James R. Markusen, and Keith E. Maskus. (2001) "Estimating the Knowledge-Capital Model of the Multinational Enterprise." *American Economic Review* 91(3), 693-708.
- Coughlin, Cletus and Eran Segev. (2000) "Foreign Direct Investment in China: A Spatial Econometric Study." *The World Economy* 23 (1), 1-23.
- Davies, Ronald B. (forthcoming) "Fragmentation of Headquarter Services and FDI," *North American Journal of Economics and Finance*.
- Eaton, Jonathan and Akiko Tamura. (1994) "Bilateralism and Regionalism in Japanese and US Trade and Direct Foreign Investment Patterns," *Journal of the Japanese and International Economies* 8(4), 478-510.
- Ekholm, Karolina, Rikard Forslid, and James R. Markusen. (2003) "Export-Platform Foreign Direct Investment," NBER Working Paper No. 9517.
- Harris, C. (1954) "The Market as a Factor in the Localization of Industry in the United States," *Annals of the Association of American Geographers* 64, 315-48.

Head, Keith, John Ries, and Deborah Swenson. (1995) "Agglomeration Benefits and Location Choice: Evidence from Japanese Manufacturing Investments in the United States," *Journal of International Economics* 38(3-4), 223-47.

Head, Keith and Thierry Mayer. (2004) "Market Potential and the Location of Japanese Investment in the European Union," *Review of Economics and Statistics*.

Helpman, Elhanan. (1984) "A Simple Theory of International Trade with Multinational Corporations," *Journal of Political Economy* 92(3), 451-71.

Helpman, Elhanan, Marc J. Melitz, and Stephen R. Yeaple. (2004) "Export Versus FDI with Heterogeneous Firms," *American Economic Review* 94(1), 300-16.

IDA Ireland (2002) *Annual Business Survey*. Dublin.

Keller, Wolfgang and Carol Shiue. (forthcoming) "The Origins of Spatial Interaction," *Journal of Econometrics*.

LeSage, James P. (1999) "The Theory and Practice of Spatial Econometrics," Working Paper.

Markusen, James R. (1984) "Multinationals, Multi-Plant Economies, and the Gains from Trade," *Journal of International Economics* 16(3-4), 205-26.

Markusen, James R. (2002) *Multinational Firms and the Theory of International Trade*. Cambridge and London: MIT Press.

Markusen, James R. and Keith E. Maskus. (2001) "Multinational Firms: Reconciling Theory and Evidence," in Magnus Blomström and Linda S. Goldberg, eds., *Topics in Empirical International Economics: A Festschrift in Honor of Robert E. Lipsey*. Chicago, IL: University of Chicago Press for National Bureau of Economic Research.

Markusen, James R. and Keith E. Maskus. (2002) "Discriminating among alternative theories of the multinational enterprise." *Review of International Economics* 10(4), 694-707.

Overman, Henry G., Stephen Redding and Anthony J. Venables. (2003) "The Economic Geography of Trade, Production, and Income: A Survey of the Empirics," in E. Kwan Choi and James Harrigan, eds., *Handbook of International Trade*. Oxford, UK: Blackwell Publishing.

Redding, Stephen, and Anthony J. Venables. (2004) "Economic Geography and International Inequality," *Journal of International Economics* 62(1), 53-82.

Yeaple, Stephen R. (2003) "The Complex Integration Strategies of Multinationals and Cross Country Dependencies in the Structure of Foreign Direct Investment," *Journal of International Economics* 60(2), 293-314.

## Appendix A

In the spatial lag models we use in this paper, the error terms are typically assumed to be normally distributed with constant variance, which implies the following log-likelihood function:

$$[8] \quad \log L = -\frac{n}{2} \log(2\pi) - \frac{1}{2\sigma^2} \sum_{i=1}^n \varepsilon_i^2 - \frac{n}{2} \log \sigma^2 + \log |I - \rho W| \quad .$$

Eq. [8] differs from a standard log-likelihood function for a linear regression model with the last term – the Jacobian of the transformation from  $\varepsilon$  to  $FDI$ . The first-order condition for  $\sigma^2$  implies that  $\hat{\sigma}^2 = n^{-1} \sum_{i=1}^n (Y_i - \rho \cdot W \cdot FDI_i - X_i \beta)^2$ , where  $X$  represents all our covariates on the right-hand side of equation [1] in the text other than the spatial lag term. Substituting this expression into [8], the log-likelihood function is

$$[9] \quad \log L = -\frac{n}{2} \log(2\pi + 1) - \frac{n}{2} \log \hat{\sigma}^2 + \log |I - \rho W| \quad .$$

The Jacobian term makes estimation difficult as calculating the determinant of the  $n$  by  $n$  matrix is computationally costly. However, estimation may be simplified by first calculating the eigenvalues of  $W$ ,  $\omega_i$ , as  $\log |I - \rho W| = \sum_{i=1}^n \log(1 - \rho \omega_i)$ . Although calculating eigenvalues of an  $n$  by  $n$  matrix is also time-consuming, the calculation need only be made once.

Letting  $Z = W \cdot FDI$ , where  $A = (I - \rho W)^{-1}$  and  $\theta = (\beta, \rho)'$ , the score vector and information matrix implied by [9] are:

$$[10] \quad \frac{\partial L}{\partial \theta} = \frac{1}{\sigma^2} \begin{pmatrix} X'u \\ Z'u - tr(AW) \end{pmatrix} = G$$

and,

$$[11] \quad -E \frac{\partial^2 L}{\partial \theta \theta'} = \frac{1}{\sigma^2} \begin{pmatrix} X'X & X'Z \\ Z'X & Z'Z + \sigma^2 tr(AWAW) \end{pmatrix} = V,$$

respectively. Standard iterative maximum-likelihood estimation procedures use these matrices to calculate the change in  $\theta$  across iterations:  $\theta_{j+1} = \theta_j + V^{-1}G$ . The presence of the  $tr(AW)$  in [10] and  $tr(AAWW)$  in [11] imply that the change in coefficients across iterations  $j$  and  $j+1$  cannot be calculated via a simple regression of  $\varepsilon$  on  $X$  and  $Z$ .