



"la Caixa"

Working Paper Series No. 05/2006

Research Department

Av. Diagonal, 629 T.I P.6

08028 Barcelona

research@lacaixa.es

The Determinants of Cross-Border Investment: A Value-Chain Analysis

Claudia Canals

Marta Noguer

December 2006

The views expressed in this working paper are those of the authors only and do not necessarily represent those of "la Caixa"

The Determinants of Cross-Border Investment: A Value-Chain Analysis*

Claudia Canals

Marta Noguer

"la Caixa"

"la Caixa"

First Draft: November 20, 2006

This Draft: January 23, 2007

Abstract

This paper contributes to the literature on the determinants of FDI. We use a new data set covering greenfield and expansion projects to examine which factors influence the decision to invest abroad. Our empirical framework is an augmented gravity model that incorporates elements of factor proportions theory. At the aggregate level, we find that distance discourages FDI, size and sharing a language encourages it, and that FDI targets relatively capital scarce countries. When we classify investment projects according to their stage in the chain of production, we observe a lot of variation across stages. Nevertheless, economic size, distance, and capital abundance are still determining factors for most value-chain stages and preserve the sign of their effects. Moreover, even though the results confirm FDI targetting capital scarce countries, we find evidence of a minimum requirement on the host country's capital endowment in all the stages of production except extraction. Finally, ease of doing business is also important, especially so for the location of regional headquarters.

Keywords: FDI, Gravity, Factor Proportions, Value Chain

JEL Classification: F21, F23, D23, F11, F12, C23

©Caja de Ahorro y Pensiones de Barcelona, "la Caixa"

©Claudia Canals, Marta Noguer

*We are grateful to Lluís Torrens, Mireia Raluy, and the research center SPSP at IESE School of Business for providing us with the Locomonitor data, and to all the participants at "la Caixa" seminar for their helpful comments. The views expressed herein are those of the authors and not necessarily those of "la Caixa". The authors take sole responsibility for the use and interpretation of the data and any errors therein.

1 Introduction

In the last decade, cross-border capital flows, and in particular Foreign Direct Investment (FDI), increased much faster than cross border trade or even than world GDP (see table 1) and they are projected to reach 1.2\$ trillion in 2006^{1,2}.

Table 1: FDI in the Global Economy

	1986-1990	1991-1995	1996-2000	2002	2003	2004	2005
FDI Inflows	21.7	21.8	40	-25.8	-9.7	27.4	28.9
GDP (current Prices)	11.1	5.9	1.3	3.9	12.1	12.1	9.1
Exports of goods and non-factor services	12.7	8.7	3.6	4.9	16.5	21	12.9

Data Source: World Investment Report 2006

This recent surge in FDI (see Figures 1 and 2), and the positive qualities associated to this sort of international capital flow, that is, its relative stability and its potential for spurring productivity and diffusing technology, have fed a growing discussion over the effects and determinants of FDI³. In particular, there is much light to shade on the factors that drive firms in a certain source country and in a certain industry to invest abroad, and to do so in a particular host country, operating a specific stage of the chain of production.

This paper contributes to the empirical literature on the *determinants* of FDI in using a recently available data set on new and expansionary investment projects to examine which home and host country characteristics influence the decision of a firm to invest abroad, and whether those factors play a different role according to the stage of production that the firm intends to operate in the foreign country.

We take first an aggregate approach, that is, we study the determinants of aggregate FDI activity between two countries in a particular year, and focus on the relative importance of factor proportion versus proximity-concentration arguments. Next, we exploit our data set at a higher level of disaggregation and classify investment projects into six different stages of the chain of production to analyze whether the factors considered as potential determinants of FDI play a different role depending on the stage of production of the investment project. These stages are extraction, manufacturing, business services, retail services, research and development (R&D),

¹See World Investment Report (2006).

²According to Economist Intelligence Unit and the Columbia Program on International Investment, and reported in The Economist, Sep. 14th 2006.

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

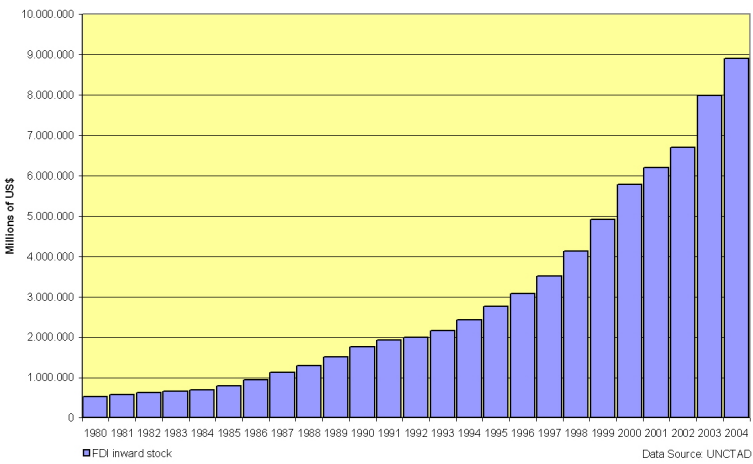


Figure 1: World FDI Inward Stock

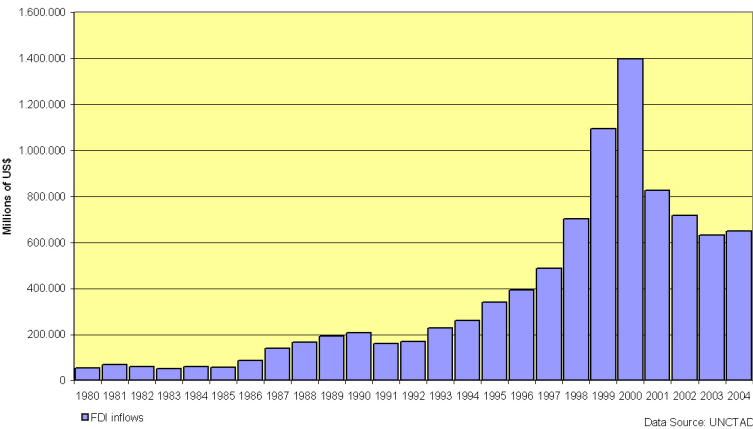


Figure 2: World FDI Inflows

and regional headquarters (HQ).

Anticipating our main results, when we analyze the aggregate number of FDI projects, we find that FDI flows to relatively capital scarce countries and that distance discourages FDI while size, and sharing a language or a border encourage it. At the disaggregate level, we observe that the effect of distance, and of all the other factors as well, vary a lot across stages. Moreover, even though the disaggregated results confirm the finding that FDI flows to relatively capital scarce countries, we find evidence of a minimum capital requirement in the foreign country for all the stages of production except extraction. Finally, ease of doing business appears to be particularly relevant for the decision of where to locate regional HQ.

The remaining of the paper is organized as follows. Section 2 reviews the literature and motivates the model. Section 3 describes the data set. Section 4 presents the empirical approach used and the results for the aggregate bilateral data. Section 5 analyzes the data at six different stages of production. Section 6 presents some robustness analysis. Finally, Section 7 concludes.

2 The Determinants of FDI: Where do we Stand?

FDI includes aspects of international trade, international flows, and information asymmetries. Hence, any study of the determinants of FDI activity, must acknowledge first the complexity involved in the decision to invest abroad.

Theoretical studies depart from the premise that a firm investing abroad will be at some initial disadvantage with respect to the local firms and that they will incur significant higher costs to operate in the foreign market relative those local firms. Therefore, there must be some sort of offsetting advantage in order for a firm to chose to become a multinational. Dunning (1977, 1981) provides a simple but systematic framework to understand what those advantages are. This framework is often referred to as the OLI (ownership, localization, internalization) framework. Dunning argues that the decision of a firm to invest abroad is based on ownership, location, and internalization considerations. First, a firm that invests must enjoy some economies of scale with respect to intangible assets such as knowledge capital or organizational know-how that can be easily exploited by investing abroad⁴.

⁴FDI is often understood as the acquisition of 10% or more of the assets of a foreign enterprise. It is usually defined as an investment involving a long-term relationship and reflecting a lasting

Second, the firm must have a reason to locate production abroad rather than concentrate it at home, especially if there are scale economies at the plant level. Hence, the decision to invest abroad will depend on the costs of investing there, the costs of operating there, and/or on the market access provided by that investment. Finally, the investing firm must have some incentive to want to exploit its ownership advantage internally, rather than licensing or selling its product or process to an unrelated foreign firm.

Hence, given that the decision to undertake an FDI project comprises many aspects, economic theory needs to connect all these ideas with firm and country characteristics in a consistent manner. The Knowledge-Capital model of FDI, which has become the workhorse of the multinational firm theory, does a serious effort in that direction, especially in the formulation of the **location** aspect of the FDI decision⁵.

According to this model, a firm will engage in FDI activity for two different reasons: market access or cheaper production. First, it may choose to produce abroad for market access purposes and in order to save on the trade costs associated to exporting. In this case, the firm will set up foreign facilities that mirror those at home (*Horizontal Expansion, HE hereafter*). In this scenario, location advantages arise when the host-country is large and when trade costs are large, and FDI would serve as a substitute for trade⁶. Hence, this hypothesis would predict that cross-border investment will take place between countries with high transportation costs and high trade barriers, low investment barriers, and low scale economies at the plant level⁷. At the same time, multinational activity would be two-way. Finally, market size at home and abroad would also have a positive impact, in the former case by increasing the market potential and in the latter by increasing the pool of potential investors abroad. Since size and trade barriers appear as fundamental determinants of horizontal FDI, an adaptation of the **gravity model**, which relates trade with size and distance, provides a good departing framework to study multinational activity.

interest and control of a resident entity in one economy (foreign direct investor or parent enterprise) in an enterprise resident in an economy other than that of the foreign direct investor (FDI enterprise or affiliate enterprise or foreign affiliate). FDI implies that the investor exerts a significant degree of influence on the management of the enterprise resident in the other economy. Conceptually, FDI is an extension of corporate control over national borders.

⁵See Carr, Markusen and Markusen (2001) among others.

⁶Krugman (1983), Horstman and Markusen (1992), Brainard (1992). A somewhat related motive would be what is referred to as *Quid pro Quo FDI*, a term coined by Bhagwati (1985) and that refers to a situation where a firm does FDI in response to a threat of protection.

⁷If scale economies at the plant level were high enough, the cost of replicating production in different locations might offset the gains from savings in transport costs.

Second, a firm may choose to fragment the production process and to locate some stages abroad in order to arbitrage differences in factor prices (*Vertical Expansion, VE hereafter*). This is also known as the factor proportions hypothesis for the location of multinational activity⁸. In that case, FDI would be equivalent to a sort of capital (factor of production) flow⁹. Paraphrasing Arndt and Kierzowski (2001), fragmentation, which is the possibility of breaking a production process into physically separable phases (stages or intermediate inputs), is not a new concept. On the contrary, it dates back to the beginning of the Industrial Revolution or even earlier than that. What is relatively new is the possibility of locating different phases of the production process in very distant places. It is in this process of splitting the production process across borders where FDI fits in. In this setting, FDI would be equivalent to a sort of capital (factor of production) flow¹⁰. Thus, when factor proportions are identical, the differential in GDP shares of the home and host countries and their joint income should be the only determinants of trade volumes, and there should be no multinational activities. At the same time, and again understanding FDI as an international movement of a production factor (capital flow), multinational activity should arise only in a single direction and between countries with large factor proportions differences¹¹. Note, as well, that this type of FDI may require shipment back to the home country to complete the production process. In sum, for a vertical firm, location advantages arise when trade costs are low, stages of production differ in factor intensities, and countries differ in relative factor endowments. Hence, in order to capture the central motivations for Vertical FDI, we shall extend the gravity framework to control for factor endowment characteristics.

In accordance to the discussion above, an extended gravity model that includes elements of factor proportions theory, is our basic framework to analyze the determinants of FDI from a macroeconomic perspective. The gravity model of trade is

⁸Markusen (1984), Helpman (1984), Helpman and Krugman (1985), Ethier and Horn (1990).

⁹As underscored in Ethier (1995), FDI does not always imply an actual movement of capital. For instance, a Japanese company may purchase a US firm by using the proceeds of selling its own US bonds. In that case, there would be no flow of capital from Japan to the US. Ethier notes that "the distinctive feature about direct investment, then, is not that one country is on balance acquiring the assets of another [...]. The distinctive feature is that firms in one country are acquiring control of business operations in another country."

¹⁰As underscored in Ethier (1995), FDI does not always imply an actual movement of capital. For instance, a Japanese company may purchase a US firm by using the proceeds of selling its own US bonds. In that case, there would be no flow of capital from Japan to the US. Ethier notes that "the distinctive feature about direct investment, then, is not that one country is on balance acquiring the assets of another [...]. The distinctive feature is that firms in one country are acquiring control of business operations in another country."

¹¹Brainard (1993, NBER) provides evidence that contradicts this line of thought.

well known as the workhorse of any empirical analysis of bilateral trade flows. Its success started in the empirical arena to later receive theoretical grounds¹². Although the theoretical foundations of the gravity equation for bilateral FDI are less clear than for trade flows it is still possible to derive a gravity-type framework to analyze bilateral FDI activity¹³. In this case, GDP of the host country captures the size of the market to be served by affiliates while the GDP of the home country is assumed to be proportional to the pool of potential investors in that country. The distance term accounts for the various transaction costs of carrying out international investment.

In the empirical arena, Eaton and Tamura (1994) provide an early example of the application of the gravity to FDI. Brainard (1993, 1997) also applies a gravity-type framework to FDI and she finds evidence on trade frictions increasing FDI, and hence in support of the horizontal firm's location advantages. Carr, Markusen, and Maskus (2001) provide the first empirical examination of the knowledge-capital model of FDI. From simulation of the model they derive a gravity-like empirical specification that relates affiliate sales in a host country to the GDP of both source and host countries, trade costs, FDI costs, plus differences in factor endowments. The authors use a panel data set on bilateral country-level US outbound and inbound affiliate sales and find evidence for both the horizontal and the vertical motivations for FDI. Yeaple (2003) also finds evidence in support of the vertical motive¹⁴.

3 The Data

Our main data set is a panel covering 6,728 investment projects (new and expansions) over 29 source countries and 131 destinations, between January 2002 and June 2005¹⁵. The source of this data is the Locomonitor database that documents 29,139 cross-border investment projects during that same period. We follow Gual and Torrens (2006), and our final sample includes only those projects undertaken by the 190 most important multinational firms. These firms are selected according to the following criteria: (1) firms that are listed in the top 100 MNEs by the UNCTAD world investment report (2004), (2) firms that are listed in the top 500 by

¹²See for instance Anderson and van Wincoop (2003). Refer to Feenstra (2004) for a review of theoretical foundations of the gravity model for trade.

¹³See Brainard (1993).

¹⁴See also Hanson, Mataloni and Slaughter (2003) and Freinberg and Keene (2001).

¹⁵For a complete list of the source and host countries, refer to Appendix A.

Fortune (2004), and (3) firms that report more than 20 investment projects during the period considered¹⁶.

For each project, the Locomonitor data base reports the following information: source country (or home, h hereafter); destination country (host or foreign, f hereafter), sector, and value-chain stage where the firm invests; whether the project represents a new investment, expansion or relocation; the name of the investing firm; and a brief description of the project. For some of the projects, it also provides information on the technology or product of the investment, the motivation that drove the investment decision, the public incentives that influenced the decision to invest in that location, or the value of the investment. Nonetheless, the latter is provided for a small number of projects only, too small to carry out a thorough analysis of the determinants of FDI based on the value of the investment. For that reason, when analyzing the determinants of FDI, we focus on the number of projects, that is, on the frequency of FDI activity, and not on the value of the investments. However, the value of the investment for this subset of projects is useful in reassuring the robustness of our results.

In the first part of the analysis we consider the number of investment projects between countries h and f in a particular year t . Given the number of investments per firm, the source, and the destination countries at each point in time, we can compute the number of projects for each country pair and year, including those observations (country pair-year) where that count is zero, that is, when no investment project takes place between two countries in a particular year. Our final sample for analysis includes a total of 14,384 observations for our dependent variable, which is defined as the number of FDI projects from firms in country h to invest in country f at time t ¹⁷.

In the second part of our study, we classify the projects according to their stage in the value chain. This further disaggregation of the data allows a better understanding of the determinants of the foreign investment decision of a firm. In that part of the analysis, we have a sample of 61,488 observations where 2,199 of them are non-zero. Similarly to the aggregate case, we include all country-value-chain observations that are zero but do engage in an FDI relationship in another stage of the value chain. Per

¹⁶For a complete description of the selection process see Gual et al. (2006).

¹⁷Note that a sample of 29 source and 131 host countries of over 4 years would correspond to a panel of 15,080 observations. However, missing data on GDP and other regressors for some of the countries forces us to drop a few observations.

contrary, we exclude country-pairs that do not engage at all in FDI. Thus, for each value-chain regression we have a sample of 6,966 observations (given the restriction just mentioned and other data restrictions like the availability of GDP data and other), with a minimum of 139 non-zero observations per value-chain stage.

In all our study, we consider two different sets of independent variables. The first set includes standard gravity regressors while the second set controls for elements of factor proportion models. The data for all these variables is briefly described below and more detail is provided in Appendix B.

3.1 Descriptive Statistics

We start by summarizing the basic features of our dependent and independent variables.

In the first part of the analysis, our dependent variable records the number of projects between two countries in a particular year. Hence, it is a count variable that takes nonnegative integer values, with great predominance of zeros as we can see in Figure 3. More precisely, it ranges from 0 to 85 and is strongly right-skewed, with 87% of the observations being zero counts (refer to Table 2 for the summary statistics of the main variables).

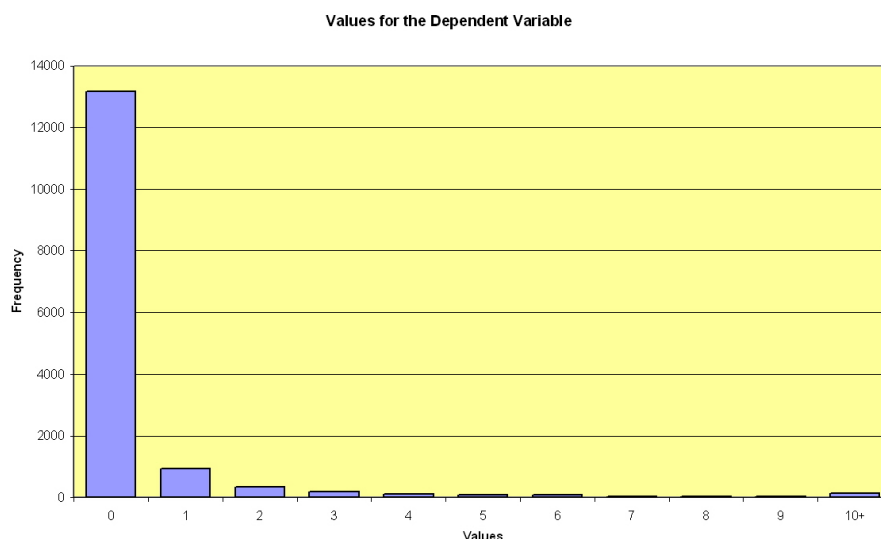


Figure 3: Overdispersion and Excess Zero Counts. SOURCE: Locomonitor data base and own calculations.

We shall note that using OLS with such a dependent variable would yield biased and inconsistent estimates. Instead, the Poisson model is considered the standard way of describing count data. Nonetheless, the Poisson distribution imposes a strong assumption on the data, namely that the conditional variance equals the mean. In our case, the mean is 0.4 and the standard deviation 2.4, and, moreover, formal regression-based tests rejected equality of the mean and variance. With such evidence of overdispersion in the data, the application of a negative binomial in generalized linear models appears as a better option to deal with the lack of normality in the data. Still, given the excess zero counts, the Zero Inflated Negative Binomial distribution (ZINB) is an even better option to deal with both, overdispersion and predominance of zeros. ZINB models are negative binomial models that allow for additional over-dispersion via a splitting process that models the probability of a zero outcome by logistic regression, while the continuous count outcome is modeled using a negative binomial error structure. In other words, ZINB assumes that two different underlying mechanisms are involved in generating zero and non-zero counts. A negative binomial process generates both zero and positive counts, and in addition, zero counts can arise separately through a logistic process.¹⁸ In our case, for example, where we have a panel containing the number of FDI projects for country-pairs hf at time t , the outcome $\# projects = 0$ may arise from two different scenarios: $\# projects = 0$ is the number of projects by a foreign firm from country h in country f in a particular year t while, at another time t' , that same firm may choose $\# projects = j > 0$. Alternatively, $\# projects = 0$ may occur when a foreign firm in country h never considers entering a country f market by FDI, regardless of the characteristics considered in the model.

In terms of the distribution of project counts across source and host countries, we observe a lot of variation with USA being the country projecting more FDI operations abroad, 547 in 2003 for instance, followed by Germany, Japan, and France (see Figure 4)¹⁹. The country receiving the largest number of projects is, by far, China, receiving half the projects, followed by USA, and Russia (see Figure 5).

Our basic set of regressors includes the product of GDPs, and the product of host

¹⁸We tested the appropriateness of the ZINB over a regular binomial using the Vuong statistic and we obtained high positive values in all our specifications, which favors zinb over negative binomial. Refer to Appendix 3 for a more detailed description. For further reference see Greene (2005).

¹⁹This figure is based on data for the year 2003 because it has the largest number of FDI projects in our sample.

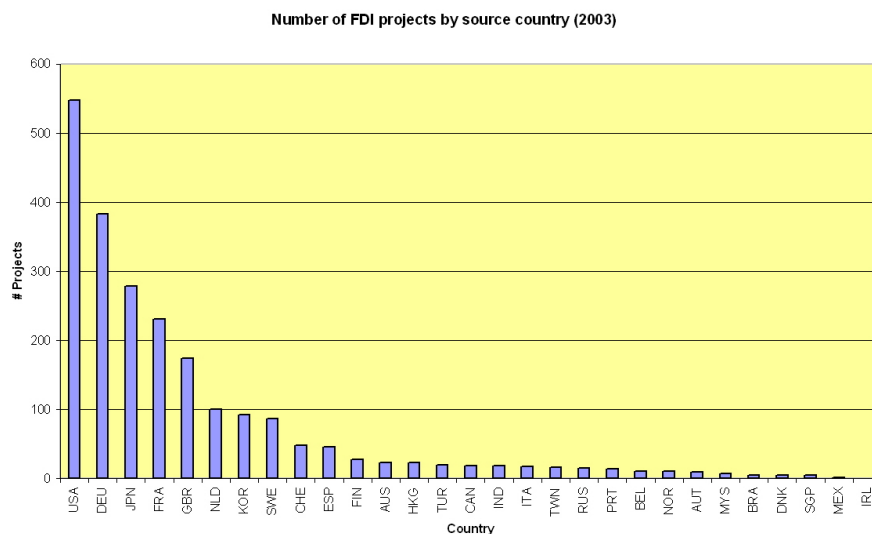


Figure 4: Number of Outward FDI Projects by Source Country (2003)

and source host country's Population, as measures of size; distance between the two countries in the pair (measured as the great circle distance between the main cities in each of the two countries) which accounts for transportation and other transaction costs influencing FDI that increase with distance; relative education between both countries, measured as the ratio of net enrollment in secondary education over net enrollment on primary education in the host relative to the source country; density of the foreign country relative to home (where density is measured as total population over land area); the capital labor ratio of the host country relative to the source country; and a measure for ease of doing business in the host country, where the more favorable the environment to set up a business in the host country is, the lower the values the measure takes. Notice that the measures for size and distance would correspond to gravity arguments, the following three variables are related to the factor proportion approach, and the latter is a proxy for other transaction costs. Table 2 reports the basic descriptive statistics for all these regressors. It is interesting to notice that, on average, host countries tend to be less "educated", in secondary over primary terms, and more labor-abundant than the source countries, a mean of -0.19 for the former versus -1.55 for the latter. In regards to simple correlations between all these variables, the largest correlation is by far that between the relative level of education and the relative capital labor ratio, which equals 0.63. This high correlation suggests that including one of these two variables while omitting the other may bias the results and attribute to the one included some or all of the

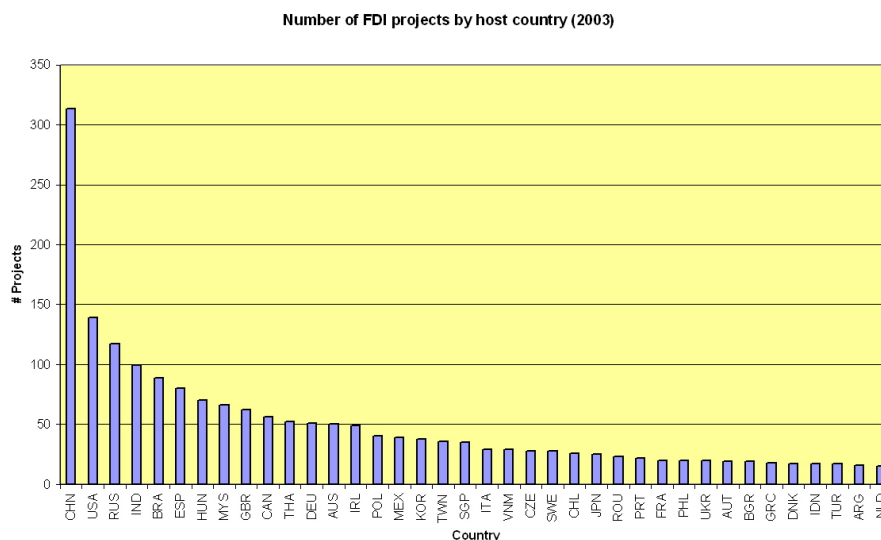


Figure 5: Number of Inward FDI Projects by Host Country (2003)

effects of the omitted one.

When turning to the dependent variable in the second part of the analysis, that is, the number of projects between two countries in a particular year and value-chain stage, we observe that the stages with the higher average number of projects are retail services and manufactures (see Table 3). Moreover, we see that the maximum number of projects per country-pair is in manufactures and between Japan and China (see column 3 to 5), and the second largest is the one from USA to China in Retail Services, which coincides with the maximum in retail services as well.

After this preliminary description of the variables, a more detailed study follows aimed at providing further insight on the determinants of FDI.

4 The Determinants of FDI: Aggregate

We use a combination of a standard **Gravity** model and **Factor Proportion** models of trade as the basic framework to analyze which factors influence the decision to invest in a particular project abroad. We take a macroeconomic perspective and consider country-pair characteristics that might drive FDI patterns. In the first part of the analysis we pool together all the investment projects recorded in our sample. In the second part, we work at a higher level of disaggregation and classify the in-

vestment projects per year and country pair, according to their stage in the value chain.

4.1 The Gravity Model

To derive our empirical specification, we start with a **BASIC GRAVITY** model where the number of investment projects by firms from country h (for home) in country f (foreign) at time t is explained by the size of countries h and f and the distance between them. Thus:

$$\begin{aligned} n_{hft} &= \alpha + \beta_1 \ln(GDP_{ht} \cdot GDP_{ft}) + \beta_2 \ln(dist_{hf}) + \epsilon_{hft} \\ h &= 1, \dots, 29; f = 1, \dots, 131; t = 2002, \dots, 2005 \end{aligned} \quad (1)$$

where n_{hft} stands for the number of projects that country h establishes in country f at time t and can take integer values greater than or equal to zero²⁰; GDP_{ht} is the gross domestic product of country h at time t ; similarly for GDP_{ft} , and $dist_{hf}$ is the distance between country h and f .

Regarding Equation (1) we expect market size to enter with positive sign, that is, a positive β_1 . In regards to the sign on β_2 , it depends on what distance is proxying for. If, on one hand, it is proxying for trade costs, the sign on β_2 would be ambiguous, and dependent on the complementarity or substitution between FDI and trade. In particular, if FDI is a substitute for trade, distance will raise export costs and the profitability of setting up affiliates in the foreign country mirroring the ones at home (*HE*), thus, implying a positive β_2 . If instead, FDI and trade are complements, higher distance will be associated with higher costs of production and lower profitability of an investment abroad. That would be the case of *vertical expansion*, since the cost of putting together the different fragments increases with distance, which implies a negative β_2 . On the other hand, if distance acts like a proxy for the overall transaction costs of doing FDI abroad, and if these costs increase with distance, then we would expect β_2 to be negative. Note, however, that some costs of doing FDI activity abroad may be totally independent from distance (ex. legal costs, information costs, etc.). Hence the sign on β_2 is again ambiguous.

²⁰Note that each of the 29 potential "investing" countries (home) does not invest in all of the 131 potential "host" countries (foreign). However, we still consider the complete set of (29home*130foreign*4year) observations which means that there is a large number of zero counts in our dependent variable.

Before discussing the results, an explanation on how to read them is warranted. When estimating Equation (1) with ZINB, we obtain two sets of results: those corresponding to the *inflated* model, and those corresponding to the *main* model. The inflated part corresponds to a logit, where the dependent variable would be the odds of a zero event, that is, the **probability of no FDI** activity between the two countries in the pair. Hence, if the coefficient on distance is positive in the inflated part it indicates that larger distance increases the probability of no FDI projects, or analogously, higher distance decreases the probability of FDI relationship between a country-pair. In the main model, the results can be read as the effect of each independent variable in increasing the **expected frequency of FDI** activity (expected number of project counts) between two countries²¹. In sum, estimated coefficients in the negative binomial part of the model (main part) are interpreted as in count models, while coefficients from the logistic part (inflated part) are related to the probability of a zero outcome.

Table 4 column (1) reports the results of estimating Equation (1) above by maximum likelihood based on a zero-inflated negative binomial distribution (ZINB). As aforementioned, we present two sets of coefficient estimates, corresponding to the logit and the negative binomial parts of the model. In the main model, the coefficient on the product of *GDPs* is positive and equals 0.71, as expected, while the one for distance is negative and equal to -0.41 , suggesting complementarity between trade and FDI, that is, vertical expansion motives or, in general, higher costs of FDI activity associated to distance. In the inflated model, the coefficient on *GDP* is negative indicating that economic size increases the probability of FDI relationship between countries f and h . Distance enters positively re-affirming the previous reasoning.

In the subsequent columns we extend the basic gravity to include other size controls, like population (L_i) or *GDP* per capita, and other geographical controls such as sharing a common border, sharing an official language, number of landlocked countries in the pair, and number of island countries in the pair. We observe that the coefficients on *GDP* and *distance* preserve their sign, however, the magnitude varies with the inclusion of the new gravity variables. In particular, the coefficient on *GDP* divides by half, (from 0.71 to 0.38), and that on distance decreases slightly in absolute terms from -0.41 to -0.31 , in the main part. Population size and shar-

²¹We consider the same factors as potential determinants in both the logistic and the negative binomial parts of the model.

ing a common language, both have a positive impact on the expected number of FDI projects. Sharing a border increases the likelihood of an FDI link (-2.82 in the inflated part), but it seems to affect negatively the expected number of projects once an FDI link is likely to occur ($-.244$ in the main). Island is never significant, and landlock is only significant in the main part, incentivating FDI.

4.2 Magnitude of the Effects

Table 5 presents the elasticities for each of the regressors, that is, the percentage change in the expected number of projects from country h in country f in year t when a particular regressor increases by one percent if the regressor is a continuous variable, and, the increase in the expected number of projects when one of the discrete regressors increases by one unit (from zero to one)²². All these elasticities and marginal effects are based on the estimates for Equation 1 and they provide a summary figure of the coefficients in both parts of the model, main and inflated.

We find that: one percent increase in the product of GDPs raises the expected number of projects by roughly one percent; one percent increase in distance decreases the expected number of projects by a magnitude ranging between 0.45 and 0.77 percent; the effect of population is not robust and sometimes not even significant, but overall seems to be positive; sharing a border or a language both slightly increase the expected number of projects, as does being landlocked. Island is not significant. The positive sign on landlock could suggest substitution between trade and FDI, given that no access to sea routes makes trade more difficult.

4.3 Introducing Factor Proportion Considerations

We turn now towards comparative advantage motivations for FDI and extend the basic gravity above to include elements from Factor Proportion models. The new estimating equation becomes

²²Note that the elasticities are evaluated at the mean of the regressors.

$$\begin{aligned}
n_{hft} = & \alpha + \beta_1 \ln(GDP_{ht} \cdot GDP_{ft}) + \beta_2 \ln(dist_{hf}) + \beta_3 \ln(L_{ht} \cdot L_{ft}) + \\
& + \beta_4 \ln\left(\frac{Sec_{ft}/Prim_{ft}}{Sec_{ht}/Prim_{ht}}\right) + \beta_5 \ln\left(\frac{Density_{ft}}{Density_{ht}}\right) + \beta_6 \ln\left(\frac{K_{ft}/L_{ft}}{K_{ht}/L_{ht}}\right) + \\
& + \beta_7 \ln(EaseBus_{ft}) + Z_{hf} + \epsilon_{hft}
\end{aligned} \tag{2}$$

where $\frac{Sec_{ft}/Prim_{ft}}{Sec_{ht}/Prim_{ht}}$ stands for the relative level of secondary education over primary education between foreign country, f , and home country, h , at time t ; similarly for $\frac{Density_{ft}}{Density_{ht}}$, and capital-labor ratio $(\frac{K_{ft}/L_{ft}}{K_{ht}/L_{ht}})$; $EaseBus_{ft}$ stands for ease of doing business in the host country, and controls for some of the transaction costs between country-pair (lower values corresponding to better environment to set up a business); finally, Z_{hf} stands for other controls such as sharing a common border or a common language, and the number of landlocked countries or island countries in the pair hf .

Table 6 presents the results of estimating Equation (2) by ZINB. The signs on the basic gravity regressors, $GDPs$ and $distance$, remain unchanged all over the different specifications. In other words, GDP increases FDI while distance discourages it, the latter suggesting, again, either a vertical motive for FDI or transaction costs of FDI that increase with distance. In Column (1), we introduce our first factor proportion control, namely, the relative ratio of **secondary educated workers over primary educated** workers abroad and at home, that proxies for the ratio of relative skilled over unskilled workers. It gets a positive coefficient in both the main (which ranges between 1.1 and 1.8 depending on the specification of controls) and the inflated model (between 2.8 and 3.9). Note that these effects work in opposite directions. The inflated part indicates that FDI activity is less likely the more abundant the foreign country is in secondary over primary educated workers relative to home, suggesting that FDI is searching for cheap unskilled labor, or that FDI travels from "North" (developed countries) to "South" (less developed countries). However, the results in the main part of the model suggest that once an FDI link is established between two countries, the more skilled the host country is relative to the source one, the higher the frequency of FDI activity between them.

The second column adds a measure of **relative density** (or relative land abundance) and the results (not significant in the main part and positive in the inflated model) indicate that the more dense the foreign country is relative to home, the less FDI activity will exist between them.

Column (3) adds a measure of **capital-labor abundance** in the foreign country relative to home. On top of it, column (4) adds the additional geographical factors aforementioned, and column (5) adds the measure of **ease of doing business** in the host country. The results in both parts of the model, inflated and main, show that the more capital abundant the foreign country is relative to home, the lower the frequency of FDI activity between them. This is consistent with the idea of FDI being a sort of capital flow, since you would expect capital to target countries with relatively higher return to capital, that is, those relatively labor-abundant or relatively capital-scarce.

Geographic variables behave similarly as they did before. Finally, **ease of doing business** gets a positive coefficient in the inflated part, thus, suggesting that countries where the process of setting up a business is easier, are more likely to attract FDI. The result in the main part is not significant, indicating no effect of that measure on the frequency of FDI.

As in the previous section, we compute the **elasticities** and present them in Table 7²³. The elasticity signs on the basic gravity regressors, *GDPs* and *distance*, remain unchanged all over the different specifications, positive for *GDPs* and negative for distance. Once again, this indicates that *GDP* increases FDI while distance discourages it, with the latter suggesting either a vertical motive for FDI or transaction costs of FDI that increase with distance. In Column (1), we introduce our first factor proportion control, namely, the relative ratio of **secondary educated workers over primary educated** workers abroad and at home, that proxies for the ratio of relative skilled over unskilled workers. It gets a negative elasticity, thus, FDI activity is less likely the more abundant the foreign country is in secondary over primary educated workers relative to home. This would suggest that FDI is searching for cheap unskilled labor, or that FDI travels from "North" (developed countries) to "South" (less developed countries). However, when we control for relative capital labor abundance, besides a significant change in the magnitude of the effect of economic size (increasing from around 0.7 to 1.0), the significance of the effect of relative skill abundance disappears. The high correlation between our measures of relative skill-abundance and relative capital-labor endowment could explain a relatively lower weight on the skill abundance variable when the capital abundance variable is included (the omission of capital abundance would attribute

²³In order to construct the elasticity table we use the estimation results in table 6 of the ongoing research with the same name ("la Caixa" Working Paper Series Nr. 5/2006).

to skill abundance the combined role of both variables). Nevertheless, it is highly unlikely that this high correlation explains the fact that skill abundance becomes insignificant once capital abundance is included. Instead, we attribute this effect to our measure of skill abundance being a noisy proxy for the level of skilled-unskilled labor abundance.

To sum up, the main results when controlling for gravity and factor proportion elements together, presented in column (5), indicate that one percent increase in the product of **GDPs** yields a 1 percent increase in the number of FDI projects. A particular pair of countries will trade around 0.65 percent more than another pair of countries that are 1 percent **farther away** from each other. Both the effects of population and relative **skilled** abundance are small and not significant. One percent increase in the relative **density** of the foreign country decreases the number of projects by 0.1 percent. One percent increase in the relative **capital** abundance of the foreign country yields 0.4 percent less projects. Hence, the more capital abundant the foreign country is relative to home, the lower the frequency of FDI activity between them. This is consistent with the idea of FDI being a capital flow, since you would expect capital to target countries with relatively higher return to capital, that is, those relatively labor-abundant or relatively capital-scarce.²⁴ **Ease of doing business** gets a negative elasticity, thus, suggesting that countries where the process of setting up a business is easier (lower regressor value), are more likely to attract FDI. As for all the **geographic variables**, their effects are similar to those in the basic gravity case. That is, border, language, and landlocked all have a positive overall effect on the expected number of FDI projects between two countries, while island is not significant.²⁵

When we control for relative capital labor abundance, the most significant changes are on the magnitude of the effect of economic size, and on the economic and statistical significance of the effects of population and relative skill abundance. On one hand, economic size is more relevant economically, with 1% increase in the product of GDPs yielding a 1% increase in the number of FDI projects, and, on the other

²⁴Please refer to footnote 9.

²⁵Note that these numbers are the result of a combination of the estimates of both, the main and the inflated coefficients in Table 6. For instance, the figures for education in the first column are negative and significant because the effect in the main part of the model is stronger and dominates the effect (in the opposite direction) in the inflated part. However, the figures for education in the last three columns are positive but not significant because the coefficient estimates in the main and the inflated part of the model go in opposite directions and they are so similar in magnitude that they cancel each other.

hand, both the effects of population and relative skilled abundance are smaller and not significant. The effects of distance and density remain practically the same, and for capital abundance, a 1% increase in the relative capital abundance of the foreign country yields 0.4 percent less projects. In terms of the variable capturing **ease of doing business** in the host country (Column 5), we observe that a business-friendly environment is associated with more FDI activity. As for all the geographic variables, their effects are similar to those in the basic gravity case. That is, border, language, and landlock all have a positive overall effect on the expected number of FDI projects between two countries.

5 The Determinants of FDI: Stages of the Value Chain

In this section we classify FDI projects according to their stage in the value chain of production. We consider six potential stages in the production of a good: extraction, manufacturing, business services, retailing services, R&D, and regional HQ.

For each stage v , we estimate Equation (2) above with the dependent variable counting the number of projects between countries h and f at time t that belong to that stage (v). Hence, we are allowing for the possibility that the β s differ across stages in the value chain, that is, for the possibility that the independent variables have a different effect on the likelihood and frequency of FDI activity, depending on the stage in the value chain the project belongs to.

Tables 8 and 9 summarize the results of estimating Equation (2) for each stage while Table 10 reports elasticities for each of them²⁶. Even when the detail provided by the first two tables helps in understanding the underlying mechanisms, we focus our attention on the last table, since it provides a summary of the overall effect of each of the regressors on the expected number of FDI projects in stage v and how they compare across stages.

We can summarize the main results as follows:

Economic size (GDPs) is a determining factor and increases the expected of FDI

²⁶For simplicity, we only present the results for the full regression, that is, with all the independent variables.

activity in all stages. The magnitude of its effect is especially important for manufacturing, business services, and retailing services. Overall, this effect ranges between 0.6 and 1.3 (1% increase in the combined economic size of two countries raises the expected number of projects between 0.6 and 1.3 percent).

- **Distance** discourages FDI activity in all but extraction. This makes sense from an economic perspective since factors driving extraction investment have little to do with distance and a lot more with the specific location of natural resources. Distance has the largest impact on the number of regional HQ projects. This indicates that investing firms would rather locate regional HQ nearby. Monitoring purposes could be an explanation.
- For **Population**, the sign, and magnitude of the effect varies a lot across stages in the value chain.
- The overall effect of **relative skill abundance** of the host country is positive on the frequency of FDI activity in all the stages of production, except in Business Services, where it gets a negative sign. Thus, in general education incentives FDI, especially in R&D and HQ. Nevertheless a note of caution is warranted with respect to these estimates since our measure of relative skill abundance is a noisy proxy for the true level of skilled-unskilled labor abundance.
- **Density** has a negative but modest effect on FDI activity, indicating that the more dense the foreign country is relative to home, the less FDI activity will exist between them. This result suggests that firms would rather do cross border investment in countries such as the USA (or even China that has a density below the mean) than in countries such as Mauritius or Malta with densities well above the mean.
- As in the aggregate analysis, the effect of relative **capital-labor-abundance** of the foreign country is always negative, although the magnitude of its effect varies considerably across stages. We interpret this as FDI targeting capital scarce countries where its rentability is higher. However, in this particular case, we can get further insight into the mechanisms behind this effect if we look at the regression results from the main and inflated part of the ZINB model (Tables 8 and 9). While the coefficient of capital-labor abundance in the main part is always negative (except for extraction), thus, suggesting that

the frequency of FDI activity increases with the relative capital abundance of the home country (or capital scarcity of the foreign), the coefficient on the inflated part is negative as well, which would indicate that FDI is less likely the more relatively capital scarce the foreign country is. This result can be interpreted as a *minimum* requirement of capital on the host country in order for and FDI link to be established. In other words, the host country must have a minimum level of capital in order to increase the likelihood of receiving an FDI project. Once this threshold is set, FDI flows to relatively-labor abundant countries. This minimum requirement, that is, the negative effect of capital scarcity in the foreign country on the likelihood of FDI, is quite important for R&D, that gets a coefficient of -3.13 in the inflated part, much larger than in all the other stages of production (except extraction) where magnitudes range between $-.33$ and $-.53$. Again, extraction behaves differently, with capital scarcity in the foreign country increasing the likelihood of FDI and exerting no significant effect on the frequency of projects. Once more, it seems plausible that the decision to invest in extraction projects is based on location of natural resources more than on any other factor.

- **Geographic** variables, exert in general a weak impact when looking stage by stage.
- Finally, **ease of doing business** gets the expected sign (negative) in all the stages of production, except again extraction. Remember that lower values of the variable correspond to more favorable environment to set up a business. Thus, the negative sign says that high values in the ease of doing business (bad environment) decrease the quantity of projects between two countries. The largest impact of the business environment is on regional HQ projects.

6 Robustness Analysis

One possible criticism to the results presented in this paper relates to the choice of the dependent variable. In particular, we refer to the use of the number of projects undertaken by large firms between two countries, instead of the usual volume of FDI flows (or inward stock) from the source to the host country.

Given that our data set provides information on the value of the investment for some

of the projects, we use of that information to test the robustness of our estimates. In particular, we show that the projects included in our data account for a large proportion of the world FDI flows. Hence, we conclude that using the number of projects should not be a big problem.

We rely on two different approaches to estimate how big the projects are. First, based on the projects for which we have information on the value of the investment, **we impute for each year the total volume of FDI** that the projects account for and compare it with world FDI figures. For instance, in 2003 and out of the 2,224 projects in our sample, 904 of them report the investment value (refer to Table 11), and add up to a total of \$20,499 millions. Then, assuming that the remaining projects, 1,320 ($1320 = 2224 - 904$), are similar in amount invested to those for which the information is reported, we conclude that our sample of projects represents a total of \$512,490 millions in FDI flows, equivalent to 83% of the total world FDI (see columns 3 and 5 respectively). Similarly for the rest of the years, we conclude that the projects considered represent around 50% of the world FDI.

Based on these results, we conclude that the projects in our sample account for a big portion of the world FDI flows. Moreover, we observe that in 2003 the percentage of the world FDI that is accounted for by the data set is larger than in other years. Hence, we re-run the gravity and the factor proportion specifications in Tables 4 and 6 for 2003. We find that our main results are robust both in sign and magnitude²⁷.

In a **second** approach, we calculate **for each year and source country the total value of the investment** for all those projects for which the information is available. Then, as we did before, we assume that the projects without investment information are similar in magnitude to those for which the information is reported. Next, we aggregate the total amount invested per source country and compare it to outflow data across countries. Using this approach, we observe that for the US and most European countries we have close to 100% of the FDI flows.

Since extraction seems to behave quite different than the rest of value-chain stages, we perform a second robustness check. We re-run equation 2 at the aggregate level but without including extraction projects. The results are quite robust and, thus, we conclude that extraction is not driving the main results. Nonetheless, we notice that the effect of the capital labor variable in the inflated part, even tough, still not significant, diminishes, tending to the negative value we observe in the disaggregated

²⁷These results are available upon request.

case for each stage, except for extraction.

7 Concluding Remarks

The recent surge in Foreign Direct Investment (FDI) activity and the increased complexity of production processes that split tasks across borders warrant an investigation on the driving factors behind this type of cross border investment.

This paper contributes to the literature on the determinants of FDI on two grounds. First, it uses a recently available data set on new and expansionary investment projects to reexamine which home and host country characteristics influence the decision of a firm to invest abroad. Second, it uses one further level of disaggregation in the data to investigate how the goal of the investment affects the influence of the different determinants of FDI. We define the goal of the investment in terms of the stage of production that the firm intends to operate in the foreign country.

At the aggregate level, our most robust findings can be summarized as follows. On the gravity side, size and sharing a common language promote FDI while distance discourages it. This negative effect of distance could be interpreted as evidence suggesting vertical motives for FDI, as opposed to horizontal motives. However, this interpretation considers distance, exclusively, as a measure of trade costs and ignores other barriers to FDI that increase with distance. If we take a broader interpretation of our "distance" measure, then we can only conclude that any barrier to FDI that increases with distance has a negative impact on FDI. For instance, it could well be that investments in nearby countries are easier to monitor than businesses in far-away countries, discouraging FDI in the latter. In regards to factor proportion differences, our results indicate that FDI targets countries that are relatively capital-scarce (or relatively labor-abundant).

When we classify investment projects according to their stage in the chain of production, that is from a disaggregate standpoint, we conclude that, in general, the effect of all the regressors varies substantially across stages. Nonetheless, there are some regularities. Economic size, as well as distance, are determining factors in explaining the number of FDI projects between countries in a particular value-chain stage, and preserve their sign with respect to the aggregate analysis. Even though the results confirm the aggregate finding that FDI targets capital-scarce countries, we find evidence of a minimum requirement on the foreign country's capital endow-

ment in all the stages of production except extraction. Ease of doing business is important for all types of FDI, especially so for regional HQ. Geographic variables have a small effect. Finally, FDI in extraction behaves in a way that differs from all the other stages and seems to get little impact from the regressors considered in our study. This finding is not at odds with economic intuition since factors driving extraction investment have little to do with distance or relative factor proportions and a lot more with the specific location of natural resources.

In a nutshell, our findings underscore an important role for both *economic size* and *transaction costs* associated with distance and language barriers in determining FDI activity. Furthermore, FDI in stages of production other than extraction targets capital-scarce countries and requires a minimum level of capital in the host country, while a favorable business environment promotes inward FDI in all stages and especially regional HQ.

References

- Anderson, J. E., and E. van Wincoop, (2003). "Gravity with Gravitas: A Solution to the Border Puzzle," *American Economic Review*, *American Economic Association*, 93(1), pp. 170-192.
- Arndt, S.W. and H. Kierzowski (2001). *Fragmentation: New Production Patterns in the World Economy*, Oxford University Press.
- Bhagwati, J., (1985). "Protectionism: Old Wine in New Bottles," *Journal of Policy Modeling*, 7, pp. 23-34.
- Brainard, S.L., (1992). "A Simple Theory of Multinational Corporations with a Trade-off Between Proximity and Concentration." *MIT Sloan Working Paper* 3492-92-EFA.
- Brainard, S.L., (1993). "An Empirical Assessment of the Factor Proportions Explanation of Multinational Sales." *NBER Working Paper* 4380.
- Brainard, S.L., (1997). "An Empirical Assessment of the Proximity-Concentration Trade-off Between Multinational Sales and Trade." *American Economic Review*, 87(4), pp. 520-544.
- Cameron, A.C., and P.K. Trivedi, (1998). *Regression Analysis of Count Data*. Cambridge University Press, Cambridge, UK.
- Carr, D.L., J.R. Markusen, and K.E. Maskus, (2001). "Estimating the Knowledge Capital Model of the Multinational Enterprise." *American Economic Review*, 93 (3), pp. 693-708.
- Dunning, J.H., (1977). "The Determinant of International Production." *Oxford Economic Papers* 25, pp. 289-330.
- Dunning, J.H., (1981). *International Production in and the Multinational Enterprise*. London, Unwin Hyman.
- Eaton, J. and A. Tamura, (1994). "Bilateralism and Regionalism in Japanese and U.S. Trade and Direct Foreign Investment Patterns," *NBER Working Papers* 4758.
- Ethier, W.J., (1995). *Modern International Economics*, Norton & Company, Inc., New York.

- Ethier, W.J., and H. Horn, (1990). "Managerial Control of International Firms and Patterns of Direct Investment." *Journal of International Economics* 28, pp. 25-45.
- Feenstra, R., (2004). *Advanced International Trade*. Princeton University Press, Princeton, NJ.
- Freinberg, S.E., and M.P. Keene, (2001). "U.S.-Canada Trade Liberalization and MNC Production Location," *Review of Economics and Statistics*, 83(1), pp. 118-132.
- Gual, J., Torrens, Ll., (2006). "L'atracció d'inversió estrangera: Factors clau perquè Barcelona i Catalunya puguin competir en el mercat internacional." *Col·leccio: Prospectiva, Pla Estratègic Metropolità de Barcelona*, 9.
- Greene, W.H., (1994). "Accounting for Excess Zeros and Sample Selection in Poisson and Negative Binomial Regression Models." *Stern School of Business WP*, EC-94-10.
- Greene, W.H., (2005). *Econometric Analysis*, Prentice Hall.
- Hall, R.E., and C. I. Jones (1999). "Why Do Some Countries Produce So Much More Output per Worker than Others?" *Quarterly Journal of Economics*, 114(1), pp. 83-116
- Hanson, G.H., R.J. Mataloni, and M.J. Slaughter, (2003). "Vertical Production Networks in Multinational Firms," *Review of Economics and Statistics*, 87, pp. 664-678.
- Hausman, J.A., B.H. Hall, and Z. Griliches, (1984). "Econometric Models for Count Data with an Application to the Patents-R&D Relationship," *Econometrica*, 52, pp. 909-938.
- Helpman, E., (1984). "A Simple Theory of Trade with Multinational Corporations." *Journal of Political Economy*, 92, pp. 451-471.
- Helpman, E., and P. Krugman, (1985). *Market Structure and International Trade*. MIT Press, Cambridge, MA.
- Horstman, I.J., and J.R. Markusen, (1992). "Endogenous Market Structures in International Trade." *Journal of International Economics*, 20, pp. 225-247.
- Krugman, P., (1983). "New Theories of Trade among Industrial Countries," *American Economic Review*, 73(2), pp. 343-347

Markusen, J.R. (1984). "Trade and the Gains from Trade with Imperfect Competition." *Journal of International Economics*, 11, pp. 331-351.

Locomonitor Data Base, (2005), <http://www.locomonitor.com/>.

Rappaport, J. (2000). "How Does Openness to Capital Flows Affect Growth?." Mimeo, Federal Reserve Bank of Kansas City, June.

Romer, P., (1993). "Idea Gaps and Object Gaps in Economic Development." *Journal of Monetary Economics*, 32(3), pp. 543-573.

Rodriguez-Clare, (1996). "Multinational, Linkages, and Economic Development." *American Economic Review*, 86, pp. 852-873.

UNCTAD, (2006). *World Investment Report 2006. FDI from Developing and Transition Economies: Implications for Development*. United Nations, New York and Geneva.

Vuong, Q.H., (1989). "Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses," *Econometrica*, 57, pp. 307-333.

Yeaple, S.R., (2003). "The role of Skill Endowments in the Structure of U.S. Outward Foreign direct Investment," *Review of Economics and Statistics*, 85(3), 726-734.

Appendix A: Country List

Home Countries (Source)

Australia	Germany	Netherlands	Switzerland
Austria	Hong Kong	Norway	Taiwan
Belgium	India	Portugal	Turkey
Brazil	Ireland	Russia	UK
Canada	Italy	Singapore	USA
Denmark	Japan	South Korea	
Finland	Malaysia	Spain	
France	Mexico	Sweden	

Host Countries (Destination)

Afghanistan	El Salvador	Liechtenstein	Serbia & Mont.
Albania	Ecu. Guinea	Lithuania	Singapore
Algeria	Estonia	Luxembourg	Slovakia
Angola	Ethiopia	Macau	Slovenia
Argentina	Finland	Macedonia	Somalia
Armenia	France	Madagascar	South Africa
Australia	Georgia	Malaysia	South Korea
Austria	Germany	Maldives	Spain
Azerbaijan	Ghana	Malta	Sri Lanka
Bahamas	Greece	Mauritius	Sudan
Bahrain	Guatemala	Mexico	Suriname
Bangladesh	Guinea	Moldova	Sweden
Belgium	Haiti	Morocco	Switzerland
Bolivia	Honduras	Mozambique	Syria
Bosnia & Her.	Hong Kong	Namibia	Taiwan
Brazil	Hungary	Netherlands	Tanzania
Brunei	Iceland	New Zealand	Thailand
Bulgaria	India	Nicaragua	Trinidad & Tob.
Cambodia	Indonesia	Nigeria	Tunisia
Canada	Iran	Norway	Turkey
Chad	Iraq	Oman	Turkmenistan
Chile	Ireland	Pakistan	UAE
China	Israel	Panama	UK
Colombia	Italy	Paraguay	USA
Costa Rica	Jamaica	Peru	Uganda
Croatia	Japan	Philippines	Ukraine
Cuba	Jordan	Poland	Uruguay
Cyprus	Kazakhstan	Portugal	Uzbekistan
Czech Republic	Kenya	Puerto Rico	Venezuela
Denmark	Kuwait	Qatar	Vietnam
Dominican Rep.	Latvia	Romania	Zambia
Ecuador	Lebanon	Russia	Zimbabwe
Egypt	Libya	Saudi Arabia	

Appendix B: Data Description and Sources

Variable	Description	Source
# of FDI Projects	Number of FDI projects in country f from firms in countries h , recorded in the Locomonitor data base.	Locomonitor data base, January 2002-June 2005.
GDP; GDP/cap	Gross domestic product, in current US\$, and Gross domestic product per capita, in current US\$.	World Economic Outlook, IMF, September 2006.
Distance	Great-circle distance in kilometers between the principal cities in countries h and f .	Own calculations based on area, latitude, and longitude data from the CIA World Factbook, 2006.
Population	Total Population.	World Bank, GDI 2006.
Border	Equals 1 if countries h and f share a common border. Zero otherwise.	Own construction based on data from CIA World Factbook, 2006.
Language	Equals 1 if countries h and f share a common (official or widely spoken) language. Zero otherwise.	Own construction based on data from CIA World Factbook, 2006.
Landlock	Number of landlocked countries in the pair hf .	Own construction based on data from CIA World Factbook, 2006.
Island	Number of island countries in the pair hf .	Own construction based on data from CIA World Factbook, 2006.

Variable	Description	Source
Secondary-Primary	Enrollment ratio in secondary education over Enrollment ratio in primary education in the host country, relative to the equivalent in the source country.	UNESCO, Statistical Yearbook. Use Net enrollment in primary education and secondary education. For missing values we use either the Gross enrollment or estimate them from gross or net enrollment rates available. For specific extrapolations refer to the authors.
Density	Population over land area in kilometers.	Population data from World Bank WDI; area data from CIA World Factbook, 2006.
Capital-Labor	Capital endowment over population. Capital accumulation is calculated using perpetual inventory method from 1978.	Population data from World Bank WDI; Investment data from Penn World Tables 6.1; initial capital from Hall and Jones (1999).
Easiness Bus.	Easiness of Doing Business measure, with lower values indicating worse scenario.	Easiness of doing business from the World Bank, http://www.doingbusiness.org

Appendix C: Count Data Model

Our dependent variable, n_{hf} (the number of investment projects in country f from firms in country h), is a count variable. Since OLS estimates would be biased and inconsistent, we use a standard count data model that specifies the probability that the j^{th} host country attracts n_{hf} investment projects from firms in country f , as a function of a vector of country-pair- hf 's specific attributes (X_{hft})²⁸:

$$\Pr(n_{hft} = n) = F(X'_{hft}\beta) \quad (A1)$$

where β is the vector of unknown parameters to be estimated. The most common way to specify such a probability function is to use a Poisson process. In that process, the integer property of the dependent variable is modeled as:

$$\Pr(n_{hft} = n) = \frac{\exp(-\mu_{hft})\mu_{hft}^n}{n!} \quad (A2)$$

where $n = 0, 1, 2, \dots$ and $\mu_{hft} = \exp(X'_{hft}\beta)$

The shortcoming of the Poisson model in Eq. A2 is its implication that μ_{ijt} is both the mean and the variance of the process. This restriction is often violated due to over- or under-dispersion of the data. In our empirical analysis, we tested and rejected this equi-dispersion assumption in all cases. The Negative Binomial model relaxes this assumption by allowing μ to vary systematically, which in turn allows the variance of the process to differ from the mean and makes it more appropriate to estimate models in the presence of over-dispersion of the data. In the NB model, the probability distribution becomes:

$$\Pr(n_{hft} = n) = \frac{\exp(-\mu_{hft})\mu_{hft}^n}{n!} \quad (A3)$$

$$\mu_{hft} = \exp(X_{hft}\beta + u_{hft}), \quad \exp(u_{hft}) \sim \Gamma(1/\alpha, \alpha)$$

α being the overdispersion parameter

$$n = 0, 1, 2, \dots$$

Unlike the Poisson model, the NB has an additional parameter α , and, therefore n_{hft} is an iid NB process with mean μ_{hft} and variance $\sigma^2 = \mu_{hft} + \alpha\mu_{hft}$. Note

²⁸See Hausman, Hall and Griliches (1984) for further details.

that when $\alpha = 0$, the NB reduces to a Poisson.

Our data not only shows signs of over-dispersion but also of excessive zero counts of FDI projects. Excessive number of zeros may arise for various reasons. First, due to adverse host-country characteristics (inherent natural features or political factors), many countries may not appeal to potential foreign investors. In that case, some countries may always attract zero counts of foreign investment projects. Second, even if a country has "good characteristics" as a potential recipient of foreign investment it may receive no investment in a particular year from a particular country. In the latter case, assuming a general probability distribution for zero and non-zero counts shouldn't be a problem. In the former situation, though, there is the possibility that the zeros follow a different process. Zero-inflated models allow each process to be modeled separately²⁹. The common strategy is to map a logit onto a count model. We adopt such a strategy and model our dependent variable as the product of two latent variables, a binary (0,1) following a logistic distribution function, and a discrete count variable following a NB as specified above (ZINB):

$$\begin{aligned}
n_{hft} &= z_{hft} n_{hft}^* \\
z_{hft} &= 0 \text{ with prob } q_{hft}, z_{hft} = 1 \text{ with prob } 1 - q_{hft} \\
q_{hft} &= (\exp(Z_{hft}\gamma)) / (1 + \exp(Z_{hft}\gamma)) \\
Z_{hft} &\text{ is a set of observables} \\
n_{hft}^* &\sim NB(\mu_{hft}, u_{hft}) \text{ as specified in Eq. A3}
\end{aligned} \tag{A4}$$

This means that the probability structure will be:

$$\begin{aligned}
\Pr(n_{hft} = 0) &= \Pr(z_{hft} = 0) + \Pr(z_{hft} = 1, n_{hft}^* = 0) = q_{hft} + (1 - q_{hft})f(0) \\
\Pr(n_{hft} = n) &= \Pr(z_{hft} = 1, n_{hft}^* = n) = (1 - q_{hft})f(n), n = 1, 2, \dots \\
q(.) &\text{ is a logit process and } f(.) \text{ is a NB as specified above}
\end{aligned} \tag{3}$$

In all our estimations we assume the vector of observables Z_{hft} to be identical to our vector of explanatory variables X_{hft} . We follow Vuong (1989) and test for the choice between the ZIP and the ZINB³⁰ and the results favor the use of a ZINB.

²⁹See Cameron and Trivedi ("Regression Analysis of Count Data", Econometric Society monographs No. 30. Cambridge: Cambridge University Press, 1998) for a thorough discussion of the different empirical approaches to account for excess zeros.

³⁰It is straight forward as models are nested. Large positive values of the Vuong test favor ZINB while large negative models favor the ZIP. See Vuong, Q.H. (1989), "Likelihood Ratio Tests

Table 2: Summary Statistics

	Obs	Mean	Std. Dev.	Min	Max
Number of Projects	15080	.40	2.40	0	85
$\ln(GDP_{ht} \cdot GDP_{ft})$	14384	9.97	2.20	4.04	17.86
$\ln(dist_{hf})$	15080	8.65	.82	4.16	9.90
$\ln(L_h \cdot L_f)$	13776	33.38	2.19	27.75	41.81
$\ln\left(\frac{Sec_f}{Prim_f} / \frac{Sec_h}{Prim_h}\right)$	15107	-.19	.49	-2.71	.86
$\ln\left(\frac{Density_f}{Density_h}\right)$	13776	.37	2.33	-7.84	7.88
$\ln\left(\frac{(K/L)_f}{(K/L)_h}\right)$	9396	-1.55	1.78	-6.01	3.55
$\ln(EaseBus_f)$	1308	3.97	1.04	0	5.15

Summary of the "aggregate" dependent variable and the main regressors in logarithmic terms.

Table 3: Number of Projects by Value Chain Stage

Value-Chain	Avg Num. Projects	Max	Home	Foreign
Extraction	.03	4	USA	United Kingdom
Manufactures	.19	54	Japan	China
Bus. Serv.	.09	11	USA	India
Retail Serv.	.22	30	USA	China
R & D	.04	18	USA	China
HQ	.02	4	Germany	USA

Year 2003. The second column shows the average number of projects by each value chain stage, the third column shows the maximum flow, thus, indicates the flow with the maximum number of projects; column fourth and fifth between which countries the maximum flow occurred.

for Model Selection and Non-Nested Hypotheses." *Econometrica* 57: 307-333 and Greene (1994), "Accounting for Excess Zeros and Sample Selection in Poisson and Negative Binomial Regression Models," Stern School of Business WP, EC-94-10, 1994.

Table 4: Basic Gravity Regression

Main (Expected # counts)	(1)	(2)	(3)	(4)
$\ln(GDP_{ht} \cdot GDP_{ft})$.71 (.02)***	.38 (.03)***		.38 (.03)***
$\ln(dist_{hf})$	-.41 (.04)***	-.31 (.03)***	-.31 (.04)***	-.32 (.04)***
$\ln\left(\frac{GDP_{pc_{ht}}}{GDP_{pc_{ft}}}\right)$.38 (.03)***	
$\ln(L_h \cdot L_f)$.28 (.02)***	.66 (.02)***	.30 (.02)***
<i>border</i>				-.24 (.13)*
<i>language</i>				.13 (.09) ¹
<i>landlock</i>				.29 (.09)***
<i>island</i>				.01 (.06)
<hr/>				
Inflated (Prob zero count)				
$\ln(GDP_{ht} \cdot GDP_{ft})$	-.67 (.07)***	-.99 (.07)***		-1.08 (.08)***
$\ln(dist_{hf})$.31 (.14)**	.80 (.12)***	.80 (.12)***	.67 (.12)***
$\ln\left(\frac{GDP_{pc_{ht}}}{GDP_{pc_{ft}}}\right)$			-1.00 (.07)***	
$\ln(L_h \cdot L_f)$.38 (.06)***	-.61 (.05)	.47 (.07)***
<i>border</i>				-2.82 (1.20)**
<i>language</i>				-1.16 (.29)***
<i>landlock</i>				-.30 (.21) ¹
<i>island</i>				.06 (.16)
N	14384	13664	13664	13664
Non-Zero	1904	1872	1872	1872

Standard errors are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels of significance

Table 5: Elasticities - Basic Gravity Regression

Expected # counts	(1)	(2)	(3)	(4)
$GDP_{ht} \cdot GDP_{ft}$.81 (.04)***	.96 (.04)***		1.00 (.04)***
$dist_{hf}$	-.45 (.03)***	-.77 (.06)***	-.77 (.06)***	-.70 (.06)***
$\frac{GDP_{pc_{ht}}}{GDP_{pc_{ft}}}$.96 (.04)***	
$L_h \cdot L_f$.05 (.025)**	1.01 (.03)***	.03 (.03)
$border^1$.07 (.02)***
$language^1$.08 (.01)***
$landlock^1$.04 (.01)***
$island^1$				-.00 (1.14)

Standard errors are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels of significance. (1) For this variable, marginal effect is reported for a discrete change of the variable from 0 to 1.

Table 6. Gravity Regressions Adding Factor Abundance Variables

Main (Expected # counts)	(1)	(2)	(3)	(4)	(5)
$\ln(GDP_{ht} \cdot GDP_{ft})$.40 (.03)***	.39 (.03)***	.49 (.04)***	.48 (.04)***	.48 (.04)***
$\ln(dist_{hf})$	-.27 (.04)***	-.27 (.03)***	-.22 (.04)***	-.28 (.04)***	-.27 (.05)***
$\ln(L_h \cdot L_f)$.33 (.03)***	.34 (.03)***	.18 (.04)***	.22 (.04)***	.22 (.04)***
$\ln\left(\frac{Sec_f}{Prim_f} / \frac{Sec_h}{Prim_h}\right)$	1.18 (.13)***	1.20 (.13)***	1.69 (.18)***	1.72 (.18)***	1.72 (.18)***
$\ln\left(\frac{Density_f}{Density_h}\right)$		-.02 (.02)	-.03 (.02)	-.03 (.02)*	-.03 (.02)
$\ln\left(\frac{(K/L)_f}{(K/L)_h}\right)$			-.37 (.07)***	-.34 (.06)***	-.35 (.07)***
$\ln(EaseBus)$					-.02 (.05)
<i>border</i>				-.41 (.15)**	-.36 (.15)**
<i>language</i>				.11 (.09)	.10 (.09)
<i>landlock</i>				.14 (.10) ¹	.17 (.10)*
<i>island</i>				-.13 (.06)	-.03 (.07)

Table 6. Gravity Regressions Adding Factor Abundance Variables (cont.)

Inflated (Prob zero count)	(1)	(2)	(3)	(4)	(5)
$\ln(GDP_{ht} \cdot GDP_{ft})$	-.83 (.06)***	-.82 (.062)***	-1.03 (.08)***	-1.16 (.09)***	-1.11 (.09)***
$\ln(dist_{hf})$	1.11 (.13)***	1.02 (.13)***	.75 (.13)***	.53 (.15)***	.72 (.17)***
$\ln(L_h \cdot L_f)$.30 (.06)***	.32 (.05)***	.30 (.06)***	.41 (.07)***	.32 (.09)***
$\ln\left(\frac{Sec_f}{Prim_f} / \frac{Sec_h}{Prim_h}\right)$	4.57 (.47)***	4.46 (.44)***	2.80 (.48)***	3.38 (.58)***	3.92 (.63)***
$\ln\left(\frac{Density_f}{Density_h}\right)$.16 (.05)***	.12 (.04)***	.12 (.05)**	.16 (.05)***
$\ln\left(\frac{(K/L)_f}{(K/L)_h}\right)$.11 (.07) ¹	.12 (.08) ¹	.19 (.09)**
$\ln(EaseBus)$.35 (.16)**
<i>border</i>				-5.32 (1.99)***	-4.76 (1.71)***
<i>language</i>				-1.27 (.29)***	-1.20 (.29)***
<i>landlock</i>				-.76 (.29)***	-.79 (.29)***
<i>island</i>				-.11 (.20)	-.00 (.21)
N	13664	9288	9288	9288	8856
Non-Zero	1872	1872	1497	1497	1484

Standard errors are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels of significance.

Table 7: Elasticities - Gravity adding Factor Abundance

Expected # counts	(1)	(2)	(3)	(4)	(5)
$GDP_{ht} \cdot GDP_{ft}$.71 (.041)***	.70 (.04)***	1.06 (.07)***	1.05 (.08)***	.97 (.08)***
$dist_{hf}$	-.69 (.06)***	-.66 (.06)***	-.64 (.06)***	-.54 (.06)***	-.59 (.06)***
$L_h \cdot L_f$.21 (.02)***	.22 (.02)***	.01 (.04)	.02 (.05)	.08 (.05)*
$\frac{Sec_f}{Prim_f} / \frac{Sec_h}{Prim_h}$	-.55 (.16)***	-.49 (.16)***	.12 (.18)	.06 (.18)	-.01 (.18)
$\frac{Density_f}{Density_h}$		-.08 (.02)***	-.09 (.02)***	-.09 (.02)***	-.10 (.02)***
$\frac{(K/L)_f}{(K/L)_h}$			-.43 (.05)***	-.39 (.06)***	-.43 (.06)***
$EaseBus_f$					-.17 (.06)***
$border^1$.06 (.03)**	.06 (.04)*
$language^1$.11 (.02)***	.10 (.02)***
$landlock^1$.07 (.02)***	.08 (.02)***
$island^1$.00 (.01)	.00 (.01)

Standard errors are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels of significance. (1) For this variable, the marginal effect is reported for a discrete change of the variable from 0 to 1.

Table 8. Value Chain Regressions (Part I)

Main (Expected # counts)	Extraction	Manufactures	Bus Serv
$\ln(GDP_{ht} \cdot GDP_{ft})$.91 (.18)***	.58 (.05)***	.44 (.12)***
$\ln(dist_{hf})$.27 (.17)	-.32 (.06)***	-.23 (.08)***
$\ln(L_h \cdot L_f)$	-.47 (.21)**	.18 (.05)***	.18 (.11)
$\ln\left(\frac{Sec_f}{Prim_f} / \frac{Sec_h}{Prim_h}\right)$	-.01 (.36)	.12 (.33)	3.78 (.49)***
$\ln\left(\frac{Density_f}{Density_h}\right)$	-.05 (.06)	-.08 (.02)***	-.06 (.03)*
$\ln\left(\frac{(K/L)_f}{(K/L)_h}\right)$.07 (.13)	-.87 (.08)***	-1.00 (.15)***
$\ln(EaseBus)$.55 (.18)***	-.22 (.06)***	-.03 (.08)
<i>border</i>	1.62 (.78)**	-.27 (.21)	-.87 (.29)***
<i>language</i>	.53 (.26)**	-.18 (.14)	.67 (.16)***
<i>landlock</i>	.71 (.55)	.59 (.14)***	.24 (.18)
<i>island</i>	-.09 (.21)	-.08 (.09)	-.08 (.12)

Table 8. Value Chain Regressions (Part I, cont.)

Inflated (Prob of zero count)			
$\ln(GDP_{ht} \cdot GDP_{ft})$	1.30 (.39)***	-1.54 (.19)***	-1.35 (.27)***
$\ln(dist_{hf})$	1.15 (.50)**	.71 (.23)***	.97 (.34)***
$\ln(L_h \cdot L_f)$	-1.92 (.45)***	.54 (.16)***	.35 (.29)
$\ln\left(\frac{Sec_f}{Prim_f} / \frac{Sec_h}{Prim_h}\right)$	-3.35 (1.71)**	-.53 (.47)	6.44 (1.44)***
$\ln\left(\frac{Density_f}{Density_h}\right)$.22 (.14)	.10 (.08)	.15 (.11)
$\ln\left(\frac{(K/L)_f}{(K/L)_h}\right)$	2.43 (1.14)**	-.36 (.17)**	-.33 (.25)
$\ln(EaseBus)$	1.35 (.41)***	.65 (.34)*	.97 (.40)**
<i>border</i>	2.59 (1.34)*	-15.99 (608.38)	-2.01 (2.00)
<i>language</i>	-.53 (.68)	-.76 (.41)*	-.81 (.53)
<i>landlock</i>	3.93 (1.35)***	-1.03 (.41)**	-1.48 (.67)**
<i>island</i>	-.57 (.53)	-.23 (.26)	-.02 (.43)
N	6642	6642	6642
Non-Zero	152	623	385

Standard errors are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels of significance. ¹ Denotes 15% level of significance.

Table 9. Value Chain Regressions (Part II)

Main (Expected # counts)	Retail Serv	R and D	HQ
$\ln(GDP_{ht} \cdot GDP_{ft})$.65 (.06)***	.40 (.09)***	.45 (.13)***
$\ln(dist_{hf})$	-.49 (.07)***	-.68 (.14)***	.17 (.14)
$\ln(L_h \cdot L_f)$.00 (.06)	.49 (.08)***	.10 (.12)
$\ln\left(\frac{Sec_f}{Prim_f} / \frac{Sec_h}{Prim_h}\right)$	1.26 (.49)**	6.82 (.76)***	-1.18 (1.21)
$\ln\left(\frac{Density_f}{Density_h}\right)$	-.09 (.03)***	.01 (.04)	.06 (.04)
$\ln\left(\frac{(K/L)_f}{(K/L)_h}\right)$	-.70 (.10)***	-2.07 (.19)***	-.68 (.25)***
$\ln(EaseBus)$.07 (.07)	-.71 (.12)***	-.45 (.14)***
<i>border</i>	-.38 (.21)*	-.53 (.39)	.48 (.59)
<i>language</i>	.27 (.16)*	.13 (.19)	-.06 (.29)
<i>landlock</i>	-.02 (.15)	.39 (.26)	.40 (.39)
<i>island</i>	-.16 (.10)*	.16 (.15)	.02 (.19)

Table 9. Value Chain Regressions (Part II, cont.)

Inflated (Prob of zero count)			
$\ln(GDP_{ht} \cdot GDP_{ft})$	-1.46 (.26)***	-1.39 (.32)***	-.62 (.63)
$\ln(dist_{hf})$.51 (.22)**	-1.63 (.42)***	2.01 (1.07)*
$\ln(L_h \cdot L_f)$.63 (.24)***	.91 (.35)**	-.92 (.81)
$\ln\left(\frac{Sec_f}{Prim_f} / \frac{Sec_h}{Prim_h}\right)$	1.19 (.79)	15.96 (3.15)***	-6.45 (4.01)
$\ln\left(\frac{Density_f}{Density_h}\right)$.10 (.11)	.12 (.17)	.30 (.22)
$\ln\left(\frac{(K/L)_f}{(K/L)_h}\right)$	-.40 (.21)*	-3.13 (.66)***	-.53 (.97)
$\ln(EaseBus)$.75 (.50)	-.93 (.49)*	2.26 (.95)**
<i>border</i>	-16.28 (377.03)	-6.78 (2.81)**	7.06 (4.28)*
<i>language</i>	-.26 (.52)	-1.41 (.79)*	-.02 (1.17)
<i>landlock</i>	.15 (.46)	-1.58 (.86)*	.93 (1.59)
<i>island</i>	-.24 (.31)	.76 (.61)	.69 (.85)
N	6642	6642	6642
Non-Zero	669	218	139

Standard errors are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels of significance. ¹ Denotes 15% level of significance.

Table 10: Elasticities - Value Chain

	(1)	(2)	(3)	(4)	(5)	(6)
Expected # counts	Extraction	Manufactures	Business	Retail	R and D	HQ
$GDP_{ht} \cdot GDP_{ft}$.60	1.27	1.26	1.34	.74	1.05
$dist_{hf}$.01	-.64	-.81	-.73	-.27	-1.79
$L_h \cdot L_f$	-.02	-.07	-.03	-.30	.26	.99
$\frac{Sec_f}{Prim_f} / \frac{Sec_h}{Prim_h}$.77	.36	-.12	.70	2.85	5.10
$\frac{Density_f}{Density_h}$	-.10	-.12	-.15	-.13	-.02	-.24
$\frac{(K/L)_f}{(K/L)_h}$	-.49	-.71	-.80	-.51	-1.29	-1.6
$EaseBus_f$.23	-.51	-.62	-.28	-.47	-2.65
<i>border</i> ¹	.02	.05	.00	.05	.00	.00
<i>language</i> ¹	.02	.01	.04	.04	.01	.00
<i>landlock</i> ¹	.00	.07	.03	-.01	.01	.00
<i>island</i> ¹	.00	.00	.00	.00	.00	.00

In the disaggregate case the delta method used to compute the standard errors does not always converge, thus, they are not reported in the table. Nonetheless, the model estimates presented in the two previous tables show that most of the results are significant. (1) For this variable, the marginal effect is reported for a discrete change of the variable from 0 to 1.

Table 11: Projects

Year	Number Projects (1)	Number Projects Info (2)	Total Investment (Millions dollars) (3)	World FDI flows (Millions dollars) (4)	Percentage % (5)
2002	1427	679	332,666	652,181	50
2003	2224	904	512,490	616,923	83
2004	2317	867	403,227	730,257	55
2005	760	282	319,356	.	.

For 2005 the data goes from January to June, thus, we assume the same number of projects for the second half of the year as for the first half.