Foreign direct investment in provinces: A spatial regression approach to FDI in Vietnam

Bulent Esiyok and Mehmet Ugur

University of Greenwich

20. December 2011

Foreign direct investment in provinces:
A spatial regression approach to FDI in Vietnam

Bulent Esiyok and Mehmet Ugur, University of Greenwich

Foreign direct investment (FDI) flows into Vietnam have increased significantly in recent years, with unequal distribution between provinces and regions. We aim to contribute to the literature on locational determinants of FDI by accounting for spatial interdependence between 62 Vietnamese provinces from 2006-2009. For this purpose, we estimate a spatial lag model using maximum likelihood estimation method. We report existence of spatial dependence between provinces as well as spatial spill-over effects. The results are robust to different specifications for weight matrices and inclusion of different explanatory variables and/or proxies. We also report that conventional determinants of FDI such as market size, domestic investment, openness to trade, labour cost, education and governance, etc. are significant and remain robust to inclusion of spatial interdependence. The sign of the spatial dependence suggests that the distribution of FDI between provinces is subject to conglomeration effects.

Key words: Foreign direct investment, spatial dependence, conglomeration, Vietnam.

University of Greenwich Business School

International Business and Economics

bulentesiyok@btinternet.com

m.ugur@gre.ac.uk

Foreign direct investment in provinces:
A spatial regression approach to FDI in Vietnam

1. Introduction

Locational determinants of foreign direct investment (FDI) have been investigated extensively, but empirical work on determinants of FDI in sub-national units is limited to few studies that concentrate mainly on China. In addition, most of the empirical work overlooks spatial interdependence between host markets - even though foreign investors' location decisions involve a choice between a number of competing host units that are related to each through physical distance among other factors. This is particularly the case when the investigation is about the distribution of FDI between sub-national units (regions or provinces) within the same jurisdiction. The distribution of FDI between sub-national units is highly likely to be influenced not only by region-specific factors (e.g., market size, labour costs, governance quality and human capital); but also by spatial interdependence between neighbouring units as the latter are affected by a common set of macroeconomic and trade policies. Therefore, understanding the patterns of such interdependence and the conglomeration/competition effects that they may generate is important in terms of research effort as well as development policy.

As Blonigen et al. (2007) have indicated, spatial econometrics provides useful techniques that can be applied to multiple countries as well as regions within a given country to account for spatial interdependence. In this article, we use the spatial lag model in order to estimate the direct and indirect effects of spatial interdependence. Direct-effect estimates capture the effects of own explanatory variables on FDI within the host spatial unit. Indirect effects, on the other hand, measure the effect of own explanatory variables on FDI within neighbouring units (LeSage and Pace, 2009; Elhorst, 2010a).

FDI in 62 Vietnam provinces constitutes a highly relevant area of application for spatial regression models not only because of data availability at the provincial level, but also because, during the period under investigation (2006-2009), the FDI/GDP ratio for Vietnam provinces has been the highest in the region with the exception of Singapore. Furthermore, provincial authorities in Vietnam have been competing to attract foreign investors, using fiscal incentives and disseminating provincial-level data on governance quality, education, labour training facilities, infrastructure, etc. Theoretically, spatial analysis allows for discovering whether multinational enterprises consider sub-national units as complements or substitutes in their investment decisions. Empirically, it enhances the reliability of inference by incorporating spatial dependence as a specific manifestation of time-invariant fixed effects, which are either ignored (as it is the case in standard OLS estimations) or subsumed under a common intercept (as it is the case in panel data estimations). Finally, spatial analysis can inform policy by providing information about the extent to which FDI inflows into sub-national units are subject to competition or conglomeration effects – and whether such affects are invariant to distance between neighbouring spatial units.

The article is organised in six sections. Section 2 reviews the literature on locational determinants of FDI, with particular attention to work on FDI in sub-national units with and without a spatial-dependence approach. Section 3 provides contextual information on FDI in Vietnam, its distribution between provinces, and geographical information on the number of 'neighbouring provinces' that a province would have at different cut-off points for distance. Section 4 introduces the spatial regression methodology and describes the data. In section 5, we present the empirical findings, which consist of a set of spatial lag estimates based on different specifications for the number of neighbouring provinces and for different cut-off values for distance between provinces. We also report results from sensitivity checks involving use of lagged values for the explanatory variables and different measures for provincial-level governance quality and secondary school enrolment. Section 6 summarises the main findings and discusses their policy and future research implications.

2. Literature review

Reliance on a two-country (or bilateral) framework that consists of one home and one host country is a potential weakness in the theoretical and empirical work on locational determinants of FDI (Blonigen et al. 2007). There are two reasons as to why this may be the case. First, FDI decisions by multinational enterprises (MNEs) may be motivated by horizontal, vertical or complex-vertical investment considerations that must take account of host-country as well as third-country characteristics. For example, in the case of horizontal FDI, distance to and/or market potential of neighbouring countries/provinces may not affect the decision to invest in a particular country/province. However, such factors are highly likely to influence the decision negatively if the investment decision is motivated by vertical integration or export-platform considerations. (See, Baltagi et al. 2007; Blonigen et al. 2007). Secondly, FDI in a host country/province may be influenced by agglomeration or competition dynamics unleashed by distance or policy spill-overs between neighbouring countries/provinces; or by capital-market imperfections that limit the amount of capital available for investment in other countries/provinces once a decision is made in favour of one host country/province (Blonigent et al. 2007). These factors imply that analysis of FDI decisions that do not account for spatial interdependence may yield biased results.

The earliest attempt at estimating the determinants of FDI by taking account of spatial interdependence is Head et al. (1995), who examine the role of agglomeration effects in determining the location of Japanese FDI in the US. They use a conditional-logit model that include interdependence of the location decisions and report that agglomeration effects between bordering states are significant. Head and Mayer's (2004), on the other hand, examine the distribution of Japanese FDI in the European Union, taking account of distance-weighted or trade-frictions-weighted GDP in adjacent regions. They report that more developed regions attract higher levels of FDI and that this effect is robust to inclusion of agglomeration measures as in Head et al. (1995). Although innovative, these studies utilise discrete choice models and as such they impose significant restrictions on the use of data for FDI levels (Blonigen et al, 2007: 1305).

In between, Coughlin and Segev (2000) use a spatial error model to estimate the determinants of US FDI across Chinese provinces. They conclude that FDI shock in one province has positive effects on FDI in nearby provinces. Furthermore, they report that market size, labour productivity, coastal location, wages, and illiteracy rates are statistically significant, while transportation cost is not a significant determinant of FDI across Chinese provinces. Coughlin and Segev (2000) represent the first departure from the discrete choice models developed by Head et al. (1995) and Head and Mayer (2004). It has also motivated two seminal contributions by Baltagi et al. (2007) and Blonigen et al. (2007), both of whom examine the impact of spatial dependence on outbound US FDI.

Baltagi et al. (2007) develops a model of FDI activity that allows for a variety of MNE motivations and spatial interactions. They report significant evidence of spatial interactions, but they cannot conclude whether export-platform or complex vertical FDI motivation is the dominant one. On the other hand, Blonigen et al. (2007) find that spatial interdependence has a significant effect on the distribution of FDI between neighbouring countries and that the estimated parameters for the traditional determinants of FDI (i.e., for the host-country characteristics) are robust to inclusion of spatial interdependence terms. Nevertheless, Blonigen et al. (2007) also report that the existence of spatial interactions does not necessarily allow for robust conclusions about export-platform or complex-vertical motivations for FDI. This is because the estimated spatial interdependence may be sensitive to sample selection.

Another innovative work in this tradition is that of Drukker and Millimet (2007), who illustrate the importance of third-country effects in the context of environmental policy spill-overs. The authors examine the patterns of spatial interdependence between US states with respect to inward FDI at the aggregate and industry levels. They report that own state attributes (including the stringency of environmental protection regulations) do not have statistically significant effects on own aggregate FDI in manufacturing but most of the neighbouring state attributes have a significant effect. In a different context, Garretsen and Peeters (2009) estimate a spatial lag model of outward Dutch FDI to 18 countries from 1984-2004 and also report that third-country effects are significant.

Mainly due to data constraints, the volume of work on sub-national distribution of FDI in developing countries is small and the number of work utilising spatial regression techniques is even smaller. Some of the work on the distribution of FDI across Chinese provinces includes Cole et al. (2009), Na and Lightfoot (2006), and Du et al. (2008). Using a panel data of 30 provinces in China over the period 1998-2003, Cole et al. (2009) report that provincial GDP per capita, government efficiency, anti-corruption effort, good road transportation networks, and surplus of unskilled labour are significant determinants of FDI across provinces. On the other hand, Na and Lightfoot (2006) use cross-section data for 30 regions in 2002 and concludes that market size, labour quality, high labour costs, and the level of infrastructure are important determinants of FDI. Finally, Du et al. (2008) confirm the significance of economic institutions, wages and infrastructure; while they further add special economic zones, coastal cities to their analysis of locational determinants of US multinationals in China.

In the case of Vietnam, Pham (2002) uses averaged data over the period 1988-1998 and provides OLS estimates of the FDI determinants in 53 provinces. The author finds that income per capita, labour quality and phone lines per capita are correlated with FDI flows. However, tax incentives do not explain the variation in FDI inflows among provinces. Another empirical work is Malesky (2007), who uses cross-section data with different measures of FDI, including new FDI projects licensed, implemented FDI as a proportion of registered FDI and additional capital for existing projects. To capture the effect of economic governance quality, the author uses a provincial competiveness index (PCI) and sub-indices of PCI. Only the composite index and private sector development policies sub index are significant for all three measurements of FDI, while the significance of sub-indices vary with the type of FDI used as dependent variable. Furthermore, his findings indicate that FDI is not related to GDP per capita, labour quality, tax incentives and FDI.

Using a panel data for 60 provinces over the period 2000-2005 Vu et al. (2007) also examines the link between FDI and tax incentives offered by provincial governments independently of the national government. In line with Malesky (2007), they measure FDI as registered and implemented FDI. The effect of tax incentives on FDI is rejected by their study for both specifications but investment climate measured by the Provincial Competitiveness Index

(PCI), infrastructure, proximity to major markets, education are found statistically significant. Furthermore, wage is found to be positively related to the implemented FDI, while GDP per capita is found to have no effect on implemented FDI. Findings by Pham (2008) lend support to the relationship between education, income per capita and FDI in 64 provinces over the period 2002-2004. Nguyen (2006) uses panel data with longer time dimension (8 years) and reports that economic growth, market size, domestic investment, exports, the skill of labour, labour cost; infrastructure, real exchange rate and regional dummy are related to FDI inflows among 61 provinces.

To our knowledge, there are no empirical studies that analyse FDI across Vietnamese regions with spatial regression techniques. In what follows, we will summarise the studies that have utilised spatial regression methods in the context of regions in other countries. As indicated above, Coughlin and Segev (2000) use a spatial error model to analyse FDI determinants in 29 Chinese provinces. They conclude that an FDI shock in one province has positive effects on FDI in nearby provinces. In contrast, Sharma et al. (2010) use the spatial lag model with aggregate (1999-2007) and industry-level FDI data (2001-2006) in different provinces in China. The authors find significant spatial interdependence between FDI in Chinese provinces, with the competition effect being dominant at province level and mixed results at industry level. At the regional level in Russia, Ledyaeva (2009) also finds weak evidence of competition between provinces for FDI and reports that that market size, the presence of big cities and sea ports, oil and gas resources, distance to the European market, political and legislation risks and FDI in neighbouring regions are important determinants of FDI in Russia. Finally, Villarde and Maza (2011) include spatially lagged independent variables in their analysis and conclude that there is no spatial dependence in the dependent variable (FDI) but they find significant effect of spatially lagged independent variables.

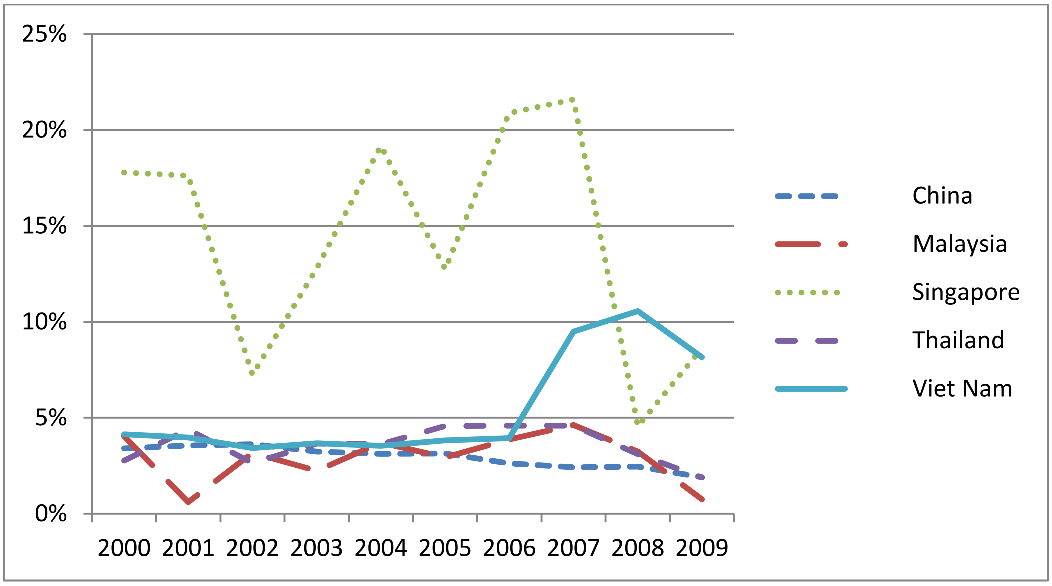
In this article, we aim to make three contributions to the emerging literature on spatial analysis of FDI across sub-national units. First, we provide a range of empirical estimates for conventional FDI determinants and spatial dependence, using spatial regression models with different specifications for weight matrices based on different numbers of neighbouring provinces and different cut-off values for distance between provinces. In doing this, we follow LeSage and Pace (2009) and Elhorst et al. (2010b) to test for the weight matrix specification that best fits the data. Secondly, we provide estimates of not only direct but also indirect effects of the spatial interdependence on FDI. The *direct effect* refers to the extent to which FDI in a host province is affected by the province-specific explanatory variables. The *indirect effect*, on the other hand, measures the extent to which a given change in explanatory variables for a host province affects FDI in all other provinces. Third, we evaluate the sign and magnitude of the spatial interdependence coefficient to establish whether conglomeration or competition effects dominate in the distribution of FDI between Vietnamese provinces; and whether the conglomeration effect diminishes with increased distance.

3. FDI in Vietnamese provinces

Liberalisation of FDI policies in Vietnam dates back to the first FDI law, which was introduced in 1987 and amended several times in 1992, 1996 and 2000 with a view to provide a better investment climate for foreign investors. In a further effort to liberalise FDI policies, the Unified Law of Investment replacing previous laws and regulations was accepted in 2006. Equal treatment of foreign and domestic investors was the major innovation in the Unified Law, which was introduced to comply with the requirements of the World Trade Organisation (WTO) membership. Liberalisation of FDI policies coupled with WTO membership in 2007 boosted FDI inflows in Vietnam.

The outstanding performance of Vietnam in attracting FDI is apparent in comparison with other top destinations in the region, as can be seen in Figure 1. While Vietnam received FDI inflows equivalent to 4% of its GDP in 2006, the corresponding ratio for 2009 was %8. In this regard, Vietnam outperformed not only China, Malaysia and Thailand from 2007 onwards, but also Singapore in 2008.

Figure 1 FDI Inflows as percentage of GDP in selected East-Asian countries



Source: Own figure based on the data from UNCTAD FDI Database

Table 1 below presents the distribution of FDI by top ten investors in Vietnam, which account for 79% of total cumulative registered FDI. It is worth noting that FDI inflows in Vietnam are dominated by regional investors. Of the latter, three are members of the Association of South East Nations (ASEAN) - namely Singapore, Malaysia and Thailand. Taiwan, Republic of Korea, Hong Kong and China are other regional investors. None of European countries has FDI commitments comparable to regional investors. Only the USA follow regional investors in Vietnam with 6.37% of total FDI commitments.

Table :1 Top-ten sources of cumulative registered FDI in Vietnam in 2009 (millions of US$)

Sources of	Registered FDI	Share of source country in total registered FDI (%)
Taiwan	21528.1	12.28%
Korea Rep. of	19843.9	11.32%
Singapore	17304.6	9.87%
Japan	18560.9	10.59%
Malaysia	17926.1	10.23%
British Virgin Islands	13690.7	7.81%
United States	11167.9	6.37%
Hong Kong SAR (China)	7597.7	4.33%
Cayman Islands	6866.4	3.92%
Thailand	5676.4	3.24%
Total	**140162.7**	**79.95%**
All countries	**175309.7**	

Source: Own calculation based on data from the General Statistics Office (GSO) of Vietnam

FDI inflows into Vietnamese provinces are concentrated mainly in North-Central, Central-Coastal, South-Eastern and the Red River regions. As Table 2 indicates, ten provinces from these regions hold 85% of cumulative FDI in 2009. Of these ten provinces, Ho Chi Minh City (HCMC), Ba Ria–Vung Tau (BRVT), Dong Nai and Binh Duong of South-Eastern regions and stand out with 51% share in total FDI. The top three provinces in terms of FDI inflows in Table 2 are also the richest provinces in Vietnam according to per capita GDP figures for 2009.

Table 2 Top-ten Vietnamese provinces with registered FDI in 2009 (millions of US$)

Region	Province	Registered FDI	Share in Total FDI
South East	HCMC	30981.6	18%
South East	BRVT	25700.2	15%
Red River	Ha Noi	22306.9	13%
South East	Dong Nai	17838.1	10%
South East	Binh Duong	13924.6	8%
North Central and Central Coastal	Ninh Thuan	10055.9	6%
North Central and Central Coastal	Ha Tinh	8068.5	5%
North Central and Central Coastal	Phu Yen	8060.8	5%
North Central and Central Coastal	Thanh Hoa	7040.3	4%
North Central and Central Coastal	Quang Nam	5190.5	3%
Total		**149167**	**85%**

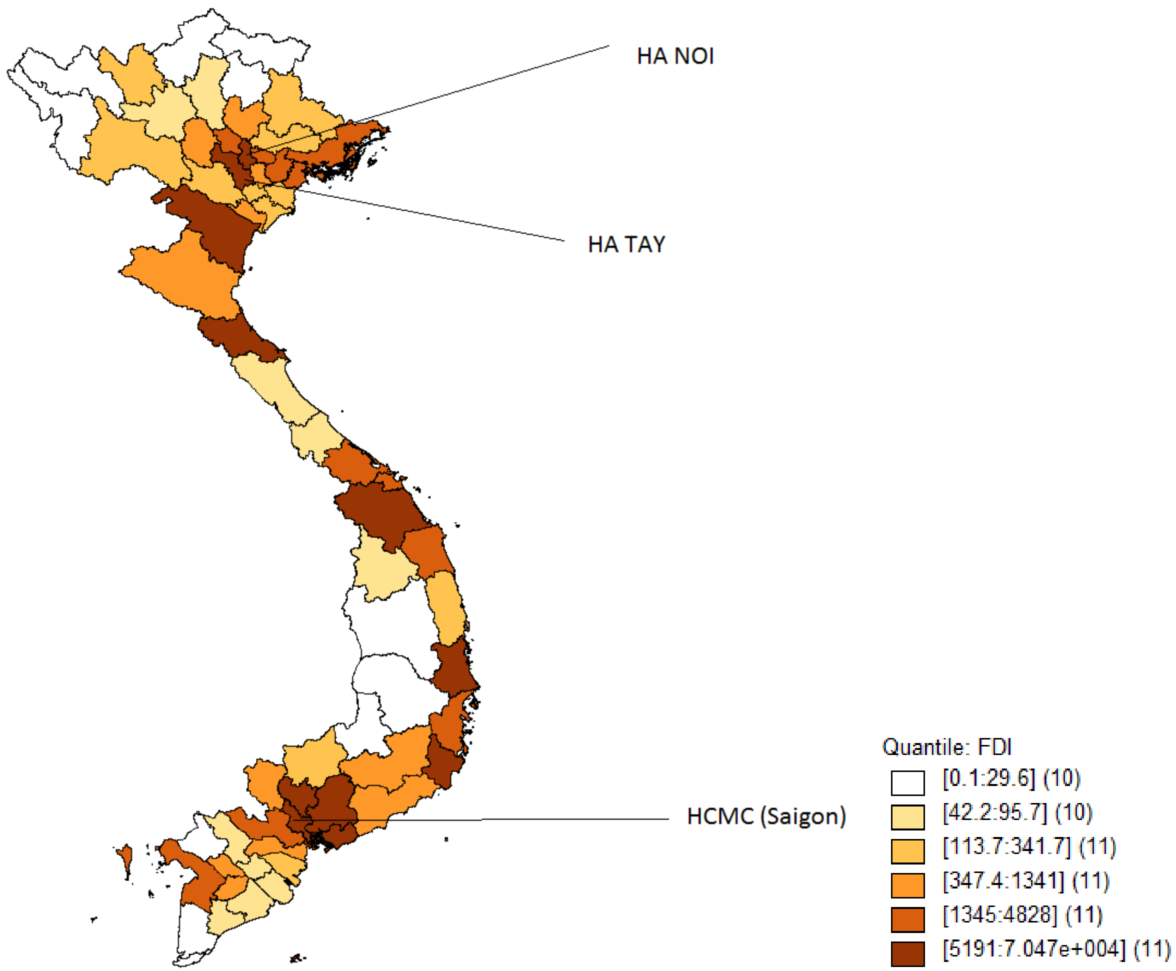
Source: Own calculation based on data from the General Statistics Office (GSO) of Vietnam

The map of Vietnam below provides an overview of cumulative FDI inflows in 2009. White areas indicate the provinces with ten lowest FDI inflows, while brown areas show the provinces with highest FDI inflows. Provinces with low FDI inflows are located together. For instance, Ha Giang, Cao Bang and Bac Kan in the North and Dak Nong, Dak Lak and Gia Lai in the South-West are neighbours. By the same token, there is a correlation in space among provinces with high FDI inflows. Four provinces with highest FDI inflows in the South-East are clustered and they are surrounded by provinces with high FDI inflows as well.

Fiscal decentralisation in 1996 and the decentralisation of FDI administration since 1987 gave power to provincial governments over investment incentives to foreign investors. To compete with provinces with relatively high cumulative FDI, provinces with low level of FDI offered extra incentives in the form of corporate income tax exemptions and VAT reductions and extended exemptions of rent - a practice known as “fence-breaking” that led to high budget deficits in provinces with low FDI inflows (Vu et al. 2007). The effectiveness of these investment incentives is still an open question. On the cost side, most of fence-breaking provinces have been running budget deficits for a long time (Vu et al. 2007). Although the central government suspended all illegal practises on investment incentives provided by 32 provincial governments in late 2005, the extent of violations and the timing of the termination of illegal investment incentives in practise are not clear. Currently 54 provinces out of 63 in Vietnam are eligible for investment incentives in various sectors designed to support areas with socio-economic difficulties as provided for in Government Decree No. 108 of 2006.

In section 4 below, we model spatial interaction between provinces using distance-weighted or neighbouring-province-weighted matrices, with different cut-off values for distance and different numbers of neighbouring provinces. We report estimation results for spatial interaction with one nearest neighbour, three nearest neighbours, 186km and 350km. The cut-off distance of 186km ensures that a province has at least 3 nearest neighbour (with an average of 12 neighbours), whereas the cut-off distance of 350km ensures that a province has at least 7 nearest neighbours (with an average of 19 neighbours).

Figure 2 Provincial distribution of cumulative FDI in Vietnam in 2009 (million of US$)



Ha Noi and Ha Tay merged in 2007. Therefore, the cumulative FDI for Ha Noi in 2009 is equally allocated to both provinces in this figure

4. Methodology and data

Locational determinants of FDI at national or sub-national levels are well specified in the literature. We choose the most frequently-used determinants, consisting of GDP per capita, domestic investment, labour cost, enrolment in lower- and upper-secondary education, budget balance, and openness to trade. In addition, we use the provincial competitiveness index (PCI) and one of its components (informal charges as a proxy for corruption) as governance quality indicators. We specify our model as follows:

$$lnFDI_{it} = \beta_1 lnPCGDP_{it} + \beta_2 lnDI_{it} + \beta_3 BB_{it} + \beta_4 lnLC_{it} + \beta_5 lnOP_{it} + \beta_6 lnEDU_{it} + \beta_7 PCI_{it} + \mu_i + \delta_t + \varepsilon_{it} \qquad (1)$$

In equation (1), subscripts i and t denote province and time respectively. The dependent variable, $lnFDI_{it}$ is the natural logarithm of the real cumulative registered foreign capital scaled by population in province i at time t; $lnPCGDP_{it}$ is the natural logarithm of real per-capita GDP with base year 2005; $lnDI_{it}$ is the natural logarithm of real domestic investment scaled by population (as it is the case in Nguyen and Nguyen, 2007); BB is budget balance as a ratio of provincial GDP; *lnLC* is the natural logarithm of real wages computed as average monthly compensation per employee; *lnOP* is the natural logarithm of trade openness defined as percentage share of exports plus imports in provincial GDP; *lnEDU* is the natural logarithm of the number of students in lower secondary school per 1000 people, a proxy for human capital; and *PCI* is the provincial competitive index (PCI) as a proxy for economic governance quality at the provincial level. μ_i captures unobservable province fixed effect that is constant over time; δ_t controls for time fixed effect that is common across provinces; and ε_{it} is the classical error term that varies across provinces and time.

Equation (1) ignores potential spatial dependence in the dependent variable (*lnFDI*). To check whether spatial dependence exists, we use the residuals of the OLS estimation to establish whether the dependence is due to spatially-lagged dependent variable or spatially-autocorrelated error term. This requires the use of Lagrange Multi plier (LM) tests proposed by Anselin (1988) and robust LM tests proposed by Anselin et al. (1996). In both tests, the null hypothesis of no spatially-lagged dependent variable or no spatially-autocorrelated error term must be rejected for the OLS estimation to be valid. The main difference between the LM and the robust-LM tests is that the latter is capable of detecting one type of spatial even if the other type of dependence exists. As such, it is more powerful in detecting spatial dependence than the standard LM tests.

As indicated above, spatial dependence may be of two types and both types have serious implications for statistical inferences. The type with less severe implications is spatial dependence due to spatial autocorrelations between the error terms and is usually known as the 'spatial error' problem, where the error terms are correlated because of correlation between neighbouring provinces in space. OLS estimates with spatially-autocorrelated error terms are still valid, but they would be inefficient. The other type is due to spatial dependence of the dependent variable and the level of such dependence is captured by the spatial autoregressive coefficient. The latter coefficient measures the extent to which FDI in a given spatial unit is affected by FDI in neighbouring spatial units. Ignoring this type of spatial dependence not only renders statistical inferences invalid but also leads to biased parameter estimates.

We model spatial interaction between observations by using a matrix of distance between spatial units, which consist of 62 provinces. The advantage of using physical distance is due to its exogeneity with respect to FDI (Anselin and Bera, 1998). Empirical studies use different specifications for distances, including the nearest neighbour, contiguous provinces, distance-based matrices, and distance-based matrices with a critical cut-off value. The choice of a distance cut-off value may depend on expected level of spatial spill-overs as a function of travel time. However, there must be a limit to adding new data points to spatial weights by increasing the cut-off distance (Anselin, 2002) due to the asymptotical feature required for obtaining consistent estimates. In the absence of clear guidance about the choice of cut-off distance, empirical studies make use of the log-likelihood and R-squared values to compare estimation results based on different weight matrices (Abreu et al. 2004 and Seldadyo et al. 2010).

We define our distance-based weights, which depend on geographical distance d_{ij} measured as great circle distance between provinces i and j as follows:

$$w_{ij} = 0 \; if \; i = j \tag{2}$$

$$w_{ij=} \; 1/{d_{ij}}^2 \, , if \; i \; \neq j \; and \; d_{ij} < d^* \tag{3}$$

$$w_{ij} = 0, if \; i \; \neq j \; and \; d_{ij} > d^*, \tag{4}$$

Here, d^* is a cut-off point. The resulting matrix **W** is a square and symmetric matrix with 62 rows and 62 columns. While diagonal elements of **W** are set to zero so that no observation of FDI predicts itself, off diagonal elements presents weights associated with provinces.

$$W = \begin{bmatrix} 0 & w_{ji} & \vdots \\ w_{ij} & \ddots & \vdots \\ \ldots & \ldots & 0 \end{bmatrix} \tag{5}$$

We further standardize weight matrix **W** so that each row sums to unity.

$$ws_{ij} = w_{ij} \Big/ \sum\nolimits_j w_{ij} \tag{6}$$

Multiplying the spatial-weight matrix **W** with the vector of the *dependent* variable $lnFDI_{it}$, we obtain $W*lnFDI_{jt}$ as a new *independent* variable that consist of distance-weighted values of the dependent variable. For robustness check, we estimate the models with different specifications for the number of neighbouring provinces and cut-off distance values, including one nearest neighbour (denoted as **W1**), three nearest neighbours (**W3**), 186km cut-off distance (**W186**) and 350km cut-off distance (**W350**).

Quite often, the spatial lag model is preferred to the spatial error model. This is because the former allows for obtaining a rich set of estimates for the effects of a given explanatory variable - including direct, indirect and feedback effects. In addition, the spatial lag model also allows for establishing whether spatial dependence is reflected as conglomeration or competition effects in the distribution of FDI between spatial units (Blonigen et al. 2007). However, the choice between the two models must be based on Lagrange Multiplier (LM) test (Anselin, 1988) or robust LM test (Anselin, 1996) – as indicated above. In this article, we follow a decision rule that is based on the result of the robust LM test. The rule is to choose the more informative spatial lag estimation under two conditions: (i) if the robust LM tests justify this choice against the spatial error model; or (ii) if the robust LM tests indicate that both spatial error and spatial lag models are appropriate.

Spatial dependence can be modelled by augmenting equation (1) as follows:

$$\begin{aligned} lnFDI_{it} = \alpha + \beta_1 lnPCGDP_{it} + \beta_2 lnDI_{it} + \beta_3 BB_{it} + \beta_4 lnLC_{it} + \beta_5 lnOP_{it} \\ + \beta_6 lnEDU_{it} + \beta_7 PCI_{it} + \rho W * lnFDI_{jt} + \mu_i + \delta_t + \psi_{it} \end{aligned} \tag{7}$$

$$\psi_{it} = \lambda W \psi_{jt} + \varepsilon_{it} \tag{8}$$

The LM test for *spatial lag* tests the hypothesis whether ρ=0 in Equation (7) and LM test for *spatial error* tests if $\lambda=0$ in Equation (8). It is apparent from Equation (7) that *W*lnFDI* is correlated with the error term ε_{it} and therefore standard OLS will fail to produce consistent estimates (Anselin, 1988). This problem is demonstrated below (dropping the subscripts for notational simplicity):

$$lnFDI = \rho W lnFDI + \alpha + X\beta + \varepsilon \quad (9)$$

Here *lnFDI* is a vector of dependent variable, *X* is the matrix of explanatory variables, ρ is the spatial lag term parameter, α is a vector of constant term parameter, β is a vector of parameters for explanatory variables and ε is the classical error term. Equation (9) can be solved for the vector of *lnFDI* with simple algebra:

$$(I - \rho W) lnFDI = \alpha + X\beta + \varepsilon \quad (10)$$

$$lnFDI = (I - \rho W)^{-1}\alpha + (I - \rho W)^{-1}X\beta + (I - \rho W)^{-1}\varepsilon \quad (11)$$

where *I* is identity matrix. Due to the spatial multiplier matrix $(I - \rho W)^{-1}$, *lnFDI* in a province *i* depend not only on its own error term, but also on the error terms of other provinces. This is because $(I - \rho W)^{-1}$ is a full inverse, which yields an infinite series that involves error terms at all provinces $(I + \rho W, \rho^2 W^2 \ldots.)\varepsilon$[1]. As a result, the spatial lag term *W*lnFDI* also depends on the error term of other provinces. A common approach to this simultaneity problem is to use maximum likelihood (ML) estimation (Blonigen et al., 2007 and Seldadyo et al., 2010), which yields consistent and efficient parameter estimates in the presence of spatially lagged dependent variable (Anselin, 1988,2006).

It is also apparent from equation (11) that *lnFDI* in a province *i* is determined by factors in province *i* as well as those of neighbours because the term $(I - \rho W)^{-1}X\beta$ is equal to the right-hand side of equation (12) below.

$$(I - \rho W)^{-1}X\beta = X\beta + \rho W X\beta + \rho^2 W^2 X\beta + \rho^3 W^3 X\beta +, \ldots\ldots, +\rho^q W^q X\beta \quad (12)$$

Increasing powers of the matrix W (W^2, W3,… etc.) present neighbours set in more and more remote contiguity (second order contiguity is one's neighbours' neighbours and third order is

[1] Note that the first term is identity matrix *I* because $\rho^0 W^0$ equals *I*.

one's neighbour's neighbour's neighbours, and so on). Since ρ is smaller than one in absolute value, each successive term in equation (12) has smaller and smaller effect. This means that distant observations exhibit less and less influence as the expansion in equation (12) continues.

Once the coefficients are estimated with spatial lag model, LeSage and Pace (2009) proposes a calculation method that decomposes the total effect into direct and indirect effects. The direct effect refers to change in the dependent variable caused by explanatory variables for a spatial unit; whereas the indirect effect, which is also known as spatial spill-over effect, refers to changes in the dependent variable for other spatial units due to change in the explanatory variable of the unit in question. According to LeSage and Pace (2009: 40), the direct effect can be calculated as the average of the product of the point estimates (β) with the diagonal elements of the unit matrix *I* in Equation (13).

$$(I - \rho W)^{-1} \approx (I + \rho W + \rho^2 W^2 + \rho^3 W^3 +, \dots\dots, +\rho^q W^q)\beta \tag{13}$$

The identity matrix *I* multiplied with β represents the *direct effect* of a given explanatory variable on FDI in a given province. This effect does not include the *feedback effects* that percolate from neighbouring provinces into the province in question because the off-diagonal elements of the matrix *I* are all zero. On the other hand, the second term in parenthesis (ρW) multiplied by β represents the *indirect effects* of the corresponding variable *on the first-order neighbours* of the province in question. Remember that the diagonal entries in matrix *W* are zero; hence the indirect effect on the spatial unit itself is zero. The remaining terms in the parenthesis in equation (13) represent *indirect effects* on *second- and higher-order* neighbours as well as *feedback effects* from those neighbours onto the spatial unit itself. The *cumulative indirect effect* is obtained by summing up the indirect effects emanating from first- and higher-order neighbours. On the other hand, the *cumulative feedback effect* is obtained by adding up the feedback effects from second- and higher-order neighbours – leaving the first-order effect as the direct effect.

Our dataset covers 62 out of 63 provinces from six regions of Vietnam for the period 2006-2009[2]. Our data is obtained from General Statistics Office of Vietnam (GSO), with the exception of the Provincial Competitiveness Index (PCI) and informal charges. These governance quality proxies are collected through collaborative effort between the Vietnam Chamber of Commerce and Industry (VCCI) and the U.S. Agency for International Development(USAID) and the Asia Foundation[3].We exclude one province (Bac Lieu) for which data is incomplete. The omission is dictated by the need to have a balanced panel as a condition for carrying out spatial regression estimations using software package in MATLAB[4].

As the dependent variable, we use the natural logarithm of real registered FDI capital (*lnFDI*) in provinces, measured in Vietnamese Dong and deflated by the GDP deflator. Our FDI measure is then scaled by the population of each province (obtained from GSO) with a view to reduce the risk of heteroscedasticity related to scale (Baum, 2006).

In line with the empirical literature on locational determinants of FDI (Cole et al, 2009; Malesky, 2007; Pham 2002 and 2008; Segev, 2000), we use the log of provincial real GDP per capita (*lnPCGDP*) to capture the effect of provincial market on FDI. We expect higher levels of GDP per-capita to lead to higher levels of registered FDI. The log of domestic investment scaled by population (*lnDI*) is used to test the hypothesis whether domestic investment crowds out FDI or support it. Provinces offered various incentives and extra-legal tax holidays (fence-breaking) to attract FDI. This resulted in long-lasting budget deficit in

---

[2] See Appendix A1 for the list of provinces covered by our sample.
[3] PCI measures overall economic governance in Vietnam at province level and consists of nine sub-indexes: entry costs; land access and security of tenure; transparency and access to information; time costs of business start-ups; proactivity or local administration; informal charges; quality of business support services; labour training services; and legal institutions. Information regarding measurement and methodology of index construction is available on www.pcivietnam.org .
[4] We used sar_panel_FE function from http://www.regroningen.nl/elhorst/software/sar_panel_FE.m for our spatial lag model estimations.

provinces (Vu *et al*., 2007). Hence, we include budget balance to test whether there is correlation between FDI and budget balance of provinces. We use budget balance (*BB*) calculated as percentage of provincial GDP.

Trade openness of provinces may also impact the decision of multinationals with respect to location. Especially export-oriented multinationals firms may prefer provinces with already established trade links. We also take the natural logarithm of openness (*lnOP*), which is defined as sum of provincial exports and imports as percentage of provincial GDP. Labour costs are assumed to be an important component of production costs and hence an important determinant of competitiveness when FDI is motivated by export-seeking MNEs. Therefore, we use compensation per employee deflated by GDP deflator as a proxy for real wage in each province. We expect higher wages in a province to have a negative effect on provincial-level FDI in that province.

As far as human capital is concerned, we use the natural logarithm of number of students in lower-secondary (*lnLS*) and upper-secondary schools (*lnUS*) per 1000 people due to incompleteness of data for other proxies such as qualification levels of people in working age. We have used both measures of education to establish whether estimation results are sensitive to the type of education measure used. Finally, we include the Provincial Competitive Index (*PCI*) to measure the impact of governance quality on FDI. Furthermore, we use a sub-component of PCI, namely informal charges, to establish whether corruption (*CORRPT*) on its own has a significant effect on registered FDI; and to check whether the estimation results are sensitive to different measures of governance quality. Higher values of *PCI* and lower values of *CORRPT* indicate better governance, which we assume to have a positive impact on registered FDI.

5. Empirical results

We first estimated Equation (7) once with province dummies and once with time dummies. However, we do not report these estimation results because of three drawbacks associated with the inclusion of fixed-effect or time-effect dummies in estimations involving spatial dependence as part of the model. The first drawback is that spatial dependence may correlate with unobserved province fixed effects. Secondly, spatial effects may be present but subsumed within province dummies (Blonigen et al. 2007). As a result, estimation of spatial dependence together with unobserved province effects is highly inefficient. Third, our time dimension is very small and it is well known that time dimension of the sample should be sufficiently large in order to get consistent estimates for fixed effects. As suspected, inclusion of province dummies resulted in insignificant spatial term ($\rho W*lnFDI$) in our estimations. Furthermore, all province dummies are found to be individually insignificant but jointly significant regardless of weight matrix choice.[5] Blonigen et al. (2007) report similar results with respect to insignificant spatial dependence after adding country dummies. Finally, all time dummies are found to be individually and jointly insignificant although the spatial term ($\rho W*lnFDI$) is robust to inclusion of time dummies. Therefore, we estimated model (7) using the Maximum Likelihood (ML) method, excluding country and time dummies.

Table 3 below presents our findings for the determinants of registered FDI in Vietnamese provinces from 2006 -2009. Panel (1) reports the OLS estimation results without spatially lagged dependent variable. Panel (2) presents the results of the Maximum Likelihood estimations of the spatial lad model (equation 7) in which the spatially-lagged dependent variable ($W*lnFDI$) is included as explanatory variable. The ML estimation results and the Lagrange Multiplier (LM) tests in Panel (2) are based on different specifications for

[5] We do not report these results here, but they are available on request.

neighbouring provinces and cut-off values for distance between provinces. In column (W1), we estimate the model with one nearest neighbour; in column (W3) with three nearest neighbours; in column (W186) with a distance cut-off value of 186km; and in column (W350) with a distance cut-off value of 350km[6]. At the bottom of the table, we first report the results of the LM and robust LM tests for checking the presence of spatial dependence and for deciding whether a spatial error or spatial lag version of model (7) is appropriate. Then, we report the R^2 value for the OLS estimation and the corrected R^2 values for the spatial lag models along with the number of observations and log likelihood values.

Another feature of Panel (2) results in Table 3 is that they differentiate between direct and indirect effects of each explanatory variable, following the procedure proposed by LeSage and Pace (2009). As we have indicated above, the direct effect refers to change in the dependent variable (*lnFDI*) caused by explanatory variables within a given province. On the other hand, the indirect effect captures the change in the dependent variable within neighbouring provinces due to the change in the explanatory variable of the province under consideration.

Finally, we must indicate that the results in Panel (2) of Table 3 are derived by estimating a spatial lag rather than a spatial error model. The choice in favour of the spatial lag estimation is justified on two grounds, First, the LM test results indicate that spatial lag is the appropriate model for estimation with weight matrices W1, W3 and W186; and both spatial lag and spatial error models are appropriate for estimation with weight matrix W350. The robust LM test results, on the other hand, indicate that both spatial lag and spatial error models are

[6] We have also used two other matrices based on two nearest neighbours and distance cut off at 500km; and the results are remain unchanged.

Table 3: Determinants of FDI with different weight matrices for spatial dependence

	Panel (1)	Panel (2)			
	OLS	ML estimation with weight matrices			
		(W1)	(W3)	(W186)	(W350)
Constant	-16.295***	-15.288**	-12.912*	-16.308**	-16.343**
t value	(-2.24)	(-2.19)	(-1.84)	(-2.29)	(-2.29)
lnPCGDP					
Point estimate	1.285 ***	1.099***	1.104***	1.189***	1.206***
t value	(3.44)	(3.03)	(2.99)	(3.22)	(3.27)
Direct effect in province *i*		1.107***	1.128***	1.196***	1.201***
Indirect effect in provinces $j \neq i$		0.142**	0.240*	0.224	0.27
lnDI					
Point estimate	1.145***	1.257***	1.135***	1.172***	1.157***
t value	(4.81)	(5.45)	(4.93)	(5.03)	(4.97)
Direct effect in province *i*		1.266***	1.141***	1.176***	1.154***
Indirect effect in provinces $j \neq i$		0.167**	0.251**	0.228	0.27
lnLC					
Point estimate	-1.367***	-1.418***	-1.469***	-1.371***	-1.400***
t value	(-2.74)	(-2.95)	(-3.05)	(-2.81)	(-2.87)
Direct effect in province *i*		-1.432***	-1.512***	-1.392***	1.400***
Indirect effect in provinces $j \neq i$		-0.187*	-0.330*	-0.265	-0.323
lnOP					
Point estimate	0.594***	0.552***	0.554***	0.559***	0.557***
t value	(5.15)	(4.89)	(4.91)	(4.88)	(4.86)
Direct effect in province *i*		0.556***	0.558***	0.560***	0.562***
Indirect effect in provinces $j \neq i$		0.071**	0.120**	0.104	0.127
BB					
Point estimate	0.002	0.001	0.002	0.001	0.002
t value	(0.55)	(0.29)	(0.53)	(0.33)	(0.41)
Direct effect in province *i*		0.001	0.002	0.000	0.002
Indirect effect in provinces $j \neq i$		0.000	0.000	0.000	0.000
PCI					
Point estimate	0.033**	0.031**	0.028*	0.030**	0.030**
t value	(2.03)	(2.02)	(1.82)	(1.98)	(1.97)
Direct effect in province *i*		0.030**	0.028*	0.030**	0.031**
Indirect effect in provinces $j \neq i$		0.003	0.006	0.005	1.152
lnLS					
Point estimate	2.109***	1.977***	1.822***	1.891***	1.878***
t value	(3.51)	(3.42)	(3.13)	(3.22)	(3.20)
Direct effect in province *i*		1.957***	1.853***	1.922***	1.851***
Indirect effect in provinces $j \neq i$		0.254**	0.402*	0.359	0.421
W*lnFDI (Spatial dependence)		0.117***	0.179***	0.155*	0.186*
t value		(2.61)	(2.69)	(1.86)	(1.88)
Observations	248	248	248	248	248
LM No Spatial Lag		7.90***	5.37**	2.62*	2.18
Robust LM No spatial Lag		24.90***	16.26***	19.17***	18.67***
LM No Spatial Error		0.0242	0.02	0.81	0.71
Robust LM No spatial Error		17.02***	10.91***	17.36***	17.19***
R^2/Corrected R^2	0.457	0.481	0.477	0.472	0.471
Log Likelihood	-465.354	-461.550	-462.353	-463.925	-464.062

Note: t values are in parenthesis. ***, **,* denotes 0.01, 0.05, 0.10 significance level respectively.

appropriate for estimation with all weight matrices. This evidence implies that spatial dependence exists and that this dependence can be modelled either as spatial lag or as spatial error. Secondly, compared with the spatial error model, the spatial lag model allows for estimating a richer set of coefficients that include point estimates, direct effect estimates, indirect effect estimates, and feedback effect estimates. Given this information-rich feature of the spatial lag model and given that its estimation is justified under both the LM and robust LM tests, we report estimation results based on the spatial lag model only. [7]

In line with previous studies on Vietnam, the ***point estimates*** of the coefficients (with the exception of budget balance – BB) are statistically significant in the OLS estimation (Panel 1). The results are robust to adding spatially lagged dependent variable (*W*lnFDI*) in Panel (2), where we also report point estimates obtained with different weight matrix (W) specifications. The coefficient of the spatially-lagged dependent variable (*W*lnFDI*) is significant and indicates a positive relationship between registered FDI in a province and that in nearest neighbours or surrounding provinces. The spatial dependence captured by *W*lnFDI* indicates that registered FDI capital in a province increases by 1.1%, 1.8%, 1.5% and 1.9% as a result of 10 % per cent increase in the registered FDI of the nearest one neighbour, three nearest neighbours, surrounding provinces within a distance of 186km and those within a distance of 350km respectively. This positive relationship confirms the positive spatial autocorrelation in *lnFDI* results obtained from the *Moran s I* test, which are reported in Table A1 of the Appendix for each year and each weight matrix specification[8].

[7] We can indicate here that, as far as point estimates for the coefficients are concerned, the spatial error model yielded similar results to that of spatial lag model.

[8] Moran's *I* statistic tests whether provinces, which are located closer together are more likely to have similar registered FDI levels than those which are further apart. The null hypothesis for this tests states that there is zero spatial autocorrelation in the variable *lnFDI*.

Comparing the log likelihood and R^2 results for estimations with different weight matrices, we can see that the matrix with one nearest neighbour (**W1**)) yields the highest log likelihood R^2 values. A Monte-Carlo study carried out by Stakhovych and Bijmolt (2009) shows that the probability of finding the true specification increases if weight matrix selection is based on goodness of fit criterion. In addition, Elhorst (2010c) demonstrates that the value of the log-likelihood function should also be taken as a criterion for goodness of fit in spatial regression models. The combination of the two criteria implies that the weight matrix **W1** is the best specification for our data. Although the R^2 and log-likelihood values for estimations with other weight matrices (**W3**, **W186** and **W350**) are quite similar to those obtained with weight matrix W1, we follow the literature and use the estimation with weight matrix W1 as the benchmark results for sensitivity checks later.

As far as conventional explanatory variables are concerned, our ***point estimates*** indicate that higher levels of GDP per capita (*lnPCGDP*) and domestic investments per inhabitant (*lnDI*) lead to higher levels of registered FDI capital (*lnFDI*). In line with expectations, provinces that are more open to international trade are more attractive destinations for FDI. Furthermore, provinces with lower real wage costs tend to receive more FDI as the coefficient of labour cost (*lnLC*) carries a negative sign. The findings with respect to openness to trade and wage cost suggest that FDI in Vietnamese provinces may be motivated by lower wage costs as a source of competitive advantage to be exploited in international trade. This interpretation is justified by the fact that around 50% of Vietnam's export during the period under investigation (2006-2009) is realised by enterprises classified as FDI entities. The governance quality indicator (*PCI*) is positively related to FDI, albeit with small magnitude. Finally, our proxy for human capital (the number of pupils in lower secondary education - *lnLS*) is positively associated with FDI, implying that provinces with higher levels of lower-secondary education tend to receive more FDI.

The ***point estimates*** discussed above are un-biased and more efficient when compared to standard OLS estimates reported in Panel (1). As indicated above, OLS estimates are inefficient when spatial dependence is due to spatial autocorrelations between the error terms; and they are biased when spatial dependence is due to spatial correlation between the dependent variable (*lnFDI*) in province *i* and its neighbouring provinces. Comparing OLS estimates with point estimates from the spatial lag model, we can see that the former tend to over-estimate the effects of provincial per-capita GDP (*lnPCGDP*), openness to trade (*lnOP*) and lower-secondary school pupils (*lnLS*); and they underestimate the effects of domestic investment (*lnDI*) and labour cost (*lnLC*).

Although the ***point estimates*** discussed above are more relevant and reliable for inference, they may under- or over-estimate the true effect of each explanatory variable – depending on whether spatial dependence also leads to ***feedback effects*** that may be positive or negative. Stated differently, the ***point estimates*** overlook the likely presence of ***feedback effects***, which can be calculated as the difference between the direct effect and the point estimate (Elhorst, 2010). In what follows, we will focus on ***direct effects*** as the true measure of effects on registered FDI within a given province in response to a given change in one of the explanatory variables. This is because direct effect estimates include not only the point estimates but also the feedback effects - i.e., the effects that pass through neighbouring provinces and back into the province that instigates the change. On the other hand, we will focus on the ***indirect effect*** as the true measure of the how much a change in explanatory variable for province *i* affects registered FDI in all provinces with subscript $j \neq i$. [9]

[9] As noted by Elhorst (2010), direct and indirect effect estimates – unlike point estimates - are the true marginal effects (i.e., the partial derivatives of model 7). For calculating direct and indirect effect estimates in a spatial lag model, we used the 'panel_effects_sar' function in Matlab developed by Le Sage and Pace; and adapted for the spatial panel models by Elhorst at http://www.regroningen.nl/elhorst/software/panel_effects_sar.m.

Comparing ***direct effect*** and ***point estimates***, we can see that the direct effect are larger than the points estimates for four explanatory variables: per-capita GDP (*lnPCGDP*), domestic investment (*lnDI*), labour cost (*lnLC*), and openness to trade (*lnOP*). Hence relying on point estimates only would lead to under-estimated inference with respect to the effect of these explanatory variables. Under-estimation would range from about 0.5% to 2.2%.[10] With respect to remaining explanatory variables (the competitiveness index and labour cost), the difference between point estimates and direct effect estimates is too small. Although the magnitude of the feedback effects is small in this particular case, it is important to indicate that the feedback effects are positive. In other words, after a change occurs in the explanatory variable within a given province, the change pass through neighbouring provinces and leads to an increase in FDI within the province that instigates the change.

Comparing ***direct effects*** with ***indirect effects***, we observe that direct effects are always larger than indirect effects. This is to be expected because the change in explanatory variables for a given province will first and foremost affect registered FDI in that province. The effect on neighbouring provinces will tend to decline as the distance between the province itself and its neighbours increases. For example, the indirect effect of per-capita GDP (*lnPCGDP*) is 12% of the direct effect in column W1, where the weight matrix includes the nearest neighbour only. When we include the three nearest neighbours (column W3), the indirect effect is 21%. However, the indirect effect is usually insignificant when we increase the distance to 186km or 350 km. Reading down Table 3, we can see that indirect effect estimates are all significant when the weight matrix consists of one nearest province (W1) or 3 three nearest provinces (W3). These findings indicate that an increase in *lnPCGDP*, *lnDI*, *lnOP* and *lnLS* in a particular province is conducive to an increase not only in the registered FDI of that province (direct effects) but also an increase in the registered FDI of its neighbours (indirect effects).

[10] The under (over) estimation is equal to the feedback effect (or the difference between direct effects and point estimates) as percentage of the point estimate.

By the same token, if wages (*lnLC*) in a province decreases, not only the registered FDI of that province itself but registered of FDI of its neighbours will also increase.

Finally, the estimation results in Panel (2) indicate that the coefficient of the spatially-weighted FDI (*W*lnFDI*) is positive and significant with different specifications for the number of neighbouring provinces and distance cut-off values. This implies that FDI in neighbouring provinces has a positive effect on FDI in a given host province. This spatial effect does not diminish as the number of neighbouring provinces increases from 1 to 3 or the distance increases from 186km to 350km. Therefore, we can conclude that the distribution of FDI between Vietnamese provinces is subject to a conglomeration effect, whereby the existence of FDI in neighbouring provinces leads to higher levels of FDI in a province.

In what follows, we use the model estimated with weight matrix **W1** to check whether our findings would remain robust to a number of sensitivity checks. First, we control for the possibility of simultaneity and dual causality in the relationship between the dependent and independent variables by using one-period lags for the explanatory variables and the weight matrix **W1** that which yields the highest R^2 and log-likelihood function values. Because of using lagged explanatory variables, our observations reduce to 186 (Table 4). In general, the sign and significance of the point estimates and the direct effect estimates remain similar to those obtained with contemporaneous values in Table 3. In terms of magnitudes, estimation with lagged values yields slightly larger point estimates and direct effect estimates for lnPCGDP and lnDI; lower point estimates and direct effect estimates for secondary education (lnLS); and similar estimates for labour cost (lnLC), openness (lnOP) and governance index (PCI). The main difference between Table 3 and Table 4 concerns two indirect effects that have the same sign as before but are now statistically insignificant - the indirect effect of per-capita GDP (lnPCGDP) and labour costs (LnLC).

Table 4: ML estimation of FDI with spatial dependence:
Lagged explanatory variables and weight matrix W1

Lagged explanatory variables	Estimates
Constant	- 14.354*
t value	(-1.75)
lnPCGDP	
Point estimate	1.045**
t value	(2.50)
Direct effect	1.054**
Indirect effect	0.138
lnDI	
Point estimate	1.150***
t value	(4.37)
Direct effect	1.148***
Indirect effect	0.156*
lnLC	
Point estimate	-1.401**
t value	(-2.50)
Direct effect	-1.365**
Indirect effect	-0.183
lnOP	
Point estimate	0.535***
t value	(3.88)
Direct effect	0.536***
Indirect effect	0.070*
BB	
Point estimate	0.002
t value	(0.46)
Direct effect	0.002
Indirect effect	0.000
PCI	
Point estimate	0.029*
t value	(1.67)
Direct effect	0.028**
Indirect effect	0.003
lnLS	
Point estimate	2.331***
t value	(3.44)
Direct effect	2.334***
Indirect effect	0.309*
W*lnFDI (spatial dependence)	0.117** (2.24)
Observations	186
Corrected R^2	0.465
Log-likelihood	-347.578

Notes: t values are in parenthesis. ***, **,* denotes 0.01, 0.05, 0.10 significance level respectively.

Next, we check whether our results remain robust to changing the proxies for explanatory variables for which alternative measures exist. Since we have established that there is no discernible difference between the estimations with contemporaneous and lagged explanatory variables, we estimate the model with contemporaneous explanatory variables. Table 6 below reports the estimated results. Column (1) reports the results when we replace the number of lower secondary students per 1000 people (*lnLS)* with upper secondary students per 1000 people (*lnUS)*. Column (2) reports the estimation results when we use informal charges *CORRPT* instead of PCI. Since informal charges are components of PCI, we do not use them together.

According to the results reported in the first column and the second column in Table 5, the explanatory power of the model slightly improves when we use *CORRPT* and *lnUS* instead of *PCI* and *lnLS*. Furthermore, there is an increase in log-likelihood value in both Column 1 and Column 2 results. Although both *CORRPT* and *lnUS* are significant at 5% level, they do not have significant indirect effects. In line with expectations, these results show that informal charges (*CORRPT*) deter FDI in provinces in Vietnam, while the number of upper secondary students per 1000 people has a positive effect on FDI. Other explanatory variables and the lagged dependent variables *W*lnFDI* are robust to changing alternative proxies. Budget balance is still significant as in other estimation results.

Table 5: ML estimation of FDI with spatial dependence and weight matrix W1: Using alternative proxies for governance and education

		Column (1)	Column (2)
	Constant	-15.391**	-9.834
	t value	(-2.40)	(-1.47)
lnPCGDP			
	Point estimate	0.969***	1.160***
	t value	(2.75)	(3.35)
	Direct effect	0.981***	1.162***
	Indirect effect	0.099	0.126*
lnDI			
	Point estimate	1.151***	1.032
	t value	(5.07)	(4.45)
	Direct effect	1.165***	1.045***
	Indirect effect	0.120*	0.118*
lnLC			
	Point estimate	-0.984**	-1.144**
	t value	(-2.04)	(-2.37)
	Direct effect	-0.997**	-1.125**
	Indirect effect	-0.103	-0.125
lnOP			
	Point estimate	0.533***	0.601***
	t value	(4.80)	(5.40)
	Direct effect	0.536***	0.607***
	Indirect effect	0.054*	0.066**
BB			
	Point estimate	-0.003	-0.000
	t value	(-0.68)	(-0.14)
	Direct effect	-0.003	-0.000
	Indirect effect	-0.000	-0.000
PCI			
	Point estimate	0.031**	
	t value	(2.08)	
	Direct effect	0.032**	
	Indirect effect	0.003	
CORRPT			
	Point estimate		-0.306**
	t value		(-2.09)
	Direct effect		-0.302**
	Indirect effect		-0.033
lnUS			
	Point estimate	1.793***	1.470***
	t value	(4.55)	(3.85)
	Direct effect	1.795***	1.485***
	Indirect effect	0.182*	0.162*
W*lnFDI (spatial dep.)		0.096**	0.100**
		(2.14)	(2.24)
Observations		248	248
Corrected R^2		0.499	0.496
Log-likelihood		-457.057	-456.979

Notes:t values are in parenthesis. ***, **,* denotes 0.01, 0.05, 0.10 significance level respectively.

Conclusions

In this article, we have conducted an empirical investigation into the determinants of registered FDI capital across 62 Vietnamese provinces between 2006 and 2006. Our aim was to contribute to the literature with novel empirical findings, drawing on recent developments in spatial regression methodology and a unique dataset at the sub-national level.

First, we have established that OLS estimation ignoring spatial dependence tends to yield under-estimated or over-estimated coefficients. To address this shortcoming, we have carried out maximum likelihood (ML) estimation with spatial dependence and obtained unbiased estimates for a number of locational determinants of FDI examined in the literature at the national and/or sub-national levels. These determinants included per-capita GDP, domestic investment, openness to trade, budget balance, labour cost, governance quality and education at the provincial level. The point estimates obtained from the ML estimation are in line with existing evidence at the national and sub-national levels; and they remain robust to inclusion of the spatially-lagged dependent variable and to different specifications for weight matrices capturing the number of neighbouring provinces or distance between provinces.

Our findings, however, contribute to existing evidence in a number of ways. First, they provide the first estimates of the spatial dependence in the distribution of registered FDI capital between Vietnamese provinces. The sign of the spatial dependence is positive and remain robust to change in the specification of the weight matrix from one nearest neighbour and 3 nearest neighbours to distance cut-off values of 186km and 350km. Although the significance of the spatial dependence decreases from 1% to 10%, the magnitude tend to increase as the number of neighbouring provinces or as distance between provinces increases. This finding indicates that the distribution of registered FDI between Vietnamese provinces is subject to conglomeration effects, whereby FDI inflows to neighbouring provinces have a positive effect on FDI flows into a given province. A 10% increase in FDI registered in neighbouring provinces tends to lead to an increase of 1.2% to 1.8% in FDI of a given province.

Secondly, we demonstrate that the point estimates for determinants of FDI conceal the potential existence of feedback effects and therefore one needs to measure direct effect estimates as true marginal effects. Our findings indicate that the point estimates tend to under-estimate the true marginal effects of per-capita GDP, domestic investment, labour cost and openness to trade. Drawing on a recently-proposed estimation procedure, not only do we highlight the limitation of the point estimates but also we provide direct effect estimates that incorporate both the point estimates and the feedback effects from neighbouring provinces. Although the feedback estimates are small, they have a positive sign and as such they are consistent with the conglomeration effect established through the coefficient of spatial dependence.

Finally, we have added to the existing evidence base by breaking down the effects of the explanatory variables on provincial-level FDI into direct and indirect effects. We have found that a one-unit change in a given explanatory variable first and foremost affects the FDI in a given host province. This is the direct effect, which includes second- and higher-order feedback effects that flow from neighbouring provinces affected by the shock in the host province back into the host province in question. Our findings indicate that direct effect estimates are larger than indirect effect estimates. Estimates of indirect effects on neighbouring provinces are smaller than direct effects within the host province, but their magnitude is significant enough to warrant special attention. Our findings indicate that indirect effects on neighbouring provinces are about 12% - 20% of the direct effects on FDI within the host province.

Appendix

Table A1: List of Provinces in the Sample

REGIONS	PROVINCES
Central Highlands	Dak Lak, Lam Dong, Dak Nong, Gia Lai, Kon Tum
Mekong River Delta	An Giang, Hau Giang, Thai Binh, Vinh Long, Soc Trang, Ca Mau, Long An, Can Tho, Kien Giang, Tra Vinh, Ben Tre, Dong Thap, Tien Giang
North Central and Central Coastal area	TT-Hue, Khanh Hoa, Quang Binh, Quang Nam, Nghe An, Ninh Thuan, Da Nang, Binh Dinh, Phu Yen, Quang Ngai, Ha Tinh, Quang Tri, Thanh Hoa, Binh Thuan
Northern midlands and mountain areas	Lai Chau, Thai Nguyen, Dien Bien, Lang Son, Cao Bang, Bac Kan, Ha Giang, Lao Cai, Yen Bai, Son La, Hoa Binh, Tuyen Quang, Phu Tho, Bac Giang
Red River	Hung Yen, Quang Ninh, Ha Nam, Nam Dinh, Hai Duong, Ninh Binh, Hai Phong, Bac Ninh, Ha Noi, Vinh Phuc
South East	Dong Nai, Binh Duong, Binh Phuoc, BRVT, Tay Ninh, HCMC

Table A2: *Moran s I* Test for Spatial Autocorrelation *lnFDI*

Moran s I test	W1	W3	W186	W350
lnFDI 2006	0.426	0.290	0.181	0.141
	(0.00)	(0.00)	(0.00)	(0.00)
lnFDI 2007	0.449	0.317	0.203	0.161
	(0.00)	(0.00)	(0.00)	(0.00)
lnFDI 2008	0.419	0.312	0.204	0.163
	(0.02)	(0.00)	(0.00)	(0.00)
lnFDI 2009	0.438	0.325	0.221	0.178
	(0.00)	(0.00)	(0.00)	(0.00)

Notes: Two-sided and under normality. P-values are in parenthesis.

Table A3: Descriptive statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
lnFDI	248	15.0655	2.14910	7.74161	19.49431
lnPCGDP	248	16.0222	0.51743	15.00167	18.64245
lnLC	248	14.2117	0.27031	13.52721	15.15096
lnDI	248	15.1399	0.51487	13.24763	16.84037
BB	248	-3.05303	25.74335	-129.34330	60.50403
lnOP	248	3.57076	1.18553	-0.06860	6.61966
PCI	248	55.38839	7.86282	36.39000	77.20000
CORRPT	248	6.43181	0.72146	4.63000	8.35000
lnLS	248	4.22414	0.19556	3.53035	4.65764
lnUS	248	3.52646	0.27551	2.64107	4.05648
W*lnFDI	248	15.49943	2.10583	7.74161	19.49431

Table A4: Correlation matrix

	lnFDI	lnGPC	lnLC	lnDI	BB	lnOP	PCI	CORRPT	lnLS	lnUS	WlnFDI
lnFDI	1.0000										
lnPCGDP	0.5501*	1.0000									
lnLC	0.2521*	0.6396*	1.0000								
lnDI	0.4338*	0.5095*	0.4136*	1.0000							
BB	0.3498*	0.5719*	0.3602*	0.1571*	1.0000						
lnOP	0.5436*	0.6373*	0.3538*	0.2710*	0.4296*	1.0000					
PCI	0.3433*	0.4745*	0.2758*	0.1635*	0.3786*	0.4384*	1.0000				
CORRPT	- 0.0683	0.1034	-0.0086	-0.1496*	0.1270*	0.1807*	0.3097*	1.0000			
lnLS	-0.1061*	-0.3909*	-0.2551*	-0.2142*	-0.1511*	-0.3266*	-0.4113*	-0.1407*	1.0000		
lnUS	0.2095*	-0.0508	-0.1709*	-0.0306	0.1999*	-0.0320	-0.1662*	-0.1405*	0.7042*	1.0000	
W*lnFDI	0.4373*	0.5080*	0.3231*	0.0549	0.4463*	0.4634*	0.3306*	0.0901	-0.1282*	0.1617*	1.0000

* denotes significance at the 10% level.

References

Abreu, M., H.L.F. de Groot and J.G.M. Florax (2004), ‘Space and Growth: A Survey of Empirical Evidence and Methods’, Tinbergen Institute Working Paper no. TI 04-129/3.

Anselin, L. (1988), ‘Spatial Econometrics, Methods and Models’, Kluwer Academic Publisher, Dordrecht.

Anselin, L., A.K. Bera, R. Florax and M.J. Yoon (1996), ‘Simple Diagnostic Tests for Spatial Dependence’, *Regional Science and Urban Economics*, **26**, 77–104.

Anselin, L. and A. Bera (1998), ‘Spatial Dependence in Linear Regression Models with an Introduction to Spatial Econometrics’ , in: A. Ullah and D. Giles (eds), Handbook of Applied Economic Statistics, Springer, Berlin.

Anselin, L. (2002), ‘Under the Hood. Issues in the Specification and Interpretation of Spatial Regression Models’, *Agricultural Economics*, **27**, 247–267.

Anselin, L. (2006), ‘Spatial Econometrics’, in: Mills, C. and K. Patterson (eds), Palgrave handbook of econometrics’, Terence Palgrave McMillan, Basingstoke.

Baltagi, B. H., P. Egger and M. Pfaffermayr (2007), ‘Estimating Models of Complex FDI: Are there Third-Country Effects?’, *Journal of Econometrics*, **140**, 1, 260–281.

Baum, F.C. (2006), ‘An Introduction to Modern Econometrics Using Stata’, Stata Press, Texas.

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**, 5, 1303-1325.

Cole, A.M., R.J.R. Elliott and J. Zhang (2009), ‘Corruption, Governance and FDI Location in China: A Province-Level Analysis’, *Journal of Development Studies*, **45**, 9, 1494–1512.

Coughlin, C.C. and E. Segev (2000), ‘Foreign Direct Investment in China: A Spatial Econometric Study’, *The World Economy*, **23**, 1, 1-23.

Drukker, D. and D.L. Millimet (2007), 'Assessing the Pollution Haven Hypothesis in an Interdependent World', working paper 0703, Southern Methodist University.

Du, J.L., Y. Lu and Z.G. Tao (2008), 'Economic Institutions and FDI Location Choice: Evidence from US Multinationals in China', *Journal of Comparative Economics*, **36**, 3, 412–429.

Elhorst, J.P. (2010a), 'Spatial Panel Data Models', in Handbook of applied spatial analysis, in M.M. Fischer and A. Getis(eds), Berlin: Springer, pp. 377-407.

Elhorst, J.P. (2010b), 'Matlab Software for Spatial Panels', Paper presented at the IVth World Conference of the Spatial Econometrics Association (SEA), Chicago, June 9-12, 2010.

Elhorst, J.P. (2010c), 'Applied Spatial Econometrics: Raising the Bar', *Spatial Economic Analysis*, **5**, 1, 9–28.

Garretsen, H. and J. Peeters (2009), 'FDI and the Relevance of Spatial Linkages: Do Third-Country Effects Matter for Dutch FDI?', *Review of World Economics*, **145**, pp. 319 – 338.

General Statistics Office (GSO) of Vietnam, www.gso.gov.vn, accessed [15 November 2011].

Head, K. and T. Mayer, (2004), 'Market Potential and the Location of Japanese Investment in the European Union', *Review of Economics and Statistics*, **86**, 4, 959–972.

Head, K., J. Ries and D. Swenson, (1995), 'Agglomeration Benefits and Location Choice: Evidence from Japanese Manufacturing Investments in the United States', *Journal of International Economics*, **38**, (3–4), 223–247.

Ledyaeva, S. (2009), 'Spatial Economic Analysis of Foreign Direct Investment Determinants in Russian Regions', *The World Economy*, 643–666.

LeSage, J.P. and R.K. Pace (2009), 'Introduction to Spatial Econometrics', Boca Raton, US: CRC Press Taylor & Francis Group.

Malesky, E. (2007), 'Provincial Governance and Foreign Direct Investment in Vietnam', 20 Years of Foreign Investment: Reviewing and Looking Forward (1987–2007), Knowledge Publishing House, 2007.

Moran, P. A. P. (1950), 'Notes on Continuous Stochastic Phenomena', *Biometrika*, **37**, 1–2, 17–23.

Na, L. and W.S. Lightfoot (2006), 'Determinants of Foreign Direct Investment at the Regional Level in China', *Journal of Technology Management in China*, **1**, 3, 262–278.

Nguyen, P.L. (2006), 'Foreign Direct Investment and its Linkage to Economic Growth in Vietnam: A Provincial Level Analysis', *mimeo*, Centre for Regulation and Market Analysis, University of South Australia.

Nguyen, N.A. and T. Nguyen (2007), 'Foreign Direct Investment in Vietnam: An Overview and Analysis the Determinants of Spatial Distribution Across Provinces', MPRA Paper No. 1921.

Pham, H.M. (2002), 'Regional Economic Development and Foreign Direct Investment Flows in Vietnam, 1988-98', *Journal of the Asia Pacific Economy*, **7**, 2, 182–202.

Pham, T.H. (2008), 'The Effects of ODA in Infrastructure on FDI Inflows in Provinces of Vietnam, 2002-2004', Working paper 089, Vietnam Development Forum, June 2008.

Provincial Competitiveness Index (PCI), www.pcivietnam.org, accessed [25 October 2011].

Seldadyo, H., J.P. Elhorst and J. de Haan (2010), 'Geography and Governance: Does Space Matter? ', *Papers in Regional Science*, **89**, 3, 625–640.

Sharma, S., M.G. Wang and M.C.S. Wong (2010), 'A Spatial Analysis of Aggregate and Industry-Level FDI in China', Presented at the Southern Economic Association 2010 Conference, Atlanta, Georgia, November 2010.

Stakhovych, S. and T.H.A. Bijmolt (2009), 'Specification of Spatial Models: A Simulation Study on Weights Matrices', *Papers in Regional Science*, **88**, 389–408.

UNCTAD FDI Database, www.unctad.org, accessed [31 October 2011].

Villaverde, J. and A. Maza (2011), 'Foreign Direct Investment in Spain: Regional Distribution and Determinants', *International Business Review*, doi:10.1016/j.ibusrev.2011.08.004

Vu, T.T.A., V.T. Le and T.T. Vo (2007), 'Provincial Extralegal Investment Incentives in the Context of Decentralisation in Viet Nam: Mutually Beneficial or a Race to the Bottom?', working paper.