Kleinert, Jörn; Toubal, Farid

Conference Paper

Distance and the margins of multinational firms' activities

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2010: Ökonomie der Familie - Session: Trade, Gravity and Distance, No. A4-V2

Provided in Cooperation with:
Verein für Socialpolitik / German Economic Association

Suggested Citation: Kleinert, Jörn; Toubal, Farid (2010) : Distance and the margins of multinational firms' activities, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2010: Ökonomie der Familie - Session: Trade, Gravity and Distance, No. A4-V2, Verein für Socialpolitik, Frankfurt a. M.

This Version is available at:
https://hdl.handle.net/10419/37165

Distance and the margins of multinational firms' activities ⋆

Jörn Kleinert [a] and Farid Toubal [b]

[a] *Department of Economics, Eberhard-Karls-University of Tübingen, Germany*

[b] *University of Angers, Paris School of Economics, and CEPII, France*

Abstract

We present a model of multinational firms that predicts that aggregated affiliates' sales fall in distance. The distance effect on foreign affiliate sales is driven by the extensive margin: distance affects the number of affiliates negatively while it has an ambiguous effect on the average affiliate sales. We derive gravity equations explaining aggregate and average foreign affiliate sales and their number from the model. We discuss the revealed endogeneity bias and propose a system estimation to cure it. To assess the relative importance of the extensive and the intensive margins of foreign affiliates activities we use a comprehensive German data set. We find the extensive margin driving the negative effect of distance on multinational firms' activities.

Keywords: Gravity equation, multinational firms, distance costs.

JEL classification: F23, F12, C21

⋆ Thanks to Bruce Blonigen, and Thierry Mayer. The hospitality of the Bundesbank and access to its firm-level database "Micro Data Direct Investment" (Midi) are gratefully acknowledged. The project has also benefited from financial support through the German Science Foundation under the project Multinational Enterprises: New Theories and Empirical Evidence from German Firm-Level Data (BU 1256/6-1). This paper has been partly written as the authors were visiting the Hitotsubashi University Tokyo, Japan. We thank Hitotsubashi University for its hospitality. We are of course solely responsible for any remaining errors.

1 Introduction

Research on multinational firms' foreign activities has emphasized the importance of distance costs for cross-border activities. A growing empirical literature uses gravity-type equations to investigate the determinants of various type of these activities. This literature finds a postive effect of home and host countries' income and a negative impact of geographical distance on foreign direct investment (De Sousa and Lochard, 2006; Eaton and Tamura, 1994; Eichengreen and Tong, 2005; Wei, 2000) or on foreign affiliates' sales (Buch *et al.*, 2005; Carr *et al.*, 2001).

The negative effect of distance on foreign affiliate sales results mainly from the extensive margin: the number of foreign affiliates that are active in a country falls with the distance to the partner country. The intensive margin, the change of the average size of an affiliates with respect to distance, does not seem to be so important (Buch *et al.* 2005). The same has been found and discussed for international trade (Eaton *et al.* 2004). However, while the effect of distance on trade is in line with economic theory, the distance effect on foreign activities of foreign affiliates poses a puzzle.

We explain this "*distance costs puzzle*" within the bilateral proximity-concentration framework, because this framework has mainly be used as theoretical foundation for the use of the gravity equation in the empirical literature. We relax the simplifying assumption that fixed costs of establishing a foreign affiliate are identical for all countries and include intermediate goods in the production

process. We show in a model of monopolistic competition with heterogenous firms and specific intermediate goods, that increasing distance costs affect negatively the number of foreign affiliates if fixed set-up costs increase in distance. We motivate the assumption that fixed costs increase with distance by the fact that distance raises upfront search costs and organization costs (Chaney, 2006; Rauch, 1999).

We derive the gravity equation for foreign affiliate sales, the number of firms abroad, and the average size of the foreign affiliate from the model. The derived gravity equations reveal an endogenity bias that results from the fact that the number of firms active abroad and the sales of their affiliates are jointly determined. This bias can be corrected by estimating a two equation system explaining the number of firms and aggregate sales jointly. The theory offers two predictions concerning the impact of distance costs on aggregate foreign affiliates' production, the number of foreign affiliates, and their average size: (i) aggregate sales of affiliates fall, because fewer firms are active in more distant countries, and (ii) the effect of distance on the average size of a foreign affiliate in contrast is ambiguous.

We test our two predictions using a comprehensive data set on German multinationals' foreign sales. We find a large and significantly negative effect of distance on aggregate affiliate sales and the number of foreign affiliates in a particular host country and no significant effect of distance on the average size of the foreign affiliate using different methodologies. Thus, our results suggest that distance works mainly through the extensive margin. A higher impor-

tance of the extensive margin is also found for international trade in Eaton *et al.* (2004).

The remainder of the paper is structured as follows. In Section 2, we present the model and derive the equations to estimate. In Section 3, we provide a discussion of the data and present the empirical strategy. In section 4, we present our main results and the robustness check. In section 5, we conclude.

2 A heterogenous firm model with distance dependent fixed costs and specific intermediate goods

2.1 The model

We consider an economy with two sectors: agriculture, which produces a homogeneous good A and manufacturing which produces a bundle M of differentiated goods. Consumers purchase A and M and have identical preferences described by a utility function defined on A and M. Consumers preferences for single varieties of the M good are described by a sub-utility function defined on the varieties. The utility function of the representative consumer from country j has the Cobb-Douglas form described in equation (1):

$$U_j = X_{Aj}^{\mu} X_{Mj}^{1-\mu} \tag{1}$$

where $0 < \mu < 1$. X_{Mj} is a sub-utility function of CES-type defined in (2)

$$X_{Mj} = \left[\int_i \int_k x_{kij}^{(\sigma-1)/\sigma} dk di \right]^{\sigma/(\sigma-1)} \tag{2}$$

x_{kij} is the consumption by an individual in country j of a single variety produced by firm k from country i. The elasticity of substitution σ is the same for any pair of product and larger than one. We assume monopolistic competition in manufacturing so that each variety of the manufacturing good is produced by only one firm.

We assume heterogeneous firms. Firms differ in productivity which they draw from a common distribution $g(\omega)$ at market entry. Differences in productivity translate into different marginal costs, different prices and different quantities sold by each firm k. We denote the productivity of a firm k by ω_k and its marginal costs by a_k. Profit maximization yields a fixed markup over the marginal costs a_k of $\rho = (\sigma - 1)/\sigma$. Thus, the price of firm k located in country i and selling in country j $p_{kij} = a_{kij}/\rho$ leads to firm specific quantities sold in j. The optimal quantity sold in country j by a firm located in country i is given by equation (3).

$$x_{kij} = p_{kij}^{-\sigma}(1-\mu)Y_j P_{Mj}^{\sigma-1} \tag{3}$$

The price index, P_{Mj}, in country j includes prices that differ in three dimensions. First, the price index is affected by the difference in the productivity among firms and thus their different prices and quantities. Second, the price index in country j depends on the home country i of the firms producing the goods sold in j. Third, the price index depends on firms' choice how to serve the foreign market. Firms that choose to serve the foreign market decide whether to export or to produce abroad. Their choice depends on

their productivity level ω_k. The price index of country j is given by $P_{Mj} = \left[\int_k \left(p^h_{kij}\right)^{1-\sigma} dk\right]^{1/(1-\sigma)}$. p^h_{kij} denotes the price of firm k from country i selling in market j and having chosen the mode of entry h. The subscript $h = Ex, MNE$ indicates respectively whether a firm exports or produces abroad. The mass of firms in home country i is denoted by M_i.

Firms produce using specific intermediate goods. These inputs are not substitutable and must therefore be imported by foreign affiliates of a firm that produces abroad. The technology of the final goods producer is described by a Cobb-Douglas cost function with wages w, the price of the intermediate good q, and the cost share ϵ, i.e. $a_{kij} = \frac{1}{\omega_k}(w_j)^{\epsilon}(q_{ij})^{1-\epsilon}$, where j might equal i. We assume perfect competition among the intermediate goods' producer and a production technology that uses one unit of labor to produce one unit of the intermediate good. The intermediate good is uses as numeraire, thus its price q_{ii} and the wage w_i equal one and marginal costs simplify to $a_{kii} = \frac{1}{\omega_k}$. Free trade in the homogenous good equalizes wages w_j across countries.

Each firm compares the profit related to each mode of entry in market j. Firms that have a higher productivity level than ω^{Ex}_{ij} are active in the foreign market and earn positive profit. Firms that have a productivity level equal to ω^{MNE}_{ij} are indifferent between exporting and producing abroad because both strategies yield the same profit. Firms with a productivity level higher than ω^{MNE}_{ij} produce in country j and have higher profits than firms with a lower productivity level that export to j. We use the zero-profit conditions to derive the minimum productivity levels (a) for a firm that produces only for its home

market i (b) for an exporting firm from i selling in j and (c) for a firm from i that also produces in j. These are given in equations (4).

$$\omega_i^{Dom} = \left(\frac{(1-\mu)Y_i(1-\rho)}{f_i^{Dom}}\right)^{1/(1-\sigma)} P_{Mi}^{-1}\rho^{-1} \tag{4a}$$

$$\omega_{ij}^{Ex} = \tau_{ij}\left(\frac{(1-\mu)Y_j(1-\rho)}{f_{ij}^{Ex}}\right)^{1/(1-\sigma)} P_{Mj}^{-1}\rho^{-1} \tag{4b}$$

$$\omega_{ij}^{MNE} = \left(\tau_{ij}^{(1-\epsilon)(1-\sigma)} - \tau_{ij}^{1-\sigma}\right)^{1/(1-\sigma)}\left(\frac{(1-\mu)Y_j(1-\rho)}{f_{ij}^{MNE} - f_{ij}^{Ex}}\right)^{1/(1-\sigma)} P_{Mj}^{-1}\rho^{-1} \tag{4c}$$

Where ω_i^{Dom} is the productivity level of the least productive (domestic) firm from country i that is active in its home market. We assume that fixed costs of exporting f^{Ex} is a fixed share ϕ of the fixed costs, f^{MNE}, associated with the production abroad with $0 < \phi < 1$, i.e. $f^{MNE} - f^{Ex} = (1-\phi)f^{MNE}$.

We use a Pareto distribution to parameterize the distribution of firms with respect to their productivity ω_k. We denote the shape parameter by κ. Aggregated affiliates' sales of all firms from country i in the foreign market j, AS_{ij}, are thus given by equation (5) which is derived in the Appendix A.

$$\begin{aligned} AS_{ij} &= \int_{\omega_{ij}^{MNE}}^{\infty}\left(\frac{\omega_k\rho}{\tau_{ij}^{1-\epsilon}}\right)^{\sigma-1} g(\omega)M_i\frac{(1-\mu)Y_j}{P_j^{1-\sigma}}d\omega_k \\ &= \left(\frac{\omega_i^{Dom}}{\omega_{ij}^{MNE}}\right)^{\kappa} M_i\left(\frac{\kappa}{\kappa-\sigma+1}\right)\left(\frac{\omega_{ij}^{MNE}\rho}{\tau_{ij}^{1-\epsilon}}\right)^{\sigma-1}\frac{(1-\mu)Y_j}{P_j^{1-\sigma}} \end{aligned} \tag{5}$$

The first term gives the probability of a firm from country i to have an affiliate in country j. Multiplied by the total mass of firms from i M_i, this gives the number of affiliates in country j that we denote by N_{ij}. The product

$\frac{\kappa}{\kappa-\sigma+1}\left(\frac{\omega_{ij}^{MNE}\rho}{\tau_{ij}^{1-\epsilon}}\right)^{\sigma-1}$ describes the average weighted price of goods from affiliates of country i firms in j. Multiplication with the last term gives the sales of the average foreign affiliate of firms from country i in country j.

The threshold productivity level ω_{ij}^{MNE} determines the minimal and the average size and the number of affiliates from country i in country j. It is easy to see in (5) that the threshold productivity level ω_{ij}^{MNE} is inversely related to aggregate affiliate sales if $\kappa > \sigma - 1$.[1] From the first term of (5), we see that the threshold productivity level is negatively related to the number of firms from i producing in j. From the second term, we see that the threshold productivity is positively related to the average size of their affiliates. The effect of the minimum productivity level on aggregate affiliate sales is then given by

$$\frac{\partial AS_{ij}}{\partial \omega_{ij}^{MNE}} = (-\kappa + (\sigma - 1))\left(\omega_{ij}^{MNE}\right)^{(-\kappa+\sigma-2)}\Lambda_j < 0$$

where $\Lambda_j = (\omega_i^{Dom})^{\kappa} M_i \left(\frac{\kappa}{\kappa-\sigma+1}\right)\left(\frac{\rho}{\tau_{ij}^{1-\epsilon}}\right)^{\sigma-1}\frac{(1-\mu)Y_j}{P_{Mj}^{1-\sigma}}$. The effect is negative, since κ is larger than $\sigma - 1$.

The number of firms from country i setting up affiliates in j $N_{ij} = \left(\frac{\omega_i^{Dom}}{\omega_{ij}^{MNE}}\right)^{\kappa} M_i$ is negatively affected by the minimum productivity level ω_{ij}^{MNE}. Average sales as_{ij} of foreign affiliates, in contrast, depend positively on the minimum productivity level.

$$\frac{\partial N_{ij}}{\partial \omega_{ij}^{MNE}} = -\kappa\left(\frac{\omega_i^{Dom}}{\omega_{ij}^{MNE}}\right)^{\kappa}\frac{M_i}{\omega_{ij}^{MNE}} < 0 \tag{6a}$$

[1] κ must be larger than $\sigma - 1$ for the integral in equation (5) to be finite.

$$\frac{\partial as_{ij}}{\partial \omega_{ij}^{MNE}} = \frac{\kappa(\sigma-1)}{\kappa-\sigma+1}\left(\omega_{ij}^{MNE}\right)^{\sigma-2}\left(\frac{\rho}{\tau_{ij}^{1-\epsilon}}\right)^{\sigma-1}\frac{(1-\mu)Y_j}{P^{1-\sigma}} > 0 \qquad (6b)$$

2.2 Distance effects

Distance D_{ij} affects trade costs τ_{ij} and the fixed costs f_{ij}. Trade costs, in turn, affect aggregate and average sales directly and jointly with distance indirectly through its effect on the minimum productivity level ω_{ij}^{MNE}. The number of foreign affiliates N_{ij} is only indirectly affected by distance and trade costs through its effect on ω_{ij}^{MNE}. The direct effect of trade costs τ_{ij} on aggregate and average sales is negative. The effect through the minimum productivity level ω_{ij}^{MNE} on aggregate sales AS_{ij} and the number of affiliates N_{ij} is negative, its effect on average sales as_{ij} is positive. The effect of distance on multinational firms' activities depends therefore on the effect of distance on the minimum productivity level $\frac{\partial \omega_{ij}^{MNE}}{\partial D_{ij}}$.

We assume that the trade cost τ_{ij} are an increasing function in distance D_{ij}. We specify trade costs using the flexible functional form $\tau_{ij} = \lambda_1 D_{ij}^{\eta_1}$ with $\lambda_1 > 0$ and $\eta_1 > 0$. Moreover, we assume that fixed costs f_{ij}^{MNE} increase also in distance D_{ij} between the two countries. We use the same functional form with possibly different parameters to specify fixed costs $f_{ij}^{MNE} = \lambda_2 D_{ij}^{\eta_2}$ with $\lambda_2 > 0$ and $\eta_2 > 0$. Using these forms, we show in the Appendix A that distance affects the minimum productivity level ω_{ij}^{MNE} positively.

Since aggregate sales are negatively related to distance D_{ij} through the mini-

mum productivity ω_{ij}^{MNE} and distance affects the minimum productivity level positively, aggregate sales are affected negatively by distance through the indirect effect. In addition, the direct effect of distance is negative as well. Thus, aggregate sales of foreign affiliates AS_{ij} are negatively affected by distance.

$$\frac{\partial AS_{ij}}{\partial D_{ij}} = -(1-\epsilon)(\sigma-1)\eta_1 \frac{AS_{ij}}{D_{ij}} - (\kappa-\sigma+1)\frac{AS_{ij}}{\omega_{ij}^{MNE}}\frac{\partial \omega_{ij}^{MNE}}{\partial D_{ij}} < 0$$
$$if \quad \frac{\partial \omega_{ij}^{MNE}}{\partial D_{ij}} > 0$$

The number of affiliates N_{ij} is only affected by distance D_{ij} through the minimum productivity ω_{ij}^{MNE}. This effect is negative. Average sales, in contrast, are negatively affected through the direct effect of trade costs τ_{ij} and positively by the indirect effect through minimum productivity. Thus, for average sales, the intensive margin of affiliates activities in a foreign country, the effect of distance is ambiguous.

$$\frac{\partial as_{ij}}{\partial D_{ij}} = -(1-\epsilon)(\sigma-1)\eta_1 \frac{as_{ij}}{D_{ij}} + (\sigma-1)\frac{as_{ij}}{\omega_{ij}^{MNE}}\frac{\partial \omega_{ij}^{MNE}}{\partial D_{ij}}$$

3 Estimating the margins of activity

3.1 Gravity equations

In this subsection, we derive gravity equations for aggregate sales AS_{ij}, average sales as_{ij}, and the number of foreign affiliates N_{ij} because all three have been used in the literature. We discuss the differences among them, point to the

endogeneity problem that the sales equations suffer from, and propose how to cure this problem.

To derive a gravity equation we re-write equation (5) as

$$AS_{ij} = \underbrace{\left[\left(\frac{\kappa}{\kappa-\sigma+1}\right)\left(\omega_i^{Dom}\rho\right)^{\sigma-1} M_i\right]}_{s_i} \underbrace{\left(\frac{\omega_i^{Dom}}{\omega_{ij}^{MNE}}\right)^{\kappa-(\sigma-1)} \tau_{ij}^{(1-\epsilon)(1-\sigma)}}_{\Phi_{ij}} * \underbrace{\frac{(1-\mu)Y_j}{P_{Mj}^{1-\sigma}}}_{m_j}. \tag{7}$$

The first term gives the *supply capacity*, s_i, of country i. It represents the sales of the average firm from country i multiplied with the number of firms M_i active in country i. The second term, which we denote by Φ, is the weighted ratio of the smallest productivity level of a domestic firm and the threshold productivity level for production in country j. As shown in Appendix A, distance affects the minimum productivity of a multinational firm ω_{ij}^{MNE} positively. Thus, distance affects the second term Φ negatively. We assume that $\Phi = \lambda(D_{ij})^{-\eta}$. This form is very flexible and exhibits the negative impact of distance on aggregate sales. The third term gives the *market capacity* m_j of country j. Thus, aggregate affiliate sales of firms from country i in country j are then given by:

$$AS_{ij} = s_i \lambda (D_{ij})^{-\eta} m_j \tag{8}$$

Log-linearizing equation (8) and adding an error term yields the gravity equation.

$$ln(AS_{ij}) = \alpha + \zeta ln(s_i) - \beta ln(D_{ij}) + \xi ln(m_j) + u_{ij} \tag{9}$$

where $\alpha = ln(\lambda)$ and $\beta = \eta$. The gravity equation implies a constraint on the estimates of parameter ζ and ξ.

The number of firms from country i that are active in country j N_{ij} is given by $N_{ij} = \left(\frac{\omega_i^{Dom}}{\omega_{ij}^{MNE}}\right)^{\kappa} M_i$. In Appendix A, we show that this equation can be transformed in $N_{ij} = A_N s_i \Phi_{ij} m_j \frac{1-\tau_{ij}^{\epsilon(1-\sigma)}}{(1-\phi) f_{ij}^{MNE}}$ where $A_N = (1-\rho)^{\frac{\kappa-(\sigma-1)}{\kappa}}$.

The income coefficients of the resulting gravity equation equal those of the regression for aggregate sales, the constant α_N differs with $\alpha_N = ln(A_N)$. The distance coefficient β_N is the same as in the regression for aggregate sales if the entry costs $\frac{1-\tau_{ij}^{\epsilon(1-\sigma)}}{(1-\phi) f_{ij}^{MNE}}$ and the trade costs for the intermediate good $\tau_{ij}^{(1-\epsilon)(1-\sigma)}$ are controlled for. Otherwise the distance coefficient differs.

Average sales as_{ij} can also be expressed as function of supply and market capacity and distance $as_{ij} = \frac{\kappa}{\kappa-(\sigma-1)} \left(\omega_i^{Dom} \rho\right)^{\sigma-1} \left(\frac{\omega_i^{Dom}}{\omega_{ij}^{MNE}}\right)^{1-\sigma} \tau_{ij}^{(1-\epsilon)(1-\sigma)} \frac{(1-\mu) Y_j}{P_j^{1-\sigma}}$. Yet, supply capacity is measured incorrectly because the number of firms M_i is missing. Thus, the supply capacity must be weighted by $1/M_i$. The gravity equations for the number of firms and their average size are thus given by

$$ln(N_{ij}) = \alpha_N + \zeta ln(s_i) - \beta_N ln(D_{ij}) + \gamma ln(entry_{ij}) + \xi ln(m_j) + u_{ij} \quad (10a)$$

$$ln(as_{ij}) = \alpha_{as} + \zeta ln(s_i) - \beta_{as} ln(D_{ij}) - \nu ln(M_i) + \xi ln(m_j) + u_{ij} \quad (10b)$$

where $entry_{ij} = \left[\frac{\tau_{ij}^{(\epsilon-\sigma)(\epsilon-1)} \left(\tau_{ij}^{\epsilon(\sigma-1)} - 1\right)}{(1-\phi) f_{ij}^{MNE}}\right]$ and $\nu = 1$.

The analysis yields reduced-form gravity equations. In particular, distance

costs affect aggregate sales, AS_{ij}, and the number of foreign affiliates N_{ij} negatively whereas its effect on the average size of the foreign affiliate as_{ij} is ambiguous. In all three gravity equations, the income variables have positive effects on the measure of activities of multinational firms.

Below, we estimate gravity equations, that explain the effect of distance costs on (i) aggregate foreign affiliates sales, (ii) the number of foreign affiliates active abroad, and (iii) average affiliate sales. We decompose *market capacity* $m_j = (1-\mu)Y_j P_{Mj}^{\sigma-1}$ into its income and its weighted price level components, Y and P. The coefficient of the *market capacity* variable is one. The *supply capacity* is proportional to home country's income in the model. We proxy the supply capacity by home country's GDP. As argued above, the coefficient of home country's GDP is constrained to one.

3.2 System Estimation

The gravity equations derived above are often used in empirical research on multinational firms. However, they deliver biased estimates since the right hand side variables include the ratio of the two endogenous minimum productivity levels $\frac{\omega_i^{Dom}}{\omega_{ij}^{MNE}}$. In the context of international trade, Helpman *et al.* (2008) proposed a correction of this bias by a two step estimation. We tackle the problem differently by using the fact that the productivity ratio is the weighted ratio of the number of firms from i active in j on the total number of firms in i $\frac{N_{ij}}{M_i}$. Rewriting (7) in terms of the number of firms abroad N_{ij}

yields (11) after log-linearization and adding the error term.

$$
\begin{aligned}
AS_{ij} &= \left[\left(\frac{\kappa}{\kappa-\sigma+1}\right)\left(\omega_i^{Dom}\rho\right)^{\sigma-1} M_i\right]\left(\frac{N_{ij}}{M_i}\right)^{\frac{\kappa-(\sigma-1)}{\kappa}} \tau_{ij}^{(1-\epsilon)(1-\sigma)}\frac{(1-\mu)Y_j}{P_j^{1-\sigma}} \\
&= s_i (N_{ij})^{\frac{\kappa-(\sigma-1)}{\kappa}} \tau_{ij}^{(1-\epsilon)(1-\sigma)} m_j M_i^{\frac{(\sigma-1)-\kappa}{\kappa}}
\end{aligned}
$$

$$
\begin{aligned}
ln(AS_{ij}) = {} & \alpha_1 + \zeta\ ln(s_i) + \gamma\ ln\left(N_{ij}\right) - \beta_1\ ln(D_{ij}) + \xi\ ln(m_j) \\
& - \nu\ ln(M_i) + u_{ij}
\end{aligned} \tag{11}
$$

where $\gamma = \frac{\kappa-(\sigma-1)}{\kappa} > 0$ with $\gamma < 1$, and $\nu = \frac{(\sigma-1)-\kappa}{\kappa} < 0$. Data on the number of firms active in country j is at least in principal available and can be used to estimate an unbiased gravity equation. Estimating aggregate sales AS_{ij} jointly with the fraction of firms active in j eliminates the bias and gives unbiased estimates for aggregate sales. The system estimator is more efficient than the two-stage estimator proposed in Helpman *et al.* (2008). The drawback is that it requires information on the number of firm which is not always available.

The more complex effect of market entry and distance costs allows identification of the number of firm equation through the $entry_{ij}$ term in the system of two jointly estimated equations explaining the number of firms from i active in j and the aggregated and average sales respectively, as given in (12).

$$
\begin{aligned}
ln\left(N_{ij}\right) = {} & \alpha_{S_1} + \zeta_{S_1}\ ln(s_i) - \beta_{S_1}\ ln(D_{ij}) + \psi_{S_1}\ ln(entry_{ij}) + \xi\ ln(m_j) \\
& + u_{S_1 ij} \\
ln(AS_{ij}) = {} & \alpha_{S_2} + \zeta_{S_2}\ ln(s_i) - \beta_{S_2}\ ln(D_{ij}) + \gamma_{S_2}^{AS}\ ln(N_{ij}) \\
& + \xi\ ln(m_j) - \nu_{S_2}\ ln(M_i) + u_{S_2 ij}
\end{aligned} \tag{12a}
$$

$$
\begin{aligned}
ln\,(N_{ij}) &= \alpha_{S_1} + \zeta_{S_1}\ ln(s_i) - \beta_{S_1}\ ln(D_{ij}) + \psi_{S_1}\ ln(entry_{ij}) + \xi\ ln(m_j) \\
&\quad + u_{S_1 ij} \\
ln(as_{ij}) &= \alpha_{S_2} + \zeta_{S_2}\ ln(s_i) - \beta_{S_2}\ ln(D_{ij}) + \gamma^{as}_{S_2}\ ln(N_{ij}) \\
&\quad + \xi\ ln(m_j) - \nu_{S_2}\ ln(M_i) + u_{S_2 ij}
\end{aligned}
\tag{12b}
$$

where $\gamma^{as}_{S_2} = \frac{1-\sigma}{\kappa} < 0$ and $\nu^{as} = \frac{(\sigma-1)-\kappa}{\kappa} = \nu^{AS} = \nu_{S_2}$. Thus, the two sales regressions differ only in the effect γ of the endogenous number of firms from country i in j. Notice that the two effects are not independent since $\gamma^{AS} - \gamma^{as} = 1$. We estimate both systems (12) for two samples of foreign affiliate sales of multinational firms.

3.3 Data

Data on bilateral activities of multinational firms are rare. We use the comprehensive German *MIDI* database (Microdatabase Direct Investment) of the *Deutsche Bundesbank* that contains information at the micro level.[2] This database comprises firm level information on foreign affiliate sales of German multinational firms. The MiDi database covers a very large share of German multinational firms, because the reporting limits are fairly low.[3] We aggregate foreign affiliates' sales and the number of foreign affiliates from each of the 16 German states in 116 countries for each year between 1989 and 2007. We restrict however our analysis to the period from 1992 to 2004. Starting with micro-level data has the advantage that we can tailor the aggregates in a

[2] For a description of the database see Lipponer 2009.

[3] Up to 2002, the activities of foreign affiliates with annual sales of more than 1 million DM (500.000 Euro in 2001) must compulsory be reported to the German Bundesbank. In 2002, the reporting thresholds were raised to annual sales of 3 million Euro.

ways commanded by the theory. Since the theory explains the decision of firms in the manufacturing sector we aggregate only foreign sales of manufacturing affiliates from multinational firms classified in manufacturing. We deflate the sales data using consumer price indexes of the host country and convert them into US dollar using the year end exchange rate of 2000.

We aggregate sales in particular host country over all firms from each of the 16 German states. Our unit of observation is a German-state-host-country combination. That creates 16 home countries which enables us to analyze both home and host country effects. We are not aware of any bias we introduce by treating the 16 states as "independent" countries. The problem we see is the low variation in some of the explanatory variables. Usually bilateral variables such a dummy variable indicating whether both countries are members of the European union or a dummy variable indication the existence of a preferential trade agreement vary at the level of the host country only because there is no variation in these variables among the 16 German states.

We are nevertheless cautious with our home countries. We therefore compared the results with results from the OECD *Measuring Globalization* database which contains information on sales of foreign affiliates and their number for 21 OECD countries and about 50 partner countries from 1983 to 2001. Unfortunately, this sample is very unbalanced. The results are very close to the results from the *MIDI* data. We do not report them here because not all of the analysis could be done with this sample. [4]

The low reporting limits of the *MIDI* database allow us to treat all non-existent state-host-country combinations as zeros in the sense that there exists no foreign affiliate of a firm from the particular state in the particular host country. A missing indicates zero activity. Thus, we add this "information" to the data.

Regarding the explanatory variables, we retrieve the GDP data in constant US dollar from the WDI database of the World Bank. The price level is taken from the OECD *Comparative Price Level* database. We convert the bilateral price level indexes into an index of countries' price level relative to the OECD average. Using OECD price level data further strongly reduces the sample, particularly the German sample, by restricting it to 21 OECD partner countries. [5] That excludes developing countries from the analysis. Distance is taken from the CEPII distance database [6] which contains the distance between the largest city of any two countries. We use the great arc distance between the largest cities of German states and their partner. The institutional variables characterizing political and economic freedom in the host country are taken from ?? The indexes ranch from zero for lowest level of freedom to 100?? indicating full freedom. Table B.1 in Appendix B provides the summary statistics of our data.

Before we interpret the results, we briefly mention two econometric issues

[4] Some results are available in an earlier version of the paper (Kleinert and Toubal (2006)), other results are available upon request.
[5] Australia, Canada, Czech Republic, Denmark, Finland, France, Hungary, Ireland, Italy, Japan, Luxembourg, Mexico, Netherlands, Norway, Poland, Portugal, Spain, Sweden, Turkey, United Kingdom, United States
[6] www.cepii.fr

concerning the specified gravity models. First, since the number of affiliates is a count variable, we use Poisson regression techniques for the equation explaining the number of foreign affiliates throughout the analysis. Silva and Tenreyro (2006) have argued that using Poisson regressions is also preferable when explaining aggregated and average sales for two reasons: (i) the inclusion of the zeros in the analysis, and (ii) a correct treatment of the error factors in the empirical analysis. We therefore compare our OLS gravity equations to Poisson regressions as proposed by Silva and Tenreyro (2006). Second, we use the Huber-White method to correct for serially correlated country pairs (Wooldridge 2001).

4 Results

We present four kind of results. First, we estimate five gravity equations as presented and discussed above. Second, we account for the critique of Silva and Tenreyro (2006) and estimate the gravity equations using zero inflated Poisson regressions. Third, we analyze the distance variable more deeply addressing Anderson and van Wincoop's (2003) insight regarding the control for multilateral resistance and non-linear effect of distance found in the theory. Forth, we present from system estimations as introduced and discussed above and compare them to two stage estimation using the methodology proposed by by Helpman *et al.* (2008).

4.1 Gravity equations

The effect of the gravity variables on foreign affiliates sales, average sales of a foreign affiliate, and the number of foreign affiliates is shown in Table 1. Specification (S1) is the gravity equation (9) explaining foreign affiliate sales. Thereby, (S1) explicitly accounts for the parameter restriction on the coefficients of the GDP of the home country ζ and the GDP of the host country ξ discussed above. Both coefficients are constrained to one. Specification (S2) is the gravity equation (9) explaining foreign affiliate sale but estimates ζ and ξ. Specification (S3) is the gravity equation explaining average affiliate sales and (S4) and (S5) are gravity equations explaining the number of foreign affiliates. Thereby, (S5) includes the entry variables.

The results in (S2) confirm earlier results from gravity equations. While home and host country GDP affect foreign affiliate sales positively, distance between the two countries affects sales negatively. All three coefficients are significant at one percent. The coefficient on home country GDP is slightly larger than one. The restriction on both coefficients in (S1) is therefore rejected at the one percent level of significance. Although the gravity equation suggests that the coefficients on both GDP variables are one, the restriction is not consistent with the data. The results regarding the Border dummy variable indicating a neighboring country are also in line with earlier finding (Barba-Navaretti and Venables 2004). Activities in neighboring countries are higher than predicted by their size and distance alone. In contrast, we find no significant additional

Table 1
Gravity Equation explaining Total foreign sales, Average foreign sales and the Number of Affiliates: OLS, poisson regressions

	Constrained Model	Unconstrained Model			
	(S1)	(S2)	(S3)	(S4)	(S5)
	AS_{ij}	AS_{ij}	as_{ij}	$n_{ij}^{\dagger}$	$n_{ij}^{\dagger}$
GDP_{home}	1.00	1.201***	0.260***	1.054***	1.058***
		(0.091)	(0.065)	(0.059)	(0.057)
GDP_{host}	1.00	0.929***	0.335***	0.653***	0.641***
		(0.079)	(0.055)	(0.044)	(0.040)
Distance	-0.603***	-0.569***	-0.096	-0.499***	-0.482***
	(0.038)	(0.102)	(0.073)	(0.060)	(0.058)
Border	0.561***	0.522	-0.032	0.444*	0.471**
	(0.148)	(0.372)	(0.227)	(0.244)	(0.215)
Intra-EU	-0.056	0.001	-0.030	-0.120	-0.111
	(0.099)	(0.215)	(0.138)	(0.144)	(0.136)
East	-1.789***	-1.518***	-0.532**	-1.217***	-1.223***
	(0.128)	(0.289)	(0.232)	(0.159)	(0.157)
Trade openness	-0.013**	-0.013*	0.002	-0.012***	-0.011***
	(0.006)	(0.007)	(0.006)	(0.004)	(0.004)
Price Level	-0.013***	-0.014	0.001	-0.014***	-0.017***
	(0.001)	(0.004)	(0.002)	(0.002)	(0.002)
Business					0.005
freedom					(0.004)
Investment					-0.002
freedom					(0.004)
Property rights					-0.005
					(0.005)
loc2					-0.416***
					(0.161)
loc3					-0.155
					(0.126)
loc4					-0.121
					(0.200)
Constant	-5.596***	-7.542***	3.532***	-13.18***	-13.02***
	(0.426)	(1.480)	(1.037)	(0.977)	(1.014)
Observations	2137	2137	2137	2137	2137
R^2		0.50	0.19		
LR-statistics	37.32***				
p-value	0.000				

The freely estimated regressions include time dummies. Robust standard errors in parentheses. Standard errors have been adjusted for clustering around country pairs.

$n_{ij}^{\dagger}$: Poisson regression.

*** denotes statistical significance at one percent level of significance.

** denotes statistical significance at five percent level of significance.

* denotes statistical significance at ten percent level of significance.

effect for Intra-EU activities. The *East* dummy is negative and significant. Firms in East Germany have started to internationalize their activities only in 1991. The coefficient of trade openness is negative indicating the alternative to production abroad. Barriers to trade increase the trade costs with favors production abroad over exports. The price index has a significantly negative effect on foreign affiliate sales.

The gravity equations related to the number of foreign affiliates (S4) and (S5) show basically the same effects as the one presented in specification (S2). The effect of distance on foreign affiliates' sales is essentially as high as (in absolute terms) than on the number of foreign affiliates. The effect of distance on average sales of foreign affiliates is negative, small and insignificant. Thus, distance affects total affiliate sales negatively through reducing the number of affiliates in a foreign country but not by changing the average size of the foreign affiliate. Finally, the entry variables in (S5) are mainly insignificant. The other explanatory variables are unchanged.

4.2 Zero inflated Poisson regressions

In Table 2 we report marginal effects from the zero inflated Poisson regressions. There is little change to the OLS results concerning the GDP coefficients. The distance coefficient is smaller (in absolute terms). The dummy variables (*border, intra EU*) are mostly insignificant. Trade openness also loses significance compared to the OLS regressions. The use zero inflated Poisson because the

Vuong test rejects the validity of Poisson regressions. The zero inflated Poisson fits additionally a probit regression to explain the outcome no activity. The results of the probit is given in the right column of each specification. The coefficients have usually the opposite sign from the coefficients explaining the level of activity and are significant. The number of observations almost doubles while the coefficients are widely unchanged.

4.3 A closer look at distance

In Table 3 we present results that center around a deeper analysis of the distance effect. In columns (1) through (3), we present the two bilateral variables in our analysis *distance* and the state *border* dummy in regressions including fixed effects for the home and for the host country as proposed by Anderson and Wincoop (2003). The distance variable is significantly negative in all three regressions. The *border* dummy in contrast is not significant anymore. The distance coefficients are larger (in absolute terms) than in the OLS and the ZIP regressions.

Columns (4) through (6) address the non-linearity that came out of the empirical analysis. There we find, that distance exerts a negative effect on the activities of multinational firms if the distance between the home and the host country exceeds a threshold level. If home and host country are closer, distance has a positive effect on multinational firms activities. To detect such threshold, we calculate quantiles with respect to distance. Quantiles are separated

Table 2
Gravity Equation explaining Total foreign sales, Average foreign sales and the Number of Affiliates: zero inflated poisson regressions

	AS_{ij}		as_{ij}		n_{ij}	
	(ZIP1)		(ZIP2)		(ZIP3)	
	Poisson	Probit	Poisson	Probit	Poisson	Probit
GDP_{home}	1.075***	-1.69***	0.164*	-1.692***	1.103***	-1.329***
	(0.139)	(0.152)	(0.100)	(0.153)	(0.062)	(0.209)
GDP_{host}	0.955***	-1.17***	0.225**	-1.174***	0.657***	-1.119***
	(0.131)	(0.112)	(0.104)	(0.112)	(0.041)	(0.203)
Distance	-0.412**	0.883***	0.037	0.882***	-0.494***	0.801***
	(0.183)	(0.145)	(0.132)	(0.145)	(0.061)	(0.228)
Border	0.482	-2.38***	-0.313	-2.383***	0.614**	-3.164**
	(0.409)	(0.595)	(0.365)	(0.595)	(0.217)	(1.399)
Intra-EU	-0.591	-0.77***	-0.114	-0.770***	-0.119	-1.319***
	(0.381)	(0.285)	(0.235)	(0.285)	(0.141)	(0.443)
East	-1.613**	2.99***	-0.340	3.524***	-2.146***	1.128
	(0.693)	(0.271)	(0.634)	(0.271)	(0.271)	(0.750)
Trade	-0.021	-0.029**	-0.002	0.029**	-0.011***	0.010
openness	(0.018)	(0.013)	(0.016)	(0.013)	(0.004)	(0.018)
Price Level	-0.008	0.036***	0.008	0.036***	-0.014***	0.042***
	(0.005)	(0.005)	(0.006)	(0.005)	(0.002)	(0.009)
Business					0.005	0.032**
freedom					(0.005)	(0.013)
Investment					-0.003	0.005
freedom					(0.004)	(0.018)
Property					-0.005	-0.018
rights					(0.005)	(0.019)
loc2					-0.422**	0.560
					(0.169)	(0.864)
loc3					-0.156	1.080
					(0.129)	(0.708)
loc4					-0.120	0.659
					(0.203)	(0.666)
Constant	-6.40**	21.16***	5.534***	21.17***	-13.65***	16.74***
	(2.920)	(2.457)	(2.052)	(2.457)	(1.074)	(3.293)
Obs., zeros	2137	1911	2137	1911	2137	1911

All regressions include time dummies. Robust standard errors in parentheses.
Standard errors have been adjusted for clustering around country pairs.
*** denotes statistical significance at one percent level of significance.
** denotes statistical significance at five percent level of significance.
* denotes statistical significance at ten percent level of significance.

Table 3
Equation explaining Total foreign sales, Average foreign sales and the Number of Affiliates: ZIP fixed effects and non-linear effects

	Fixed Effects Regression			Non-linear Distance effects		
	AS_{ij}	as_{ij}	n_{ij}	AS_{ij}	as_{ij}	n_{ij}
Distance	-0.999**	-0.430*	-0.746**	-0.800	0.503	-0.510**
	(0.314)	(0.259)	(0.360)	(0.575)	(0.424)	(0.215)
Border	0.361	-0.310	0.364	0.551	-0.011	0.510**
	(0.369)	(0.407)	(0.463)	(0.463)	(0.380)	(0.219)
Quantile 2				0.048	0.011	0.004
				(0.060)	(0.046)	(0.024)
Quantile 3				0.117	-0.074	0.031
				(0.100)	(0.065)	(0.043)
Quantile 4				0.116	-0.118	-0.017
				(0.149)	(0.110)	(0.063)
Intra-EU				-0.682*	-0.051	-0.234*
				(0.381)	(0.269)	(0.135)
Trade				-0.029*	0.002	-0.013***
openness				(0.017)	(0.016)	(0.004)
Observations	4048	4048	4048	4048	4048	4048

All regressions include time dummies. Fixed effects regression include home and host country fixed effects. Robust standard errors in parentheses. Standard errors have been adjusted for clustering around country pairs. For non-linear effects regression only coefficients reported that differ from those in Table 2.
*** denotes statistical significance at one percent level of significance.
** denotes statistical significance at five percent level of significance.
* denotes statistical significance at ten percent level of significance.

at 900km, 1570km, and 9170km. We include dummy variables for the second, third, and forth quantile interacted with the distance in the regressions. We do not detect a non-linearity and conclude that the introduction of possibly varying distance effects does not alter the results. Two other explanatory variables *Intra Eu* and *Trade openness* gain in significance while the *distance* variable loses.

4.4 System estimations

In Table 4, we present the results from the system estimation as given in (12). Identification through fixed costs and trade cost variables works well. The two equation system requires to control for the weighted number of firms in country i in the sales equations. We use the number of firms in the particular state taken from *German Statistical Office*. The endogenous number of firms N_{ij} is significant and has the expected sign in both systems. Moreover, the coefficient of the number of firms in the two sales regressions meet the restriction $\gamma^{AS} - \gamma^{as} = 1$. In line with the theory, the sales regressions in the system estimation differ only in the coefficients of the endogenous variable. The system estimation confirms that distance affects affiliate sales only through the number of firms. Finally, the price level has the correct sign and is statistically significant in the system estimation.

In columns (4) through (6) we present results for zero inflated poisson regression that include the correction for the correlation of the explanatory variables with the residual as proposed by Helpman *et al.* (2008). We include the residual from the ZIP equation explaining the number of foreign affiliates in the regressions explaining aggregated and average sales to control for the correlation of the explanatory variables and the residual.

Consistent with the theory, the number of home country's firms has the same negative effect on aggregate and average sales. The coefficient of the endogenous number of affiliates have the correct sign, are significant and meet the

Table 4
Equation explaining Total foreign sales, Average foreign sales and the Number of Affiliates: System approaches

	LS System			ZIP with correction		
	AS_{ij}	$n_{ij}^{\dagger}$	as_{ij}	AS_{ij}	$n_{ij}^{\dagger}$	as_{ij}
n_{ij}	0.530*		-0.470			
	(0.328)		(0.328)			
GDP_{home}	0.9075***	0.941***	0.907***	1.175***	1.103***	0.246**
	(0.261)	(0.019)	(0.261)	(0.139)	(0.062)	(0.098)
GDP_{host}	0.620***	0.610***	0.620***	0.922***	0.657***	0.261***
	(0.200)	(0.016)	(0.200)	(0.104)	(0.041)	(0.090)
Distance	-0.320**	-0.470***	-0.320**	-0.362***	-0.494***	0.020
	(0.158)	(0.023)	(0.158)	(0.168)	(0.061)	(0.132)
Border	0.237	0.515***	0.237	0.646*	0.482**	-0.305
	(0.212)	(0.074)	(0.212)	(0.357)	(0.217)	(0.359)
Intra-EU	0.002	0.053	0.002	-0.408	-0.119	0.091
	(0.083)	(0.050)	(0.083)	(0.289)	(0.141)	(0.234)
East	-0.836**	-0.977***	-0.836**	-2.782***	-2.146***	-0.032
	(0.383)	(0.068)	(0.383)	(0.675)	(0.271)	(0.508)
Trade	-0.007	-0.013***	-0.007	-0.024**	-0.011***	-0.003
openness	(0.007)	(0.003)	(0.007)	(0.016)	(0.004)	(0.016)
Price Level	-0.006	-0.017***	0.006	-0.009***	-0.014***	0.007
	(0.005)	(0.001)	(0.005)	(0.006)	(0.002)	(0.006)
No. of home	-0.160**		-0.160**			0.042***
firms	(0.075)		(0.075)			
Residual				1.066***		0.297**
				(0.006)		(0.139)
Business		0.003**			0.005	
freedom		(0.002)			(0.005)	
Investment		-0.002			-0.003	
freedom		(0.002)			(0.004)	
Property		-0.001			-0.005	
rights		(0.018)			(0.005)	
loc2		-0.306			-0.422**	
		(0.227)			(0.169)	
loc3		-0.302***			-0.156	
		(0.082)			(0.129)	
loc4		-0.518***			-0.120	
		(0.118)			(0.203)	
Observations	2001	2001	2001	4048	4048	4048
R^2	0.65	0.68	0.12			

All regressions include time dummies. Robust standard errors in parentheses.
Standard errors have been adjusted for clustering around country pairs.
$n_{ij}^{\dagger}$: Poisson regression.
*** denotes statistical significance at one percent level of significance.
** denotes statistical significance at five percent level of significance.
* denotes statistical significance at ten percent level of significance.

restriction. The other coefficients in the sales equation are equal. The direct distance effect on sales is negative but not significant. Identification works well. The state-border dummy variable has a significant and positive effect on the number of affiliates in a particular partner country. The other variables identifying $entry_{ij}$ *Intra EU*, *Protection*, and *FDI Restrictions*, in contrast, are not significant. The East German dummy variable is significantly negative at one percent level in all three regressions. East German firms have less and smaller foreign affiliates than firms from West Germany. That stems from the late start of their internationalization process.

In sum, the results from the German sample confirm that aggregated sales of foreign affiliates fall in distance. That this fall is mainly due to the smaller number of affiliates that are active in more distant countries.

5 Conclusion

In this paper, we present a model of multinational firms that predicts that aggregated affiliates' sales fall in distance. We derive a gravity equation explaining aggregate foreign affiliate sales from the model. We are particularly interested in the distance effect on foreign affiliate sales. In the model, the distance effect on foreign affiliate sales is driven by the extensive margin: distance affects the number of affiliates negatively while it has an ambiguous effect on the average affiliate sales, i.e. the intensive margin. The theory also reveals an endogeneity problem in the gravity equations explaining aggregate

and average sales of foreign affiliates. Moreover, the theory proposes how to cure this endogeneity problem. By jointly estimating the number of foreign affiliates and their sales unbiased estimates can be obtained.

For the empirical assessment of the relative importance of the extensive and the intensive margin of activities, we use German data set from a compulsory survey of multinational firms that we aggregate to the level of German states. It fits the theory, because we restrict activities to those by manufacturing affiliates of German multinational firms from manufacturing. The low reporting limits of the survey allows to treat non-existing state-host country combinations as non-activity combinations and use the information of zero activity.

Our results demonstrate the importance of the extensive margin of activities. The number of foreign affiliates of firms from a particular home country in a particular host country decreases in the distance between the two countries. Additionally, neighboring countries receive an over-proportional share of foreign affiliates. The fall in the number of affiliates in more distant foreign countries explains a very large fraction of the fall in total affiliate sales in these countries. In contrast, distance does not significantly affect the average size of foreign affiliates.

References

Anderson, James A., van Wincoop, Eric (2003). Gravity with Gravitas: A Solution to the Border Puzzle. *American Economic Review* 93, 170-192.

Barba Navaretti, Giorgio and Venables, Anthony, (2004). Multinational Firms in the World Economy. *Princeton University Press*.

Brainard, Lael S. (1997). An empirical assessment of the proximity-concentration trade-off between multinational sales and trade. *American Economic Review* 87, 520-544.

Buch, Claudia M., Kleinert, Joern, Lipponer, Alexander, and Toubal, Farid, (2005). Determinants and Effects of Foreign Direct Investment: Evidence from German Firm-Level Data. *Economic Policy* 41 , 52–110.

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.

Chaney, Thomas, (2008). Distorted Gravity: the intensive and extensive margins of international trade. *American Economic Review* 98 1707–1790.

De Sousa, José and Lochard, Julie, (2006). Does the Single Currency Affect Foreign Direct Investment? A Gravity-Like Approach. Mimeo, University of Paris 1, Panthéon-Sorbonne

Eaton, Jonathan, Kortum, Samuel, and Francis Kramarz (2004). Dissecting Trade: Firms, Industries and Export Destinations. *American Economic Review* 94(2): 150–154.

Eaton, Jonathan and Tamura, Akiko, (1994). Bilateralism and Regionalism in Japanese and U.S. Trade and Direct Foreign Investment Patterns. *Journal of the Japanese and International Economies* 8, 478-510.

Eichengreen, Barry and Tong, Hui, (2005). Is Chinas FDI Coming at the Expense of Other Countries? *Journal of the Japanese and International Economies* 21(2), 153–172.

Head, Keith and Ries, John, (2008). FDI as an Outcome of the Market for Corporate Control: Theory and Evidence. *Journal of International Economics* 74 (1), 2–20.

Helpman, Elhanan, Melitz, Marc J., and Rubinstein, Yona (2008). Estimating Trade Flows: Trading Partners and Trading Volumes *Quarterly Journal of Economics* CXXIII (2), 441–487.

Lipponer, Alexander (2009). Microdatabase Direct Investment - MiDi: A Brief Guide. http://www.bundesbank.de/download/vfz/fdi.

Marin, Dalia and Schnitzer, Monika, (2004). Global versus Local: The financing of foreign Direct Investment, University of Munich, mimeo

Neary, Peter, (2009). Trade Costs and Foreign Direct Investment. *International Review of Economics and Finance*.

Rauch, James, (1999). Networks versus Markets in International Trade. *Journal of International Economics* 48(1) , 7–35.

Silva, J. M. C. Santos, Tenreyro, Silvana (2006). The log of Gravity. *Review of Economics and Statistics* 88, 641–658.

Wei, Shan-Jin (2000). How Taxing is Corruption on International Investors.

Review of Economics and Statistics 1, 1–11.

Wooldridge, Jeffrey, (2001). The Econometric Analysis of Cross Section and Panel Data. Cambridge: *The MIT Press*.

Appendices

A Derivations

A.1 Derivation of Equation (5)

We use the pareto distribution of productivity with the shape parameter κ and the scale parameter ω_{ij}^{Dom} to aggregate over affiliate sales of firm k.

$$\begin{aligned}
AS_{ij} &= \int_{\omega_{ij}^{MNE}}^{\infty} \left(\frac{\omega_k \rho}{\tau_{ij}^{1-\epsilon}}\right)^{\sigma-1} g(\omega) M_i \frac{(1-\mu)Y_j}{P_j^{1-\sigma}} d\omega_k \\
&= \frac{\rho^{\sigma-1}(1-\mu)Y_j}{P_j^{1-\sigma}} \tau_{ij}^{(1-\epsilon)(1-\sigma)} M_i \int_{\omega_{ij}^{MNE}}^{\infty} \omega_k^{\sigma-1} \frac{\kappa}{\omega_k} \left(\frac{\omega_i^{Dom}}{\omega_k}\right)^{\kappa} d\omega_k \\
&= \left[0 - \left(\frac{\kappa}{\sigma-\kappa-1}\right)\left(\omega_{ij}^{MNE}\right)^{\sigma-1-\kappa}\right]\left(\omega_i^{Dom}\right)^{\kappa} M_i \left(\frac{\rho}{\tau_{ij}^{1-\epsilon}}\right)^{\sigma-1} \frac{(1-\mu)Y_j}{P_j^{1-\sigma}} \\
&= \left(\frac{\omega_i^{Dom}}{\omega_{ij}^{MNE}}\right)^{\kappa} M_i \left(\frac{\kappa}{\kappa-\sigma+1}\right) \left(\frac{\omega_{ij}^{MNE}\rho}{\tau_{ij}^{1-\epsilon}}\right)^{\sigma-1} \frac{(1-\mu)Y_j}{P_j^{1-\sigma}}
\end{aligned}$$

A.2 Distance Costs' Effect on the Minimum Productivity Threshold

We use equation (4c) to derive the effect of distance on the critical level of productivity. We assume that fixed costs are a function of distance in a similar way as variable distance costs. Hence, $(1-\phi)f_{ij}^{MNE} = \lambda_2 D_{ij}^{\eta_2}$ and $\tau_{ij} = \lambda_1 D_{ij}^{\eta_1}$. Substituting this functional forms into equation (4c) gives:

$$\begin{aligned}\omega_{ij}^{MNE} &= \left[\left(\lambda_1 D_{ij}^{\eta_1}\right)^{(1-\epsilon)(1-\sigma)} - \left(\lambda_1 D_{ij}^{\eta_1}\right)^{(1-\sigma)}\right]^{\frac{1}{1-\sigma}} \Omega^{\frac{1}{1-\sigma}} \left(\lambda_2 D_{ij}^{\eta_2}\right)^{\frac{1}{\sigma-1}} \\ &= \left[\left(\lambda_1 D_{ij}^{\eta_1}\right)^{(1-\sigma)} \left(\left(\lambda_1 D_{ij}^{\eta_1}\right)^{\epsilon(\sigma-1)} - 1\right)\right]^{\frac{1}{1-\sigma}} \Omega^{\frac{1}{1-\sigma}} \left(\lambda_2 D_{ij}^{\eta_2}\right)^{\frac{1}{\sigma-1}} \\ &= \lambda_1 D_{ij}^{\eta_1} \left[\left(\lambda_1 D_{ij}^{\eta_1}\right)^{\epsilon(\sigma-1)} - 1\right]^{\frac{1}{1-\sigma}} \Omega^{\frac{1}{1-\sigma}} \left(\lambda_2 D_{ij}^{\eta_2}\right)^{\frac{1}{\sigma-1}}\end{aligned}$$

where $\Omega = \frac{(1-\mu)Y_j(1-\rho)}{P_j^{1-\sigma}\rho^{1-\sigma}}$.

We derive the effect of distance on the critical level of productivity as

$$\begin{aligned}&\frac{\partial \omega_{ij}^{MNE}}{\partial D_{ij}} = \lambda_1 D_{ij}^{\eta_1} \left[\left(\lambda_1 D_{ij}^{\eta_1}\right)^{\epsilon(\sigma-1)} - 1\right]^{\frac{1}{1-\sigma}} \Omega^{\frac{1}{1-\sigma}} \left(\lambda_2 D_{ij}^{\eta_2}\right)^{\frac{1}{\sigma-1}} \\ &* \left[\frac{\eta_1}{D_{ij}} - \frac{1}{\sigma-1}\left[\left(\lambda_1 D_{ij}^{\eta_1}\right)^{\epsilon(\sigma-1)} - 1\right]^{-1} \epsilon(\sigma-1)\eta_1 \lambda_1^{\epsilon(\sigma-1)} D_{ij}^{\eta_1\epsilon(\sigma-1)-1} + \frac{1}{\sigma-1}\frac{\eta_2}{D_{ij}}\right]\end{aligned}$$

This simplifies to

$$\begin{aligned}\frac{\partial \omega_{ij}^{MNE}}{\partial D_{ij}} &= \omega_{ij}^{MNE} \left[\frac{\eta_1 + \frac{1}{\sigma-1}\eta_2}{D_{ij}} - \frac{\epsilon\eta_1 \lambda_1^{\epsilon(\sigma-1)} D_{ij}^{\eta_1\epsilon(\sigma-1)}}{D_{ij}\left[\left(\lambda_1 D_{ij}^{\eta_1}\right)^{\epsilon(\sigma-1)} - 1\right]}\right] \\ &= \omega_{ij}^{MNE} \left[\frac{\left(\eta_1 + \frac{1}{\sigma-1}\eta_2\right)\left(\lambda_1 D_{ij}^{\eta_1}\right)^{\epsilon(\sigma-1)} - \epsilon\eta_1\left(\lambda_1 D_{ij}^{\eta_1}\right)^{\epsilon(\sigma-1)} - \left(\eta_1 + \frac{1}{\sigma-1}\eta_2\right)}{D_{ij}\left[\left(\lambda_1 D_{ij}^{\eta_1}\right)^{\epsilon(\sigma-1)} - 1\right]}\right] \\ &= \omega_{ij}^{MNE} \left[\frac{\left[(1-\epsilon)\eta_1 + \frac{1}{\sigma-1}\eta_2\right]\left(\lambda_1 D_{ij}^{\eta_1}\right)^{\epsilon(\sigma-1)} - \left(\eta_1 + \frac{1}{\sigma-1}\eta_2\right)}{D_{ij}\left[\left(\lambda_1 D_{ij}^{\eta_1}\right)^{\epsilon(\sigma-1)} - 1\right]}\right]\end{aligned}$$

$$\frac{\partial \omega_{ij}^{MNE}}{\partial D_{ij}} > 0 \quad iff \quad D_{ij} > \left[\frac{\eta_1 + \frac{1}{\sigma-1}\eta_2}{(1-\epsilon)\eta_1 + \frac{1}{\sigma-1}\eta_2}\right]^{\frac{1}{\epsilon(\sigma-1)\eta_1}} \left[\frac{1}{\lambda_1}\right]^{1/\eta_1)}$$

A.3 The number of firms from i *active in* j

$$
\begin{aligned}
N_{ij} &= \left(\frac{\omega_i^{Dom}}{\omega_{ij}^{MNE}}\right)^{\kappa} M_i \\
&= \left(\frac{\omega_i^{Dom}}{\omega_{ij}^{MNE}}\right)^{\kappa-(\sigma-1)} \left(\frac{\omega_i^{Dom}}{\omega_{ij}^{MNE}}\right)^{\sigma-1} M_i \\
&= \left(\omega_i^{Dom}\right)^{\sigma-1} M_i \left(\frac{\omega_i^{Dom}}{\omega_{ij}^{MNE}}\right)^{\kappa-(\sigma-1)} \left(\omega_{ij}^{MNE}\right)^{1-\sigma} \\
&= \left(\omega_i^{Dom}\rho\right)^{\sigma-1} M_i \left(\frac{\omega_i^{Dom}}{\omega_{ij}^{MNE}}\right)^{\kappa-(\sigma-1)} \frac{\tau_{ij}^{(1-\sigma)^2}\left(\tau_{ij}^{\epsilon(\sigma-1)}-1\right)}{(1-\phi)f_{ij}^{MNE}} \frac{(1-\mu)Y_j(1-\rho)}{P_j^{1-\sigma}} \\
&= (1-\rho)\frac{\kappa-(\sigma-1)}{\kappa} \underbrace{\frac{\kappa}{\kappa-(\sigma-1)}\left(\omega_i^{Dom}\rho\right)^{\sigma-1} M_i}_{s_i} \underbrace{\left(\frac{\omega_i^{Dom}}{\omega_{ij}^{MNE}}\right)^{\kappa-(\sigma-1)} \tau_{ij}^{(1-\epsilon)(1-\sigma)}}_{\Phi_{ij}} \\
&\quad * \underbrace{\frac{1-\tau_{ij}^{\epsilon(1-\sigma)}}{(1-\phi)f_{ij}^{MNE}}}_{entry_{ij}} \underbrace{\frac{(1-\mu)Y_j}{P_j^{1-\sigma}}}_{m_j} \\
&= A_N s_i \Phi_{ij} m_j \frac{1-\tau_{ij}^{\epsilon(1-\sigma)}}{(1-\phi)f_{ij}^{MNE}}
\end{aligned}
$$

where $A_N = (1-\rho)\frac{\kappa-(\sigma-1)}{\kappa}$.

B Appendix B

Table B.1
Summary Statistics

Variable	Obs.	Mean	Std. Dev.
	OECD Sample		
ln Foreign Affiliates Sales	713	7.625	1.903
ln Average Sales	713	3.868	0.982
Number of Affiliates	713	3.757	1.390
ln GDP Home	713	27.768	1.271
ln GDP Host	713	27.000	1.228
ln Distance	713	7.861	1.199
Price Index	713	92.888	30.801
Border	713	0.123	0.329
Former Colony	713	0.052	0.222
	German Sample		
ln Foreign Affiliates Sales	6782	10.998	2.206
ln Average Sales	6782	9.727	1.371
Number of Affiliates	6782	8.832	16.405
ln GDP Home	6782	11.887	0.846
ln GDP Host	6567	11.832	1.649
ln Distance	6782	8.072	1.140
Price Index	3168	87.719	26.638