# ACES – EU CENTER
# EU CENTER OF EXCELLENCE

## RESEARCH GRANTS
## &
## WORKING PAPER SERIES

## #7

## AY 2007-08

# Horizontal and Vertical Interdependence in the Multinational Subsidiary Networks

Maggie X. Chen*

George Washington University

July 2008

*(preliminary and incomplete)*

**Abstract**

Existing theoretical studies have predicted that a multinational firm's location choices are interdependent across countries. Little has, however, been done to test the hypothesis at individual subsidiary level. This paper uses a detailed French multinational subsidiary dataset and estimates how a firm's decision to invest in a foreign country is not only conditional on the characteristics of that country but also the firm's existing subsidiary network. Using a structural spatial econometric model, the paper finds evidence of both horizontal and vertical interdependence in multinationals' location decisions. The results indicate that while multinational firms have little incentive to duplicate their production in countries with low bilateral trade costs, they are motivated to build a vertical subsidiary network in these countries — especially when the countries have complementary comparative advantages.

Key words: multinational firm, location decision, horizontal interdependence, vertical interdependence, input-output linkage

JEL codes: F23, D24

*2115 G Street, NW, #367, Washington, DC 20052; Email: xchen@gwu.edu; Phone: (202) 994-1092.

# 1 Introduction

Multinational corporations (MNCs) often operate a network of subsidiaries across countries. An average French multinational, for example, invests in 3.5 countries, with some of the largest such as Peugeot and Renault investing in more than a dozen. While theoretical economists have addressed this phenomenon and the cross-country interaction in multinationals' location choices (see important contributions in this area by Motta and Norman, 1996; Yeaple, 2003a; Ekholm, Rikard and Markusen, 2007), most empirical analyses have assumed that a multinational firm's decision to invest in one foreign country is independent of its locations in other foreign countries.[1]

This paper seeks to examine the interdependence in the multinational subsidiary networks and estimate two types of linkage between subsidiaries. First, the paper considers a horizontal linkage. Theoretical studies led by Markusen (1984) and Markusen and Venables (2000) have posited that in the presence of large trade costs and small plant-level scale economy firms are motivated to invest abroad and expand horizontally. This strategy has been referred to as the market-access or tariff-jumping motive in MNCs' investment decisions and suggests trade costs have a positive effect on FDI. The paper also considers a vertical linkage. The strand of theoretical literature led by Helpman (1984) has pointed out that firms whose production consists of several stages may find it more profitable to separate the production stages in various countries and locate each stage in a country where the required production cost is lowest. This strategy has been referred to as the comparative-advantage motive and is considered to lead to a vertical type of FDI. The role of trade costs is, however, reversed in this case relative to the case of horizontal FDI: Firms are more likely to locate their production in countries with low bilateral trade costs. Influential papers, such as Brainard (1997), Carr, Markusen and Maskus (2001), Yeaple (2003b), and Helpman, Melitz and Yeaple (2005), have found evidence of both market-access and comparative-advantage motives in multinationals' investment activities.

Let us now take a look at the geographic distribution of some of the largest French MNCs. Figure 1 plots Renault's production locations in the world. Two observations are noteworthy. First, Renault separates its production stages across countries. It produces components in countries such as Argentina, South Korea and Spain (indicated by the darker area of the figure) but performs the end processes in a larger set of countries (indicated by the sum of darker and lighter areas). Second, countries in which Renault only performs the end processes (indicated by the lighter area) are almost always geographically close to Renault's component production subsidiaries. A similar pattern is also found for Peugeot

[1] See below for a more elaborate discussion of the literature and the notable exceptions in the area including Head and Mayer (2004), Baltagi, Egger and Pfaffermayr (2007), and Blonigen, Davies and *et al.* (forthcoming), which formally investigate third-country effects in multinationals' location decisions.

(Figure 2) and Essilor (Figure 3), two French MNCs that specialize, respectively, in motor and lens products.

[Figures 1-3 about here]

This paper seeks to explain the structure of these subsidiary networks. In particular, it asks: Given the existence of a subsidiary in one country, how likely will a firm invest in another market? Does this likelihood always increase with the trade costs of importing from the firm's existing subsidiaries? Is the effect of trade costs consistent with the theory of horizontal FDI or the theory of vertical FDI? To address these questions, the paper employs a dataset that reports detailed information on French manufacturing MNCs' foreign subsidiaries. This dataset, by listing each included French MNC's subsidiary location abroad and their respective primary products, identifies for each firm and each country (i) the trade costs of importing (intermediate inputs or final product) from an existing subsidiary and (ii) the input-output relationship. The paper then applies a novel spatial autoregressive estimator developed by Kelejian and Prucha (1998) and Kapoor, Kelejian and Prucha (2007) to estimate the interdependence in French MNCs' subsidiary networks.

This study is one of the first attempts to estimate intra-firm interdependence *across space* and, to my knowledge, the first to do so at individual subsidiary level. Past studies have mostly focused on the relationship between home and host production and treated firms' foreign activities as uncorrelated across host countries. The potential substituting or complementary relationship between a firm's foreign production locations has been largely ignored. The following few studies took the first step to estimate third-country effects in MNCs' location decisions and are most relevant to this analysis. Specifically, Head and Mayer (2004) point out that third-country market demand can play a significant role in a country's ability to attract multinational firms. They find that Japanese multinationals are more likely to locate their production in regions proximate to large markets, a result that is consistent with the export-platform FDI theory modeled in Ekholm, Rikard and Markusen (2007). Baltagi, Egger and Pfaffermayr (2007) investigate a broader set of third-country characteristics, including third countries' factor endowment, and find most of the characteristics affect the level of bilateral FDI. Blonigen, Davies and *et al.* (forthcoming) emphasize on the direct interdependence of FDI flows and find investments in third countries can have a negative effect on a host country's receipt of FDI.

Similar to these studies, in particular, Baltagi, Egger and Pfaffermayr (2007) and Blonigen, Davies and et al. (forthcoming), this paper uses a (albeit different) spatial econometric approach to examine the interaction of MNCs' location decisions across host countries. It, however, differs in three areas. First, in contrast with the analysis of Baltagi, Egger and Pfaffermayr (2007) and Blonigen, Davies and *et al.* (forthcoming) which uses either country- or sector-level FDI data, this paper uses a subsidiary-level dataset that consists of individual MNCs' subsidiary information in the world. The use of this disaggregate data makes it possible to isolate intra-firm spatial interaction from other types of interaction, such as those between firms. It also allows the paper to obtain subsidiary specific information and draw conclusive evidence on the significance of horizontal and vertical FDI, which has been largely missing in the literature.

The paper also differs by introducing an input-output linkage between foreign production locations. This is important as multinationals' location choices are likely to vary significantly with the nature of the relationship between subsidiaries. For subsidiaries with duplicate production, multinationals have little incentive to locate them in countries where the trade costs of importing from one another are low. For subsidiaries with complementary production processes, however, multinationals benefit from building the facilities in countries with geographic proximity and low bilateral tariffs. While the latter hypothesis has been tested with respect to the relationship between parent firms and foreign affiliates (see a recent contribution in this area by Hanson, Mataloni and Slaughter, 2005), it is largely ignored for the relationship between foreign affiliates. This paper takes into account the input-output linkage between MNCs' foreign production locations and allows the effect of trade costs to vary between substituting and complementary subsidiaries.

Last, the existing studies have used distance as the main spatial weight of third-country effects. Two issues are associated with this approach. First, distance is not the only trade cost faced by multinationals. Tariffs imposed by a foreign country on each MNC are also expected to exert a similar and significant effect on multinationals' location decisions. Moreover, as discussed above the role of trade costs is different between horizontal and vertical FDI. It is therefore important to differentiate the cost of importing final products from the cost of importing intermediate inputs. These issues are addressed in this paper using information on the level of trade costs between each pair of subsidiaries.

To proceed, I first build a theoretical framework where both the number and location of subsidiaries are endogenously determined. Based on this framework, I derive a structural spatial econometric model and estimate a firm's probability of investing in one country as a function of the characteristics of that country and the firm's existing subsidiaries. More specifically, I allow for two types of interdependence in the econometric model: (i) spatially lagged dependent variable and (ii) spatial autoregressive errors. Furthermore, in addition to distance — the spatial weighting matrix considered in the previous studies, I take into account three other factors: host-country tariff rates (on intermediate inputs and final goods), host-country differences in factor endowment, and input-output linkage between subsidiaries. If a multinational firm is driven by the market access motive and engages in horizontal FDI, its probability of investing in a country should increase with both the transport cost and the tariff of importing the final good from its existing subsidiaries. This is especially likely when countries are similar in factor endowment. In the case of comparative advantage motive and vertical FDI, the multinational firm's incentive to invest in a country should decrease with the trade costs of importing intermediate inputs from upstream subsidiaries especially when countries are complementary in comparative advantage and the input-output relationship between subsidiaries is close. The availability of subsidiary-level data allows me to obtain detailed information to test the above hypotheses.

I find evidence of both horizontal and vertical interdependence between French multinationals' foreign production locations. Specifically, I find that multinationals have little incentive to duplicate their production in countries that are proximate to one another. Similarly, they are less likely to invest in countries that set a low tariff on the firms' final-good imports from existing subsidiaries. The effect of trade costs is, however, reversed for multinationals with vertical subsidiary networks. This type of multinationals tend to locate their vertically linked subsidiaries in countries with low bilateral trade costs, especially when there are also complementary comparative advantages.

In the sensitivity analysis, I address the potential concern of unaccounted trade costs and consider alternative weighting matrices. In particular, I use the sectoral trade flow between each pair of host countries (in the MNC subsidiaries' primary product) as an alternative measure of between-subsidiary trade cost (or openness) and find it does not lead to any significant change in the results. I also control for unobserved host-country sectoral characteristics, such as market structure, and firm heterogeneities, such as product quality, with country-industry and firm fixed effects and find the results again remain qualitatively similar.

The rest of the paper is organized as follows. In Section 2, I adopt a three-country model in which firms' location decision is endogenously determined. Based on this model, I derive several testable hypotheses and a structural spatial econometric model. I then describe the econometric methodology in Section 3 along with the relevant data sources. In Section 4, I examine the distribution of French MNCs' subsidiaries and discuss its geographic attributes. I present the main econometric results and sensitivity analysis in Section 5 and conclude the paper in Section 6.

## 2 The Model

To motivate the empirical analysis, I build a model that is related to Motta and Norman (1996). Suppose the world consists of three countries, $H$, $A$ and $B$. The representative consumer in each country allocates a certain amount of her expenditure, denoted as $Y_j$ ($j = H, A$ and $B$), to the industry of differentiated products. Within this industry, the consumer has a utility function that exhibits constant elasticity of substitution (CES). Maximizing the CES utility function subject to the consumer's expenditure level yields the consumer demand function for each representative variety: $q_{ij} = A_j p_{ij}^{-\sigma}$, where $q_{ij}$ is the quantity of the differentiated product produced by firms in country $i$ and sold to destination country $j$, $A_j \equiv Y_j / \sum_r p_{rj}^{1-\sigma}$ is the demand level in country $j$ with $r$ representing the set of varieties, $p_{ij}$ the price of this product, and $\sigma$ the constant elasticity of substitution. Note that $p_{ij} = \phi_{ij} \cdot p_i$, where $p_i$ is country $i$'s product market price and $\phi_{ij} \geqslant 1$ is the iceberg trade cost of exporting from country $i$ to country $j$ (with $\phi_{ii} = 1$).

For the purpose of this paper, the model assumes that country $H$'s firms, denoted as $h$, can locate their production in any of the three countries while country $A$ and $B$'s firms produce only at home and serve the foreign markets via exports. The model also considers two types of production. First, it assumes that production consists of only one stage and firm $h$ may potentially duplicate their production across countries and engage in horizontal FDI. Then, it considers the case in which production consists of two separable stages and firms may locate each stage in a different country and undertake vertical FDI. A structural estimation equation is derived at the end of the section and incorporates both of these possibilities. Let us now begin with the horizontal case.

## 2.1 horizontal FDI

When production consists of only one stage, firm $h$ sets the profit-maximizing price at $p_i = [\sigma/(\sigma - 1)]c_i$, where $c_i$ is the marginal cost of producing the differentiated good in country $i$. The quantity of the good produced in country $i$ and sold to country $j$ is therefore

$$q_{ij} = \frac{(\sigma - 1)(c_i \phi_{ij})^{-\sigma} A_j}{\sigma}, \tag{1}$$

which indicates that the gross profit firm $h$ earns by producing in country $i$ and selling to destination country $j$ is

$$\pi_{ij} = \frac{c_i^{1-\sigma} \phi_{ij}{}^{-\sigma} A_j}{\sigma}. \tag{2}$$

Note that $\pi_{ij}$ is an increasing function of country $j$'s demand (i.e., $A_j$) and a decreasing function of marginal production cost (i.e., $c_i$) and trade cost (i.e., $\phi_{ij}$).[2]

Now denote $d_i$ as firm $h$'s location decision in country $i$, where $i = H, A$, and $B$; $d_i = 1$ if firm $h$ establishes a local production facility in country $i$ and 0 otherwise. Because firm $h$ can produce in more than one location, $\sum_i d_i \geqslant 1$. If $d_i = 1$ for $i \neq H$, it means that firm

[2] Following the literature, the demand level, $A_j$, is exogenous from an individual firm's point of view.

$h$ has a plant in country $A$ or $B$ and is considered as a multinational firm. For locations where $d_i = 1$, firm $h$ also determines $q_{ij}$ for all $j$ which includes the quantity it sells to local market (when $j = i$) and the quantity exported to foreign markets (when $j \neq i$ and $d_j = 0$).[3] In addition to the variable cost of production, I assume each plant requires a fixed cost $F_i$. The total amount of fixed cost for firm $h$ is therefore $\sum_i F_i \cdot d_i$.

There is a total of 7 possible location configurations for firm $h$, denoted as $[d_H, d_A, d_B]$. These possible configurations, assuming firm $h$ produces in at least one location (i.e., $\max_i \{d_i\} > 0$), can be divided to three categories: (1) $\sum_i d_i = 1$, (2) $\sum_i d_i = 2$, and (3) $\sum_i d_i = 3$.[4] First, let us consider the case in which firm $h$ only produces in one location $i$, i.e., $d_i = 1$ and $d_j = 0$ for all $j \neq i$. Its gross profit in this case is

$$\Pi\left(\sum_r d_r = 1\right) = \pi_i - F_i = \sum_j \pi_{ij} - F = \frac{c_i^{1-\sigma}}{\sigma} \sum_j \phi_{ij}{}^{-\sigma} A_j - F_i. \tag{3}$$

By comparing equation (3) across all three potential locations, firm $h$ chooses the location that maximizes the profit (if it decides to have only one plant).

If firm $h$ chooses to produce in two locations instead, say countries $i$ and $j$ (i.e., $d_i = d_j = 1$), and export to the third country $k$ (i.e., $d_k = 0$) from country $i$, its gross profit is

$$\Pi\left(\sum_r d_r = 2\right) = \sum_{r \neq j} \pi_{ir} + \pi_{jj} - F_i - F_j = \frac{c_i^{1-\sigma}}{\sigma} \sum_{r \neq j} \phi_{ir}{}^{-\sigma} A_r + \frac{c_j{}^{1-\sigma}}{\sigma} A_j - F_i - F_j. \tag{4}$$

Note the above equation assumes that $c_i^{1-\sigma} \phi_{ik}^{-\sigma} \geqslant c_j^{1-\sigma} \phi_{jk}^{-\sigma}$, which ensures that firm $h$ would export to $k$ from $i$ instead of $j$. To select the optimal pair of locations, firm $h$ ranks the above profit function for each pair of countries and chooses the pair that yields the maximum.

Firm $h$ may also decide to produce in all three countries and supply each country through local production, i.e., $d_i = 1$ for all $i$. In this case, its gross profit becomes

$$\Pi\left(\sum_r d_r = 3\right) = \sum_i \pi_i - \sum_i F_i = \sum_i \frac{c_i{}^{1-\sigma}}{\sigma} A_i - \sum_i F_i. \tag{5}$$

[3] I assume that firm $h$ would supply a market from only one production location. If firm $h$ has a plant in a country, it supplies this country's consumers through local production. If firm $h$ does not have a plant in the country, it exports to this country from a location that maximizes its gross profit.

[4] This paper focuses on firm $h$'s location choices and supply strategies and hence assumes that the parameters of the model, such as the market size of each country, $Y_j$, satisfy the nonnegative profit condition, that is, firm $h$ would always supply each market. The decision faced by firm $h$ is therefore to choose the optimal strategy to serve each market, i.e., local production versus exports, and the optimal production locations.

Given all the possible location choices above, firm $h$ picks the optimal location configuration $[d_H^*, d_A^*, d_B^*]$ that satisfies

$$\Pi\left(d_H^*, d_A^*, d_B^*\right) \geqslant \Pi\left(d_H, d_A, d_B\right), \ \forall\left[d_H, d_A, d_B\right], \tag{6}$$

where $\Pi\left(d_H, d_A, d_B\right)$ is defined by equations (3), (4) and (5). The optimal number of plants is thus $\sum_r d_r^* = d_H^* + d_A^* + d_B^*$.

Now without loss of generality let us consider firm $h$'s decision to invest in country $i$ given an existing plant in country $j$. Firm $h$ would produce in both country $i$ and country $j$ if and only if $\Pi\left(d_i = 1, d_j = 1\right) > \Pi\left(d_i = 0, d_j = 1\right)$ and $\Pi\left(d_i = 1, d_j = 1\right) > \Pi\left(d_i = 1, d_j = 0\right)$. Given equations (3) and (4), the necessary conditions for these inequalities to hold simultaneously are[5]

$$\begin{aligned} \left({c_i}^{1-\sigma} - c_j^{1-\sigma}{\phi_{ji}}^{-\sigma}\right) A_i - \sigma F_i &> 0 \\ \left({c_j}^{1-\sigma} - c_i^{1-\sigma}{\phi_{ij}}^{-\sigma}\right) A_j - \sigma F_j &> 0. \end{aligned} \tag{7}$$

which require $c_i$ and $c_j$ to be sufficiently similar and $\phi_{ij}$ and $\phi_{ji}$ to be sufficiently large. This result is consistent with the existing studies, which predict horizontal FDI is more likely to arise in countries with similar factor prices (or endowments) and large trade costs. Because of the symmetry, I focus here on the first condition. A sufficient condition for this inequality to hold is

$$(1-\sigma)\ln c_i + \ln A_i + \ln\left[1 - {\phi_{ji}}^{-1}\right] - \ln \sigma F_i > 0. \tag{8}$$

This condition suggests that the likelihood of firm $h$ investing in country $i$ given an existing plant in country $j$ is a decreasing function of $c_i$ and $F_i$ and an increasing function of $A_i$ and $\phi_{ji}$.

## 2.2 vertical FDI

Now let us consider a two-stage production. Specifically, I assume that firm $h$ must produce an intermediate input before manufacturing the final product.[6] Like the final product, the

[5] Because firm $h$ may also have a plant in the third-country $k$ and export to country $i$ (and analogously country $j$) from there, (7) serves as necessary conditions.

[6] Given this paper's focus on intra-firm linkages, I do not consider here the option of purchasing intermediate inputs from unaffiliated suppliers. The latter possibility and its role in firms' location decision is, however, an interesting reseach question in its own right and has a large scope for future empirical research. For seminal theoretical studies in this area, see, for example, Krugman and Venables (1996), Venables

intermediate input can be produced in any of the three countries. This implies that firm $h$'s location decision is now characterized as $[u_H, u_A, u_B, d_H, d_A, d_B]$, which consists of the location choices for upstream and downstream production. For simplicity, I assume that the plant-level fixed cost required for the intermediate input, denoted as $G_k$, is sufficiently large such that firm $h$ will only locate this stage of production in one country. I also assume that the upstream subsidiary will sell the inputs to the downstream subsidiaries at $m_k \phi_{ki}$, where $m_k$ is the marginal cost of producing the intermediate input and $\phi_{ki}$ is the cost of exporting the intermediate input from country $k$ to country $i$.

As a result, the gross profit firm $h$ earns by producing the intermediate input in country $k$ and final good in country $i$ and selling to destination country $j$ is

$$\pi_{ij} = \frac{(c_i + m_k \phi_{ki})^{1-\sigma} {\phi_{ij}}^{-\sigma}}{\sigma} A_j. \tag{9}$$

Similar to the previous case, $\pi_{ij}$ is an increasing function of country $j$'s demand (i.e., $A_j$) and a decreasing function of the final good marginal cost (i.e., $c_i$) and the trade cost to ship the final good from country $i$ to country $j$ (i.e., $\phi_{ij}$). However, because production now consists of more than one stage, $\pi_{ij}$ is also a decreasing function of the marginal cost used to produce the intermediate input (i.e., $m_k$) and the trade cost to ship the input to the final good production location (i.e., $\phi_{ki}$).

Now suppose that firms have chosen country $k$ as the location to produce intermediate inputs. The total profit function will then be

$$\Pi\left(\sum_r d_r = 1, u_k = 1\right) = \frac{(c_i + m_k \phi_{ki})^{1-\sigma}}{\sigma} \sum_j {\phi_{ij}}^{-\sigma} A_j - F_i - G_k \tag{10}$$

if firm $h$ decides to concentrate the second-stage production in one location, say $i$, or

$$\Pi\left(\sum_r d_r = 2, u_k = 1\right) = \frac{(c_i + m_k \phi_{ki})^{1-\sigma}}{\sigma} \sum_{r \neq j} {\phi_{ir}}^{-\sigma} A_r + \frac{(c_j + m_k \phi_{kj})^{1-\sigma}}{\sigma} A_j - F_i - F_j - G_k \tag{11}$$

if firm $h$ decides to locate the second stage in two countries, $i$ and $j$, or

$$\Pi\left(\sum_r d_r = 3, u_k = 1\right) = \sum_i \frac{(c_i + m_k \phi_{ki})^{1-\sigma}}{\sigma} A_i - \sum_i F_i - G_k \tag{12}$$

(1996), and Puga and Venables (1997). The empirical analysis of this paper attempts to control for these factors using host-country-industry and firm fixed effects as firm-level data that identifies intermediate input suppliers is largely missing (Section 5.3).

if firm $h$ produces the final good in all three countries.

There are totally 21 possible location configurations. Firm $h$ will choose the optimal configuration, i.e., $[u_H^*, u_A^*, u_B^*, d_H^*, d_A^*, d_B^*]$, that satisfies $\Pi(u_H^*, u_A^*, u_B^*, d_H^*, d_A^*, d_B^*) \geqslant \Pi(u_H, u_A, u_B, d_H, d_A, d_B)$, $\forall [u_H, u_A, u_B, d_H, d_A, d_B]$.

Now consider firm $h$'s decision to produce the final good in country $i$ given its upstream plant in country $k$.[7] Here I assume firm $h$ also produces final product in country $k$.[8] Firm $h$ would invest in country $i$ only if

$$\left[(c_i + m_k\phi_{ki})^{1-\sigma} - (c_k + m_k)^{1-\sigma}\phi_{ki}{}^{-\sigma}\right] A_i - \sigma F_i > 0. \tag{13}$$

Denoting $\gamma_{ki} \equiv c_k/c_i$ and $\rho \equiv m_k/c_k$, the above inequality can be further transformed to

$$(1-\sigma)\ln c_i + \ln A_i + \ln\left[(1+\rho\gamma_{ki}\phi_{ki})^{1-\sigma} - (1+\rho)^{1-\sigma}\gamma_{ki}^{1-\sigma}\phi_{ki}{}^{-\sigma}\right] - \ln\sigma F_i > 0. \tag{14}$$

Note when $\rho = 0$ (i.e., $m_k = 0$), the above inequality is identical to inequality (7), the condition of horizontal expansion. A sufficient condition for the above inequality to hold is

$$(1-\sigma)\ln c_i + \ln A_i + \ln\left(1 - \phi_{ki}{}^{-1}\right) + (1-\sigma)\ln(1+\rho\gamma_{ki}\phi_{ki}) - \ln\sigma F_i > 0. \tag{15}$$

It is not difficult to see that the role of trade costs is now twofold. On the one hand, it encourages firm $h$ to expand its downstream production across countries as in the case of horizontal linkage (reflected in the term $\ln(1-\phi_{ki}{}^{-1})$). But on the other hand, it raises the cost of shipping intermediate inputs to downstream plants and discourages vertical FDI (reflected in the term $(1-\sigma)\ln(1+\rho\gamma_{ki}\phi_{ki})$ where $1-\sigma < 0$). The latter effect is especially strong when (i) the input-output relationship captured in $\rho$ is great and (ii) two countries' difference in cost structures captured in $\gamma_{ki}$ is large.

To summarize, I incorporate the conditions specified in the horizontal and vertical linkage cases, i.e., (8) and (15), in the following estimation specification:

[7] The location of upstream plants is arguably also endogenously determined. Here I focus on the location choice of downstream plants because the location choice of upstream subsidiaries (such as R&D divison) is more likely to be driven by factors such as the availability of natural or human resources and firms have less incentive to duplicate this stage of production across countries.

[8] This assumption is in alignment with annecdotal evidence including those presented in Figures 1-3. All the individually researched multinational firms, such as Renault, Peugeot and Essilor, are found to conduct some extent of final processing in upstream production locations but do not always perform intermediate-input production in downstream production sites.

$$\Pr(y_i = 1|y_j) = \Phi\left[\alpha + X_i\beta + (\lambda_1 \ln \phi_{ji} + \lambda_2 \ln(1 + \rho\gamma_{ji}\phi_{ji})) \cdot y_j + \varepsilon_i\right] \tag{16}$$

where $y_i = 1$ if $\max(u_i, d_i) = 1$ and 0 if $\max(u_i, d_i) = 0$, $\Pr(y_i = 1|y_j)$ is the probability of the firm investing in country $i$ given its location decision in country $j$, $\Phi(.)$ represents the cumulative probability function, $X_i\beta \equiv \beta_1 \ln c_i + \beta_2 \ln A_i + \beta_3 \ln c_i$ represents a vector of country $i$'s characteristics, $(\lambda_1 \ln \phi_{ji} + \lambda_2 \ln(1 + \rho\gamma_{ji}\phi_{ji})) \cdot y_j$ captures the relationship between $y_i$ and the firm's existing subsidiary captured by $y_j$, and $\varepsilon_i$ denotes the error term. In particular, I expect $\lambda_1 > 0$ in the presence of horizontal linkage and $\lambda_2 < 0$ in the presence of vertical linkage.

# 3 Econometric framework

## 3.1 Methodology

To estimate equation (16), I adopt the following cross-sectional (first-order) spatial autoregressive model with (first-order) autoregressive disturbances:

$$\begin{aligned} y_n &= X_n\beta + \lambda_1 W_n y_n + \lambda_2 V_n y_n + u_n, \qquad |\lambda| < 1, \\ u_n &= \rho M_n u_n + \varepsilon_n. \qquad |\rho| < 1. \end{aligned} \tag{17}$$

In the above equations, $y_n$ is an $n \times 1$ vector of observations on the dependent variable, $X_n$ is an $n \times m$ matrix of observations on $m$ exogenous variables, $W_n$, $V_n$ and $M_n$ are three $n \times n$ spatial weighting matrices, $\beta$ is a $m \times 1$ vector of parameters, $\lambda_1$, $\lambda_2$ and $\rho$ are scalar autoregressive parameters, $u_n$ is an $n \times 1$ vector of regression disturbances, $\varepsilon_n$ is an $n \times 1$ vector of innovations. The variables $W_n y_n$, $V_n y_n$ and $M_n u_n$ are respectively the spatial lags of $y_n$ and $u_n$. For notational simplicity, I denote $Z_n = (X_n, W_n y_n, V_n y_n)$ and $\delta = (\beta, \lambda_1, \lambda_2)'$. Applying a Cochrane-Orcutt transformation to this model yields

$$y_n^*(\rho) = Z_n^*(\rho)\delta + \varepsilon_n, \tag{18}$$

where $y_n^*(\rho) = y_n - \rho M_n y_n$ and $Z_n^*(\rho) = Z_n - \rho M_n Z_n$.

I adopt the approach introduced in Kelejian and Prucha (1998, 1999) to estimate the above model. This approach is more computationally feasible than the alternative estimator, i.e., the maximum likelihood estimator, especially in large samples where the latter estimator requires the weighting matrices to have simplifying features such as spareness and symmetry. Given the weighting matrices used in this paper are relatively large $(9,717 \times 9,717)$ and do not always have features like symmetry (e.g., in the tariff weighting matrix discussed in the next section), I use the former method.

This method consists of three steps. First, I estimate equation (18) by a two-stage Least-Square estimator (2SLS). In particular, I estimate $W_n y_n$ and $V_n y_n$ using $X_n$ and $H_n$ and obtain:

$$\begin{aligned} \hat{\delta}_n &= (\hat{Z}_n' \hat{Z}_n)^{-1} \hat{Z}_n' y_n, \\ \hat{u}_n &= y_n - Z_n \hat{\delta}_n \end{aligned} \tag{19}$$

where $\hat{Z}_n = (X_n, \widehat{W_n y_n}, \widehat{V_n y_n}) = P_{H_n} Z_n$, $\widehat{W_n y_n} = P_{H_n} W_n y_n$, $\widehat{V_n y_n} = P_{H_n} V_n y_n$, $P_{H_n} = H_n (H_n' H_n)^{-1} H_n'$ and $H_n$ is the matrix of instruments formed by a subset of linearly independent columns of $(X_n, W_n X_n, W_n^2 X_n, V_n X_n, V_n^2 X_n)$.

In the second step, I estimate the spatial autoregressive parameter $\rho$ using the residuals obtained from the first step and the generalized moments procedure introduced in Kelejian and Prucha (1998). Specifically, the estimators of $\rho$ and $\sigma_\varepsilon^2$, i.e., $\tilde{\rho}$ and $\tilde{\sigma}_\varepsilon^2$, are defined as the nonlinear Least-Square estimators that minimize

$$\left[ g_n - G_n \begin{bmatrix} \rho \\ \rho^2 \\ \sigma_\varepsilon^2 \end{bmatrix} \right]' \left[ g_n - G_n \begin{bmatrix} \rho \\ \rho^2 \\ \sigma_\varepsilon^2 \end{bmatrix} \right] \tag{20}$$

where

$$G_n = \frac{1}{n} \begin{bmatrix} 2\hat{u}_n' \widehat{\hat{u}}_n & -\widehat{\hat{u}}_n' \widehat{\hat{u}}_n & n \\ 2\hat{u}_n' \widehat{\widehat{\hat{u}}}_n & -\widehat{\hat{u}}_n' \widehat{\hat{u}}_n & Tr(M_n' M_n) \\ \hat{u}_n' \widehat{\hat{u}}_n + \hat{u}_n' \widehat{\widehat{\hat{u}}}_n & -\widehat{\hat{u}}_n' \widehat{\hat{u}}_n & n \end{bmatrix}, \quad g_n = \frac{1}{n} \begin{bmatrix} \hat{u}_n' \hat{u}_n \\ \widehat{\hat{u}}_n' \widehat{\hat{u}}_n \\ \hat{u}_n' \widehat{\hat{u}}_n \end{bmatrix},$$

and $\widehat{\hat{u}}_n = M_u \hat{u}_n$ and $\widehat{\widehat{\hat{u}}}_n = M_u^2 \hat{u}_n$.

In the third step, I estimate equation (18) using 2SLS and obtain

$$\tilde{\delta}_n = \left[ \hat{Z}_n'^{*}(\tilde{\rho}_n) \hat{Z}_n^{*}(\tilde{\rho}_n) \right]^{-1} \hat{Z}_n'^{*}(\tilde{\rho}_n) y_n^{*}(\tilde{\rho}_n), \tag{21}$$

where $\hat{Z}_n'^{*}(\tilde{\rho}_n) = P_{H_n} Z_n^{*}(\tilde{\rho}_n)$, $Z_n^{*}(\tilde{\rho}_n) = Z_n - \tilde{\rho}_n M_n Z_n$, and $y_n^{*}(\rho) = y_n - \tilde{\rho}_n M_n y_n$.

## 3.2 Spatial weighting matrices

Now let us define the spatial weighting matrices that will be used to estimate the interdependence of multinationals' locations. To proceed, I construct two $N \times N$ matrices, $w_h$ and $v_h$, for each firm in the sample (where $N$ denotes the number of countries in the data and $h$ denotes the firm). In particular, the cells in $w_h$ are defined as follows:

$$\omega_{hij} = \begin{cases} \ln \phi_{hij} \text{ for } i \neq j \\ 0 \text{ for } i = j \end{cases} . \tag{22}$$

Recall $\phi_{hij}$ is the trade cost firm $h$ faces when importing the final good from country $i$ to country $j$. The cells in $v_h$ are defined as

$$v_{hij} = \begin{cases} \ln(1 + \rho_{hj}\gamma_{hji}\phi^{m}_{hji}) \text{ for } \rho_{hj} > 0 \\ 0 \text{ for } \rho_{hj} = 0 \end{cases} . \tag{23}$$

In this definition, $\rho_{hj}$ is the input-output coefficient between products produced in country $j$ and final product, $\gamma_{hji}$ is countries' difference in production costs, and $\phi^{m}_{hji}$ is the trade cost of firm $h$ importing its products produced in country $j$ to country $i$.

Based on $w_h$ and $v_h$, I then construct the aggregate weighting matrices, $W_n$ and $V_n$, which consist of $w_h$ and $v_h$, respectively, along the diagonal and 0 everywhere else. Formally,

$$W_{N\times K} = \begin{bmatrix} w_1 & 0 & 0 & 0 \\ 0 & w_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & w_K \end{bmatrix} \text{ and } V_{N\times K} = \begin{bmatrix} v_1 & 0 & 0 & 0 \\ 0 & v_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & v_K \end{bmatrix} \tag{24}$$

where $w_h \equiv \{\omega_{hij}\}$, $v_h \equiv \{v_{hij}\}$, $N$ represents the number of countries and $K$ represents the number of firms. Last, I assume that the weighting matrix used to estimate the autocorrelation in disturbances, $M_n$, is identical to $W_n$.

### 3.3 Data description

I employ a dataset of French manufacturing MNCs to estimate the determinants of multinational subsidiary networks. This dataset is supplied by the AMADEUS, a comprehensive database that contains the financial and ownership information of public and private European firms. AMADEUS is collected by information providers at each national official public body (e.g., Institut National de la Propriete Industrielle (National Institute for Industrial Property) in the case of France) and has a particularly good coverage for countries including France, which partly motivated the use of French firms for this analysis.

Before discussing explanatory variables, let us first define the dependent variable of this analysis, $y_{hi}$. $y_{hi}$ is a dummy variable that is equal to 1 if firm $h$ has at least one subsidiary in country $i$ at 2006 and 0 otherwise. I include all the countries that have 10 or more French MNC subsidiaries as potential hosts, resulting in a total of 41 countries,[9] and all the French multinationals that report financial and subsidiary data in AMADEUS and invest in at least one of the included countries.

As discussed in Section 3, three subsidiary-level variables are used to construct the weighting matrices. These variables include (i) the distance between each pair of countries (as a proxy for transport cost), (ii) host-country tariff rates on firms' final product and goods produced in each subsidiary, and (iii) the input-output coefficient between the firm's final product and the subsidiary's primary good. The tariff data are taken from the WITS database and the input-output table is available from the INSEE. To capture host countries' differences in comparative advantage (an element of the weighting matrix defined in equation (23)), I follow the existing literature and use countries' differences in factor endowment measured by capital-labor ratio as a proxy.[10] The construction of factor abundance ratio is described below.

In addition to the spatial lag terms, conventional host-country characteristics are also included to explain firms' location choices. First, I use host countries' real gross domestic product measured in 2000 U.S. dollars as a proxy for market demand. Second, I include host countries' capital-labor ratio to control for their comparative advantage. A country's capital-labor ratio is the ratio of capital stock relative to the size of total labor force, where the level of capital stock is constructed by using the perpetual inventory method outlined in Leamer (1984) and assuming a depreciation rate of 7%. The data of GDP, investment and labor force are all taken from the World Development Indicators. I also take into account host countries' tax policy by including the maximum corporate tax rate.[11] This data is available from the U.S. Office of Tax Policy Research. Furthermore, I include the cost of starting business in each potential host country, a variable provided by the World Bank's World Development Indicators, as a proxy for the fixed cost of establishing foreign subsidiaries.

I also follow the existing literature and separately account for the trade costs between France and host countries.[12] Specifically, I include the distance and the existence of a border between a potential host and France, host-country tariffs on France and French tariff rates on the hosts. Note, however, that the first two variables are also correlated with the fixed cost of investing in a foreign country (e.g., monitoring cost) and have therefore

[9] This criteria also helps to keep the weighting matrices within a computationally feasible size.

[10] I also used countries' differences in sectoral unit labor costs to construct the weighting matrix. While the results were qualitatively similar, the sample size was reduced because of the missing values in the unit labor cost data.

[11] Ideally, I would like to use host countries' applied corporate tax rates. But such data consist of a large number of missing values for the countries included in the sample.

[12] An alternative is to treat France the same as any other country where French MNCs have subsidiaries and include it in the construction of spatial lag terms. This was found not to change the results qualitatively.

an ambiguous effect on MNCs' location decision. All the tariff data used in the paper are applied sectoral tariff rates at the NACE 4-digit level and reflect the preferential trade relationships that pertain to the sample countries. The source of this data is the World Integrated Trade Solution (WITS) database.

Finally, the paper takes into account two firm-level characteristics. First, motivated by the recent literature that relates firm productivity with the decision to invest in foreign countries (see, for example, Helpman *et al.*, 2005), I control for each MNC's total factor productivity (TFP). Specifically, I use the financial data reported in the AMADEUS database and the approach introduced in Levinsohn and Petrin (2003) to obtain estimates of TFP. According to Helpman *et al.* (2005), firms with a greater TFP are more likely to engage in FDI than their less efficient counterparts. I also include firms' labor intensity (i.e., the ratio of labor expenditure in value added) and expect firms with a higher labor intensity to have a greater likelihood to invest in labor-abundant countries. To test this prediction, I include labor intensity both independently and in an interaction term with host-country capital-labor ratio. The source and summary statistics of the above variables are provided in Table A.1.

# 4 Geographic attributes of French MNC subsidiary networks

Before turning to the econometric analysis, let us first take a look at the geographic attributes of French MNC subsidiary networks. As indicated by Figure 4, there is a large variation across French multinationals in the number of invested countries. The majority of French MNCs concentrate their foreign production activities in less than 3 countries while some spread to as many as 63. It is, however, clear that the density of firms falls with the number of invested countries (for all firms that invest in at least two countries).

[Figure 4 about here]

Now what about the geographic density of subsidiaries? In Figure 5, I plot the level of distance between each pair of subsidiaries owned by the same French MNCs. As shown in the graph, the closest two subsidiaries is 66 kilometers apart (located in Austria and Slovakia) and the furthest is 19,845 kilometers (in Estonia and New Zealand). The majority of subsidiaries are less than 6,126 kilometers apart, while the average distance is about 6,000. Also noteworthy in this graph is the shape of the density curve, in particular, the two humps at around 1,200 and 9,100 kilometers. This implies that a large percentage of French MNC subsidiaries are either clustered in neighboring countries (such as the EU members) or located relatively distant from each other. This is further confirmed in Figure 6 where the average subsidiary distance is calculated for each firm: While a significant fraction of French

MNCs have a dispersed subsidiary network, many of them concentrate their subsidiaries geographically.

[Figures 5 and 6 about here]

Next, I examine the tariff level between subsidiaries. As shown in Figure 7, more than 30 percent of subsidiary pairs do not have tariff between each other and more than 50 percent have 7% or lower tariff rates. Figure 8 further shows that more than 15 percent of French MNCs locate their subsidiaries in countries where tariffs have been removed for each other and 50% percent face an average of 6% or lower tariff when exporting from one subsidiary to another. This seems to suggest that French MNCs are not always driven by the tariff-jumping motive when they choose their foreign production locations; a large percentage of them invest in countries where they can export to without paying tariff.

[Figures 7 and 8 about here]

So what motivates these multinationals to build a low-trade-cost subsidiary network? Is it the comparative-advantage motive which leads multinational firms to separate their production stages in countries with different comparative advantages and engage in intra-firm trade between subsidiaries? To answer this question, I compare each subsidiary pair's difference in their countries' factor abundance, i.e., capital-labor ratio. I find, in Figure 9, that while a large number of subsidiary pairs are located in countries with similar comparative advantage — as one would expect for multinationals whose main motive to invest abroad is to avoid trade costs, more than 50 percent of subsidiaries are set up in countries whose difference in capital-labor ratio exceeds 127%. This result remains similar when I examine subsidiaries' average K/L difference at the firm level in Figure 10. While 30 percent of French multinationals set up their subsidiaries in countries with 20% or smaller difference in factor abundance, there are also 30 percent of MNCs investing in countries whose capital-labor ratio difference exceeds 200%.

[Figures 9 and 10 about here]

I then repeat the above step for subsidiary pairs with relatively high trade costs and those whose trade cost, measured by either distance or tariff, is below average. The comparison is presented in Figures 11 and 12. I find, in Figure 11, that the density curve of K/L differential is shifted slightly rightward for subsidiaries that are relatively proximate to each other, suggesting a greater difference in these subsidiaries' comparative advantage. This pattern is similarly found when I compare high-tariff subsidiary pairs with those whose tariff is lower than average. Not only do the latter appear to have a greater K/L differential, the difference is also more significantly pronounced. These results suggest that the role of trade cost in multinationals' location decisions can vary by firms' incentive to invest abroad and countries' complementarity in comparative advantage. While a greater trade cost can stimulate horizontal (or market-access driven) investment especially in countries with similar factor endowment, multinationals that seek to build a vertical production network and take advantage of countries' different comparative advantages may prefer a lower trade cost within their subsidiary network.

[Figures 11 and 12 about here]

## 5 Econometric results

Let us now turn to the econometric analysis. I proceed by first estimating a conventional MNC location model without considering any spatial interdependence, i.e., assuming $\lambda_1$, $\lambda_2$, and $\rho$ of equation (17) all equal to 0. The results are largely consistent with the existing literature and are reported in Table 1.[13]

[Table 1 about here]

Firms are significantly more likely to invest in countries with a larger GDP. The distance between France and a potential host adversely affects French MNCs' probability of investing in the foreign country, suggesting the role of distance in raising the fixed cost of investment. Entry cost, as expected from the theory, significantly deters MNCs from establishing a subsidiary in a foreign country. Tariffs set by the host country also affect French MNCs' decision to invest abroad as suggested by the tariff-jumping motive hypothesis: Firms are more likely to invest in countries where tariff rates are high. French tariff rates on the host country, however, reduce MNCs' probability of investing in the country, implying that some French multinationals supply their home consumers with production abroad. Firms' labor intensity is also significantly correlated with the probability of investing in foreign countries. Those with a higher labor intensity are more likely to produce overseas. The effect of TFP on firms' decision to invest abroad is also confirmed in the data and consistent with the literature. More productive French firms are more likely than their less efficient counterparts to set up foreign subsidiaries.

Next let us examine the potential interdependence in multinationals' location choices. To do so, I estimate the (first-order) autoregressive spatial model outlined in equation (16) and (17).

[13] The $H_0$ column in Table 1 (and the following tables) summarizes the hypotheses on the effect of explanatory variables that are predicted by either the model or the literature.

## 5.1 horizontal interdependence

I proceed by first focusing on horizontal linkage, i.e., estimating $\lambda_1$ and $\rho$ and assuming $\lambda_2 = 0$. Table 2 reports the estimates obtained from the last step of Kelejian and Prucha's procedure using distance as the spatial weighting matrix. This step itself consists of two stages. In the first stage, the spatially-lagged dependent variable $W_n y_n^*$ (i.e., MNCs' spatially-weighted number of subsidiaries in third countries) is estimated as a function of explanatory variables $X_n^*$ and instrumental variables $W_n X_n^*$ and $W_n^2 X_n^*$ (i.e., the spatially-weighted third-country characteristics).[14] As suggested by the results in the left columns, most of the third-country variables exert a significant and expected effect on MNCs' location decisions. Multinationals are more likely to invest in third countries rather than a particular host when the third countries (i) have a larger market size, (ii) are more remote from France, (iii) require a smaller entry cost, (iv) impose a higher import tariff, and (v) are subject to a lower tariff when exporting to France. Labor-intensive multinationals are also more likely to have subsidiaries in labor-abundant countries (indicated by the parameter of the interaction term, $K/L$ ratio$\times L$ intensity), a result that is predicted in the comparative-advantage motive hypothesis.

[Table 2 about here]

I then obtain the fitted values of the spatially-lagged dependent variable $\widehat{W_n y_n^*}$ based on the first-stage estimates and estimate its effect on an MNC's decision to invest in a host country. As indicated by the parameter of $\widehat{W_n y_n^*}$, MNCs' probability of investing in a foreign country decreases with the proximity of their subsidiaries in third countries. This result is consistent with the theory of horizontal FDI in which multinationals' main motive to invest abroad is to avoid trade costs. The effect of other explanatory variables remains largely similar to the results in Table 1 while the overall performance of the estimation rises.

I repeat the above estimations using the tariff rates between host countries (on MNCs' final product) to construct the weighting matrix. The results are reported in Table 3. As indicated by the parameter of $\widehat{W_n y_n^*}$ in the right panel of Table 3, multinationals' incentive to invest in a foreign country increases with that country's tariff rates on the firms' subsidiaries in third countries.[15] Put differently, firms are less likely to produce in a country where the tariff of importing from existing production locations is relatively low.

[14] Note that a Cochrane-Orcutt transformation has been applied to all the variables using the estimated spatial autoregressive parameter $\widetilde{\rho}$ obtained from the first and second steps of Kelejin and Prucha's procedure. For example, $y_n^* = y_n - \widetilde{\rho}_n M_n y_n$.

[15] First-stage estimates on one of the instrument vectors $W_n X_n$, however, become mostly insignificant when tariff is used to construct $W_n$.

[Table 3 about here]

## 5.2 vertical interdependence

The evidence so far has indicated a significant horizontal interdependence between French MNCs' foreign subsidiaries. But is there any evidence of vertical interaction? To answer this question, I allow the effect of third-country subsidiaries to depend on not only trade costs but also countries' complementarity in comparative advantage and subsidiaries' input-output linkage. Specifically, I estimate $\lambda_1$, $\lambda_2$, and $\rho$, the weights of spatial matrices defined in equation (16) and (17) by including the two spatial weighting matrices defined in equations (22) and (23), $W_n$ and $V_n$. Table 4 reports the results.

[Table 4 about here]

Similar to Table 2, I find that an MNC's likelihood of investing in a foreign country falls when it has subsidiaries in proximity (indicated by the positive parameter of $\widehat{W_n y_n^*}$ reported in the right panel of Table 4). However, at any given level of proximity, the probability increases with the country's difference in capital-labor ratio from the third-country subsidiaries (indicated by the negative parameter of $\widehat{V_n y_n^*}$). It also increases with the input-output coefficient between subsidiaries (an element of the weighting matrix $V_n$), suggesting multinationals' incentive to reduce trade costs between vertically linked subsidiaries, in particular, the trade costs of importing intermediate inputs from upstream production location. These results indicate that the geographic density of subsidiaries can vary significantly across multinationals and across countries: Market-access driven multinationals tend to set up a dispersed subsidiary network whereas MNCs with separable production stages are likely to build a vertical production network in proximate countries.

[Table 5 about here]

The role of tariff between host countries also plays a similar role. As shown in Table 5, while multinationals are overall more likely to invest in countries that set a high tariff on the firms' existing subsidiaries, the effect is reversed for multinationals that seek to build a vertical production network. These multinationals are more likely to invest in countries that have a relatively low tariff on firms' upstream subsidiaries and are sufficiently different in capital-labor ratio. This is again consistent with multinationals' incentive to reduce the intra-network trade costs of transporting intermediate inputs.

## 5.3 sensitivity analysis

*alternative weighting matrices*

The above results have been obtained by using either distance or tariff as measures of trade cost. While transport cost and tariff are possibly the two most prominent forms of trade barriers, they do not capture all the trade costs faced by multinationals. I hence consider in this section an alternative measure in the construction of weighting matrices. Specifically, I use the level of trade flows as a proxy for host countries' openness toward one another. For example, for firm $h$ and subsidiary pair $i$ and $j$, I use the trade flow from country $j$ to country $i$ in firm $h$'s primary product in country $j$ as the proxy for country $i$'s openness to country $j$. The trade flow data are obtained from COMTRADE.

[Table 6 about here]

The third column of Table 6 reports the results based on this measure. The evidence indicates significant horizontal and vertical interdependence across subsidiaries. While multinationals have less incentive to duplicate their production in countries with close bilateral trade relationships, they are motivated to locate their upstream and downstream productions in these countries if they have distinct comparative advantages. This tendency also increases with the input-output relationship between subsidiaries.

*omitted variables*

The empirical analysis so far has included conventional host-country and firm characteristics in the estimations. The issue of omitted variables, however, can still arise. For example, host countries' sectoral market structure in both the final-product and intermediate-input industries can have a significant effect on multinationals' location decision. The latter information is especially important for firms that rely on unaffiliated upstream suppliers but is often unobservable for some countries and industries. Firm heterogeneities such as product quality can also affect multinationals' incentive to invest abroad and bias the estimated effect of other variables. But these data again are difficult to obtain. To address these concerns, I include in this section country-industry and firm fixed effects and check the robustness of the results. These fixed effects, by controlling for all the host-country sectoral characteristics and firm attributes, help me mitigate the influence of omitted variable and isolate the effect of spatial lag terms.

The results are summarized in the last column of Table 6. As shown, controlling for potential omitted variables does not qualitatively change the results. The estimated interdependence in the French MNC subsidiary networks remains largely similar and significant.

## 6 Conclusion

This study is one of the first attempts to estimate the interdependence between multinationals' foreign production locations. Using a detailed French multinationals dataset and a structural spatial econometric model, I find evidence of both horizontal and vertical linkages in French multinationals' subsidiary networks. The results indicate that while trade cost motivates multinationals to expand their subsidiary networks across countries, it discourages them from building a vertical production network. Firms are more inclined to locate vertically linked subsidiaries in countries where the cost of importing intermediate inputs is low and the comparative advantages are complementary. These results are also robust to the choice of weighting matrices in the spatial autoregressive model and control of unobserved host-country characteristics and firm heterogeneities.

This paper conveys important policy implications with regard to the effect of regional economic integration on individual MNCs' location decision. A decline in trade costs between a North and a South country through, for example, the formation of a preferential trade agreement can raise firms' incentive to invest in both participating countries and set up vertically linked subsidiaries. Regional integration between North and North countries can also increase the FDI by outside multinationals — because of the improved market access within the region, but the effect is not uniform. Multinationals would have little incentive to have subsidiaries in both participating countries and would prefer instead to concentrate geographically. Countries with more favorable investment terms, in this case, would gain multinationals at the expense of their preferential trading partners.

## References


[1] Baltagi, Badi H., Peter Egger, and Michael Pfaffermayr. (2007) "Estimating Models of Complex FDI: Are There Third-Country Effects?" *Journal of Econometrics* 140 (1), 260-281.

[2] Blonigen, Bruce, Ronald Davies, Glen Waddell, and Helen Naughton (forthcoming) "FDI in Space: Spatial Autoregressive Lags in Foreign Direct Investment", *European Economic Review*.

[3] Brainard, S. Lael (1997). "An Empirical Assessment of Proximity-Concentration Trade-off between Multinational Sales and Trade". *American Economic Review* 87: 520-544.

[4] Carr, D., J. Markusen, and K.E. Maskus (2001). "Estimating the Knowledge-Capital Model of the Multinational Enterprise". *American Economic Review* 91: 691-708.

[5] Ekholm, Karolina, Rikard Forslid, and James Markusen (2007). "Export-Platform Foreign Direct Investment". *Journal of European Economic Association* 5(4): 776-95.

[6] Hanson, G., R. Mataloni, and M. Slaughter (2005). "Vertical Production Networks in Multinational Firms". *Review of Economics and Statistics* 87(4): 664-78.

[7] Head, Keith, and Thierry Mayer (2004). "Market Potential and the Location of Japanese Investment in the European Union". *Review of Economics and Statistics* 86 (4): 957-72.

[8] Helpman, E. (1984) "A Simple Theory of Trade with Multinational Corporations", *Journal of Political Economy* 92: 451-71.

[9] Helpman, E., M. Melitz, and S. Yeaple (2004). "Export versus FDI", *American Economic Review*, vol. 94, pp. 300–16.

[10] Hines, James, Jr., and Eric Rice (1994). "Fiscal Paradise: Foreign Tax Havens and American Business". *Quarterly Journal of Economics* 109 (1): 149-182.

[11] Kelejian, Harry H. and Ingmar R. Prucha (1998). "A Generalized Spatial Two-Stage Least Squares Procedures for Estimating a Spatial Autoregressive Model with Autoregressive Disturbances". *Journal of Real Estate Finance and Economics* 17: 99-121.

[12] Kelejian, Harry H. and Ingmar R. Prucha (1999). "A Generalized Moments Estimator for the Autoregressive Parameter in a Spatial Model". *International Economic Review* 40: 509-533.

[13] Krugman, Paul and Anthony Venables (1996). "Integration, Specialization, and Adjustment". *European Economic Review* 40: 959-67.

[14] Levinsohn, J. and A. Petrin (2003). "Estimating Production Functions Using Inputs to Control for Unobservables". *Review of Economic Studies* 70(2), No. 243, pp. 317-342.

[15] Markusen, James (1984). "Multinationals, Multi-Plant Economies, and the Gains from Trade". *Journal of International Economics* 16: 205-26.

[16] Markusen, James, and Anthony Venables (2000). "The Theory of Endowment, Intra-industry and Multi-national Trade". *Journal of International Economics* 52: 209-234.

[17] Motta, Massimo, and George Norman (1996). "Does Economic Integration Cause Foreign Direct Investment?". *International Economic Review* 37: 757-783.

[18] Puga, Diego and Anthony Venables (1997). "Preferential Trading Agreements and Industrial Location". *Journal of International Economics* 43: 347-68.

[19] Venables, Anthony (1996). "Equilibrium Locations of Vertically Linked Industries". *International Economic Review* 37(2): 341-59.

[20] Yeaple, Stephen (2003a). "The Complex Integration Strategies of Multinational Firms and Cross-Country Dependencies in the Structure of Foreign Direct Investment". *Journal of International Economics* 60(2): 293-314.

[21] Yeaple, Stephen (2003b). "The Role of Skill Endowments in the Structure of U.S. Outward Foreign Direct Investment". *Review of Economics and Statistics* 85 (3): 726-734.

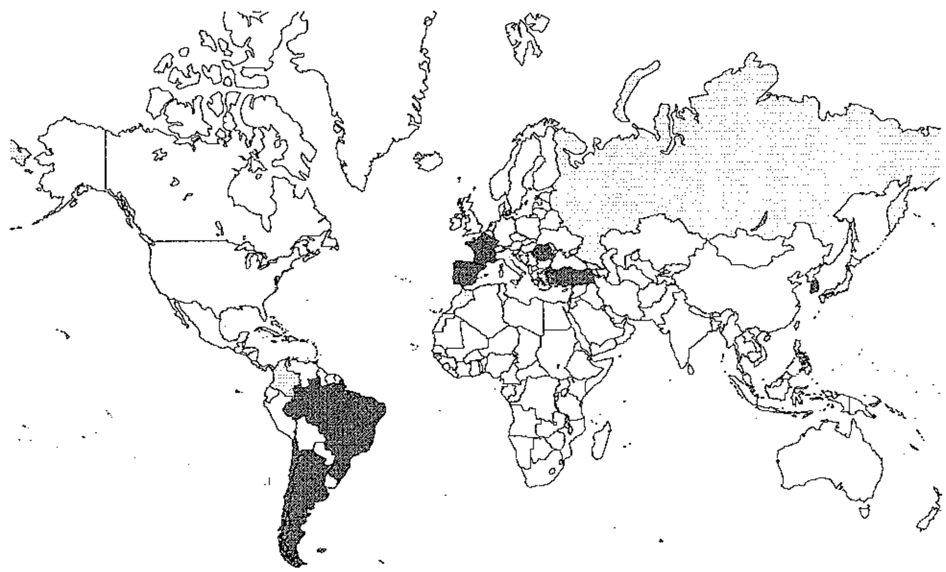

Figure 1: Renault's subsidiary network (the darker area represents intermediate-input (and final-product) production locations; the lighter area represents final-product producation locations)

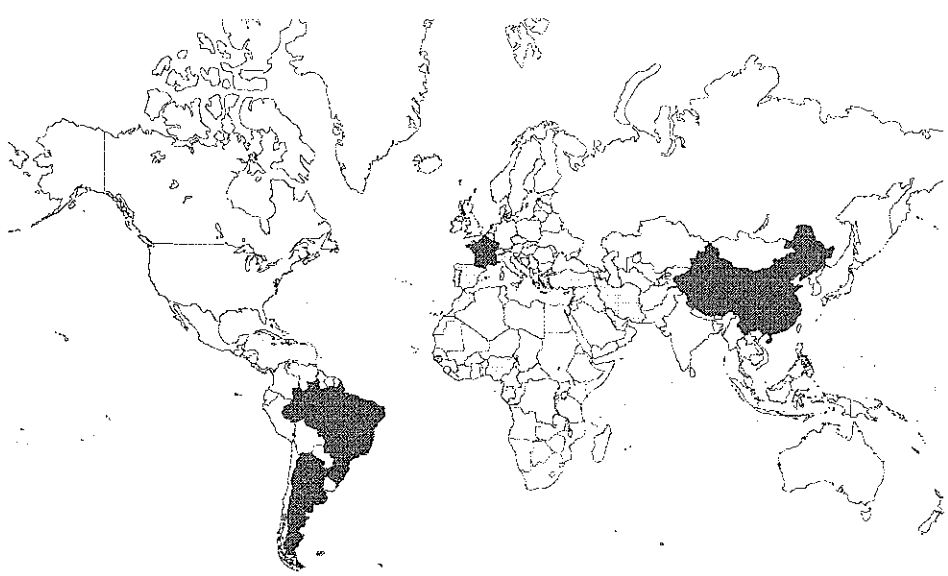

Figure 2: Peugeot's subsidiary network (the darker area represents intermediate-input (and final-product) production locations; the lighter area represents final-product producation locations)

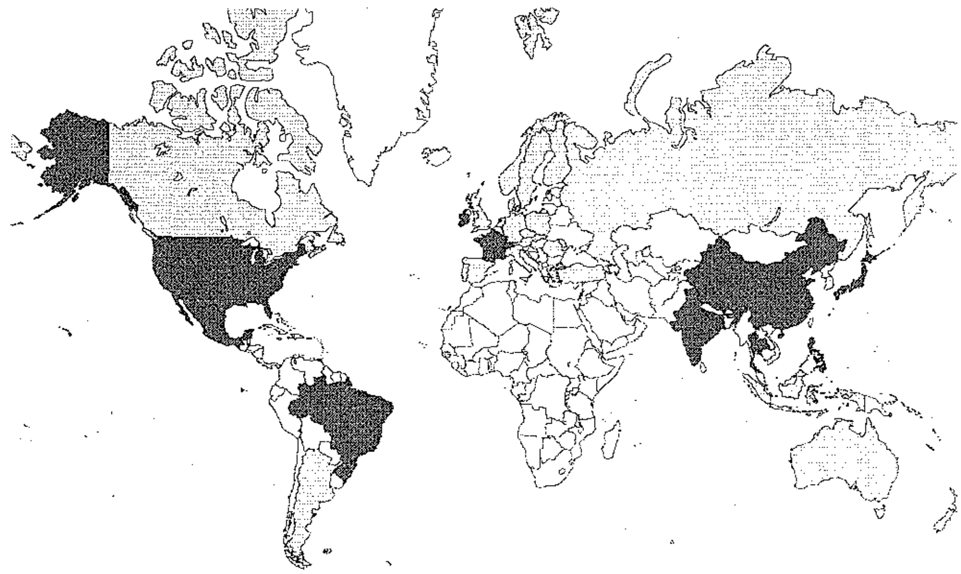

Figure 3: Essilor's subsidiary network (the darker area represents intermediate-input (and final-product) production locations; the lighter area represents final-product producation locations)

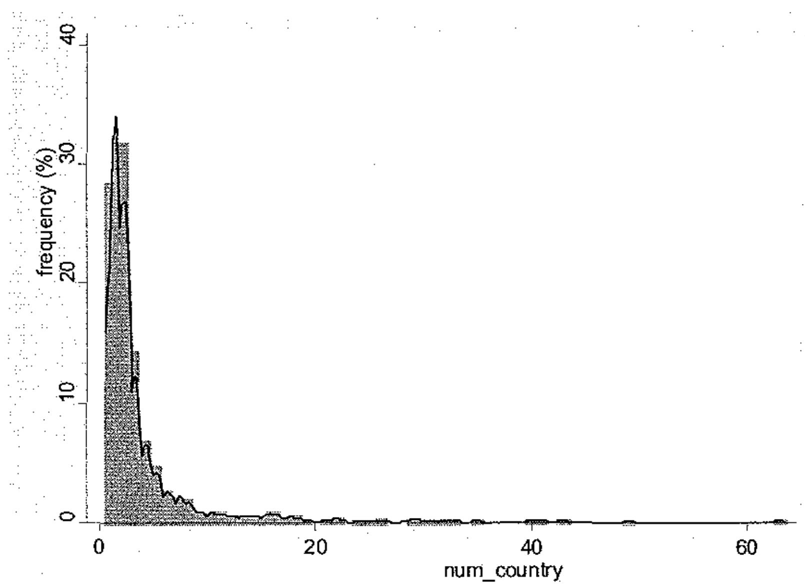


Figure 4: The distribution of French multinational firms by the number of invested foreign countries (kernel density estimates are represented by the curve)

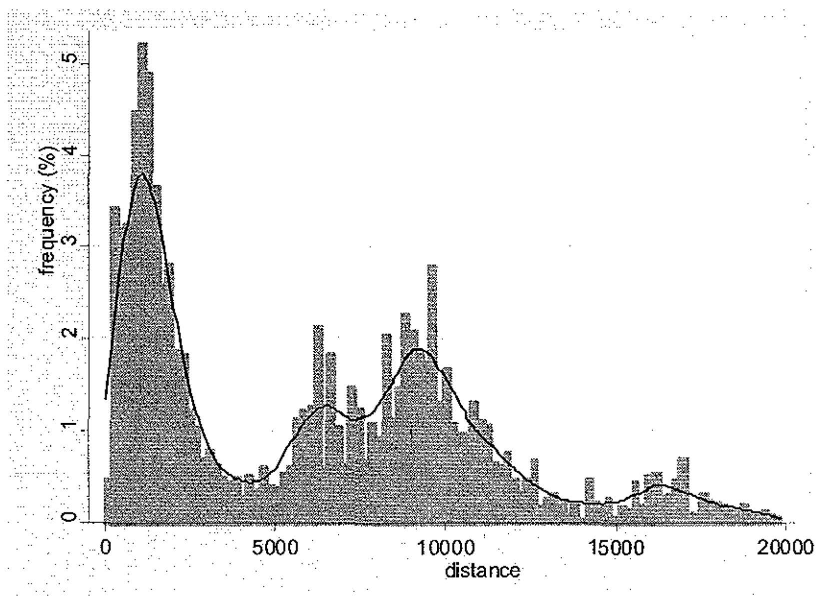


Figure 5: The distribution of French MNC subsidiaries by the distance between subsidiaries (kernel density estimates are represented by the curve)

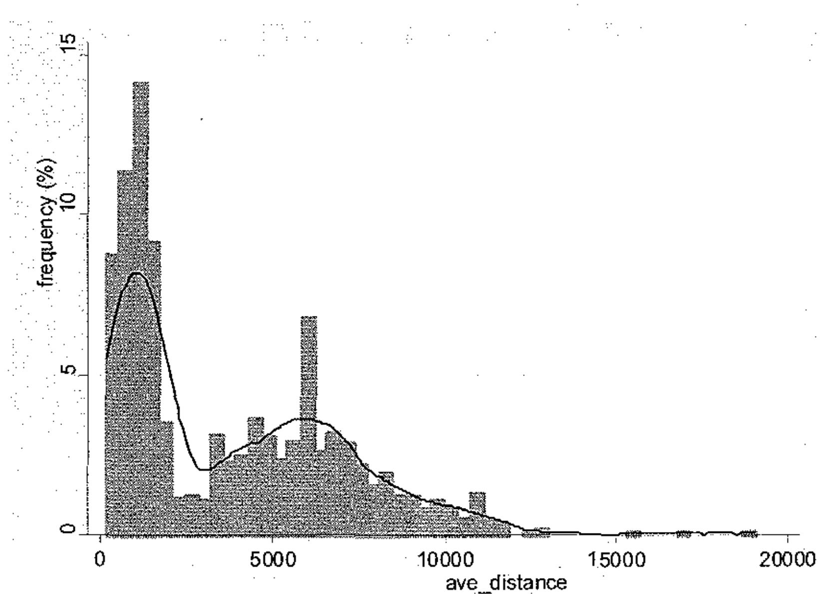


Figure 6: The distribution of French MNCs by the average distance between subsidiaries (kernel density estimates are represented by the curve)

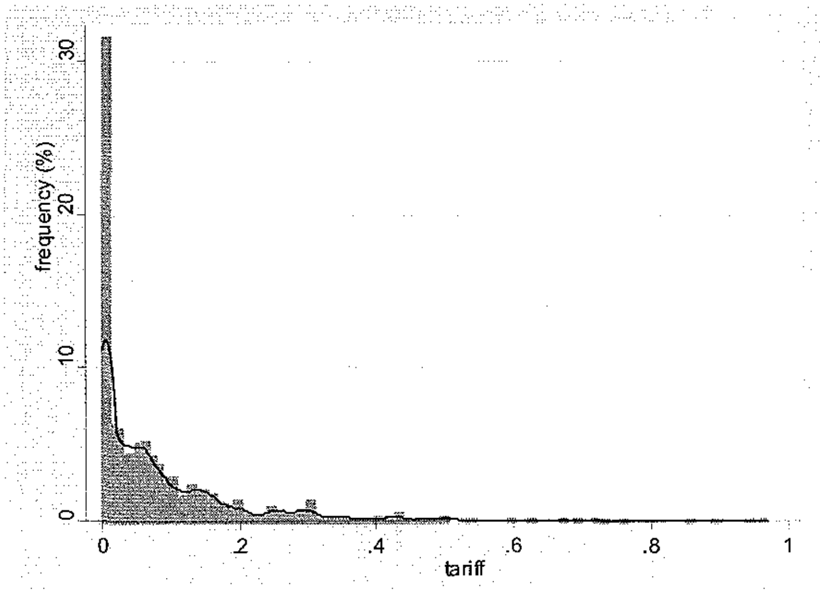


Figure 7: The distribution of French MNC subsidiaries by the tariff rate between subsidiaries (kernel density estimates are represented by the curve)

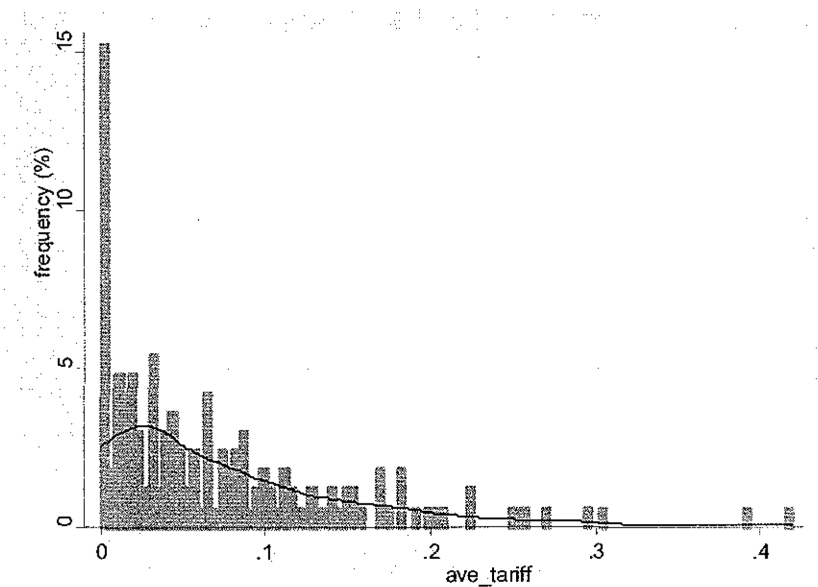


Figure 8: The distribution of French MNCs by the average tariff rate between subsidiaries (kernel density estimates are represented by the curve)

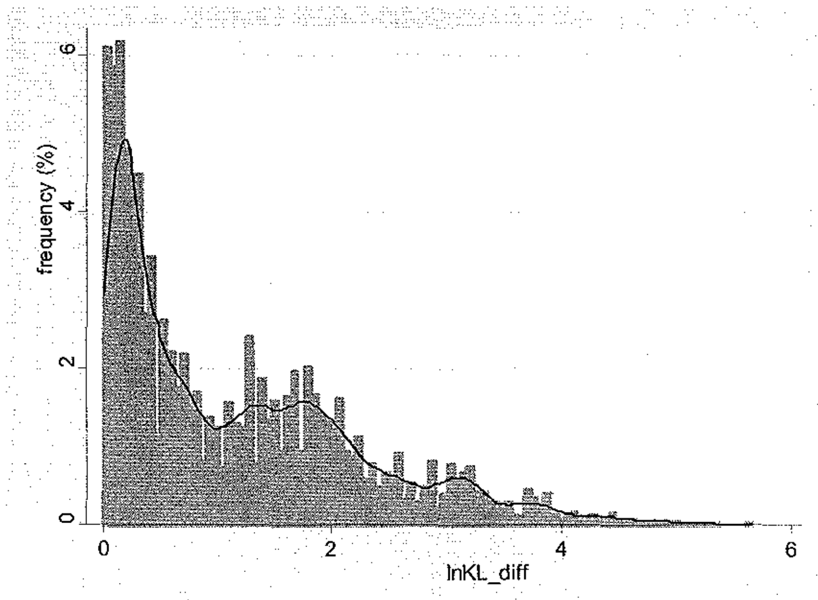


Figure 9: The distribution of French MNC subsidiaries by subsidiary countries' difference in capital-labor ratio (kernel density estimates are represented by the curve)

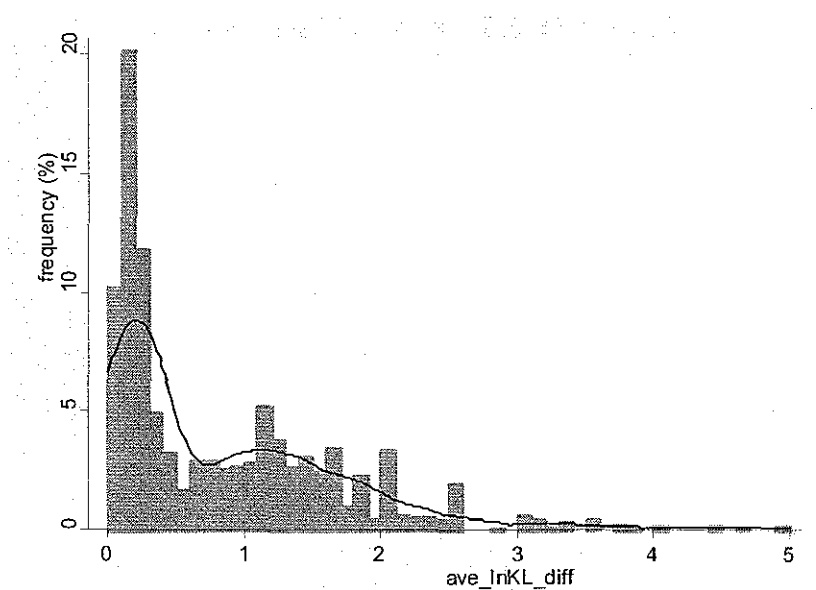


Figure 10: The distribution of French MNCs by subsidiary countries' average difference in capital-labor ratio (kernel density estimates are represented by the curve)

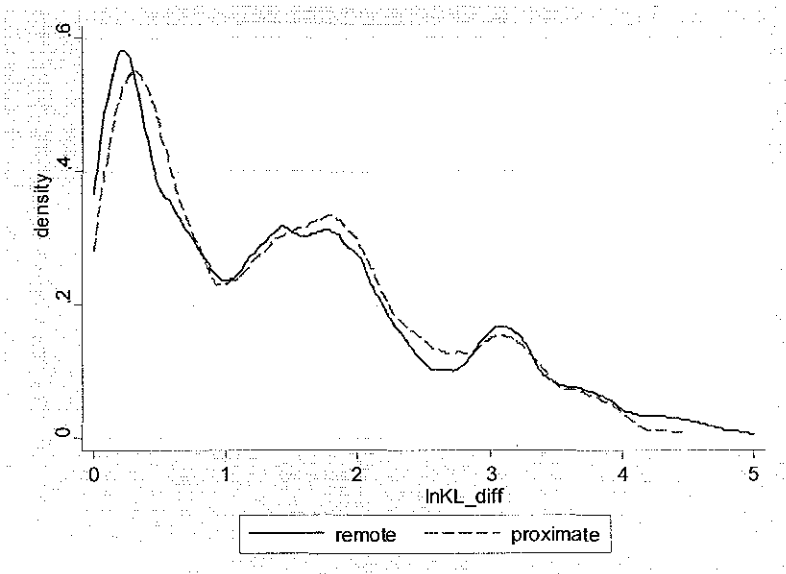


Figure 11: The distribution of French MNC subsidiaries by subsidiary countries' difference in capital-labor ratio (a comparison of kernel density curve between distant and proximate subsidiary pairs)

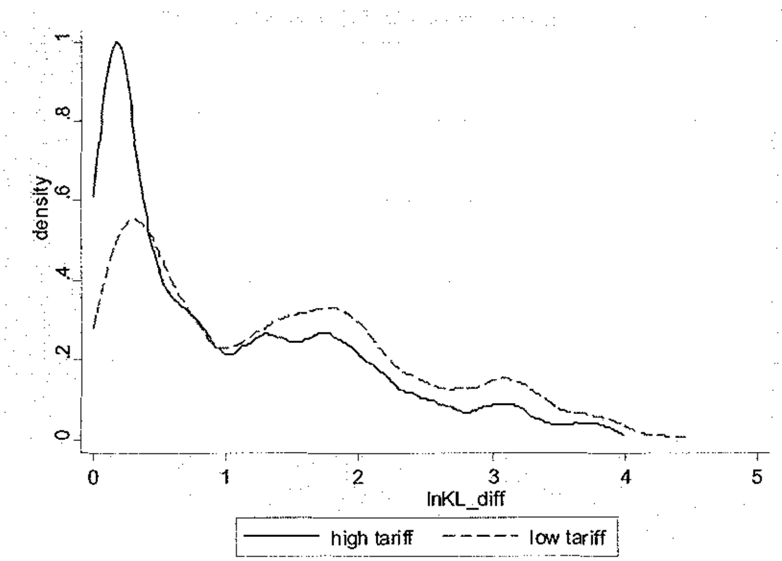


Figure 12: The distribution of French MNC subsidiaries by subsidiary countries' difference in capital-labor ratio (a comparison of kernel density curve between high- and low-tariff subsidiary pairs)

Table 1: Estimating the location decision without spatial factors

| Dependent variable: *$subsidiary_{kj}$* | $H_0$ | (1) coef. | s.e. |
|---|---|---|---|
| GDP | + | 0.09*** | (0.003) |
| K/L ratio | − | 0.03 | (0.08) |
| K/L ratio×L intensity | − | -0.05 | (0.08) |
| corporate tax | − | -0.004 | (0.01) |
| distance to home | − | -0.04*** | (0.04) |
| border | + | 0.19*** | (0.01) |
| entry cost | − | -0.006* | (0.004) |
| host country tariff | + | 0.008** | (0.004) |
| home country tariff | − | -0.03*** | (0.006) |
| firm labor intensity | + | 0.29*** | (0.08) |
| firm tfp | + | 0.07*** | (0.007) |
| No. of observations | | 9717 | |
| R square | | 0.25 | |
| Root MSE | | 0.40 | |

Note: (i) All variables except subsidiary, border and firm labor intensity are measured in natural logs; (ii) Standard errors are reported in the parentheses; (iii) ***, ** and * represent significance at 1%, 5% and 10%, respectively.

Table 2: Spatial model: distance as the weighting matrix

| Stage 1<br>Dependent: $W\cdot\ subsidiary^*_{k(j\neq i)}$ | $H_0$ | coef. | s.e. | Stage 2<br>Dependent: $subsidiary^*_{ki}$ | $H_0$ | coef. | s.e. |
|---|---|---|---|---|---|---|---|
| | | | | $W\cdot\ subsidiary^*_{k(j\neq i)}$ | + | 0.03** | (0.01) |
| $W\cdot$ GDP* | + | 0.06*** | (0.00) | GDP* | + | 0.09*** | (0.00) |
| $W\cdot$ K/L ratio* | – | 0.27*** | (0.03) | K/L ratio* | – | 0.04 | (0.08) |
| $W\cdot$ (K/L ratio×L intensity)* | – | -0.26*** | (0.03) | K/L ratio×L intensity* | – | -0.06 | (0.08) |
| $W\cdot$ corporate tax* | – | 0.06*** | (0.00) | corporate tax* | – | 0.00 | (0.01) |
| $W\cdot$ distance to home* | – | -0.10*** | (0.00) | distance to home* | – | -0.05*** | (0.01) |
| $W\cdot$ border* | + | 0.17*** | (0.00) | border* | + | 0.18*** | (0.01) |
| $W\cdot$ entry cost* | – | -0.01*** | (0.00) | entry cost* | – | -0.01* | (0.00) |
| $W\cdot$ host country tariff* | + | 0.02*** | (0.00) | host country tariff* | + | 0.01*** | (0.00) |
| $W\cdot$ home country tariff* | – | -0.01*** | (0.00) | home country tariff* | – | -0.03*** | (0.01) |
| $W\cdot$ L intensity* | | -0.62*** | (0.05) | L intensity* | + | 0.28*** | (0.09) |
| $W\cdot$ tfp* | | -0.05*** | (0.01) | tfp* | + | 0.07*** | (0.01) |
| $W^2\cdot$ GDP* | | -0.01*** | (0.00) | | | | |
| $W^2\cdot$ K/L ratio* | | -0.28*** | (0.11) | | | | |
| $W^2\cdot$ (K/L ratio×L intensity)* | | 0.30*** | (0.10) | | | | |
| $W^2\cdot$ corporate tax* | | 0.03*** | (0.01) | | | | |
| $W^2\cdot$ distance to home* | | 0.03*** | (0.00) | | | | |
| $W^2\cdot$ border* | | -0.01 | (0.01) | | | | |
| $W^2\cdot$ entry cost* | | 0.01*** | (0.00) | | | | |
| $W^2\cdot$ host country tariff* | | -0.01*** | (0.00) | | | | |
| $W^2\cdot$ home country tariff* | | 0.01 | (0.01) | | | | |
| $W^2\cdot$ L intensity* | | -0.10 | (0.13) | | | | |
| $W^2\cdot$ tfp* | | 0.01 | (0.02) | | | | |
| second-stage controls | | yes | | | | | |
| No. of observations | | 9717 | | No. of observations | | 9717 | |
| R square | | 0.76 | | R square | | 0.32 | |
| Root MSE | | 0.18 | | Root MSE | | 0.36 | |

Note: (i) All variables except subsidiary, border and firm labor intensity are measured in natural logs; (ii) Standard errors are reported in the parentheses; (iii) ***, ** and * represent significance at 1%, 5% and 10%, respectively.

Table 3: Spatial model: tariff as the weighting matrix

| Stage 1<br>Dependent: $W\cdot$ *subsidiary*$^*_{k(j\neq i)}$ | $H_0$ | coef. | s.e. | Stage 2<br>Dependent: *subsidiary*$^*_{ki}$ | $H_0$ | coef. | s.e. |
|---|---|---|---|---|---|---|---|
| | | | | $W\cdot$ *subsidiary*$^*_{k(j\neq i)}$ | + | 0.01** | (0.01) |
| $W\cdot$ GDP* | + | 0.02 | (0.02) | GDP* | – | -0.09*** | (0.00) |
| $W\cdot$ K/L ratio* | – | 0.36 | (1.67) | K/L ratio* | – | 0.04 | (0.08) |
| $W\cdot$ (K/L ratio×L intensity)* | – | -0.43 | (1.66) | K/L ratio×L intensity* | – | -0.06 | (0.08) |
| $W\cdot$ corporate tax* | – | 0.21 | (0.22) | corporate tax* | – | 0.00 | (0.01) |
| $W\cdot$ distance to home* | – | -0.05 | (0.06) | distance to home* | – | -0.04*** | (0.00) |
| $W\cdot$ border* | + | 0.32 | (0.28) | border* | + | 0.19*** | (0.01) |
| $W\cdot$ entry cost* | – | 0.02 | (0.07) | entry cost* | – | -0.01* | (0.00) |
| $W\cdot$ host country tariff* | + | -0.05 | (0.04) | host country tariff* | + | 0.01*** | (0.00) |
| $W\cdot$ home country tariff* | – | 0.02 | (0.04) | home country tariff* | – | -0.03*** | (0.01) |
| $W\cdot$ L intensity* | | 0.99 | (1.04) | L intensity* | + | 0.21** | (0.12) |
| $W\cdot$ tfp* | | 0.34*** | (0.11) | tfp* | + | 0.07*** | (0.03) |
| $W^2\cdot$ GDP* | | 0.00 | (0.01) | | | | |
| $W^2\cdot$ K/L ratio* | | 0.17*** | (0.06) | | | | |
| $W^2\cdot$ (K/L ratio×L intensity)* | | | | | | | |
| $W^2\cdot$ corporate tax* | | 0.08 | (0.13) | | | | |
| $W^2\cdot$ distance to home* | | 0.07*** | (0.03) | | | | |
| $W^2\cdot$ border* | | -0.41*** | (0.17) | | | | |
| $W^2\cdot$ entry cost* | | -0.19*** | (0.05) | | | | |
| $W^2\cdot$ host country tariff* | | -0.00 | (0.00) | | | | |
| $W^2\cdot$ home country tariff* | | 0.00 | (0.00) | | | | |
| $W^2\cdot$ L intensity* | | -1.87*** | (0.67) | | | | |
| $W^2\cdot$ tfp* | | -0.01*** | (0.00) | | | | |
| second-stage controls | | yes | | | | | |
| No. of observations | | 9717 | | No. of observations | | 9717 | |
| R square | | 0.79 | | R square | | 0.18 | |
| Root MSE | | 0.17 | | Root MSE | | 0.34 | |

Note: (i) All variables except subsidiary, border and firm labor intensity are measured in natural logs; (ii) Standard errors are reported in the parentheses; (iii) ***, ** and * represent significance at 1%, 5% and 10%, respectively.

Table 4: Spatial model: distance and K/L difference as the weighting matrix

| Stage 1 | | system of simultaneous equations | | | | Stage 2 | | | |
|---|---|---|---|---|---|---|---|---|---|
| Dependent variable: | | $W \cdot y_n^*$ | | $V \cdot y_n^*$ | | Dependent: | | | |
| $W(V) \cdot$ $subsidiary^*_{k(j \neq i)}$ | $H_0$ | coef. | s.e. | coef. | s.e. | $subsidiary^*_{ki}$ | $H_0$ | coef. | s.e. |
| | | | | | | $W \cdot$ $subsidiary^*_{k(j \neq i)}$ | + | 0.47*** | (0.11) |
| | | | | | | $V \cdot$ $subsidiary^*_{k(j \neq i)}$ | – | -0.42*** | (0.07) |
| $W(V) \cdot$ GDP* | + | 0.09*** | (0.01) | 0.08*** | (0.01) | GDP* | + | 0.10*** | (0.00) |
| $W(V) \cdot$ K/L ratio* | – | 0.66 | (0.56) | -0.42 | (0.47) | K/L ratio* | – | -0.02 | (0.08) |
| $W(V) \cdot$ (K/L ratio×L int.)* | – | -0.72 | (0.55) | 0.36 | (0.46) | K/L ratio×L intensity* | – | -0.02 | (0.08) |
| $W(V) \cdot$ corporate tax* | – | 0.06 | (0.07) | 0.16 | (0.12) | corporate tax* | – | 0.00 | (0.01) |
| $W(V) \cdot$ distance to home* | – | -0.24*** | (0.03) | -0.19*** | (0.04) | distance to home* | – | -0.04*** | (0.01) |
| $W(V) \cdot$ border* | + | 0.19*** | (0.07) | 0.28*** | (0.07) | border* | + | 0.16*** | (0.01) |
| $W(V) \cdot$ entry cost* | – | -0.01 | (0.01) | -0.06*** | (0.03) | entry cost* | – | -0.02*** | (0.00) |
| $W(V) \cdot$ host country tariff* | + | -0.01 | (0.01) | -0.02*** | (0.01) | host country tariff* | + | 0.002 | (0.00) |
| $W(V) \cdot$ home country tariff* | – | 0.03 | (0.02) | 0.03*** | (0.01) | home country tariff* | – | -0.03*** | (0.01) |
| $W(V) \cdot$ L intensity* | | 0.58** | (0.28) | 0.52*** | (0.20) | L intensity* | + | 0.39*** | (0.09) |
| $W(V) \cdot$ tfp* | | 0.03 | (0.03) | -0.01 | (0.01) | tfp* | + | 0.07*** | (0.01) |
| $W^2(V^2) \cdot$ GDP* | | -0.04*** | (0.01) | -0.08*** | (0.02) | | | | |
| $W^2(V^2) \cdot$ K/L ratio* | | -5.27*** | (2.54) | 2.24 | (2.57) | | | | |
| $W^2(V^2) \cdot$ (K/L ratio×L int.)* | | 5.25* | (2.52) | -2.20 | (2.54) | | | | |
| $W^2(V^2) \cdot$ corporate tax* | | 0.75* | (0.37) | 0.21 | (0.21) | | | | |
| $W^2(V^2) \cdot$ distance to home* | | 0.31*** | (0.07) | 0.37*** | (0.07) | | | | |
| $W^2(V^2) \cdot$ border* | | 0.11 | (0.15) | 0.66*** | (0.18) | | | | |
| $W^2(V^2) \cdot$ entry cost* | | -0.31*** | (0.07) | -0.18*** | (0.08) | | | | |
| $W^2(V^2) \cdot$ host country tariff* | | -0.03*** | (0.00) | -0.02*** | (0.00) | | | | |
| $W^2(V^2) \cdot$ home country tariff* | | 0.01 | (0.01) | 0.02 | (0.01) | | | | |
| $W^2(V^2) \cdot$ L intensity* | | 0.58 | (0.52) | -0.46 | (0.52) | | | | |
| $W^2(V^2) \cdot$ tfp* | | 0.02 | (0.02) | 0.04*** | (0.01) | | | | |
| second-stage controls | | yes | | yes | | | | | |
| No. of observations | | 9717 | | 9717 | | No. of observations | | 9717 | |
| R square | | 0.45 | | 0.42 | | R square | | 0.21 | |
| Root MSE | | 0.15 | | 0.20 | | Root MSE | | 0.36 | |

Table 5: Spatial model: tariff and K/L difference as the weighting matrix

| Stage 1 | | system of simultaneous equations | | | | Stage 2 | | | |
|---|---|---|---|---|---|---|---|---|---|
| Dependent: | | $W \cdot y_n^*$ | | $V \cdot y_n^*$ | | Dependent: | | | |
| $W(V)\cdot$ $subsidiary^*_{k(j\neq i)}$ | $H_0$ | coef. | s.e. | coef. | s.e. | $subsidiary^*_{ki}$ | $H_0$ | coef. | s.e. |
| | | | | | | $W\cdot$ $subsidiary^*_{k(j\neq i)}$ | + | 0.03*** | (0.00) |
| | | | | | | $V\cdot$ $subsidiary^*_{k(j\neq i)}$ | – | -0.02*** | (0.00) |
| $W(V)\cdot$ GDP* | + | 0.03 | (0.02) | 0.01 | (0.02) | GDP* | + | 0.09*** | (0.00) |
| $W(V)\cdot$ K/L ratio* | – | 2.04 | (1.60) | 2.44** | (1.22) | K/L ratio* | – | 0.08 | (0.08) |
| $W(V)\cdot$ (K/L ratio×L int.)* | – | -2.15 | (1.59) | -2.31*** | (1.19) | K/L ratio×L intensity* | – | -0.09 | (0.08) |
| $W(V)\cdot$ corporate tax* | – | 0.09 | (0.21) | -0.10 | (0.14) | corporate tax* | – | 0.00 | (0.01) |
| $W(V)\cdot$ distance to home* | – | -0.05 | (0.06) | 0.05 | (0.06) | distance to home* | – | -0.05*** | (0.01) |
| $W(V)\cdot$ border* | + | 0.09 | (0.21) | -0.07 | (0.29) | border* | + | 0.20*** | (0.01) |
| $W(V)\cdot$ entry cost* | – | -0.04 | (0.06) | 0.03 | (0.04) | entry cost* | – | -0.01*** | (0.00) |
| $W(V)\cdot$ host country tariff* | + | -0.01 | (0.04) | -0.005 | (0.00) | host country tariff* | + | 0.01 | (0.00) |
| $W(V)\cdot$ home country tariff* | – | -0.02 | (0.04) | -0.04*** | (0.01) | home country tariff* | – | -0.02*** | (0.01) |
| $W(V)\cdot$ L intensity* | | 0.53 | (0.98) | -2.78*** | (0.56) | L intensity* | + | 0.26 | (0.12) |
| $W(V)\cdot$ tfp* | | 0.32*** | (0.09) | -0.05*** | (0.01) | tfp* | + | 0.05** | (0.02) |
| $W^2(V^2)\cdot$ GDP* | | -0.01 | (0.01) | -0.01*** | (0.00) | | | | |
| $W^2(V^2)\cdot$ K/L ratio* | | 0.18*** | (0.15) | 0.52*** | (0.12) | | | | |
| $W^2(V^2)\cdot$ (K/L ratio×L int.)* | | | | -0.51*** | (0.12) | | | | |
| $W^2(V^2)\cdot$ corporate tax* | | -0.28*** | (0.10) | -0.15*** | (0.02) | | | | |
| $W^2(V^2)\cdot$ distance to home* | | 0.03 | (0.02) | 0.01*** | (0.00) | | | | |
| $W^2(V^2)\cdot$ border* | | -0.51*** | (0.14) | -0.00 | (0.02) | | | | |
| $W^2(V^2)\cdot$ entry cost* | | -0.13*** | (0.04) | -0.01* | (0.00) | | | | |
| $W^2(V^2)\cdot$ host country tariff* | | -0.00 | (0.00) | 0.001*** | (0.00) | | | | |
| $W^2(V^2)\cdot$ home country tariff* | | 0.00 | (0.00) | -0.001*** | (0.00) | | | | |
| $W^2(V^2)\cdot$ L intensity* | | -1.88*** | (0.55) | -0.11*** | (0.02) | | | | |
| $W^2(V^2)\cdot$ tfp* | | -0.01*** | (0.00) | 0.002*** | (0.00) | | | | |
| second-stage controls | | yes | | yes | | | | | |
| No. of observations | | 9717 | | 9717 | | No. of observations | | 9717 | |
| R square | | 0.58 | | 0.37 | | R square | | 0.18 | |
| Root MSE | | 1.22 | | 2.96 | | Root MSE | | 0.34 | |

Table 6: Sensitivity analysis: alternative weighting matrices and control of omitted variables

| Stage 2<br>Dependent: $subsidiary^*_{ki}$ | $H_0$ | coef. | s.e. | coef. | s.e. |
|---|---|---|---|---|---|
| $W \cdot subsidiary^*_{k(j \neq i)}$ | + | 0.001*** | (0.00) | 0.002*** | (0.00) |
| $V \cdot subsidiary^*_{k(j \neq i)}$ | − | -0.002*** | (0.00) | -0.003*** | (0.00) |
| GDP* | + | 0.09*** | (0.00) | | |
| K/L ratio* | − | 0.17 | (0.10) | | |
| K/L ratio×L intensity* | − | -0.20*** | (0.10) | -0.30*** | (0.11) |
| corporate tax* | − | 0.01 | (0.01) | | |
| distance to home* | − | -0.05*** | (0.01) | | |
| border* | + | 0.20*** | (0.01) | | |
| entry cost* | − | -0.01* | (0.00) | | |
| host country tariff* | + | 0.01*** | (0.00) | | |
| home country tariff* | − | -0.02*** | (0.00) | | |
| L intensity* | + | 0.31*** | (0.10) | | |
| tfp* | + | 0.09*** | (0.01) | | |
| country-industry fixed effect | | No | | Yes | |
| firm fixed effect | | No | | Yes | |
| No. of observations | | 9717 | | 9717 | |
| R square | | 0.32 | | 0.56 | |
| Root MSE | | 0.37 | | 0.22 | |

Note: (i) All variables except subsidiary, border and firm labor intensity are measured in natural logs; (ii) Standard errors are reported in the parentheses; (iii) ***, ** and * represent significance at 1%, 5% and 10%, respectively.

Table A.1: Summary statistics

| Variables | Source | Mean | Std. dev. | Min | Max |
|---|---|---|---|---|---|
| subsidiary | AMADEUS | 0.19 | 0.39 | 0 | 1 |
| GDP | WDI | 26.24 | 1.35 | 23.85 | 30.00 |
| K/L ratio | WDI | 6.73 | 2.23 | 0.07 | 12.22 |
| corporate tax | Office of Tax Policy Research | -1.25 | 0.29 | -2.40 | -0.91 |
| distance to home | City Distance Calculator | 7.81 | 1.19 | 5.57 | 9.73 |
| border | — | 0.14 | 0.35 | 0 | 1 |
| entry cost | WDI | 2.31 | 1.10 | 0 | 4.92 |
| host-country tariff | COMTRADE | 1.07 | 1.28 | 0 | 5.30 |
| home-country tariff | COMTRADE | 0.36 | 0.71 | 0 | 3.80 |
| firm labor intensity | AMADEUS | 0.63 | 0.19 | 0.01 | 0.99 |
| firm tfp (estimated) | AMADEUS | 0.59 | 0.65 | -3.17 | 3.54 |

Note: All variables except subsidiary, border and firm labor intensity are measured in natural logs.