

Escobar, Octavio

Conference Paper

The location pattern of FDI in Mexico after NAFTA

50th Congress of the European Regional Science Association: "Sustainable Regional Growth and Development in the Creative Knowledge Economy", 19-23 August 2010, Jönköping, Sweden

Provided in Cooperation with:

European Regional Science Association (ERSA)

Suggested Citation: Escobar, Octavio (2010) : The location pattern of FDI in Mexico after NAFTA, 50th Congress of the European Regional Science Association: "Sustainable Regional Growth and Development in the Creative Knowledge Economy", 19-23 August 2010, Jönköping, Sweden, European Regional Science Association (ERSA), Louvain-la-Neuve

This Version is available at:

<https://hdl.handle.net/10419/118997>

Standard-Nutzungsbedingungen:

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

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

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

Terms of use:

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

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

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

The location pattern of FDI in Mexico after NAFTA

Octavio R. Escobar-Gamboa*†

Abstract

This paper studies the location pattern of foreign direct investment (FDI) in Mexico for the period 1994-2004. An empirical gravity type model is specified based on recent FDI theories and empirical papers. It is then estimated using state-level data by FDI source country. An extended gravity-type model with elements of multinationals theory is specified, and a methodology is introduced to take into account the zero-flows observations. Results suggest that not taking into account these observations lead to biased results. In addition, they suggest the influence of geographical location and wages as major determinants on the FDI distribution across Mexican states.

Keywords: FDI; Mexico; gravity models;

1 Introduction

In recent years, many developing countries have tried to attract foreign direct investments (FDI) in order to compensate their lack of capital for financing their economic activity. These countries also consider FDI beneficial as a source of access to markets, technologies and other assets that are not available in the local economy [UNCTAD (2006)]. Thus, countries compete to propose the most attractive production conditions (i.e. legal environment and economic policies) and their policy-makers take measures with the aim of attracting foreign capital.

To improve its attractiveness, Mexican authorities implemented some major liberalization reforms¹. Firstly, Mexico signed the North American Free Trade Agreement (NAFTA) with Canada and the US which came into effect in January 1994. Secondly, obstacles to FDI were significantly removed with reforms to FDI regulations in 1989 and a new regulation in 1993. After the implementation of these reforms, Mexico became an attractive country for foreign investors to locate their operations. Mexico's principal advantage over other developing countries is the free access to Canadian and US markets. As a result, according to Mexican authorities, between 1994 and 2004 Mexico received about 147 US billion dollars in FDI.

In addition to NAFTA, Mexican authorities pursue free trade agreements (FTAs) with other countries. Mexico has entered into eleven free trade agreements with 41 countries and an economic partnership with Japan. The expectation for Mexican authorities was that FTAs would increase export diversification (reducing Mexico's dependence on the US economy), stabilize the economy

*Université Paris-Dauphine, LEDa, F-75016 Paris, France (e-mail: octavio.escobar@dauphine.fr)

†IRD, UMR225-DIAL, F-75010 Paris, France

(after the 1982 and 1994 economic crises), and improve the country's risk assessment and visibility (to attract FDI) [Ibarra-Yunez (2003)].

Important empirical literature has analyzed FDI effects in Mexico. Some of these studies point out the fact that FDI stimulates the total factor productivity in Mexico [Aitken et al. (1997); Lopez-Cordoba (2003); Tornell et al. (2004)]. Studies also find that economic activity with a strong foreign presence has better performances and that the productivity levels of national firms converge to the levels achieved by foreign-owned firms in their sectors. However, FDI inflows in Mexico are also a source of negative effects. Indeed, FDI focuses in the export-oriented sector, having little contact with national activity, and thus limiting Mexico's industrialization process [Mattar and Moreno-Brid (2002)]; moreover, FDI has contributed to a rise in income inequality [Feenstra and Hanson (1997)].

Even if a common belief is that foreign investments try to benefit from cheap labor in Mexico [Love and Lage-Hidalgo (2000)] to serve the North American market rather than to supply the Mexican market [Graham and Wada (2000)], the literature studying the determinants of FDI inflows in Mexico is scarce. Love and Lage-Hidalgo (2000) study the determinants of FDI flows from the US to Mexico and find that labor costs as well as local demand play an important role. However, their paper covers a period of study prior to NAFTA and differences between Mexican regions are not studied. These differences are explored by Mollick et al. (2006) who analyzed the determinants of FDI inflows into twenty-two Mexican states in the period of 1994–2001. These authors conclude that improvement in infrastructure (measured as telephone line density) is the main factor for attracting FDI in Mexican regions. Nevertheless, these authors limit their analysis to the twenty-two highest-ranking states by FDI amount. Moreover, differences across economic sectors are not analyzed.

The purpose of this paper is to study a broad issue regarding inward FDI in Mexico that has not yet received much attention. It empirically examines economic theory on multinational enterprises to determine: (1) the motivations for FDI in Mexico, and (2) the determinants of the geographical distribution of FDI at a regional level. In particular, it employs a gravity model to analyze the interaction between Mexican regions characteristics, FDI motivations and spatial linkages to explain the location pattern of FDI in Mexico. This is the first study to my knowledge that analyzes spatial linkages for FDI across the Mexican states. These spatial interactions seem to be important for other countries, as other studies on FDI and spatial econometrics suggest. Baltagi et al. (2007) and Blonigen et al. (2007) study the cases of outward US FDI, while Coughlin and Segev (2000) analyze inward FDI in Chinese regions. All of them found a source of spatial spillovers across multinational activity.

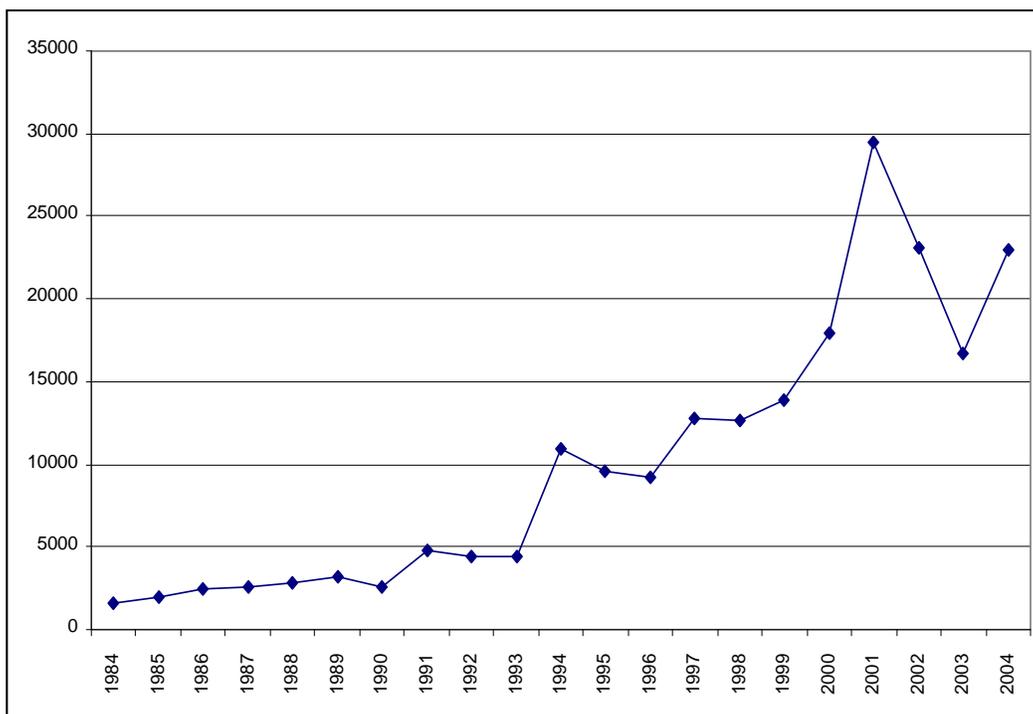
This paper is organized as follows. The next section briefly introduces the geographical and industrial configuration of inbound FDI in Mexico. Section three describes the theoretical background and it is followed by a discussion on the econometrical issues in FDI empirical studies in section four. Indeed, recent works such as Head and Ries (2008) and Kleinert and Toubal (2010) demonstrate that multinational enterprises (MNEs) theoretical models can be modeled by a gravity-type model. However this kind of model is not exempt of econometrical issues. Among these issues, the inclusion or not of zero-flows observations has opened a debate about the estimation techniques to be applied. Section five presents the empirical specification to handle the zero-flows observations. In this section, following Semykina and Wooldridge (2005) and Razin et al. (2003) works, the study of FDI inflows are decomposed into two equations: (1) an equation measuring the volume of FDI or intensive margin; and (2) an equation measuring the profits of MNEs or extensive margin. The specification work out to a main gravity equation, which needs to take into account the probability

for a country to invest and for a region to receive FDI to avoid estimation bias. It then proposes to follow the estimation procedures developed by Semykina and Wooldridge (2005) to correct for this bias. These authors propose a methodology that allows for correcting selection bias in the presence of heterogeneity, endogeneity and serial correlation which are current issues in the empirical analysis of FDI flows. Section six presents the variables definition followed by the empirical analysis in section seven. Results suggest the presence of selection bias. After correcting this bias, results suggest differences in wages, bilateral distance, neighboring FDI activity and (unobserved) setup costs as the principal FDI determinants. Finally, conclusions are drawn in section eight.

2 FDI distribution in Mexico

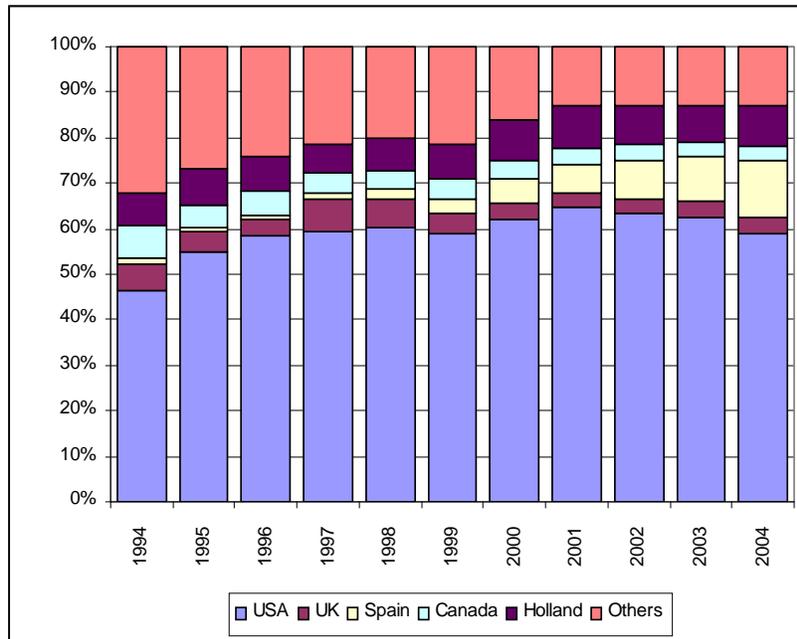
NAFTA and reforms on FDI regulation seem to be successful in attracting FDI since, as shown by figure 1, FDI inflows increased importantly after their implementation. According to figure 1, average yearly FDI inflows have more than doubled since 1994 compared to 1984–1993. However, this increase should not only be attributed to liberalization reforms since a change in FDI definitions was made in 1994. Indeed, prior to 1994, investments related to transfers of stocks from nationals to foreigners, imports of fixed assets by maquila firms and transfers between subsidiary and parent companies were not accounted as FDI [INEGI].

Figure 1: FDI inflows (US million dollars)



The main source of FDI from 1994–2004 was the US, as shown in figure 2. Moreover, the share of this country as a source of FDI increased during this period by more than 10%. Spain has also considerably increased its share of FDI inflows. Spain accounted for about 1.4% of inbound FDI in 1994, while in 2004 its share of accumulated FDI during 1994–2004 was almost 12.7%. In the meantime, other important sources of FDI in Mexico as Canada and Holland maintained its share, while the share of the United Kingdom, Germany, France and Japan has decreased.

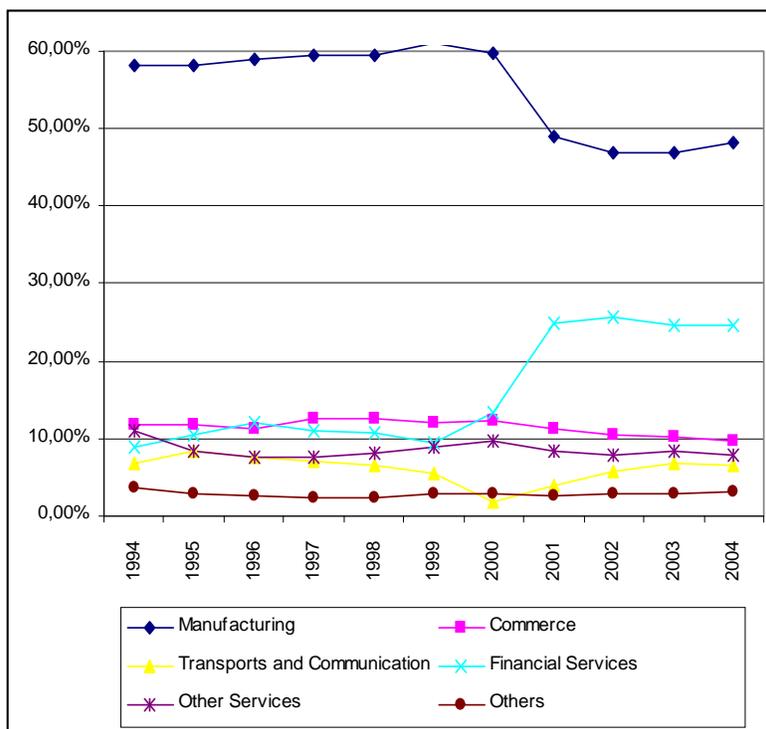
Figure 2: Accumulated FDI inflows by country of origin



Concerning the destination of FDI by economic sector in Mexico, as figure 3 suggests, most of these investments go to the manufacturing sector. This sector share of FDI inflows during 1994–2004 is about 48%. However, as figure 3 shows, the importance of this sector has decreased in favour of the financial sector. The increase of FDI in the financial sector can be attributed to a FDI deregulation implemented in 1999 that allows a 100% of foreign participation in the financial sector. The financial sector attained its highest share in 2001 with the acquisition of the biggest national commercial bank by a foreign group².

Looking at the spatial distribution of FDI inflows in Mexico, FDI is not equally distributed in the country. During 1994–2004 more than 60% of the accumulated FDI inflows went to the Federal District, which is Mexico’s capital as well as the country’s richest entity. FDI is also unequally distributed among the rest of the Mexican states. Figure 4 illustrates the density distribution of FDI inflows during 1994–2004, as well as the share of the main FDI source countries. This map shows that FDI tended to concentrate around the Federal District. It also points out the existing

Figure 3: Share of accumulated FDI by economic sector



differences in FDI volume and source between the states located in the north of Mexico and those located in the south. We can see that, except for the state of Nuevo Leon, the US FDI share dominates by far the accumulated FDI by the rest of the countries in the northern border states. This is not the case for the southern states, where, except for the states of Guerrero, Tabasco and Yucatan, the share of the US is inferior to that of the rest of the countries.

The map analysis suggests that FDI location patterns vary across country source. Moreover, it also suggests that FDI is concentrated in states located mainly near big markets, such as the Federal District and the US. However, in order to find out why FDI concentrates in these states, as well as to identify the motivations for FDI, a regression analysis based on economic theory is conducted.

3 Determinants of FDI location: theoretical background

There has been considerable progress in recent years in terms of theory regarding multinational enterprises (MNEs) and FDI location behavior. Dunning (1973) proposed that ownership, location and internationalization (OLI) advantages encouraged firms to undertake foreign investment. In this framework, location is perceived as an advantage that firms can obtain from locating production abroad (i.e. economies of scale). However, OLI's framework was not built on a formal setting. Since

Markusen (1984) and Helpman (1984), Dunning's ideas were incorporated into general equilibrium theory of trade leading to the "knowledge-capital" (KC) model.

Markusen (2002) suggests two different patterns of FDI: (1) horizontal and (2) vertical. Horizontal MNEs produce similar goods and services in different countries. Their main motivation is to place production close to consumers and thereby avoid trade costs when these costs are high. Multinationals have a plant and a headquarters in a home (source) country and other plants in each host country. So each production facility supplies each individual market. The main factors which affect the incentives for horizontal FDI are trade cost and market size of host country. When trade costs rise, exporters would encounter a higher marginal cost. Hence, they have higher incentives to build a plant in the host country to sell their products directly. If the market size of the host country expands, firms would have more incentives to build a production facility there because the associated fixed cost would be covered by the revenue generated. In the other hand, if trade costs decrease, multinationals with higher fixed costs prefer to concentrate their activity in one country and to develop trade flows with partner countries rather than to open production plants.

Vertical foreign investments refer to a fragmentation of the production process into stages of production, each of them produced in different locations. After producing (intermediate or final) goods in a lower cost country, firms with vertical FDI import them to supply their home consumers. Hence, vertical FDI is motivated by the differences in factor prices, especially by differences in labor costs, and by trade costs. As trade costs increase, firms with vertical FDI face higher costs to import goods from the FDI host country. Consequently, vertical FDI will increase as trade costs decrease and as differences in factor prices increase.

"Complex FDI" versions of the KC model have been developed more recently. The literature mainly distinguishes two types of these models: (1) complex horizontal FDI or export platform FDI; and (2) complex vertical FDI or vertical specialization FDI. In an export platform FDI model [Ekholm et al. (2007); Baltagi et al. (2007)] a home country firm would set up a production plant in a region that benefits from better access and lower production costs than the home country, to serve as a production platform for exports to a group of "neighboring" regions. Finally, in complex vertical models, MNEs separate their production process into multiple vertical activities and put them in locations offering the lowest (trade and/or production) costs. Baltagi et al. (2007) suggest that, with these complex modes of MNEs organization, host country characteristics are not the only determinants attracting FDI, the host neighbors' characteristics could also play an important role.

Concerning firm location decisions, new economic geography (NEG) models incorporate location theories into a formal model³. Moreover, most of them analyze how economic integration affects the location of economic activity. NEG literature points out the importance of agglomeration and dispersion forces to attract or to discourage firms to locate in a given region. On one hand, a firm has the incentive to locate in proximity to other firms, especially if these are its suppliers or its customers, to take advantage of positive externalities as technological spillovers, access to infrastructure and to specialized labor, forward and backward linkages, etc. On the other hand, congestion effects, as well as increases in wages and land prices, reduce the attractiveness of a region. Egger et al. (2007) as well as Hoffman and Markusen (2007) incorporate the KC model in a NEG framework to show how liberalization reforms influence FDI location decisions. These studies highlight the importance of factor prices and market size as major determinants of location decisions.

Bergstrand and Egger (2007) suggest that, in general equilibrium, bilateral multinational enterprise activity such as FDI or foreign affiliate sales depend on the same exogenous determinants

as bilateral trade flows. Thus, it is not surprising that the literature (i.e. Egger (2008), Razin et al. (2003), among others) has employed international trade gravity models in the study of FDI. Kleinert and Toubal (2010) suggest that the success of the gravity equation results from the fact that it can be derived from various theoretical models. These authors derive a gravity equation for horizontal and vertical models of multinationals firms. Expected results vary with FDI motivation, however for both models, the distance coefficient is expected to be negative. Head and Ries (2008), considering FDI in the form of cross-border mergers and acquisitions, develop a model that also yields an equation for bilateral FDI in the form of a gravity equation used to analyze bilateral trade.

In this line, to study the Mexican states determinants to attract FDI by country of origin, an augmented gravity-type model is employed. In other words, in addition to the "classical" size and distance variables included in the gravity models, elements from the NEG and from the "new" theories of FDI are included in the model. Before presenting the model and the estimation strategy, some of the major econometrical issues regarding the gravity-type models are introduced in the following section.

4 Econometrical issues on FDI studies

Before presenting the estimation strategy and data description, a set of econometrical issues when conducting empirical analysis on FDI determinants are introduced.

In addition to the well known heteroskedasticity problem present in multinational activity analysis [Egger (2008)], there are three major econometrical issues when estimating a gravity-type equations: (1) the zero problem, (2) heterogeneity, and (3) endogeneity.

4.1 Zero FDI flows

A common issue concerning gravity models estimation is the fact that a log-linearized version is usually employed. Given that the log of zero is undefined, taking logarithms of zero flows between country i and state j drops such observations from the sample. Dropping observations means not taking into account potentially useful information. Indeed, we might be able to learn something about why some countries prefer to invest in some regions and not in others. In addition, estimates could be biased when employing only a portion of the available data. It seems then important to take zeros into account since statistics on FDI exhibit more zero bilateral observations than trade data [Egger (2008)].

Recent literature on the gravity model has paid an increasing attention to take these zero observations into account. We can distinguish four different approaches: (1) add a small number to flows; (2) the Tobit model; (3) the Poisson model; and (4) the sample selection models.

Given that $\log(0)$ is undefined, but $\log(0 + 1)$ is not, some authors -i.e. Eichengreen and Irwin (1995); Levy et al. (2003)- decided to add a small number (usually 1) to take zero observations into account. However, when zero observations are important -i.e., Helpman et al. (2008) as well as Santos Silva and Teneyro (2006) find that about half of the country pairs in their sample do not trade with each other at all- most of the data will be censored at below zero. Tobin (1958) posits that he needed to take account of the concentration of observations at zero when testing his hypotheses and he develops the Tobit model.

The gravity literature has applied the tobit model to handle zero observations bias, but the value of to be added to zero before a logarithmic transformation has become a subject of debate. Eaton and Tamura propose a variation of the tobit model developed by where the logarithmic transformation of FDI takes the following form: $\ln(a_v + FDI_{ij})$, where the choice of the parameter a_v is endogenized. Santos Silva and Tenreyro (2006) criticize log-linear models and provides an alternative approach to estimation. According to these authors, log linear models cannot be expected to provide unbiased estimates of mean effects when the errors are heteroskedastic. Santos

Silva and Tenreyro (2006) then propose the Pseudo Poisson Maximum Likelihood model (PPML) which has already been widely adopted, i.e. see Egger (2008) and Head and Ries (2008). Advantages of the PPML estimator are that it is consistent under weak assumptions, the data do not need to be distributed as Poisson and, since the dependent variable is in value, not logarithm, the PPML model enables us to estimate models which includes the zeros⁴. However, the PPML model implicitly assumes that there is nothing special about zeros. It only gets them into the estimation sample.

An alternative approach to the problem is in terms of sample selection. Heckman (1979) proposes to two-step estimators. In our case, a first set of covariates determine the likelihood that a country decides to invest in a given Mexican state (extensive margin) and a second set of covariates determine the volume of this investment (intensive margin). One way of thinking of sample selection is as an omitted variables problem. By dropping zeros from the sample, the dependent variable is no longer FDI flows, but FDI flows contingent on an existing FDI relationship. Thus, an important variable left out of the model is the probability of being included in the sample (of having a FDI relationship). To the extent that the probability of selection is correlated with economic characteristics or to bilateral distance, then it has the potential to bias OLS estimates. The Heckman (1979) consists to estimate, in a first time, a Probit model in which the dependent variable is a dichotomous variable, say s , that determines whether ($s = 1$) or not ($s = 0$) a given observation is in the sample; and a measure of the probability of being in the sample (the inverse Mill's ratio) is derived from the Probit estimates and it is included into the main model. The corrected model is then estimated by OLS. Helpman et al. (2008) propose a sample selection gravity model based on the Melitz (2003) heterogeneous firms approach. They suggest that productivity is not the same for all firms and that firms self-select to enter into trading relationships based on their productivity, i.e. only the most productive firms export. In the FDI field, Razin et al. (2003) also apply a Heckman correction to a gravity-type model. They assume that FDI flows are observed only when their profitability exceeds a given threshold. Hence, the selection equation is the likelihood of surpassing this threshold.

As we can see, the gravity literature is currently undecided as to the better way to deal with zeros. Moreover, there are other econometric issues that should be considered in order to select the estimation method to study bilateral FDI relationships

4.2 Heterogeneity

Heterogeneity is associated to individual characteristics that are not observable known as individual effects. Heterogeneity bias occurs when some of the included explanatory variables are correlated with the unobserved individual effects. Thus, simple estimation methods, such as pooled OLS, will produce inconsistent estimators [Wooldridge (2002)]. In our empirical specification, the presence of this heterogeneity bias is expected, i.e. individual unobservable effects that attract FDI will also be correlated with the number of firms implanted in the state. In the presence of hetero-

geneity biases, the inclusion of an individual effect (allowed to be correlated with the explanatory variables) will produce consistent estimates.

There are two kinds of specification to capture specific effects when studying bilateral FDI relationships. The first and the most common one is to introduce a dummy variable (equal to 1 or 0) that denotes the host economy and a second dummy variable for the source economy. Both of them control for the characteristics of an economy, i.e. any economy unobserved characteristic that affects its attractiveness (capacity) to receive (to invest) FDI. The second specification is to introduce a dummy variable to control for characteristics of pair of economies that could influence FDI decisions, i.e. bilateral distance or free trade agreements.

In this paper, following Mátyás (1997), the model is specified as a three-way model with FDI source country effects, FDI host state effects and time effects. Even if this specification does not allow us to estimate time-invariant economy specific characteristics, it is preferred to specification that controls for bilateral characteristics. Indeed, in the later, it is not possible to estimate time-invariant bilateral specific characteristics, such as bilateral distance which seem to be more relevant to detect FDI motivations.

4.3 Endogeneity

Concerning endogeneity issues, an explanatory variable is considered as endogenous if it is correlated with the error term. Wooldridge (2002) distinguished three different reasons for endogeneity: (1) omitted variables, (2) measurement error, and (3) simultaneity. Endogeneity by omitted variables appears when an omitted variable (usually because of data unavailability) is correlated with any explanatory variable. In other words, the omitted variable will be captured by the error term and, as a consequence, the error term will be correlated with the explanatory variable. Endogeneity by measurement error can be found when employing a proxy variable (imperfect measure) to estimate the model. Depending on the relationship between the proxy and the theoretical variable, the error term and the proxy may or may not be correlated. Finally, simultaneity is found when an explanatory variable is determined simultaneously along with the dependent variable. Thus, if an explanatory variable is a function of the dependent variable, then the dependent variable and the error term are generally correlated.

The three-way model specification (with FDI source country effects, FDI host state effects and time effects) allows us to handle omitted variables bias. The remaining endogeneity issues are tackled employing the instrumental variables (IV) methods. However, the application of IV carries a difficulty: the choice of the appropriate instruments. Indeed, the size of the IV bias is increasing in the number of instruments [Hahn and Hausman (2002)].

An appropriate instrument must satisfy two requirements: (1) it must be correlated with the included endogenous variable, and (2) orthogonal to the error process. To evaluate the validity and the relevance of the instruments a set of tests are implemented. Most of these tests relate to the explanatory power of excluded instruments in the first-stage regressions.

The first test consists to evaluate the partial R-squared of excluded instruments. An alternative test is the F-test of the excluded instruments. A rule of thumb is that for a single endogenous regressor, an F-statistic below 10 is cause of concern [Staiger and Stock (1997)]. However, in some cases, these indicators may not be sufficiently informative about instruments relevance. For example, supposing a model with two endogenous regressors and two instruments, we can find partial R-squared and F values suggesting instruments relevance when one of the excluded instruments is highly correlated with the two endogenous regressors and the second excluded instrument is just

noisy [Shea (1997)]. Shea (1997) proposes an R-squared statistic that takes into account the correlations between the instruments for multivariate models. Instruments are considered as lacking sufficient relevance to explain endogenous regressors, if an estimated equation yields a large value of the standard partial R-squared and a small value of the Shea’s partial R-squared, and the model can be considered as unidentified. These tests are robust to heteroskedasticity and individual serial correlation.

Stock and Yogo (2005) reported critical values⁵ for weak instruments that are compared to first stage Cragg and Donald (1993) statistic g_{\min} . More precisely, if Cragg and Donald (1993) statistic is higher than the Stock and Yogo (2005) critical value, then the null of weak instruments is rejected. These authors distinguish two alternative definitions of weak instruments. The first one is that instruments are weak if the bias of the IV estimator, relative to the OLS one, could exceed a given threshold in percentage (5%, 10%, 20% or 30%). The second definition is that the instruments are weak if an $\alpha = 5\%$ Wald test based on IV statistics has an actual size that could exceed a threshold $r = 10\%$. A limit of this test is that it is not robust to heteroskedasticity and results should be interpreted with caution. However, the test is reported in my paper as a complement of the aforementioned underidentification tests.

5 Empirical specification

This section presents the empirical specification taking into account the econometrical issues presented in the previous section. This specification employs Semykina and Wooldridge (2005) work about panel data models in the presence of endogeneity and selection bias. Following gravity-type models, it is assumed that FDI flows from country- i to Mexican state- j in a year $t = 1, \dots, T$ present the following linear form:

$$y_{ijtP} = x_{ijtP}\beta + u_{ijtP} \quad (1)$$

where x_{ijtP} is a $1 \times K$ vector of explanatory variables (some of which can be endogenous); β is a $K \times 1$ vector of parameters to be estimated; and u_{ijtP} is a vector that groups both the time invariant heterogeneity (c_{ijP}) and the idiosyncratic error (ε_{ijtP}) which is pair wise specific, hence $u_{ijtP} = c_{ijP} + \varepsilon_{ijtP}$; the subscript P indicates that variables belong to the primary equation. In addition, it is assumed that it exists a $1 \times L$ vector of instruments z_{ijt} . Instruments z_{ijt} can be classified into two groups: included instruments, $z1_{ijt}$; and (2) excluded instruments, $z2_{ijt}$. Included instruments are all exogenous variables included in x_{ijtP} . Excluded instruments are not present in the vector x_{ijtP} , and they should be at least as many variables as endogenous variables so ($L \geq K$) [Wooldridge (2002)].

Suppose that MNEs investment decisions are a function of expected profits. Thus, a profit function is specified to indicate the profit from a multinational from country- i investing in Mexican state- j as:

$$s_{ijtS}^* = z_{ijtS}\delta + u_{ijtS} \quad (2)$$

where z_{ijtS} is a $1 \times L$ vector of variables; δ is a $L \times 1$ vector of parameters to be estimated; u_{ijtS} represents both the time invariant heterogeneity (c_{ijS}), and the idiosyncratic error which is pair wise specific defined (ε_{ijtS}); the subscript S is employed to differentiate selection from primary equation. However, MNEs profits data are not observable, but it is expected that a MNE from country- i may invest in a Mexican state only if the expected profits are positive. Hence, a selection indicator s_{ijtS} is generated by means of a latent variable s_{ijtS}^* such that:

$$s_{ijtS} = 1[s_{ijtS}^* > 0] = 1[z_{ijtS}\delta_S + u_{ijtS} > 0] \quad (3)$$

where $1[\cdot]$ is the indicator function, so that s_{ijt2} follows an unobserved effects probit model. In other words, it equal to one if i invest in j , and zero otherwise. It is assumed that the instruments, z_{ijt} are always observed, while (y_{ijt1}, x_{ijt1}) are only observed when the selection indicator is unity. The literature usually links decisions to invest with firms' productivity and setup costs. Helpman et al. (2003) model suggest that FDI is more profitable than exports for high-productivity firms because of the higher fixed costs when opening a plant abroad. Razin et al. (2003) posit that more productive regions offer lowest relative costs of investment to explain why rich-rich FDI flows are more important than rich-poor. They also suggest that the larger the productivity gap between the source and the host economy is, the lower is also the setup costs of FDI for the advanced country investor. In other words, if i is more productive than j , it is less expensive for an investor from i to invest in j than for an investor from j to invest in i . Following these reasoning, selection equation 3 would be helpful to determine differences in productivity and in setup costs across cross-sections, which may have an influence on the volume of FDI creating a selection bias.

In the absence of selection bias, the expected value of FDI from equation 1, in the presence of both endogenous variables and heterogeneity, is: $E(y_{ijtP}|x_{ijtP}) = x_{ijtP}\beta + c_{ijP} + E(\varepsilon_{ijtP}|z_{ijtP}, c_{ij})$; where $E(\varepsilon_{ijtP}|z_{ijtP}, c_{ij}) = 0$. However, in the presence of selection bias, this equation becomes:

$$E(y_{ijtP}|x_{ijtP}, s_{ijtS} = 1) = x_{ijtP}\beta + c_{ijP} + E(\varepsilon_{ijtP}|z_{ijtP}, c_{ijP}, s_{ijtS} = 1) \quad (4)$$

Selection bias occurs when $E(\varepsilon_{ijtP}|x_{ijtP}, c_{ijP}, s_{ijtS} = 1) \neq 0$. In other words, there is a selection bias when the error terms of both the selection and the primary equation are correlated. This can be represented more clearly substituting equation 3 into equation 4:

$$E(y_{ijtP}|x_{ijtP}, s_{ijtS} = 1) = x_{ijtP}\beta + c_{ijP} + E(\varepsilon_{ijtP}|z_{ijt}, c_{ijP}, u_{ijtS}) \quad (5)$$

Note that it is assumed that selection equation includes only exogenous variables that are instruments in the primary equation. Also note that the error term of the selection equation, u_{ijtS} , includes time invariant heterogeneity, c_{ijS} , that could be correlated with z_{ijt} . To allow c_{ijS} to be correlated with z_{ij} , Semykina and Wooldridge (2005) follow Mundlak (1978) and assumes that this correlation acts only through the time averages of the exogenous variables. Hence, the unobserved effect can then be modeled as:

$$c_{ijS} = \eta_S + \bar{z}_{ij}\xi_S + a_{ijS} \quad (6)$$

where \bar{z}_{ij} is a vector of individual exogenous variables averaged across periods of time, and $a_{ijS}|z_{ij}$ has a zero mean normal distribution. Thus, the selection indicator can be rewritten as:

$$s_{ijtS} = 1[z_{ijtS}\delta_S + \eta_S + \bar{z}_{ij}\xi_{tS} + v_{ijtS} > 0] \quad (7)$$

$$v_{ijtS}|z_{ij} \sim N(0, 1), t = 1, \dots, T \quad (8)$$

where $v_{ijtS} = a_{ijS} + \varepsilon_{ijtS}$. In addition, it is supposed that $(\varepsilon_{ijtP}, v_{ijS})$ is independent of (z_{ij}, c_{ijP}) , where $v_{ijS} = (v_{ij1S}, \dots, v_{ijTS})'$, and $(\varepsilon_{ijtP}, v_{ijtS})$ is independent of $(v_{ij1S}, \dots, v_{ij,t-1,S}, v_{ij,t+1,S}, \dots, v_{ij,TS})$. If $E(\varepsilon_{ijtP}|v_{ijtS})$ is linear, then:

$$E(\varepsilon_{ijtP}|z_{ij}, c_{ijP}, v_{ijtS}) = E(\varepsilon_{ijtP}|v_{ijtS}) = \rho_P v_{ijtS} = \rho_P E(v_{ijtS}|z_{ij}, s_{ijtS}), t = 1, \dots, T \quad (9)$$

Under the previous assumptions, we can write the primary equation 1 as:

$$y_{ijtP} = x_{ijtP}\beta_P + c_{ijP} + \rho_P E(v_{ijtS}|z_{ij}, s_{ijtS}) + e_{ijtP} \quad (10)$$

where e_{ijtP} is an idiosyncratic error term verifying $E(e_{ijtP}|z_{ij}, c_{ijP}, s_{ijS}) = 0$ by construction. To estimate equation 10, we need to know $E(v_{ijtS}|z_{ij}, s_{ijtS} = 1)$ for the available observations (that is for $s_{ijtS} = 1$). This can be obtained from the usual probit estimation:

$$E(v_{ijtS}|z_{ij}, s_{ijtS} = 1) = \lambda(\eta_S + z_{ijt}\delta_S + \bar{z}_{ij}\xi_{tS}) \quad (11)$$

where $\lambda(\cdot)$ denotes the inverse Mill's ratio. Semykina and Wooldridge (2005) propose to use the probit model to estimate, for each t , the equation $\Pr(s_{ijt2} = 1|z_{ij}) = \Phi(\eta_2 + z_{ijt}\delta_2 + \bar{z}_{ij}\xi_{t2})$. Then, use the resulting estimates to compute the inverse Mills ratios as:

$$\hat{\lambda}_{ijt2} = \frac{\phi(\hat{\eta}_S + z_{ijt}\hat{\delta}_S + \bar{z}_{ij}\hat{\xi}_{tS})}{\Phi(-\hat{\eta}_S - z_{ijt}\hat{\delta}_S - \bar{z}_{ij}\hat{\xi}_{tS})} \quad (12)$$

where ϕ and Φ are the unit normal density and distribution functions respectively. Plugging the computed $\hat{\lambda}_{ijt2}$ into the primary equation yields:

$$y_{ijtP} = x_{ijtP}\beta_P + c_{ijP} + \rho_P\hat{\lambda}_{ijtS} + e_{ijtP} \quad (13)$$

Equation 13 can be estimated 2SLS controlling for heterogeneity, serial correlation and heteroskedasticity. To add more flexibility to the model, it is possible to interact $\hat{\lambda}_{itS}$ with time dummies to allow the coefficient ρ to be different across t . In this case, a Wald statistic is used to test the joint significance of the T coefficients t . This procedure is helpful to detect selection bias; however, as a correction procedure, it generally does not lead to consistent estimators. Indeed, this procedure is useful to correct for contemporaneous selection (correlation between s_{ijtS} and e_{ijtP}), but not inter-temporal selection (correlation between s_{ijrS} and e_{ijtP} for $r \neq t$). Note that for consistency, this estimator requires that $E(e_{ijtP}|z_{ij}, c_{ijP}, s_{ijS}) = 0$; thus, if the error term e_{ijtP} is correlated with the selection indicator in other time periods s_{ijr} , the estimator is biased. In addition, the assumption of independence between v_{ijS} and c_{ijP} would not be a good assumption if, as expected, v_{ijS} contains an unobserved specific effect [Semykina and Wooldridge (2005)].

5.1 Correcting for selection bias

Semykina and Wooldridge (2005) propose a procedure to correct for both contemporaneous and inter-temporal selection bias. They made the following set of assumptions that allow to obtain a valid selection correction:

Assumptions:

- i. z_{ijt} is always observed while (x_{ijtP}, y_{ijtP}) is observed when $s_{ijtS} = 1$;
- ii. Selection occurs according to equations 7 and 8;
- iii. $E(\varepsilon_{ijtP}|z_{ij}, v_{ijtS}) = E(\varepsilon_{ijtP}|v_{ijtS}) = \rho_{tP}v_{ijtS}$, $t = 1, \dots, T$;
- iv. $c_{ijP} = \eta_P + \bar{z}_{ij}\xi_P + a_{ijP}$, where $E(a_{ijP}|z_{ij}, v_{ijtS}) = E(a_{ijP}|v_{ijtS}) = \phi_{tP}v_{ijtS}$, $t = 1, \dots, T$.

In other words, assumptions i and ii formalizes the selection form; assumption iii implies that the instruments z_{ij} are independent of $(\varepsilon_{ijtP}, v_{ijtS})$; and assumption iv applies Mundlak (1978) linearity assumption about conditional expectations on c_{iP} . Under the previous assumptions about the selection rule and the unobserved effects, the primary equation of interest 1 can be rewritten as:

$$y_{ijtP} = x_{ijtP}\beta_P + \eta_P + \bar{z}_{ij}\xi_P + \gamma_{tP}E(v_{ijtS}|z_{ij}, s_{ijtS}) + e_{ijtP} \quad (14)$$

$$y_{ijtP} = x_{ijtP}\beta_P + \eta_P + \bar{z}_{ij}\xi_P + \gamma_{tP}\lambda_{ijtS} + e_{ijtP} \quad (15)$$

$$E(e_{ijtP}|z_{ij}, s_{ijts}) = 0; t = 1, \dots, T. \quad (16)$$

Note that, comparing with the test for contemporaneous selection bias presented previously, this approach allows to ignore the correlation between the idiosyncratic error, e_{ijtP} , and the selection in other time periods, s_{ijrS} for $r \neq t$.

The estimation procedure for estimating the regressors is similar to those presented in to detect the presence of selection bias. It consists to apply 2SLS estimator with a variance matrix robust to serial correlation and heteroskedasticity to equation 15 after replacing λ_{itS} by the estimated inverse Mill's ratio $\hat{\lambda}_{itS}$. Note that γ_{tP} implies different coefficients for $\hat{\lambda}_{itS}$ in each time period. Also note that we need to have at least one instrument for each endogenous variable to induce exogenous variation in the primary equation, plus at least one additional instrument to induce exogenous variation in selection. Hence, if there are M endogenous variables, then we should have at least $M + 1$ excluded instruments.

6 Variables definition

This paper uses a panel of annual data on the inbound FDI of thirty-two Mexican states by country source for the period of 1994–2004. 1994 was chosen as the starting year for two major reasons: (1) FDI regulation reforms in late 1993 make the states-data level available and comparable; (2) NAFTA implementation in 1994. The dependent variable is defined as $y_{ijt} = \ln(\text{fdi flows}_{ijt})$, where fdi flows_{ijt} are the yearly real FDI flows from country- i to Mexican state- j . FDI yearly flows are reported by the Mexican Ministry of Economy in US current dollars, this data is then transformed to 2000 constant US dollar.

As presented in the previous sections, in order to identify and to quantify the determinants of FDI location, an econometric specification needs to take into account the sample of data, the potential endogeneity of the variables, as well as the existence of unobservable bilateral characteristics which might or might not be correlated with the explanatory variables. In addition, in the presence of selection bias, two equations are required to correct for this bias: (1) the selection equation (extensive margin); and (2) the primary equation (intensive margin). In other words, MNEs first select whether to invest (selection equation), and then how much to invest (primary equation). In this line, Razin et al. (2003) and Razin et al. (2008) employ a gravity model for both equations. Following the gravity model approach, as well as previous section notation, the latent variable and the primary equation representing FDI from country- i to Mexican state j is specified as:

$$y_{ijtP} = \alpha_P + \zeta_P GDP_{it} + \vartheta_P GDP_{jt} + \varsigma_P RFE_{ijt} + \varrho_P \phi_{ij} + \beta_P x_{jt} + u_{ijtP} \quad (17)$$

$$\begin{aligned} s_{ijts} &= 1[s_{ijts}^* > 0] \\ &= 1[\alpha_S + \zeta_S GDP_{it} + \vartheta_S GDP_{jt} + \varsigma_S RFE_{ijt} + \varrho_S \phi_{ij} + \delta_S z_{jt} + u_{ijts} > 0] \end{aligned} \quad (18)$$

where GDP_{it} and GDP_{jt} are the gross domestic product of country- i and state- j , respectively at time $t = 1, \dots, T$; $RFE_{i,j,t}$ is a measure of relative factor endowment; ϕ_{ij} represents the accessibility of market j to firms from i . x_{jt} is the $K \times 1$ vector of variables influencing FDI, z_{jt} is the $L \times 1$ vector of strictly exogenous variables influencing FDI; u_{ijt} represents both the c_{ij} time invariant

heterogeneity and the idiosyncratic error (ε_{ijt}) which is pair wise specific: $u_{ijt} = c_{ij} + \varepsilon_{ijt}$. Subscripts S and P correspond to selection and primary equations, respectively. s_{ijt} is a binary variable that takes the value of 1 if MNEs from country- i invest in the state- j at time t , and zero otherwise; s_{ijt}^* represents the latent variable which can be interpreted as the MNEs benefits. Hence, it is assumed that MNEs invest in state- j if and only if MNEs are competitive enough to overtop setup costs in state- j ; or if state- j setup costs are lower enough to bring benefits to MNEs from country- i .

Using gravity models in the analysis of regional FDI distribution presumes that economy size (measured by the GDP) and distance can be considered as important FDI determinants. In principle, theoretical and empirical literatures support this view. However, there can be important differences in the mode of operation of these variables according to FDI motivation.

Concerning the market factor, two aspects must be highlighted. Firstly, according to KC models, we can distinguish horizontal FDI which is indeed market-oriented from vertical FDI which is endowment oriented. The size of market could not have any influence in the vertical FDI. However, NEG models suggest that the size of the economy matters since FDI will be attracted by potential market demand and by agglomeration incentives. Secondly, in the case of "complex" FDI models, the markets of foreign affiliates may often reach beyond the host region and extend at least to neighboring regions. Among the third region effects, NEG and export platform FDI models underline the importance of market potential beyond the host economy, especially in integration areas such as the Mexican states. Following Harris (1954) market potential formula, Crozet et al. (2004) measured the market potential for a region- j in time t as own GDP plus the distance-weighted GDP of other k regions: $MP_{jt} = GDP_{jt} + \sum_{k=1}^K (GDP_{kt} / distance_{jk})$, for $j \neq k$. Head et al. (1995) proposes a similar measure, but reduces the number of k regions to neighboring ones. To take into account third regions market effect, following Crozet et al. (2004) and Head et al. (1995), our market potential measure takes only the neighboring regions. In other words, $MP_{jt} = GDP_{jt} + \sum_{k=1}^K (GDP_{kt} / distance_{jk})$ for $j \neq k$ where state- k is a neighbor of state- j . When including third regions effects, the market size measure (GDP_{jt}) in equation 17 is replaced by the MP_{jt} one.

According to KC models, differences in factor endowments favor vertical FDI. Relative factor endowment ratio of the source and the host economies is supposed to be positive, since it determines difference in factor prices between the economies. In the absence of comparable data concerning technological, capital and skill-labor endowments, the relative factor endowment variable is constructed as the logarithm of the source/host real GDP per capita ratio: $RFE_{ijt} = \ln(GDPpc_{it} / GDPpc_{jt})$.

Following the literature of gravity models, the vector ϕ_{ij} includes two variables: (1) bilateral distance; and (2) the entry in force of a bilateral free trade agreement (FTA). An important point concerning gravity models for FDI analyses is the interpretation with regard to the theoretical basis of the distance factor and of the role of FTA. While in international trade models, trade is clearly impeded by the bilateral distance [Head and Mayer (2002)], for FDI models distance can be an impediment as well as an incentive. In one hand, it is an impediment for vertical FDI since coordination, transaction and transport costs increase with distance. On the other hand, distance can be an incentive for horizontal FDI which would aim to avoid transportation costs or to overcome other trade barriers by local production. Such investments would take place as long as the advantages of forward and backward linkages exceed the costs of operating at a distance and the costs of exporting from the home. Since FTA also diminished trade costs, a similar reasoning can be apply to this variable. Given the presented ambiguity of distance and FTA variables, it is obvious that empirical testing is confronted with severe problems. If these variables come out to be

insignificant, this does not necessarily mean that they are unimportant. Insignificance may simply result from a positive effect on some investors and a negative one on others.

FTA variable is a dummy variable which takes the value of one if the source country has a FTA agreement with Mexico in year t , and zero otherwise. Except for the distance to the US, bilateral distance is measured as the great circle formula using the latitude and the longitude of source-country and host-state capitals. In the case of distance between state- j and the US the distance is measured as the great circle distance between state's capital and the nearest border-crossing point. This was done, in order to take into account specificities of the geography of the US and Mexico that could not be captured by distances between capitals. Indeed, when measuring distance between states and the US capitals, northern border states (Baja California and Sonora) located in the northwest was considered more faraway to the US than southeastern Mexican states (Yucatan and Quintana Roo).

In addition to variables presented previously, a set of variables are included in the x_{jt} vector following theoretical and empirical approaches. Among these variables, we can distinguish two groups: (1) third region-related variables, and (2) endowment-related variables.

Recent literature suggests the importance of third regions effects on FDI [Baltagi et al. (2007); Blonigen et al. (2007)]. Indeed, FDI inflows in a given region may also be affected by FDI inflows in other locations, especially by neighboring locations. Ignoring this spatial dependence may cause biased results [Anselin (2001)]. In spatial econometrics literature, there are two classes of spatial links: (1) a spatial-lag and (2) a spatial-error. In a spatial-lag model, inward FDI to a region depends on its characteristics but it also depends on the inward FDI to its neighboring regions; this type of model is an appropriate tool for capturing the substitution or the complementarity of FDI across regions [Blonigen et al. (2007)]. In spatial-error model the disturbance term of some observations is influenced by the error term of other observations [Coughlin and Segev (2000)]. However, it is complicated to conduct a spatial econometric analysis in our sample, because of the unbalanced panel. Indeed, spatial econometric analysis requires, for any $t = 1, \dots, T$, the construction of an $N \times N$ spatial weighting matrix of known constants (W_t) to establish spatial linkages. In a balanced panel, given that location is time-invariant, $W_t = W_r$ for $r \neq t$, but in an unbalanced panel this is no longer the case [Egger et al. (2005)]. As an alternative, following Blonigen et al. (2007), it is assumed that spatial links are found in the dependent variable. It is then supposed that once a MNE decides to invest in Mexico, the location of its investment would be influenced by the presence or absence of other MNE in the neighboring regions. To capture this idea, the selection indicator ($s_{ijt} = 1$, when a country- i invest in the state- j , and zero otherwise) is employed. This selection indicator is then an $N \times 1$ vector that can be easily multiplied to an $N \times N$ spatial weighting matrix (W_t) for any t . In addition, to test robustness of spatial linkages and to take the flow values into account, an $N \times 1$ vector is constructed by as: $\ln(fdi\ flows_{ijt} + 1)$. The inclusion of these variables carries some problems of endogeneity. Indeed, both of them can be considered as proxies of spatially lagged dependent variable which are typically correlated with the error term [Anselin (2001)]. To correct this bias spatial the econometrics literature proposes to employ maximum likelihood (ML) or instrumental variables (IV) estimation. An IV approach advantage over ML estimation is that it does not require an assumption of normality and it avoids the computational problems associated with ML for large data sets [Anselin (2001)]. Kelejian and Prucha (1998) propose to employ as instruments spatial lags of exogenous variables. Following these authors the first spatial lags of exogenous variables related to Mexican states (Wz_j) are employed. In addition a quasi-instrument is constructed for each spatial endogenous variable following the three groups' method proposed by Kennedy (2003). These quasi-instruments take the values of -1,

0 or 1 depending on whether or not the endogenous variable is in the upper, middle and lower third when placed in a rank order, respectively.

To conduct the spatial analysis the construction of a weighting matrix is needed. For any year $t = 1, \dots, T$, a weight matrix W_t is defined as:

$$W_t = \begin{bmatrix} 0 & w_{jk} & w_{jl} \\ w_{kj} & 0 & w_{kl} \\ w_{lj} & w_{lk} & 0 \end{bmatrix} \quad (19)$$

where $w_{jk} = (1/\text{distance}_{jk})$ if the states j and k are neighbors, and zero otherwise. Matrix W_t is standardized, so that the sum of every row of the matrix is one.

As endowment-related variables wages, agglomeration forces and available infrastructure in the host economy are included. Wages are usually employed by the literature as a measure of production costs. Thus, if the MNEs motivation is from the vertical type, it would be expected that wages are negatively related to FDI. However, KC models also suggest that MNEs require a certain level of skilled workers in order to produce; if wages reflects the availability of skilled workers, wages would have a positive influence attracting FDI. Empirical literature has found both positive [Head et al. (1995)] and negative [Crozet et al. (2004); Coughlin et al. (1991)] relations between wages and FDI. Havrylchyk and Poncet (2007) point out the fact that endogeneity bias could lead wages variables to have the wrong sign. After controlling for wages endogeneity, these authors find that wages coefficient becomes negative, while without endogeneity correction they found a positive influence. Following Havrylchyk and Poncet (2007), wages are controlled for endogeneity using the lagged time wage values as instrument.

As mentioned in previous sections NEG literature highlights the importance of agglomeration economies as a determinant factor in the location decision of a firm. Empirical literature usually finds a positive effect of agglomeration on FDI. However, as suggested by Head et al. (1995) and Head and Mayer (2004), these results can be due to the omission of relevant location variables rather than to agglomeration forces. In addition, there is not a consensus, because of data availability, about how to measure agglomeration. For example, the literature employs measures of population, industrial, capital and firms concentration [Coughlin et al. (1991); Crozet et al. (2004); Havrylchyk and Poncet (2007); He (2002); among others]. Head et al. (1995) suggest to include different measures of agglomeration according to MNEs industry and nationality. They show that Japanese firms locate near to other Japanese firms in the United State; moreover, this agglomeration extends to US firms. Applying a similar methodology to the French case, Crozet et al. (2004) found that MNEs are attracted by the presence of both French and MNEs, being the proximity to French firms the dominant factor. For the Mexican case, Mollick et al. (2006) employ the manufacturing activity as share of state GDP as a measure of agglomeration. They do not find robust results concerning this variable, moreover it can be considered as an industrialization measure rather than agglomeration one. To control for agglomeration in the Mexican, the ratio number of industrial firms installed before the year of FDI in the state to state's surface is employed. This variable has the disadvantage that firms cannot be classified by nationality, but to the author's knowledge is the only yearly data available.

The selection indicator of FDI from the same country to neighboring states can control for same country firm concentration. In addition, FDI from other countries to the state and its neighbors is included to capture agglomeration. The idea is that locations receiving FDI from a high number of countries may be a signal that the location is a good one.

Improvements in infrastructure can be seen as a reduction in transport costs. From a NEG

perspective, this reduction in transport costs benefits agglomeration in rich regions, in spite of poor regions. Performances in infrastructure could create a negative effect for poor regions as FDI will be located in regions offering economies of scale. Since telephone lines facilitate international communications, following Martin and Rogers (1995), Mollick et al. (2006) consider this variable as a measure for international infrastructure. Thus, improvements in international infrastructure would reduce international trade costs, providing a positive incentive for vertical FDI motivation and a negative one for horizontal FDI. Telephone line density is used as a first measure of infrastructure. Mollick et al. (2006) employs the square kilometers of the state's routes as a secondary measure of infrastructure. Since road network facilitates domestic trade, these authors posit this variable measure the domestic infrastructure. Following Mollick et al. (2006), a variable to measure domestic infrastructure is constructed as: $road = (Km^2 \text{ paved roads}/Area)$.

Finally, as a proxy of investment environment or local governance, given the lack of other data for the period of study, the delinquency rate is employed. This variable is supposed to play a negative role attracting FDI for any FDI motivation.

Table 1 shows the summary statistics and appendix A presents a description and the sources of the dataset. As we can notice, non zero FDI flows only represents 17.8% of total observations. In other words, most of the countries investing in Mexico select a state or a group of states to place their operations or invest only in given years. The appendix B shows the times that each country has invested in each Mexican state, being 11 the maximum possible. Once again, we can see that the most active country is the US, but even this country does not invest in all states for all years. In addition, comparing to the flows level, the weight of the US as a source economy is considerably lower. The same can be said concerning the Federal District as the host of FDI.

Concerning the size of the economy or GDP levels, table 1 shows that source countries are in average bigger than host ones. However, this does not mean that source economies are necessarily bigger, in average, than Mexico as a country. We can also see that differences among source countries are bigger than differences between Mexican states as the standard deviation statistics suggest. Source countries have also, in average, better relative factor endowments than Mexican states. Given the RFE measure taken, this suggests that they are more developed than Mexico's states; but the differences in RFE are also very important among cross-sections.

7 Empirical analysis

In order to evaluate the importance of third regions influence, in a first time these effects are omitted into the analysis and results are compared with those including these effects.

We have seen that the econometrical issues concerning FDI regression analysis were endogeneity, heterogeneity and selection bias. Thus, first of all, endogeneity among regression variables is looked for. A set of regressions were then conducted to test for exogeneity of included regressors by the following form. A 2SLS regression (controlling for time, FDI source country and FDI host state specific effects to avoid omitted variables bias) employing the time lagged of the variable tested as well as the other exogenous variables as instruments. Then regressors' endogeneity is tested using the Wu-Hausman and the Durbin-Wu-Hausman tests under the null of exogeneity. Only the tests for the wage variable rejected the null of exogeneity. A second set of 2SLS regressions were then conducted in similar way, but controlling for endogeneity in wage variable, and as reported in table 2 no variable rejects the null of exogeneity. Havrylchuk and Poncet (2007) suggest that endogeneity in wages is due to the simultaneity between FDI and wages and not controlling for this conduct to contradictory results. For the Mexican case, the literature has pointed out the relationship between

Table 1: Descriptive Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
FDI_{ijt}	4255	12.549	3.674	-3.404	23.552
S_{ijt}	23936	0.178	0.382	0.000	1.000
GDP_{it}	23936	25.151	1.801	20.720	29.994
GDP_{jt}	23936	23.114	0.855	21.569	25.584
RFE_{ijt}	23936	0.023	1.505	-5.190	3.051
$Wage_{jt}$	23936	2.444	0.172	2.026	3.101
$Agglo_{jt}$	23936	-2.813	1.199	-5.219	1.086
$Phone_{jt}$	23936	2.261	0.521	0.935	3.687
$Road_{jt}$	23936	-1.633	0.642	-3.009	-0.414
$Delinquency_{jt}$	23936	0.713	0.426	-0.245	1.595
$Distance_{ij}$	23936	8.963	0.659	5.899	9.791
FTA_{ijt}	23936	0.274	0.446	0.000	1.000
MP_{jt}	23936	23.166	0.828	21.654	25.593
$W s_{ijt}$	23936	0.177	0.271	0.000	1.000
$W (1+FDI_{ijt})$	23936	2.247	3.762	0.000	20.251
$W Agglo_{jt}$	23936	-2.825	0.778	-4.429	-0.853
$W Phone_{jt}$	23936	2.233	0.366	1.251	3.028
$W Road_{jt}$	23936	-1.591	0.445	-2.562	-0.538
$W Delinquency_{jt}$	23936	0.630	0.295	-0.082	1.419
$Q-I W S_{ijt}$	23936	0.002	0.671	-1.000	1.000
$Q-I W (1+FDI_{ijt})$	23936	0.002	0.671	-1.000	1.000

Table 2: Regressors endogeneity tests

H0: Regressor is exogenous				
Variable	Wu-Hausman		Durbin-Wu-Hausman	
	$F(1.382)$	P-value	$\chi^2(1)$	P-value
GDP_{it}	0.679	0.410	0.700	0.403
GDP_{jt}	1.049	0.306	1.081	0.298
RFE_{ijt}	0.006	0.939	0.006	0.938
$Wage_{jt}$	4.015	0.045	4.135	0.042
$Agglo_{jt}$	0.216	0.642	0.223	0.637
$Phone_{jt}$	0.825	0.364	0.850	0.356
$Road_{jt}$	0.143	0.705	0.147	0.701
$Delinquency_{jt}$	0.049	0.825	0.051	0.822

Table notes:

Tests obtained after a regression including time, source country and host state specific effects and controlling for accepted endogenous regressors. Excluded instrument employed was the time lagged of the variable tested.

FDI and wages. Dussel Peters (2000), Feenstra and Hanson (1997) and Hanson (1998a) observed that in the regions where FDI was most concentrated experienced a higher increase in total wages. Thus, controlling for wages endogeneity is the first measure to avoid biases.

Table 3 reports regression results for the specification without third region's influence. A set of different 2SLS regressions were conducted. Wage variable is controlled for endogeneity employing as excluded instrument its time lagged variable. The regression, in column (1), does not control for country neither state specific effects. We can see that most of the regressors are significant and most of them have the expected sign. However, these results can be biased if the remaining econometrical issues (heterogeneity and selection biases) are present.

Firstly, specification is controlled by heterogeneity, as shown in column (2) all variables lose their significance. However results suggest source and host heterogeneity, since the Wald test of joint significance for both source country and host state specific rejects the null of non significance. These results drive us to reject results on specification (1).

The remaining econometrical issue is the abundance of zero FDI flows. As shown previously, most of the countries do not invest in all regions, this fact can hide a relative unobservable advantage of some countries (states) to invest (to receive) FDI flows that needs to be taken into account. The selection bias test proposed by Semykina and Wooldridge (2005) is then applied to the model. In a first time, a probit regression is conducted for each time period and the inverse Mill's ratio ($\widehat{\lambda}_{ijt}$) is then computed from the results. However, an additional instrument, not included into the primary equation is needed. Following Razin et al. (2008), the time lagged of the selection indicator ($s_{ij,t-1}$) is included as instrument. These authors suggest that this instrument reflects lowest investment barriers for investor country when different to zero. Indeed, unobserved linkages could be created between host and source economies due to previous FDI. Once the $\widehat{\lambda}_{ijt}$ is computed, it is added to the model. In addition, an interaction between $\widehat{\lambda}_{it}$ and time dummies is permitted to allow correlation between selection and primary equation to be different across t . The model is then estimated by 2SLS controlling for time and economies specific effects. Results are presented in table 3 column (3). Remember that column (3) regression is helpful to detect selection bias, but

not to correct it. Regression results suggest the presence of selection bias, since $\hat{\lambda}_{ijt}$ is significant, as well as the joint significance of its interaction with time dummies.

The next step is then to correct for selection bias. Following the procedure proposed by Semykina and Wooldridge (2005), regression is conducted including $\hat{\lambda}_{ijt}$ and its interaction with time dummies, and replacing specific effects by the Mundlak (1978) approximation. Results are presented in table 3 column (4). We can see that, after correcting for selection bias, bilateral distance, free trade agreements, wages and $\hat{\lambda}_{ijt}$ are significant. Results suggest a negative relationship between distance and FDI, which is a common result on gravity-type models (see Head and Ries (2008); Razin et al. (2008); results among others). Negative correlation of distance can then be interpreted as an impediment for vertical FDI since trade costs increase with distance. This result is supported by the positive and significant role played by the FTA. Wage variable is significant and positively related to FDI flows. This can be interpreted in two different ways. First, wages can reflect the availability of skilled workers, since MNEs need skilled workers to produce [Markusen (2002)] high wages states attract more FDI. Another interpretation on wages can be that it reflects the purchasing power of the states, which can influence market seeking FDI. Concerning $\hat{\lambda}_{ijt}$ a negative and significant relation to FDI is found, and its coefficient value is quite similar to Razin et al. (2003) one. Following Razin et al. (2003) and Razin et al. (2008), $\hat{\lambda}_{ijt}$ is interpreted as a coefficient that accounts for differences in unobserved productivity across FDI source countries and in unobserved setup costs between Mexican states. Given the estimation results, this variable can be interpreted as an inverse of competitiveness of the source country, as well as an indicator of setup costs in Mexico's states. We can thus expect that obtained values of $\hat{\lambda}_{ijt}$ by probit estimations reflect this fact. Table 4 shows the average $\hat{\lambda}_{ijt}$ for the 32 Mexican states and for the 32 countries that reported the lowest average. In addition to $\hat{\lambda}_{ijt}$ values, the number of positive FDI flows as well as the share over total FDI is reported. The first four columns present the country values. We can see that the five most competitive countries are the US, Canada, Germany, Japan and the United Kingdom. However among this group of countries Japan and the UK do not belong to the group of five most active FDI countries in Mexico. The five Mexican states offering the lowest setup costs are the Federal District, Jalisco, Quintana Roo, Baja California and Nuevo Leon. Note that in this group, only two states share a common border with the US. Moreover, other Northern border states rank 19 and 20, which suggests that location near to the US is not a guaranty of lower setup costs.

Neighboring influences are added to the model to control for third region influences. Two variables are added to the model to control for these factors: (1) W_{sijt} , and (2) the MP_{jt} variable. The first variable would be helpful to detect complementarity or competition among Mexican states to attract FDI, while the second one would be helpful to add market influences. The same steps as for the previous specification were followed and results are presented in table 5. MP variable is significant when not controlling for specific effects, but once these effects are introduced as well as the selection bias correction, it loses its significativity. On the other hand, $W S$ is significant when introducing specific effects and when correcting for selection bias. Thus, if a country invests in neighboring states, the FDI flows received by the state would increase. Blonigen et al. (2007) also find positive spatial linkages among OECD countries. These authors interpret this result as an indicator of FDI vertical specialization.

Wage, bilateral distance and $\hat{\lambda}_{ijt}$ continue to be significant, the former positive, while the two later are negative. However, FTA agreement lost its significativity. This result shows that omitting spatial linkages could lead us to wrong conclusions.

An alternative measure of FDI activity in neighboring states is included to the model and reported results (see table 6) are pretty similar to those obtained with the previous specification. In

Table 3: Regression results

Dependent variable: FDI_{ijt}				
	(1)	(2)	(3)	(4)
GDP_{it}	0.706 (0.057)**	0.027 (1.982)	-0.734 (1.972)	-0.119 (2.182)
GDP_{jt}	0.589 (0.149)**	1.816 (1.661)	2.484 (1.647)	1.895 (1.743)
RFE_{ijt}	0.723 (0.099)**	0.433 (1.880)	0.889 (1.865)	-0.072 (2.067)
$Distance_{ij}$	-1.142 (0.164)**	-1.036 (0.675)	-0.601 (0.657)	-0.901 (0.178)**
RTA_{ijt}	0.344 (0.177)	0.249 (0.182)	0.232 (0.179)	0.359 (0.179)*
$Wage_{jt}$	2.745 (0.786)**	2.506 (1.441)	2.164 (1.418)	2.888 (1.439)*
$Agglo_{jt}$	-0.001 (0.122)	-1.059 (0.925)	-1.393 (0.911)	-1.417 (0.958)
$Phone_{jt}$	1.012 (0.358)**	-0.430 (1.063)	-0.937 (1.035)	-0.863 (1.056)
$Road_{jt}$	-0.203 (0.185)	0.135 (0.497)	-0.011 (0.504)	0.055 (0.521)
$Delinquency_{jt}$	-0.253 (0.282)	-0.261 (0.451)	-0.141 (0.452)	-0.163 (0.453)
$\widehat{\lambda}_{ijt}$			-0.832 (0.242)**	-0.840 (0.276)**
Constant	-20.382 (4.072)**	-29.395 (44.165)	-26.151 (44.141)	-20.919 (5.052)**
Observations	3936	3936	3936	3936
F-statistic	43.28**	19.16**	19.77**	25.20**
R^2	0.357	0.459	0.468	0.370
Country and State specific effects	No	Yes	Yes	Mundlak
Time fixed effects	Yes	Yes	Yes	Yes
Wald Test	63.45**	53.88** ^s	42.33**	37.25**
		483.41** ^c		

Table notes:

* significant at 5%; ** significant at 1%. Time, country and state specific effects not reported. Mundlak (1978) transformation was applied to specific effects to correct for selection bias in regression (4). Variables on the interaction of $\widehat{\lambda}_{ijt}$ with time dummies are also included in regressions (3) and (4) but not reported.

$\widehat{\lambda}_{ijt}$ values (inverse Mills ratios) employed in regressions (3) and (4) were obtained, as suggested by Semykina and Wooldridge (2005), from a probit estimation for each period. Wald test evaluates the joint significance of: time heterogeneity in regression (1); host state (^s) and source country (^c) specific effects in regression (2); and $\widehat{\lambda}_{ijt}$ interaction with time dummies in columns (3) and (4).

Table 4: Average lambda values

Country	lambda	obs.	% FDI	Rank	State	lambda	obs.	% FDI
USA	0,131	329	60,81%	1	DF	0,699	613	60,76
CAN	0,543	253	2,93%	2	JAL	1,281	348	2,51
DEU	0,731	221	2,90%	3	QTR	1,297	334	0,40
JPN	0,784	136	2,12%	4	BCN	1,309	312	4,77
GBR	0,837	184	3,39%	5	NLN	1,327	244	8,95
FRA	0,867	188	1,37%	6	MEX	1,524	324	6,06
ITA	0,918	195	0,26%	7	BCS	1,577	215	0,56
ESP	0,955	235	10,93%	8	COA	1,666	128	1,06
NLD	1,094	184	9,15%	9	GTO	1,738	159	0,60
CHE	1,258	140	1,64%	10	YUC	1,756	169	0,26
KOR	1,316	142	0,45%	11	PUE	1,758	158	1,89
BRA	1,375	92	0,10%	12	GRO	1,822	140	0,15
ARG	1,392	125	0,04%	13	MOR	1,843	84	0,50
BEL	1,420	86	0,32%	14	VER	1,877	143	0,31
SWE	1,438	65	0,78%	15	MIC	1,878	75	0,11
DNK	1,469	88	0,72%	16	TAM	1,884	144	2,35
COL	1,488	102	0,06%	17	QRO	1,907	150	0,82
CHN	1,506	99	0,06%	18	CHI	1,937	170	4,14
AUS	1,512	88	0,03%	19	SON	1,978	163	1,23
AUT	1,526	67	0,02%	20	CHS	1,990	38	0,02
VEN	1,563	53	0,03%	21	HGO	2,056	46	0,14
NOR	1,578	27	0,04%	22	SIN	2,075	88	0,25
GTM	1,599	52	0,01%	23	AGS	2,081	77	0,45
IRL	1,620	69	0,11%	24	SLP	2,098	80	0,68
FIN	1,673	35	0,32%	25	NAY	2,180	105	0,18
SLV	1,740	23	0,00%	26	DGO	2,234	47	0,26
PRT	1,750	38	0,01%	27	OAX	2,252	87	0,01
CRI	1,781	44	0,02%	28	COL	2,347	60	0,10
HKG	1,786	48	0,02%	29	TLA	2,417	47	0,12
CHL	1,792	71	0,09%	30	CAM	2,480	61	0,09
PAN	1,794	96	0,16%	31	ZAC	2,590	29	0,06
POL	1,818	20	0,00%	32	TAB	2,771	37	0,20

Table notes:

Lambda represents the average of the inverse Mill's ratio computed from probit regression; obs. indicates the number of positive investments; and % FDI refers to the share of total FDI for the period 1994-2004.

Lambda values computed for the 68 FDI source countries in the sample, but only reported the first 32 ranked from low to high lambda.

Table 5: Regression results

Dependent variable: FDI_{ijt}				
	(1)	(2)	(3)	(4)
W_{sijt}	0.496 (0.305)	0.788 (0.277)**	0.958 (0.280)**	0.777 (0.312)*
GDP_{it}	0.655 (0.066)**	-0.192 (1.982)	-0.909 (1.974)	-0.253 (2.181)
MP_{jt}	0.623 (0.158)**	1.835 (1.714)	2.591 (1.707)	2.061 (1.787)
RFE_{ijt}	0.697 (0.099)**	0.498 (1.875)	0.944 (1.862)	-0.008 (2.057)
$Distance_{ij}$	-1.075 (0.167)**	-0.916 (0.670)	-0.518 (0.658)	-0.844 (0.182)**
RTA_{ijt}	0.310 (0.177)	0.241 (0.182)	0.214 (0.179)	0.305 (0.179)
$Wage_{jt}$	2.566 (0.783)**	2.733 (1.381)*	2.367 (1.367)	2.912 (1.375)*
$Agglo_{jt}$	-0.022 (0.124)	-1.039 (0.930)	-1.383 (0.917)	-1.408 (0.950)
$Phone_{jt}$	1.037 (0.355)**	-0.604 (1.063)	-1.081 (1.036)	-0.966 (1.052)
$Road_{jt}$	-0.224 (0.181)	0.134 (0.500)	0.003 (0.508)	0.053 (0.528)
$Delinquency_{jt}$	-0.281 (0.278)	-0.241 (0.453)	-0.126 (0.456)	-0.150 (0.457)
$\widehat{\lambda}_{ijt}$			-0.739 (0.240)**	-0.681 (0.272)*
Constant	-20.313 (4.162)**	-25.601 (44.220)	-25.270 (44.311)	-21.063 (5.989)**
Observations	3936	3936	3936	3936
F-statistic	42.81**	19.20**	19.76**	23.44**
R^2	0.364	0.461	0.468	0.378
Country and State specific effects	No	Yes	Yes	Mundlak
Time fixed effects	Yes	Yes	Yes	Yes
Wald Test	65.36**	58.27** ^s 399.07** ^c	36.84**	31.04**

Table notes:

* significant at 5%; ** significant at 1%. Time, country and state specific effects not reported. Mundlak (1978) transformation was applied to specific effects to correct for selection bias in regression (4). Variables on the interaction of $\widehat{\lambda}_{ijt}$ with time dummies are also included in regressions (3) and (4) but not reported.

$\widehat{\lambda}_{ijt}$ values (inverse Mills ratios) employed in regressions (3) and (4) were obtained, as suggested by Semykina and Wooldridge (2005), from a probit estimation for each period. Wald test evaluates the joint significance of: time heterogeneity in regression (1); host state (^s) and source country (^c) specific effects in regression (2); and $\widehat{\lambda}_{ijt}$ interaction with time dummies in columns (3) and (4).

other words significant positive FDI determinants are average wage and FDI activity in neighboring states, while a negative and significant relationship between distance and FDI flows is found. Note that for the three specifications, once we controlled for selection bias, the coefficient value of distance variable decreases. In other words, once including unobserved setup costs into the model the importance of trade costs diminishes. Chaney (2008) and Helpman et al. (2008) suggest that not taking differences in competitiveness across countries into account lead to overestimate the role of distance in the international trade gravity-type models. Thus, it seems that this observation is also applied to the FDI case for the Mexican regions. However, this conclusion cannot be extended to all FDI cases since Razin et al. (2003) reports the opposite results.

Among the significant variables, wages and FDI activity in neighboring regions are endogenous. Thus, the first-stage regressions could be helpful to deepen the analysis on FDI determinants. In other words, variables such as market potential and agglomeration can have some significant influence on wages and affect FDI location through this channel. In addition, looking at this stage results, we can also evaluate the validity of excluded instruments. First-stage regressions and instruments validity tests are presented into the appendix C. These results only concerns regressions that correct for selection bias for each one of the three previous specifications. Results suggest that less developed states (in terms of GDP per capita) pay lower wages. A negative and significant correlation between wages and the delinquency rate is also found. The size of the market is significant and positively related to wages when it is measured as the GDP of the host region as well as by the MP measure. This is a current result found by the literature studying the wages differences [?]. For the Mexican case, Hanson (1998b) shows that Mexican states benefiting from better market potential pay higher wages.

We can see that results presented in this study do not agree with Mollick et al. (2006) who suggest that telephone density is one of the main determinants of FDI distribution in Mexico. However these authors do not control for differences in wages across Mexican states. From first-stage results, we can also observe that telephone density influence FDI through wages is significant. Moreover this influence is more important when the phone density in neighboring states is added to the model.

FDI activity in neighboring regions is determined by the size and the development of the source country, as well as by the distance and the presence of a FTA between the source country and Mexico. We can thus infer that FDI activity in neighboring states is also an indicator of source country capacity and ease to export FDI to Mexico. Concerning exogenous spatial lagged variables employed as instruments, it is surprising to see that they have more impact on wages than on FDI in neighboring regions. Moreover, the negative impact of spatial lagged telephone density on FDI activity in neighboring regions constitutes a puzzle to interpretation. Finally, we can see that quasi-instruments are strongly significant to spatial lagged FDI activity. To the validity and the relevance of these instruments a set of tests presented in chapter 1 are implemented. We can see, also in appendix C that there is no a significant difference between the Shea (1997) and the partial R-squared of excluded instruments, as well as an F statistic always superior to 10. Finally, the Cragg and Donald (1993) statistic is always higher than the Stock and Yogo (2005) critical value, then the null of weak instruments is rejected. Finally, the independency of the excluded instruments from the error is verified by the Hansen J statistic, since the null of independency is hardly rejected. We can then conclude that the excluded instruments are valid and relevant according to the tests presented.

One cause of concern is the large fractions of zeros in the data. As mentioned before this fraction represents 82.2% of the total observations, while for Helpman et al. (2008) and Razin et al. (2003)

Table 6: Regression results

Dependent variable: FDI_{ijt}				
	(1)	(2)	(3)	(4)
W ($1+FDI_{ijt}$)	0.051 (0.024)*	0.084 (0.022)**	0.097 (0.022)**	0.084 (0.025)**
GDP_{it}	0.620 (0.066)**	-0.092 (1.987)	-0.798 (1.980)	-0.161 (2.170)
MP_{jt}	0.631 (0.157)**	1.754 (1.712)	2.516 (1.704)	2.013 (1.773)
RFE_{ijt}	0.679 (0.099)**	0.351 (1.878)	0.784 (1.866)	-0.118 (2.042)
$Distance_{ij}$	-1.012 (0.168)**	-0.883 (0.665)	-0.474 (0.651)	-0.724 (0.180)**
RTA_{ijt}	0.282 (0.176)	0.225 (0.182)	0.196 (0.179)	0.263 (0.177)
$Wage_{jt}$	2.478 (0.776)**	2.738 (1.373)*	2.342 (1.357)	2.890 (1.363)*
$Agglo_{jt}$	-0.029 (0.122)	-0.996 (0.931)	-1.340 (0.918)	-1.381 (0.947)
$Phone_{jt}$	1.032 (0.352)**	-0.628 (1.063)	-1.122 (1.034)	-1.013 (1.050)
$Road_{jt}$	-0.232 (0.179)	0.137 (0.499)	0.005 (0.507)	0.052 (0.524)
$Delinquency_{jt}$	-0.291 (0.275)	-0.184 (0.454)	-0.062 (0.457)	-0.090 (0.459)
$\widehat{\lambda}_{ijt}$			-0.742 (0.239)**	-0.690 (0.267)**
Constant	-19.989 (4.097)**	-26.614 (44.178)	-26.696 (44.273)	-20.505 (5.842)**
Observations	3936	3936	3936	3936
F-statistic	43.64**	19.80**	19.83**	24.26**
R^2	0.370	0.463	0.471	0.387
Country and State specific effects	No	Yes	Yes	Mundlak
Time fixed effects	Yes	Yes	Yes	Yes
Wald Test	62.89**	61.96** ^s 339.92** ^c	38.62**	32.99**

Table notes:

* significant at 5%; ** significant at 1%. Time, country and state specific effects not reported. Mundlak (1978) transformation was applied to specific effects to correct for selection bias in regression (4). Variables on the interaction of $\widehat{\lambda}_{ijt}$ with time dummies are also included in regressions (3) and (4) but not reported.

$\widehat{\lambda}_{ijt}$ values (inverse Mills ratios) employed in regressions (3) and (4) were obtained, as suggested by Semykina and Wooldridge (2005), from a probit estimation for each period. Wald test evaluates the joint significance of: time heterogeneity in regression (1); host state (^s) and source country (^c) specific effects in regression (2); and $\widehat{\lambda}_{ijt}$ interaction with time dummies in columns (3) and (4).

it represents around 50%, and for Razin et al. (2008) 62%. Even if there is no threshold to fraction of zeros in the data⁶, to test the robustness of previous results the zero observations were reduced in the sample. More precisely, only countries investing at least 88 times (which represents 25% of the maximum possible) are included into the sample. This reduces the number of source countries to 18 and the fraction of zeros to 54.3%. Descriptive statistics of new data set is presented into table 7.

Comparing to data summary presented in table 1, we can see that these countries not only invest more times than the excluded ones, they also invest, in average, larger volumes. We can also see that, in average, they are bigger in terms of GDP, they are more developed (see RFE variable), they are closer to Mexican states and they have more FTA in force with Mexico.

Table 8 presents the results for the specification not employing third regions variables. As for the case of previous sample, when taken into account country and states specific effects, regression (2), many variables significant in (1) become non significant. However in this case, bilateral distance, wage and agglomeration variables are significant at 95%. Selection bias test, reported in column (3) suggest the presence of selection bias, while estimators values and significance remain quite similar to those reported in column (2). After correcting for selection bias, column (4), we can see that bilateral distance, wage and agglomeration variables are still significantly different to zero. Comparing to previous sample results, bilateral distance and wage effects on FDI remain similar, but agglomeration becomes significant and negatively related to FDI. We can interpret this result as dominance of centrifugal forces over centripetal ones. At this subject, Hanson (1998a) found a rapid growth of manufacturing employment in Northern Mexico and a contraction of this activity in the Federal District manufacturing belt. In addition, northern states are territorially bigger than south ones, which lead to an agglomeration variable lower than those for southern states.

Finally, the spatial effects are introduced to the model. The first three regressions results reported in table 9 concern the specification where FDI activity in neighboring regions is measured as the number of FDI investments from host economy in neighboring states, while the last three draws results for specification including FDI flows in neighboring region. The 2SLS results controlling for specific effects, , reported in column (1) and (4), are quite similar to those presented for the same specification of whole sample (see table 5.) In other words, bilateral distance and wages are significantly correlated to FDI, as well as FDI in neighboring states. Note that in the case of whole sample results, controlling for FDI in neighboring regions leads to a lost in significance for FTA variable, and in the case of reduced sample is the variable agglomeration that becomes non significant at 95%. Selection bias tests, columns (2) and (4), suggest the presence of this bias for both specifications. And columns (3) and (6) illustrate the regression results corrected for selection bias. Once again, reported results are similar to those presented when using the whole sample.

8 Conclusions

Mexico is a country characterized by strong regional differences to attract FDI. This paper studies the determinants of FDI flows to Mexican states, taking into account the characteristics of source countries and host Mexican states. To conduct this study, a specification based on gravity models and previous literature is followed. Indeed, the literature agrees that gravity-type approach is a good specification to model MNEs activity, since it can be easily linked to MNEs theory Kleinert and Toubal (2010)]. However, this gravity-type specification carries some econometrical issues such as selection bias, endogeneity and heterogeneity, which are even more important in the case of

Table 7: Alternative sample descriptive statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
FDI_{ijt}	2897	13.314	3.602	-3.404	23.552
S_{ijt}	6336	0.457	0.498	0.000	1.000
GDP_{it}	6336	27.046	1.457	22.932	29.994
GDP_{jt}	6336	23.114	0.855	21.569	25.584
RFE_{ijt}	6336	0.975	1.109	-3.062	2.814
$Wage_{jt}$	6336	2.444	0.172	2.026	3.101
$Agglo_{jt}$	6336	-2.813	1.199	-5.219	1.086
$Phone_{jt}$	6336	2.261	0.521	0.935	3.687
$Road_{jt}$	6336	-1.633	0.642	-3.009	-0.414
$Delinquency_{jt}$	6336	0.713	0.426	-0.245	1.595
$Distance_{ij}$	6336	8.922	0.511	7.260	9.555
FTA_{ijt}	6336	.3586	0.480	0.000	1.000
MP_{jt}	6336	23.166	0.828	21.654	25.593
$W s_{ijt}$	6336	0.460	0.314	0.000	1.000
$W (1+FDI_{ijt})$	6336	6.179	4.785	0.000	20.251
$School\ years_{jt}$	6336	1.981	0.137	1.540	2.303
$W Agglo_{jt}$	6336	-2.825	0.778	-4.429	-0.853
$W Phone_{jt}$	6336	2.233	0.366	1.251	3.028
$W Road_{jt}$	6336	-1.591	0.445	-2.562	-0.538
$W Delinquency_{jt}$	6336	0.630	0.295	-0.082	1.419
$Q-I\ W\ S_{ijt}$	6336	-0.009	0.834	-1.000	1.000
$Q-I\ W\ (1+FDI_{ijt})$	6336	-0.009	0.834	-1.000	1.000

Table 8: Regression results

Dependent variable: FDI_{ijt}				
	(1)	(2)	(3)	(4)
GDP_{it}	0.594 (0.076)**	-2.675 (2.447)	-3.674 (2.477)	-3.718 (2.554)
GDP_{jt}	0.631 (0.175)**	2.180 (1.800)	2.849 (1.822)	2.759 (1.850)
RFE_{ijt}	0.876 (0.146)**	2.426 (2.316)	2.960 (2.327)	2.972 (2.388)
$Distance_{ij}$	-1.682 (0.180)**	-2.713 (1.075)*	-2.128 (1.078)*	-1.435 (0.197)**
RTA_{ijt}	0.471 (0.205)*	0.399 (0.213)	0.292 (0.211)	0.282 (0.202)
$Wage_{jt}$	3.564 (0.916)**	3.425 (1.629)*	3.299 (1.623)*	3.371 (1.613)*
$Agglo_{jt}$	0.080 (0.142)	-2.092 (1.031)*	-2.136 (1.032)*	-2.110 (1.047)*
$Phone_{jt}$	1.782 (0.384)**	-0.017 (1.197)	-0.334 (1.203)	-0.198 (1.273)
$Road_{jt}$	-0.305 (0.213)	0.829 (0.510)	0.687 (0.523)	0.599 (0.555)
$Delinquency_{jt}$	-0.433 (0.342)	-0.571 (0.509)	-0.493 (0.501)	-0.674 (0.511)
$\widehat{\lambda}_{ijt}$			-0.923 (0.327)**	-1.320 (0.327)**
Constant	-16.994 (4.881)**	44.287 (53.316)	51.958 (54.007)	-17.820 (5.738)**
Observations	2666	2666	2666	2666
F-statistic	44.67**	25.80**	23.68**	28.70**
R^2	0.414	0.492	0.500	0.432
Country and State specific effects	No	Yes	Yes	Mundlak
Time fixed effects	Yes	Yes	Yes	Yes
Wald Test	54.65**	72.92** ^s	26.22**	42.15**
		121.85** ^c		

Table notes:

* significant at 5%; ** significant at 1%. Time, country and state specific effects not reported. Mundlak (1978) transformation was applied to specific effects to correct for selection bias in regression (4). Variables on the interaction of $\widehat{\lambda}_{ijt}$ with time dummies are also included in regressions (3) and (4) but not reported.

$\widehat{\lambda}_{ijt}$ values (inverse Mills ratios) employed in regressions (3) and (4) were obtained, as suggested by Semykina and Wooldridge (2005), from a probit estimation for each period. Wald test evaluates the joint significance of: time heterogeneity in regression (1); host state (^s) and source country (^c) specific effects in regression (2); and $\widehat{\lambda}_{ijt}$ interaction with time dummies in columns (3) and (4).

Table 9: Regression results

Dependent variable: FDI_{ijt}						
	(1)	(2)	(3)	(4)	(5)	(6)
W s_{ijt}	0.790 (0.303)**	0.869 (0.304)**	0.694 (0.328)*			
W (1+ FDI_{ijt})				0.077 (0.024)**	0.083 (0.024)**	0.075 (0.027)**
GDP $_{it}$	-2.961 (2.447)	-3.682 (2.479)	-3.695 (2.554)	-2.963 (2.449)	-3.703 (2.483)	-3.719 (2.544)
MP $_{jt}$	2.315 (1.863)	2.945 (1.878)	2.810 (1.877)	2.283 (1.859)	2.918 (1.874)	2.771 (1.873)
RFE $_{ijt}$	2.536 (2.302)	2.867 (2.318)	2.818 (2.371)	2.449 (2.300)	2.788 (2.318)	2.732 (2.360)
Distance $_{ij}$	-2.529 (1.065)*	-2.061 (1.067)	-1.388 (0.197)**	-2.506 (1.059)*	-2.030 (1.060)	-1.247 (0.203)**
RTA $_{ijt}$	0.387 (0.213)	0.298 (0.212)	0.231 (0.207)	0.357 (0.212)	0.263 (0.211)	0.187 (0.207)
Wage $_{jt}$	3.605 (1.571)*	3.365 (1.580)*	3.352 (1.578)*	3.616 (1.562)*	3.367 (1.571)*	3.364 (1.568)*
Agglo $_{jt}$	-2.006 (1.043)	-2.036 (1.040)	-2.001 (1.040)	-1.971 (1.045)	-1.995 (1.041)	-1.936 (1.040)
Phone $_{jt}$	-0.223 (1.199)	-0.463 (1.206)	-0.329 (1.251)	-0.221 (1.199)	-0.492 (1.208)	-0.401 (1.255)
Road $_{jt}$	0.837 (0.514)	0.697 (0.533)	0.627 (0.567)	0.849 (0.513)	0.717 (0.532)	0.649 (0.565)
Delinquency $_{jt}$	-0.570 (0.517)	-0.517 (0.508)	-0.626 (0.512)	-0.539 (0.517)	-0.491 (0.508)	-0.603 (0.514)
$\widehat{\lambda}_{ijt}$		-0.777 (0.309)*	-1.201 (0.312)**		-0.730 (0.309)*	-1.130 (0.308)**
Constant	46.959 (53.059)	49.229 (54.286)	-22.001 (7.090)**	47.546 (52.937)	50.266 (54.158)	-21.975 (6.961)**
Observations	2666	2666	2666	2666	2666	2666
F-statistic	25.36**	23.17**	25.64**	26.19**	23.78**	26.64**
R ²	0.493	0.500	0.438	0.495	0.501	0.444
Country and State specific effects	Yes	Yes	Mundlak	Yes	Yes	Mundlak
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Wald Test	76.47*** ^S 111.43*** ^C	22.36*	37.50**	79.05*** ^S 96.91*** ^C	22.00*	37.02**

Table notes:

* significant at 5%; ** significant at 1%. Time, country and state specific effects not reported. Mundlak (1978) transformation was applied to specific effects to correct for selection bias in regressions (3) and (6). Variables on the interaction of $\widehat{\lambda}_{ijt}$ with time dummies are also included in regressions (2), (3), (5) and (6) but not reported.

$\widehat{\lambda}_{ijt}$ values (inverse Mills ratios) employed in regressions (2), (3), (5) and (6) were obtained, as suggested by Semykina and Wooldridge (2005), from a probit estimation for each period.

Wald test evaluates the joint significance of: host state (^S) and source country (^C) specific effects in regressions (1) and (3); and $\widehat{\lambda}_{ijt}$ interaction with time dummies in regressions (2), (3), (5) and (6).

FDI analysis than in the case of trade one [Egger (2008)]. The Heckman (1979) type methodology proposed by Semykina and Wooldridge (2005) is then employed to handle these issues, since a set of tests suggest their presence.

Results show the "classical" negative correlation between FDI and distance. This suggests that the costs of operating at a distance seem to be higher than the advantage of being close to distant markets. In other words, we can infer that there is a motivation for vertical FDI. Results also suggest a positive correlation between FDI and wages. This result can be explained if we consider wages as an indicator of productivity or skilled-labor or if MNEs prefer to locate their operations in markets with better purchasing power. Combining this result with that on distance regressors, the former hypothesis is more plausible: the average wage is an indicator of productivity and/or skilled-labor endowment. The result on wages contradicts Love and Lage-Hidalgo (2000) results who found a negative correlation between Mexican hourly real wages and FDI flows from the US. Given that Love and Lage-Hidalgo (2000) study is at a national level and mine is at a regional one, we can think that once MNEs decide to invest in Mexico because of labor costs relative advantages, they select the region offering the more productive and skilled-labor abundant state to invest, even if the host state average labor costs are the highest in Mexico. For future research, it would be convenient to introduce a more appropriate unit labor costs measurement that takes into account labor productivity and labor compensation; unfortunately, to the author's knowledge, there is no data available at present at a regional level to conduct this research.

Results also suggest the presence of selection bias. Once the selection bias is corrected, the inverse Mill's ratio is significant and negatively correlated to FDI flows. Following Razin et al. (2003) and Razin et al. (2008) this variable was considered as a proxy of (unobserved) setup costs. Indeed, when this value is computed, it was found that developed countries faces lower setup costs, as proposed by Razin et al. (2008) to justify the Lucas paradox. Among the Mexican states there is not a big difference between the northern and the southern ones; however there are important differences between states to facilitate FDI.

To test the existence of spatial dependence attracting FDI, third region variables were added to the model. These variables measure the size of the economy of neighboring regions as well as the same source country multinational presence in neighboring states. The inclusion of these variables dropped the significativity of other variables, which may lead us to wrong conclusions. Hence, the inclusion of spatial spillovers seems to be pertinent in the case of FDI activity for Mexican regions. Results suggest a positive relationship or complementarity between the FDI received and the FDI activity by the same country in neighboring states. For vertical FDI, Blonigen et al. (2007) suggest that spatial complementariness are a sign of the presence of "complex" vertical or vertical specialization (fragmentation) FDI motivations. In other words, MNEs would produce the final goods via multiple stages located in multiple regions, thus the proximity to MNEs already producing a stage would be a positive determinant. This conclusion is also suggested by Hanson et al. (2005) who claim the importance of vertical production networks as a driving force of MNEs activity. They suggest vertical production networks between Mexico and the US as the most obvious example of vertical specialization where labor costs differences and low trade wages favors input processing trade. However, we cannot consider vertical specialization as the only FDI motivation. Instead, we can interpret this as a result of the weight of maquila industry and the US on Mexican states FDI flows. Indeed, for the period 1994 to 2000, the US share of Mexico's FDI was about 50% and the US maquila investment accounted on average for almost 27% of the US FDI into Mexico [Gerber (2002)].

Notes

¹Mexican liberalization reforms started in 1985; these reforms are usually separated into two waves: (1) before NAFTA and (2) after NAFTA. For more details about Mexican liberalization process see Tornell and Esquivel (1995).

²According to UNCTAD (2002), this transaction accounted for more than 50% of 2001 FDI.

³See Fujita et al. (1999) for a review of NEG models.

⁴Note that the independent variables still enter in logarithms, thus the coefficients can be interpreted as elasticities.

⁵Note that Stock and Yogo (2005) only tabulate values up to 3 endogenous regressors and 100 excluded instruments.

⁶Helpman et al. (2008) conclude that their methodology can be applied to data sets including an important fraction of zeros.

References

Aitken, B., A. Harrison and R.E. Lipsey, 1997. "Wages and Foreign Ownership: A Comparative Study of Mexico, Venezuela and the United States." NBER Working Paper No. 5102.

Anselin, L. 2001. "Spatial Econometrics." In B. Baltagi (ed.), *A Companion to Theoretical Econometrics*. Oxford: Basil Blackwell: 310-330.

Anselin L. and S. Rey, 1991. "Properties of Tests for Spatial Dependence in Linear Regression Models." *Geographical Analysis* 23: 112-131.

Baltagi, B. H., P. Egger and M. Pfaffermayr, 2007. "Estimating Models of Complex FDI: Are There Third-Country Effects?" *Journal of Econometrics* 40(1): 260-281.

Bergstrand, J. H. and P. Egger, 2007. "A knowledge-and-physical-capital model of international trade flows, foreign direct investment, and multinational enterprises." *Journal of International Economics* 73(2): 278-308.

Blonigen, B. A. and R. B. Davies, 2004. "The Effects of Bilateral Tax Treaties on US FDI Activity." *International Tax and Public Finance* 11(5): 601-22.

Blonigen, B.A., R.B. Davies, G.R. Waddell and H.T. Naughton, 2007. "FDI in Space: Spatial Autoregressive Relationships in Foreign Direct Investment." *European Economic Review* 51: 1303-1325.

Chaney, T., 2008. "Distorted Gravity: The Intensive and Extensive Margins of International Trade." *American Economic Review* 98(4): 1707-21.

- Coughlin, C.C., J.V. Terza and V. Arromdee, 1991. "State characteristics and the location of foreign direct investment within the United States." *Review of Economics and Statistics* 73: 675-683.
- Coughlin, C. and E. Segev. 2000. "Foreign Direct Investment in China: A Spatial Econometric Study." *The World Economy*, 23(1), 1-23.
- Cragg, J. G., and S. G. Donald. 1993. "Testing identifiability and specification in instrumental variables models." *Econometric Theory* 9: 222-240.
- Crozet, M., T. Mayer and J. Mucchielli, 2004. "How do firms agglomerate? A study of FDI in France." *Regional Science and Urban Economics* 34: 27- 54
- Dunning, J., 1973. "The Determinants of International Production." *Oxford Economic Papers* 25(3): 289-336.
- Dussel Peters, E., 2000. *Polarizing Mexico: The Impact of Liberalization Strategy*. Ed. Lynne Rienner Publishers Inc.
- Jonathan E. and A. Tamura, 1996. "Japanese and U.S. Exports and Investment as Conduits of Growth." NBER Chapters, in: *Financial Deregulation and Integration in East Asia*, NBER-EASE 5: 51-75
- Egger, P., 2008. "Bilateral FDI Potentials for Austria." FIW Research Reports series 11, FIW.
- Egger, P., S. Gruber, M. Larch and M. Pfaffermayr, 2007. "Knowledge-capital meets new economic geography." *The Annals of Regional Science* 41(4): 857-875.
- Egger, P., M. Pfaffermayr and H. Winner, 2005. "An unbalanced spatial panel data approach to US state tax competition." *Economics Letters* 88(3): 329-335.
- Eichengreen, B. and D. Irwin. 1995. "Trade Blocs, Currency Blocs and the Reorientation of Trade in the 1930s." *Journal of International Economics* 38 (1-2): 1-24.
- Ekholm, K, R. Forslid and J.R. Markusen, J. R., 2005. "Export-Platform Foreign Direct Investment." Updated working paper <http://spot.colorado.edu/~markusen/workingpapers.html>
- Ekholm, K, R. Forslid and J.R. Markusen, J. R., 2007. "Export-Platform Foreign Direct Investment." *Journal of the European Economic Association* 5(4): 776-795.
- Feenstra, R.C., and G.H. Hanson, 1997. "Foreign Direct Investment and Relative Wages: Evidence from Mexico's Maquiladoras." *Journal of International Economics* 42: 371-393.
- Fujita, M., P. Krugman and A. Venables, 1999. *The spatial economy. Cities, regions and international trade*. Cambridge, MIT Press.
- Gerber, J., 2002. "The Structure of US Outward Foreign Direct Investment in Mexico's Export Processing Industry." *Asian Journal of Latin American Studies* 15(1): 43-66.
- Graham, E. and E. Wada, 2000. "Domestic Reform, Trade and Investment Liberalisation, Financial Crisis, and Foreign Direct Investment into Mexico." *World Economy* 23(6): 777-797.
- Hadi, A. S., 1994. "A Modification of a Method for the Detection of Outliers in Multivariate Samples." *Journal of the Royal Statistical Society, Series (B)*, 56, 393-396.

- Hahn, J. and J. Hausman, 2002. "Notes on bias in estimators for simultaneous equation models." *Economics Letters* 75(2): 237-241.
- Hanson, G.H., 1998. "North American Economic Integration and Industry Location." NBER Working Paper 6587.
- Hanson G.H., 1998. "Regional adjustment to trade liberalization." *Regional Science and Urban Economics* 28: 419-444.
- Hanson, G.H., R.J. Mataloni and M.J. Slaughter. "Vertical Production Networks in Multinational Firms." *Review of Economics and Statistics*: 664-678,
- Harris, C., 1954 "The Market as a Factor in the Localization of Industry in the United States." *Annals of the Association of American Geographers* 64: 315-48.
- Havrylchyk, O. and S. Poncet, 2007. "Foreign Direct Investment in China: Reward or Remedy?" *The World Economy* 30: 1662-1681.
- He, C., 2002. "Information costs, agglomeration economies and the location of foreign direct investment in China." *Regional Studies* 36 (9): 1029-1036.
- Head, C.K. and T. Mayer, 2002. "Illusory Border Effects: Distance Mismeasurement Inflates Estimates of Home Bias in Trade." CEPII Working Paper No. 2002 - 01.
- Head, C.K., and T. Mayer, 2004. "Market Potential and the Location of Japanese Firms in the European Union." *Review of Economics and Statistics*, 86(4): 959-972.
- Head, C.K. and J. Ries, 2008. "FDI as an outcome of the market for corporate control: Theory and evidence." *Journal of International Economics* 74(1): 2-20.
- Head, C.K., J.C. Ries and D.L. Swenson, 1995. "Agglomeration benefits and location choice: evidence from Japanese manufacturing investments in the United States." *Journal of International Economics*, 38: 223-247.
- Head, C.K., J.C. Ries and D.L. Swenson, 1999. "Attracting foreign manufacturing: Investment promotion and agglomeration." *Regional Science and Urban Economics*, 29(2): 197-218.
- Heckman, J., 1979. "Sample selection bias as a specification error." *Econometrica* 47: 153-61.
- Helpman, E., 1984. "A Simple Theory of Trade with Multinational Corporations." *Journal of Political Economy* 92: 451-471.
- Helpman, E., M. J. Melitz and Y. Rubinstein, 2008. "Estimating trade flows: Trading partners and trading volumes". *Quarterly Journal of Economics* 123: 441-487.
- Helpman, E., M. J. Melitz and S. R. Yeaple, 2003. "Export Versus FDI with Heterogeneous Firms." *American Economic Review* 94(1): 300-316.
- Hoffman, A. and J.R. Markusen, 2007. "Investment Liberalization and the Geography of Firm Location." Forthcoming in S. Brakman and H. Garretsen (eds.), *Foreign Direct Investment and the Multinational Enterprise*, MIT Press.

A. Ibarra-Yunez, 2003. "Spaghetti Regionalism or Strategic Foreign Trade: Some Evidence for Mexico." *Journal of Development Economics*, 72(2): 567-584.

Instituto Nacional de Estadística, Geografía e Informática. Banco de Información Económica. <http://www.inegi.gob.mx/>

Kelejian, H.H. and I.R. Prucha, 1998. "A generalized spatial least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances." *Journal of Real Estate Finance and Economics*, 17: 99-121.

Kennedy, P., 2003. *A Guide to Econometrics*. Fifth Edition, Blackwell, Oxford.

Kleinert, J. and F. Toubal, 2010. "Gravity for FDI." *Review of International Economics* 18(1): 1-13.

Levy Yeyati, E., E. Stein and C. Daude, 2003. "Regional Integration and the Location of FDI." Inter-American Development Bank (IDB), Research Department, Working Paper No. 492.

Lopez-Cordova, E. 2003. "NAFTA and Manufacturing Productivity in Mexico." *Economía: Journal of the Latin American and Caribbean Economic Association* 4(1): 55-88.

Love J, and F. Lage-Hidalgo, 2000. "Analysing the Determinants of US Direct Investment in Mexico." *Applied Economics* 32: 1259-1267.

Máttar J., J.C. Moreno-Brid and W. Peres, 2002. "Foreign Investment in Mexico after Economic Reform." CEPAL: Serie Estudios y Perspectivas.

Mátyás, L., 1997. "Proper Econometric Specification of the Gravity Model." *The World Economy* 20 (3): 363-368.

Markusen, J.R., 1984. "Multinationals, Multi-Plant Economies, and the Gains from Trade." *Journal of International Economics* 16: 205-226.

Markusen, J.R. (2002). "Multinational Firms and the Theory of International Trade." The MIT press, Cambridge.

Martin, P. and C.A. Rogers, 1995. "Industrial location and public infrastructure." *Journal of International Economics* 39: 335-351.

Melitz, M. J., 2003. "The Impact of Trade on Intra-industry Reallocations and Aggregate Industry Productivity." *Econometrica*, 71: 1695-1725.

Mollick, A.V., R. Ramos-Duran and E. Silva Ochoa, 2006. "Infrastructure and FDI Inflows into Mexico: A Panel Data Approach." *Global Economy Journal* 6(1).

Mundlak, Y., 1978. "On the pooling of time series and cross section data." *Econometrica* 46: 69-85.

Razin, A., Y. Rubinstein and E. Sadka, 2003 "Which Countries Export FDI, and How Much?" NBER Working Paper No. 10145

Razin, A., E. Sadka and H. Tong, 2008. "Bilateral FDI Flows: Threshold Barriers and Productivity Shocks." *CESifo Economic Studies* 54(3): 451-470.

- Redding S. and A.J. Venables, 2004b. "Economic Geography and International Inequality." *Journal of International Economics*, 62(1): 53-82.
- Semykina, A. and J.M. Wooldridge, 2005. "Estimating Panel Data Models in the Presence of Endogeneity and Selection: Theory and Application". Michigan State University Working Paper.
- Shea, J., 1997. "Instrument Relevance in Multivariate Linear Models : A Simple Measure." *The Review of Economics and Statistics*, vol. 79, no 2 (mai), p. 348-352.
- Santos Silva, J.M.C. and S.Tenreiro, 2006. "The log of gravity." *The Review of Economics and Statistics* 88: 641-58.
- Staiger, D., and J. H. Stock, 1997. "Instrumental variables regression with weak instruments." *Econometrica* 65(3): 557–86.
- Stock, J. H. and M. Yogo, 2005. "Testing for weak instruments in linear IV regression." In *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, ed. D. W. K. Andrews and J. H. Stock, chap. 5, 80-108. Cambridge University Press.
- Tobin, J., 1958. "Estimation for relationships with limited dependent variables." *Econometrica* 26 (1): 24–36
- Tornell, A. and G. Esquivel, 1995. "The Political Economy of Mexico's Entry to NAFTA." NBER working paper No. 5322.
- Tornell, A., F. Westermann and L. Martinez, 2004. "NAFTA and Mexico's Less- Than-Stellar Performance." NBER Working Paper No. 10289.
- United Nations Conference on Trade and Development (UNCTAD), 2002. "World Investment Report 2002: Transnational Corporations and Export Competitiveness."
- United Nations Conference on Trade and Development (UNCTAD), 2006. "World Investment Report 2006: FDI from Developing and Transition Economies: Implications for Development."
- Wei, S.J., 2000. "How Taxing is Corruption to International Investors?" *Review of Economic Statistics*, 82: 1-11.
- Wooldridge, J. M., 2002. *Econometric Analysis of Cross Section and Panel Data*. Cambridge (Mass.): MIT Press, 752 p.

Appendices

A Variables definition and data sources

Preliminary notes.

1. Exempting s_{ijt} and \mathbf{FTA}_{ijt} , all variables are introduced in logarithmic transformation.
2. FDI yearly flows are reported by the Mexican Ministry of Economy in US current dollars, this data is then transformed to 2000 constant US dollar employing chain dollar deflator from the US Bureau of Economic Analysis (BEA).
3. Variables indicated by (*) concerns data reported in Mexican pesos. This data was transformed to 2000 constant USD employing the real exchange rate Mexico Peso / US Dollar reported by the Centro de Estudios de las Finanzas Públicas de la H. Cámara de Diputados.

Variables definition and data sources

A.0.1 Dependent variable

\mathbf{FDI}_{ijt} . FDI flows from country i to state j in 2000 constant USD. Sample includes only countries that invest at least 4 times in the sample period, as well as countries whose data on GDP and population was available for the sample period. A total of 68 countries are then included in the sample. These countries, with their iso3 code in parentheses, are: Argentina (ARG), Australia (AUS), Austria (AUT), Belgium (BEL), Bolivia (BOL), Brazil (BRA), Canada (CAN), Switzerland (CHE), Chile (CHL), China (CHN), Colombia (COL), Costa Rica (CRI), Cyprus (CYP), Czech Republic (CZE), Germany (DEU), Denmark (DNK), Dominican Republic (DOM), Algeria (DZA), Ecuador (ECU), Egypt (EGY), Spain (ESP), Finland (FIN), France (FRA), United Kingdom (GBR), Greece (GRC), Guatemala (GTM), Hong Kong (HKG), Honduras (HND), Hungary (HUN), Indonesia (IDN), India (IND), Ireland (IRL), Iran (IRN), Iceland (ISL), Israel (ISR), Italy (ITA), Jamaica (JAM), Japan (JPN), Republic of Korea (KOR), Luxembourg (LUX), Malaysia (MYS), Nicaragua (NIC), Netherlands (NLD), Norway (NOR), New Zealand (NZL), Pakistan (PAK), Panama (PAN), Peru (PER), Philippines (PHL), Poland (POL), Portugal (PRT), Paraguay (PRY), Romania (ROM), Russian Federation (RUS), Singapore (SGP), El Salvador (SLV), Serbia (SRB), Slovakia (SVK), Sweden (SWE), Togo (TGO), Thailand (THA), Turkey (TUR), Uruguay (URY), United States (USA), Venezuela (VEN), South Africa (ZAF), Zaire (ZAR), Zambia (ZMB). *Data source:* Secretaría de Economía - Dirección General de Inversión Extranjera (DGIE).

A.0.2 Explanatory variables

s_{ijt} . Selection indicator that takes the value of 1 if country- i invests in state- j , and zero otherwise. *Data source:* See \mathbf{FDI}_{ijt} .

\mathbf{GDP}_{it} . Country i 's real GDP in 2000 constant USD. *Data source:* World Bank - WDI.

\mathbf{GDP}_{jt} . (*) State j 's real GDP in 2000 constant USD. *Data source:* Instituto Nacional de Estadística y Geografía (INEGI).

\mathbf{RFE}_{ijt} . Ratio of country i per capita real GDP to state j 's per capita real GDP (*). *Data source:* Countries data from World Bank - WDI; State's data from INEGI.

Wage_{jt}. (*) Average daily wage reported to the social security in 2000 constant USD. *Data source*: Secretaría del Trabajo y Previsión Social (STPS).

Agglo_{jt}. Ratio of industrial sector consumers of electricity_{t-1} / state's area. *Data source*: Comisión Fédéral de Electricidad (CFE).

Phone_{jt}. Telephone line density (for every hundred inhabitants). *Data source*: Comisión Federal de Telecomunicaciones (COFETEL)

Road_{jt}. Ratio of squared kilometers of paved roads to state's area. *Data source*: INEGI

Delinquency_{jt}. The delinquency rate is the number of registered delinquents for every hundred inhabitants. *Data source*: INEGI.

Distance_{ij}. Bilateral distance is measured as the great circle formula using the latitude and the longitude of source-country and host-state capitals. In the case of distance between state-j and the US the distance is measured as the great circle distance between state's capital and the nearest border-crossing point. Included border-crossing points by US state are:

1. Arizona: Yuma, Douglas, Lukeville, Naco, Nogales, San Luis, Sasabe.
2. California: Calexico, San Diego, San Ysidro, Tecate.
3. New Mexico: Columbus, Santa Teresa.
4. Texas: Brownsville, Del Rio, Eagle Pass, El Paso, Fabens, Hidalgo, Laredo, Presidio, Progreso, Rio Grande City, Roma.

Data source: Countries capital latitude and longitude data was obtained from the CEPII distances database; State's data comes from INEGI; border crossing coordinates were obtained from world-gazetteer data set.

FTA_{ijt}. Dummy variable which takes the value of 1 if country-i has a trade agreement with Mexico in year t, and zero otherwise. *Data source*: Secretaría de Economía - Acuerdos y negociaciones comerciales.

MP_{jt}. $MP_{jt} = GDP_{jt} + \sum_{k=1}^K (GDP_{kt} / distance_{jk})$ for $j \neq k$ where state-k is a neighbor of state-j. *Data source*: See GDP_j.

W s_{ijt}. Investment decision from the same source country MNEs to neighboring states. *Data source*: See FDI_{ijt}.

W (1+FDI_{ijt}). 1 + FDI flows from the same source country to neighboring states in year t. *Data source*: See FDI_{ijt}.

A.0.3 Excluded instruments

Wage_{jt-1}. (*) One year lagged of the wage variable Wage_{jt}. *Data source*: See Wage_{jt} variable.

W Agglo_{jt}. Agglomeration in neighboring states. *Data source*: See Agglo_{jt} variable.

W Phone_{jt}. Telephone density in neighboring states. *Data source*: See Phone_{jt} variable.

W Road_{jt}. Road density in neighboring states. *Data source*: See Road_{jt} variable.

W Delinquency_{jt}. Delinquency rate in neighboring states: *Data source*: See Delinquency_{jt} variable.

Q-I W s_{ijt}. Quasi-instrument of W s_{ijt} variable. It takes the values of -1, 0 or 1 depending on whether or not the endogenous variable is in the upper, middle and lower third. *Data source*: See FDI_{ijt}.

Q-I W (1+FDI_{ijt}). Quasi-instrument of W (1+FDI_{ijt}) variable. See W s_{ijt} and FDI_{ijt}.

B Observations by country and state

C Validity tests for the excluded instruments

First stage results

endogenous variable:	Table reporting 2nd-stage results				
	Table 3 (4)	Table 5 (4)		Table 6 (4)	
	Wage _{jt}	Wage _{jt}	W s _{ijt}	Wage _{jt}	W (1+FDI _{ijt})
GDP _{it}	0.144 (0.022)**	0.045 (0.020)*	0.357 (0.133)**	0.044 (0.020)*	4.786 (1.933)*
GDP _{jt}	0.143 (0.019)**				
RFE _{ijt}	-0.146 (0.022)**	-0.040 (0.020)*	-0.286 (0.138)*	-0.039 (0.020)*	-4.097 (2.086)*
Agglo _{jt}	0.015 (0.008)*	0.003 (0.009)	-0.064 (0.057)	0.003 (0.009)	-0.289 (0.843)
Phone _{jt}	0.071 (0.010)**	0.066 (0.012)**	0.065 (0.069)	0.066 (0.012)**	1.555 (0.981)
Road _{jt}	0.013 (0.004)**	0.003 (0.005)	0.003 (0.023)	0.003 (0.005)	0.129 (0.331)
Delinquency _{jt}	-0.030 (0.003)**	-0.025 (0.003)**	0.006 (0.024)	-0.025 (0.003)**	-0.448 (0.374)
MP _{jt}		0.219 (0.017)**	-0.197 (0.122)	0.219 (0.017)**	-3.547 (1.836)
Distance _{ijt}	0.000 (0.001)	0.000 (0.001)	-0.112 (0.010)**	0.000 (0.001)	-2.324 (0.165)**
FTA _{ijt}	0.001 (0.001)	-0.001 (0.001)	0.063 (0.014)**	-0.001 (0.001)	1.152 (0.213)**
$\hat{\lambda}_{ijt}$	-0.001 (0.004)	-0.003 (0.004)	-0.051 (0.015)**	-0.003 (0.004)	-0.317 (0.217)
Wage _{jt-1}	0.794 (0.012)**	0.735 (0.011)**	-0.041 (0.070)	0.736 (0.011)**	0.890 (0.984)
W Agglo _{jt}		0.042 (0.023)	0.032 (0.143)	0.042 (0.023)	-0.886 (2.193)
W Phone _{jt}		0.233 (0.021)**	-0.308 (0.128)*	0.233 (0.021)**	-4.451 (1.945)*
W Road _{jt}		-0.065 (0.007)**	0.005 (0.041)	-0.065 (0.007)**	-0.035 (0.537)
W Delinquency _{jt}		0.044 (0.008)**	0.059 (0.046)	0.044 (0.008)**	0.418 (0.661)
Q-I W s _{ijt}		-0.001 (0.001)*	0.233 (0.006)**		
Q-I W (1+FDI _{ijt})				-0.001 (0.001)*	3.072 (0.086)**
Constant	-0.207 (0.024)**	-0.090 (0.026)**	-0.590 (0.367)	-0.091 (0.026)**	-9.555 (5.817)

First stage results

endogenous variable:	Table reporting 2nd-stage results				
	Table 8 (4)	Table 9 (3)		Table 9 (6)	
	Wage _{jt}	Wage _{jt}	W s _{ijt}	Wage _{jt}	W (1+FDI _{ijt})
GDP _{it}	0.182 (0.031)**	0.080 (0.029)**	0.723 (0.192)**	0.079 (0.029)**	10.974 (2.483)**
GDP _{jt}	0.122 (0.026)**				
RFE _{ijt}	-0.187 (0.031)**	-0.078 (0.028)**	-0.534 (0.200)*	-0.077 (0.028)**	-7.950 (2.596)**
Agglo _{jt}	0.011 (0.010)	-0.006 (0.012)	-0.028 (0.069)	-0.006 (0.012)	0.106 (1.056)
Phone _{jt}	0.056 (0.012)**	0.061 (0.015)**	0.070 (0.090)	0.060 (0.015)**	1.834 (1.226)
Road _{jt}	0.010 (0.005)	0.003 (0.006)	-0.006 (0.026)	0.003 (0.006)	-0.072 (0.358)
Delinquency _{jt}	-0.019 (0.004)**	-0.019 (0.004)**	0.003 (0.030)	-0.019 (0.004)**	-0.041 (0.450)
MP _{jt}		0.199 (0.024)**	-0.372 (0.167)*	0.200 (0.024)**	-5.973 (2.316)*
Distance _{ijt}	0.003 (0.001)	0.002 (0.001)	-0.100 (0.013)**	0.002 (0.001)	-2.603 (0.183)**
FTA _{ijt}	0.003	0.000 (0.001)	0.110 (0.017)**	0.000 (0.001)	1.692 (0.267)**
$\hat{\lambda}_{ijt}$	-0.013 (0.007)	-0.016 (0.008)*	-0.076 (0.023)**	-0.016 (0.008)*	-1.069 (0.293)**
Wage _{jt-1}	0.791 (0.014)**	0.746 (0.013)**	-0.078 (0.082)	0.746 (0.013)**	0.035 (1.121)
W Agglo _{jt}		0.010 (0.027)	-0.087 (0.163)	0.011 (0.027)	-3.325 (2.326)
W Phone _{jt}		0.182 (0.026)**	-0.547 (0.155)**	0.182 (0.026)**	-7.728 (2.063)**
W Road _{jt}		-0.064 (0.009)**	0.087 (0.050)	-0.064 (0.009)**	0.436 (0.612)
W Delinquency _{jt}		0.029 (0.010)**	0.057 (0.058)	0.029 (0.010)**	0.009 (0.783)
Q-I W s _{ijt}		-0.001 (0.001)	0.250 (0.007)**		
Q-I W (1+FDI _{ijt})				-0.001 (0.001)*	3.479 (0.098)**
Constant	-0.162 (0.028)**	-0.080 (0.032)	0.388 (0.406)	-0.081 (0.032)*	8.239 (6.516)

Instrument validity tests

endogenous variable:	Table reporting 2nd-stage results				
	Table 3 (4)	Table 5 (4)		Table 6 (4)	
	Wage _{jt}	Wage _{jt}	W s _{ijt}	Wage _{jt}	W (1+FDI _{ijt})
Shea partial R ²	0.625	0.657	0.534	0.658	0.464
Partial R ²	0.625	0.660	0.536	0.659	0.464
F-statistic	54745.46**	1450.35**	292.74**	1443.67**	214.67**
gmin		630.16		558.64 > 20.79	
Hansen J P-val		0.971		0.918	

endogenous variable:	Table reporting 2nd-stage results				
	Table 8 (4)	Table 9 (3)		Table 9 (6)	
	Wage _{jt}	Wage _{jt}	W s _{ijt}	Wage _{jt}	W (1+FDI _{ijt})
Shea partial R ²	0.611	0.633	0.589	0.633	0.542
Partial R ²	0.611	0.634	0.591	0.635	0.543
F-statistic	3353.15**	1097.67**	247.27**	1096.53**	210.90**
Cragg-Donald gmin		611.61 > 20.79		510.51 > 20.79	
Hansen J P-val		0.818		0.824	

Table notes:

* significant at 5%; ** significant at 1%. Standard errors robust to serial correlation and heteroskedasticity in parentheses under coefficient estimates.

First-stage regressions for estimators correcting for selection bias. Reported results refers to a table and the column of second stages results. See the table and column of second-stage results for more informations.

The Cragg-Donald eigen value (gmin) is compared to the critical values proposed by Stock and Yogo (2005) at 5%.

Hansen J P-val refers to the p-value of Hansen J statistic to test for overidentification. This statistic is not reported when the equation is exactly identified.