Spatial Spillover Effects of Foreign Direct Investment Flows to Arab Countries Based on Static and Dynamic Spatial Panel Data Models: A Spatial Panel Modelling Study

تأثيرات الانتشار المكانى لتدفقات الإستثمار الأجنبى المباشر إلى الدول العربية بالإعتماد على نماذج بيانات البانل المكانية الساكنة والديناميكية: دراسة نمذجة البانل المكانية

Wael Saad Hsanein El-doakly
Lecture
Statistic, Mathematics & Insurance Department
Faculty of Business, Ain Shams University
Waelsaad573@yahoo.com
01001478841

المستخلص:

يفترض نموذج البانل الديناميكى أن كل وحدة مشاهدة مستقلة عن بعضها البعض. ولكن فى بعض الأحيان يُنتهك هذا الافتراض لوجود تأثيرات مكانية فى النموذج. ويُعد الارتباط الذاتى المكانى أحد أساليب التحليل المكانى لتحديد أنماط العلاقة أو الترابط بين المواقع. وتُعد هذه الطريقة بالغة الأهمية للحصول على معلومات حول أنماط التشتت المميزة لمنطقة ما والروابط بين المواقع. كما يُشار إلى الارتباط بين المواقع بمصفوفة الأوزان المكانية، والتى تصف بنية المناطق المجاورة وتعكس التأثير المكانى. ويُعد اختيار دالة الأوزان أحد العوامل المحددة لنتائج التحليل المكانى. هدفت هذه الدراسة إلى نمذجة الاستثمار الأجنبى المباشر فى الدول العربية باستخدام النماذج المكانية خلال الفترة ٢٠١٥-٢٠٢٤. وتتمثل النماذج المستخدمة فى هذه الدراسة فى: الانحدار الذاتى المكانى (SAR)، ونماذج الخطأ المكانى (SEM)، ونموذج دوربين المكانى (SDM)، ونموذج دوربين المكانى الديناميكى (DSDM). تم تقييم النماذج الأربعة باستخدام معيار معلومات أكايكى (AIC) ومعيار المعلومات البايزى (BIC)، ومعامل التحديد المعدل ($Adj\text{-}R^2$). حيث

أظهرت النتائج وجود تداخل مكانى فى جذب الاستثمار الأجنبى المباشر إلى المنطقة المذكورة. وفى الوقت نفســه تؤثر عوامل مثل الإنفتاح التجارى وحجم الســوق والنمو السـكانى والبنية التحتية وتجمع الشـركات على تدفقات الاسـتثمار الأجنبى المباشـر على المديين القصــير والطويل، وبشــكل خاص يزيد الإنفتاح التجارى لدولة ما من تدفقات الاســـتثمار الأجنبى المباشـــر إليها وإلى الدول المجاورة، كما يزيد تجمع الشــركات فيها من تدفقات الاســتثمار الأجنبى المباشــر إليها وإلى الدول المجاورة. وأظهرت نتائج أخرى أن نموذج (DSDM) تمكن من تفســـير التنوع فى النمـاذج بنســبة ٩٢.٥٪. وبالتالى فإن نموذج (DSDM) هو النموذج الأمثل لنمذجة بيانات البانل الديناميكى الخاصـة بالاستثمار الأجنبى المباشر فى الدول العربية خلال الفترة ٢٠١٥-٢٠٢٤.

الكلمات المفتاحية

التحليل المكانى الديناميكى،تأثيرات الإنتشـــار المكانى،نماذج بيانات البانل المكـانيـة،الإنحـدار الـذاتى المكـانى،نمـاذج بيـانـات الخطـأ المكـانى،نموذج دوربين المكانى،نموذج دوربين المكانى الديناميكى،تبعية القطاع المستعرض،مصفوفة الأوزان المكانية،تقدير الإمكان الأكبر،الإستثمار الأجنبى المباشر.

Abstract

The dynamic panel model assumes that each observation unit is independent of each other. But sometimes this assumption is violated, so there are spatial effects in the model. As spatial autocorrelation is one of spatial analysis to identify patterns of relationship or correlation between locations. This method is very important to get information on the dispersal patterns characteristic of a region and linkages between locations. The link among location indicated by a spatial weight matrix. It describe the structure of neighboring and reflects the spatial influence. Selection weighting function is one determinant of the results of the spatial analysis. This study aimed to make modeling of foreign direct investment (FDI) in Arab countries using the Spatial Models during the period 2015-2024. The models used in this study are Spatial Auto Regressive (SAR), Spatial Error Models (SEM), Spatial Durbin Model (SDM), Dinamic Spatial Durbin Model (DSDM). The four models were evaluated using the Akaike Information Criterion (AIC),Bayesian Information Criterion(BIC), and Adj-R^2. The results show that there is a spatial interaction in FDI attraction to the above mentioned region. At the same time, factors such as trade openness, market size, population growth, infrastructure and enterprise agglomeration affect FDI inflows in the short and long term. In particular, the trade openness of a country increases FDI inflows of that country and neighboring countries, and enterprise agglomeration of a country increases the FDI inflows of that country and in adjacent countries. Other results showed that (DSDM) was able to explain the diversity in models at 92.4%. Thus, the result of the study proved that (DSDM) is the best model for modelling dynamic FDI panel data in Arab countries during 2015-2024.

Key words

Dynamic spatial analysis, Spatial spillover effects, Spatial panel data models, Spatial Auto Regressive (SAR), Spatial Error Models (SEM), Spatial Durbin Model (SDM), Dynamic Spatial Durbin Model (DSDM), Cross-sectional dependence, Spatial weights matrix, Maximum Likelihood Estimation (MLE), Foreign Direct Investment.

1- Introduction

Spatial regression is the development of simple linear regression, in which there exists a spatial effect on the data to be analyzed. The spatial effects that arise on the spatial regression are spatial heterogeneity and spatial dependence. The spatial heterogeneity shows the differences in characteristics from one region to another, while the spatial dependence shows the dependence between adjacent areas (LeSage and Pace,2017)[21]. In general, there are several types of spatial models, namely Spatial Autoregressive Model (SAR), Spatial Error Model (SEM), Spatial Durbin Model (SDM), and Dynamic Spatial Durbin Model (DSDM).

According to different forms of error shock on an observation's spatial autocorrelation, spatial econometric models can be divided into the Spatial Error Model (SEM) and the Spatial Autoregressive Model (SAR). In this research, the (SEM) assumes that spatial autocorrelation stems from the error shock of neighboring regions on the dependent variables and examines the effects of neighboring regions under observation, while the (SAR) assumes that spatial autocorrelation stems from dependent variables and examines the effects of Foreign Direct Investment (FDI) in neighboring regions on local FDI for the Arab countries. According to the principles of model construction, the (SEM) reflects the indirect effects caused by error terms, whereas the (SAR) controls the direct effects. LeSage and Pace(2009)[18] further establish the Spatial Durbin Model (SDM) on the basis of the (SEM) and the (SAR). The (SDM) includes spatial lag terms from dependent variables and independent variables to capture the spillover effects deriving from different variables.

The static Spatial Durbin Model (SDM) and dynamic Spatial Durbin Model (DSDM)will also be constructed to investigate how economic decentralization affects regional FDI for the Arab countries. Meanwhile, in order to reflect the FDI's possible inertia characteristics in reality, the lag item (FDI_{t-1}) of FDI will be introduced to the independent variables to measure the influence of self-reinforcing mechanism.

The SAR model only considers the effect of the spatial lag on the independent variables. Meanwhile, the (SDM) model considers the effect of the spatial lag on the independent variables and the dependent variable. The SDM therefore has been widely applied, such as (Feng and Chen,2018)[14] which discusses environmental regulation, green innovation, and green industrial development, (Feng et al.,2018)[15] which discusses the effects of air pollution control on urban development quality in Chinese cities based on the (SDM), (Triki and Maktouf ,2012)[33] who studies the factors associated with the emergence of banking crises during the process of financial liberalization, (Putri, 2022)[30] model the factors that influence the Gender Development Index (GDI) on the island of Sulawesi. The results obtained are factors that significantly influence the GDI on Sulawesi Island using the (SDM) namely life expectancy, per capita expenditure, average length of schooling, and labor force participation rate.

The main advantage of The Dynamic Spatial Durbin Model (DSDM) is its ability to test the existence of endogenous; as well as exogenous interaction effects in the short term along with the long term (LeSage and Sheng, 2014).[19] It was not only used to examine the effect of the independent variables on the dependent variable in the local and surrounding regions but also to test the spatial dependence, temporal dependence, and spatiotemporal dependence of the dependent variable (Debarsy et al., 2012; [7] LeSage and Sheng, 2014; [19] Elhorst, 2014a, b[12, 13]). Lesage and Pace (2009) [18] noted that estimation of spatial models by least squares can be lead to inconsistent estimates of the regression parameters for models with spatially lagged dependent variables, inconsistent estimation of the spatial parameters and inconsistent estimation of standard errors. In contrast, Maximum Likelihood Estimation (MLE) is consistent for these models. In addition several references have written about parameter estimation of spatial autoregressive model. There was Maximum Likelihood Estimation (MLE) which noted by Anselin and Rey (2010).[3] Baltagi and Bresson (2011)[6] has been study about MLE and lagrange multiplier tests for panel Seemingly Unrelated Regressions with spatial lag and spatial errors. In spatial analysis, it need weighting to show the neighboring among locations. This weighting is expressed in a spatial weighting matrix, namely W. This matrix gives an important role in any analysis of spatial data, because it shows the location, neighborhoods, and the relationship about distance among locations nearby. Selection of the type of matrix is also a major concern in spatial analysis. Specifications of the weighting matrix is represent the information and intensity of the spatial effect of a unit space locations in the geographic system. Wang et al.(2013)[34] state that adjacency relationships, distance relationships and comprehensive factors relationships are summarized to characterize the definitions of spatial weight matrix. According to the spatial data, type of weighting matrix can be divided into two types: point type (distance) and the neighborhood area (contiguity). The distance weighting matrix involve an element of distance between the locations whose value is continuous in building the weighting matrix, so that each location will receive weighted according to that distance. Meanwhile, contiguity weighting involve an element of area or neighborhoods. The location that neighboring directly shows the higher spatial dependency than the locations far apart.

Despite a growing literature on FDI, there is no' study analyzing FDI regarding spatial effects between neighboring Arab countries. Therefore ,this study uses spatial econometrics to analyze the spatial dependence and main determinants of FDI between 17 Arab countries. In addition, the spatial models are used to estimate the short-term and the long-term direct, indirect, and total effects of trade openness, market size, population growth, infrastructure and enterprise agglomeration on FDI.

In this context, the objective of this study is to examine the determinants of FDI in the Arab countries. Therefore, this study has the following important contributions to the existing literature.

First, The significance of introducing a dynamic spatial panel model is as follows:

1- Spatial factors are introduced to reflect the spatial correlation and spatial effects of FDI in Arab countries .

2- The approach can be used to analysis not only the long-term impact but also the short-term impact of trade openness, market size, population growth, infrastructure and enterprise agglomeration on FDI in Arab countries.

3- Introducing lagging explanatory variables as independent variables into the model and incorporating those that are not included in the econometric model to strip the impact of hidden factors on FDI in Arab countries make the model's estimates more accurate.

Second, we determine the difference between the direct and indirect effects of trade openness, market size, population growth, infrastructure and enterprise agglomeration on FDI in Arab countries in terms of the short-term and the long-term. Regarding this contribution of this study, we investigate the direct and indirect effects of factors affecting the FDI in the short-term and the long-term horizons using the spatial models.

Third, the paper uses a dynamic spatial model to analysis the factors affecting FDI inflows. This is because economic dynamics have many to do with geospatial, even more so when it is enhanced by the location of economic resources. At the same time, the application of spatial analysis can assess the spatial spillover of economic phenomena (Pan et al., 2021;[28] Ponce et al., 2020).[31] However, the use of the static spatial model only analyses the spatial correlation between a locality and neighboring localities in FDI attraction. It emphasises the importance of factors influencing FDI but does not provide sufficient spatial spillover effects in the short and long term. This omits the stability or spatial cointegration of the determinants of FDI. If different contexts or dynamics change, the former determinants of FDI produce different effects (Paul and Jadhav, 2019).[29] Therefore, the dynamic spatial model not only determines the spatial interaction of FDI inflows but also provides a novel perspective on the key factor in attracting FDI in Arab countries. This issue both increases understanding of FDI and provides a complete assessment of how economies operate in Arab countries.

2- Methodology and data

2-1 Study area

This study uses spatial econometrics to analyze the spatial dependence and main determinants of FDI between 17 Arab countries (According to available data), which includes Emirates, Bahrain ,Kuwait, Saudi Arabia, Yemen, Qatar, Oman, Jordan, Syria, Lebanon, Iraq, Sudan, Egypt, Libya, Tunisia, Algeria and Morocco. In addition, data were collected using the World Bank database during the period 2015-2024, and the data were analyzed using the statistical package R.

2-2 Dynamic Spatial Panel Model Specifications

Spatial econometrics is a discipline of economics that studies spatial aspects, specifically the relationship between regions and variations in regional structures (Elhorst et al. 2013; LeSage 2015; Arogundade et al. 2022).[10, 20, 5] Following the first law of geography, everything is related to everything else, but near things are more connected than things that are far away (Anselin, 2007)[2]. Additionally neighboring countries and countries located outside the region can be interrelated. Spatial dependence occurs because these countries are important trading partners (Arogundade et al. 2022)[5].

Spatial dependence models are typically developed first by defining a spatial connectivity matrix (Anselin 2007; Li et al. 2023).[2, 24] The matrix depicts location dependencies. Hence, the size of the spatial weight significantly affects the estimation of the spatial dependence models. There are several ways to determine the spatial weight. The spatial weighting matrix is determined based on the proximity of geographical relationships (contiguity) or distance weights (Arogundade et al. 2022).[5]

A spatial contingency weighting matrix is based on neighboring relationships. The weighting matrix using the inverse distance method was determined based on the actual distance between locations (Elhorst et al. 2014; Li et al. 2023).[11, 24] The inverse matrix gives larger weight values for shorter distances and smaller ones for longer distances.

2-3 Spatial Weight Matrix

Spatial Weight Matrix (W_{ij}) is the core element of spatial econometric models to reflect the spatial relationship of countries. To comprehensively analyze the spatial correlation characteristics of foreign direct investment, this study adopted two spatial weight matrices based on the existing literature (Feng et al., 2020[16]; Quito et al., 2023)[32]. The first is Queen contiguity spatial weights matrix (W_q) or matrix of intersections of sides and corners that defines $w_{ij} = 1$ for areas that are side by side (common side) or where the corners meet the area of concern, and $w_{ij} = 0$ for other areas that do not share a common side (Ashari et al., 2020)[4]. The second is an inverse distance-based spatial weights matrix (W_d). In this matrix, geographical distance criterion is referenced to determine the degree of spatial autocorrelation, namely, the lower the geographical distance, the higher the spatial autocorrelation will be. The forms of the two matrices are as follows:

$$W_q = \begin{cases} 1, \text{if country i shares a common boundary} \\ \text{or a common vertex with country j.} \\ 0, \text{otherwise} \end{cases}$$

$$W_d = \begin{cases} W_{ij} = \dfrac{1}{d_{ij}} & , i \neq j \\ W_{ij} = 0 & , i = j \end{cases}$$

2-4 Spatial Analysis

At the time when the observed data contains location and time information, so it tends to be spatially correlated (spatial autocorrelation). Spatial autocorrelation is condition where the value of observations at a location or region is influenced by the observed value of its closest neighbor. Spatial autocorrelation could be detected with Moran's I test[1]. At the time when spatial autocorrelation is indicated, it is needed a model that includes spatial autocorrelation that is spatial panel data model. Spatial panel data is compound of observations that contain location information (postal code, region, state, country, etc.) and observed in several periods. Spatial panel data model is model that describes the interaction of location in several periods. The spatial panel data model controls spatial interactions at location and time[25]. Spatial panel data model is divided into four, namely spatial lag panel data model, spatial error panel data model, spatial Durbin model and dynamic spatial Durbin model.

2-5 Spatial Lag Panel Data Model

Spatial lag panel data model or spatial autoregressive model (SAR) is model that the dependent variable depends on the dependent variable observed in neighboring units (observed local characteristics) or the dependent variables correlate between location[9]. The following is the equation for the spatial lag panel data model[11]:

$$Y_{it} = \rho\sum_{j=1}^{N} w_{ij}Y_{jt} + x_{it}\beta + \mu_i + \varepsilon_{it};\ i=1,2,\ldots N,\ t=1,2\ldots,T \quad (1)$$

where i is index for location or spatial unit with i = 1, 2, ….., N,
t is index for time periods with t = 1, 2, …, T.
Y_{it} is observation on dependent (response) variable at i and t,
x_{it} is (1 x K) vector of explanatory variable at i and t,
ρ is spatial autoregressive coefficient,
w_{ij} is element of spatial weight matrix W,
μ_i is spatial specific effect at i, $\varepsilon_{it} \sim NIID\ (0\ ,\ \sigma^2)$.
The estimation of parameters in the spatial lag panel data model is using the Maximum Likelihood (ML) method.

2-6 Spatial Error Panel Data Model

Spatial error panel data model (SEM) is model that has a spatial dependence on correlated errors between location[8]. The spatial dependence in the panel data spatial model error is in the error or nuisance dependence. The following is the equation for the spatial panel data model error[11].

$$Y_{it} = x_{it}\beta + \mu_i + \phi_{it}\ ,\ \text{where}\ \phi_{it} = \rho\sum_{j=1}^{N} w_{ij}\phi_{it} + \varepsilon \quad (2)$$

With ϕ_{it} is spatially autocorrelated error term,
and ρ is spatial autocorrelation coefficient.
The estimation of parameters in the spatial error panel data model is using the Maximum Likelihood (ML) method.

2-7 Spatial Durbin model (SDM)

The (SAR) model contains a spatially lagged dependent variable (Elhorst, 2014)[11]. The (SEM) model incorporates spatial autocorrelation in the error term (Lee and Yu, 2010; Zhou et al., 2023)[23, 43]. The (SAR) and (SEM) models do not consider spatially lagged independent variables in explaining the dependent variable, leading to specification bias (Elhorst, 2010; Jiang et al., 2018)[9, 17]. Hence, the (SDM) model contains both the spatially lagged independent and dependent variables. In addition, the SDM model produces consistent and unbiased estimates (Elhorst, 2014)[11]. For this reason, the (SDM) model has been widely used in various studies in the field of environment. (Wang et al., 2023a,b,c; Zhao and Sun, 2022)[35, 36, 37, 44]. The SDM model is as follows:

$$Y_{it} = \rho WY_{it} + \beta X_{it} + \theta WX_{it} + \mu_i + \lambda_t + \varepsilon_{it} \qquad (3)$$

where ρ denotes the spatial lag coefficient of the dependent variable.

W denotes the spatial weight matrix.

X_{it} represent the independent variables in country i in year t.

β represents the influence of the independent variables on dependent variable.

θ is the spatial lag coefficients of the independent variables.

μ_i and λ_t denote the space fixed effect and time fixed effect, respectively.

and ε_{it} denotes the random error vector.

2-8 Dynamic Spatial Durbin Model (DSDM)

The short-term effect of the independent variables cannot be calculated with static (SDM) (Zhou et al., 2023)[43]. In the environmental system, the effect of independent variables on the dependent variable needs to take a period of time and is often difficult to complete in a short time. Therefore, a continuous dynamic process is necessary in analysis of dependent variable. Hence, the spatial model needs to consider both spatial effect and dynamic characteristics (Zhao and Sun, 2022)[44]. The dynamic (SDM) (DSDM) can investigate the spatial effects of dependent variable (FDI) from both the short-term and the long-term perspectives while reducing the problem of endogeneity caused by omitted variables (LeSage and Pace, 2009; Wu et al., 2023)[18, 40]. For these reasons, previous empirical studies (Wang et al., 2013; Zambrano-Monserrate et al., 2020; Zhao and Sun, 2022)[38, 42, 44] have emphasized on the use of (DSDM) in environmental investigations. Thus, the (DSDM) is used to reflect the changes of spatial effects over time in this study. The (DSDM) can be expressed as follows (LeSage and Sheng, 2014; Elhorst, 2014a,b)[19, 12, 13].

$$Y_{it} = \tau Y_{i,t-1} + \eta WY_{i,t-1} + \rho WY_{it} + \beta X_{it} + \theta WX_{it} + \mu_i + \lambda_t + \varepsilon_{it} \quad (4)$$

where τ and η represent the temporal lag coefficient and the spatiotemporal lag coefficient of the dependent variable (FDI), respectively. All other parameters and variables are defined in Eq.(4) (Elhorst, 2014)[19] has argued that the estimation of a dynamic spatial panel data model gets more complicated when the condition $\tau+\eta+\rho<1$ is not satisfied, the model is unstable. LeSage and Pace (2009) stated that the estimated coefficients in spatial models do not indicate the marginal effects of independent variables on dependent variable and consequently lead to wrong conclusions. The decomposition method proposed by LeSage and Pace (2009) can identify direct and indirect effects in response to changes in the independent variables, In this study, we calculate the direct and indirect effects in the short-term and the long-term to further increase the credibility of the findings. Therefore, we rewrite the (DSDM) in vector form as follows:

$$Y_{it} = (I-\rho W)^{-1}(\tau I+\eta W)\,Y_{i,t-1} + (I-\rho W)^{-1}(\beta X_{it} + \theta W X_{it}) + (I-\rho W)^{-1}(\mu_i + \lambda_t + \varepsilon_{it}) \qquad (5)$$

where , I denotes an identify matrix. All other parameters and variables are defined in Eqs. (3) and (4). The direct effect is the average value of diagonal elements of the matrix, which reflects the effect of the independent variable of country i on the dependent variable (FDI) in country i, whereas the indirect effect is the average value row sum of non-diagonal elements of the matrix, which represents the effect of independent variables in neighboring countries on FDI in country i (Elhorst, 2014). The sum of direct and indirect effects is the total effect. The equations of direct and indirect effects in the short-term and the long-term are shown in Table (1).

Table (1): Direct and indirect effects in the short-term and the long-term

	Short-term	Long-term
Direct	$\left[(I_N-\rho W)^{-1}(\beta_k I_N+\theta_k W)\right]^{\bar{d}}$	$\left[((1-\tau)I_N-(\rho+\eta)W)^{-1}(\beta_k I_N+\theta_k W)\right]^{\bar{d}}$
Indirect	$\left[(I_N-\rho W)^{-1}(\beta_k I_N+\theta_k W)\right]^{\overline{rsum}}$	$\left[((1-\tau)I_N-(\rho+\eta)W)^{-1}(\beta_k I_N+\theta_k W)\right]^{\overline{rsum}}$

where I_N is an identity matrix, and $\bar{d}$ and $\overline{rsum}$ denote respectively two operators that allow calculating both the mean diagonal element and the mean row sum of the non-diagonal elements of a matrix (Elhorst 2014a,b).

Table (1) Shows that the total (short-term) effect and be broken down into direct (own-country) and indirect (spillover) effects (LeSage and Pace, 2009). The direct effect captures the impact within a particular country of the unit change in explanatory variable, thus reflecting the impact on the dependent variable that results from a change in the k^{th} regressor x_k in country i in the short-term. The direct effect stems from the own-partial derivatives along the diagonal in the matrix can be expressed as $(I_N - \rho W)^{-1}\left[\beta_k I_N\right]$. The indirect effect, known as spatial spillover, is reported as the average of the row sums of non-diagonal elements of the matrix $(I_N - \rho W)^{-1}\left[\theta_k W\right]$.

Given that both direct and indirect effects may persist for a long time, and analysis of their stability and relevance over the longer term is seen as important. While for the short-term effects we ignore the parameters τ and η , and these can be interpreted as mainly pure spatial feedback effects, the longer term effects contain space-time feed-backs passing from on country to another, thus capturing potential impacts that go beyond a short-term or temporary shock.

Therefore, by examining the relationship between foreign direct investment (FDI) and its macro-correlates (trade openness, market size, population growth, infrastructure and enterprise agglomeration), we can disentangle direct and indirect effect into short and long term effects. Since we leave aside the spatial interaction effects among the error terms, our specification is consistent with a (DSDM). As LeSage and Pace (2009) note, the cost of ignoring spatial dependence in the dependent and/or independent variables is relatively high, because if one or more relevant explanatory variable gets omitted from a regression equation, the estimator of the coefficients for the remaining variables is inherently biased and inconsistent. In contrast, ignoring spatial dependence in the error term, if present, only leads to a loss of efficiency.

3- Empirical Results

3-1 Descriptive Statistics

All collected variables are logarithmic to avoid abnormal phenomena in the data and form a panel data with statistical results describing the study variables. The dependent variable used in this study was the foreign direct investment (lnY). Whereas the independent variables were trade openness (lnX_1), market size (lnX_2), population growth (lnX_3), infrastructure (lnX_4), and enterprise agglomeration (lnX5). The data on FDI (lnY) in Table (2) have a relative fluctuation between maximum value, minimum value and standard deviation. It is noted that the standard deviation of each variable is low with exception of (lnY) and (lnX_4). The skewness, which is the coefficient of symmetry of each variable, is equally low and mildly skewed. The Kurtosis which is the coefficient of flatness in each variable is below 3 with the exception of (lnY) and (lnX_4) which confirms near normality.

Table (2): Descriptive statistics of the variables

	lnY	**lnX_1**	**lnX_2**	**lnX_3**	**lnX_4**	**lnX_5**
Mean	12.627	1.826	6.182	0.211	3.927	1.685
Median	13.803	1.932	6.305	0.197	4.125	1.713
Maximum	16.573	2.254	7.708	0.703	5.018	1.719
Minimum	7.739	0.963	5.729	-0.364	3.026	1.573
Std. Dev.	2.038	0.295	0.721	0.308	1.352	0.021
Skewness	-0.305	0.826	-0.395	0.173	-0.109	0.086
Kurtosis	4.108	-1.025	-0.437	0.825	3.253	0.169
Observations	170	170	170	170	170	170

Source: Author’s calculation.

3-2 Tests for Multicollinearity

To test for multicollinearity a co-relational study was conducted on the dataset and the results are summarized in Table (3). Moreover, the variance inflation Factor (VIF) was calculated and presented in Table (4) to assess the effect of correlation among variables on each other and ensure that the model is efficient.

Table (3): Correlation Matrix

	lnY	**lnX_1**	**lnX_2**	**lnX_3**	**lnX_4**	**lnX_5**
lnY	1					
lnX_1	0.2163*	1				
lnX_2	0.1258	0.2753*	1			
lnX_3	0.5874*	0.4177*	0.2817*	1		
lnX_4	0.1857*	0.3162*	0.0913	0.4025*	1	
lnX_5	0.3259*	0.4025*	0.3962*	0.5028*	0.2103*	1

(*) Significance at $\alpha = 5\%$

Source: Author's calculation.

Upon testing for multicollinearity in Table (3), it was found that there is a moderate correlation between the following variables: Corr (lnX_3, lnY) = 0.5874, Corr (lnX_3, lnX_1) = 0.4177, Corr (lnX_4, lnX_3) = 0.4025, Corr (lnX_5, lnX_1) = 0.4025, Corr (lnX_5, lnX_3) = 0.5028, inclusion is not expected to not violate the OLS assumption. However, to exam in the effect of the multicollinearity and ensure that it has not affected the efficiency of the model, the variance Inflation Factor (VIF) was calculated. The results are presented in Table (4). It is clear that all values of the coefficient of VIF are less than 10. Hence, we are not concerned with the effects of multicollinearity with the inclusion of these variables in the model.

Table (4): Variance Inflation Factor (VIF) of the Independent Variables

Independent Variable	**lnX_1**	**lnX_2**	**lnX_3**	**lnX_4**	**lnX_5**
VIF	2.0824	1.7358	2.8935	3.2571	1.1356

Source: Author's calculation using the statistical package R.

3-3 Cross-sectional Dependence (CD) Test

Panel tests will be applied to examine the incidence of unit roots in the variables time series. Panel-bassed approaches offer the advantages of more comprehensive dataset, more degrees of freedom, better variability, larger efficiency in estimation and reduced multicollinearity problem. However, many panel data methods do not provide for the possibility of cross-sectional dependence among the members of a panel. The presence of cross-sectional dependence affects the reliability of the panel unit root tests (Pasaran, 2021)[26]. In order to check for the presence of cross-sectional dependency (CD), (Pesaran, 2021) proposed a scaled version of the Lagrange Multiplier (LM) test which follows the classical normal distribution, and specified as follows:

$$CDLM=\left(\frac{1}{N(N-1)}\right)^{1/2}\sum_{i=1}^{N-1}\sum_{k=i+1}^{N}(T\hat{\rho}_{ik}^{2}-1) \qquad (6)$$

$\hat{\rho}_{ik}^{2}$ is the statistics denoting the correlation of the pooled OLS residuals from the individual regressions. The null hypothesis assumes that the cross-sectional dependency. However, the CDLM test is subject to distortions when N is large but T is finite. (Pesaran, 2021) proposed a more robust test for cross-sectional dependency (CD test), that is usable in large panels. The CD test is analysed as follows:

$$CD=\sqrt{\frac{2T}{N(N-1)}\left(\sum_{i=1}^{N-1}\sum_{k=i+1}^{N}\hat{\rho}_{ik}\right)} \qquad (7)$$

The null hypothesis states that there is cross-sectional independence and the CD test is assumed to follow the normal distribution and it is also efficient. It is suitable for panel models that are dynamic and also the models which have error variances and multiple breaks in slope coefficients, given that the averages of the series are time-invariant have symmetric distributions.

3-4 Panel Unit Root Test

With the presence of cross-sectional dependence in the members of the panel, there is a need to use unit root tests that accommodate cross-sectional dependence including the (Pesaran, 2007) test[27]. In explaining the Pesaran test; we start with the assumption that FDI or y_{it} is a variable that is generated from the equation below:

$$Y_{it}=(1-\phi_i)\mu_i+\phi_i Y_{i,t-1}+u_{it,}i=1,....,N;\ \ t=1,....,T \qquad (8)$$

Supposing that the error u_{it} has a single factor structure u_{it} = $Y_i f_t+\varepsilon_{it}$ so that f_t is the unobserved common effect, and ε_{it}s the individual-specific (idiosyncratic) error, Eq. (8) can be transformed into the following more convenient equation:

$$\Delta Y_{it} = \alpha_i + \beta_i Y_{i,t-1} + Y_i f_t + \varepsilon_{it} \qquad (9)$$

Thus, the null of nonstationarity $(\phi_i = 1)$ can be expressed as $H_0 : \beta_i = 0$ for all i. Consistent with (Pesaran, 2007); the common factor is represented by the cross-section mean of Y_{it}, which is $\overline{Y}_t = N^{-1} \sum_j^N Y_{it}$, and its lagged value (s), $\overline{Y}_{t-1}, \overline{Y}_{1-2}$ .... if N is adequately large. The unit root hypothesis is based on the t-statistics of the OLS estimate of b_i in the cross-sectionally augmented Dickey-Fuller (CADF) expression:

$$\Delta Y_{it} = \alpha_i + b_i Y_{i,t-1} + c_i \overline{Y}_{t-1} + d_i \Delta \overline{Y}_t + e_{it} \qquad (10)$$

The statistics specified above can be generalized for the more typical case of the panel unit roots tests. The augmented variant of the Im et al., (Pesaran, 2007) test can be specified as follows:

$$CIPS(N,T) = N^{-1} \sum_{i=1}^{N} t_i (N,T) = N^{-1} \sum_{i=1}^{N} CADF_i \qquad (11)$$

Table (5): Cross-sectional Independence Tests

Variables	Pesaran (CD) test		Pesaran (CDLM) test	
	Statistic	ρ Value	Statistic	p-Value
Y	15.729*	0.001	79.622*	0.001
X_1	38.625*	0.001	93.063*	0.001
X_2	29.713*	0.001	81.631*	0.001
X_3	46.828*	0.001	179.258*	0.001
X_4	32.689*	0.001	87.627*	0.001
X_5	23.602*	0.001	71.294*	0.001

(*) significant at α = 5%

Source: Author's calculation using the statistical package R.

Table (5) show that there is strong evidence to reject the null hypothesis of cross-sectional independence for all of the analysed variables. In other words, all variables have cross-sectional dependence. Therefore, this study applies the CADF and the CIPS panel unit root tests to each of the individual panel time-series data.

Table (6): Results of Panel Unit Root Tests

Variables	CADF		CIPS	
	Level	**1st Difference**	**Level**	**1st Difference**
Y	-1.389	-3.016*	-1.691	-3.529*
X_1	-0.917	-2.635*	-1.636	-3.474*
X_2	-1.801	-3.812*	-2.037	-4.208*
X_3	-1.082	-2.935*	-1.739	-3.981*
X_4	-1.617	-3.229*	-1.216	-3.286*
X_5	-0.835	-2.294*	-0.935	-2.895*

(*) Significance at α = 5%

Source: Author's calculation using the statistical package R.

Table (6) reports the results of Pesaran's CADF and CIPS panel unit root tests. We find that all variables are not stationary at their levels. However, they are stationary process at their first differences since we have enough evidence to reject the null hypothesis of unit root at 5% level of significance. The panel time series should either be stationary at their levels or have an established long-run relationship in order to have economically meaningful and reliable estimates of explanatory variables. As these time-series data contain unit root at their levels, this study performs as panel cointegration test to see whether or not the analysed variables have a long-run relationship.

3-5 Panel Cointegration Test

In order to exploit the possible presence of cointegration among the analysed variables, we apply a second-generation panel cointegration, namely, the LM boot-strap panel cointegration test due to Westerlund and Edgerton (2007)[39] that considers cross-sectional dependence. This test is based on LM test of McCoskey and Kao and uses bootstrap technique relying on sieve scheme. Some advantages of the LM boot-strap panel cointegration test are that it works well in the existence of cross-sectional dependence, it has good small sample properties, it reduces asymptotic test distorting using the bootstrap method; it performs the null hypothesis of cointegration. Given that the employed data have a cross-sectional dependence and are relatively small, this test is expected to produce a reliable output for the cointegration relation among the analysed variables.

Table (7): Results from the LM Bootstrap Panel Cointegration Test

Test	Constant		Constant and trend	
	Test statistic	**Bootstrap ρ value**	**Test statistic**	**Bootstrap p-value**
LM bootstrap	67.827	0.284	109.351	0.429

Note: The bootstrap test is based on (Westerlund and Edgerton, 2007)[39] and run using 1000 replications. This test performs the null hypothesis of cointegration in the panel for all units against the alternative hypothesis of no cointegration in the panel (at least for a cross-sectional unit).

As shown in Table (7), the results obtained from the LM bootstrap panel cointegration test indicate that the null hypothesis of cointegration among the analyzed variables against the alternative hypothesis of no cointegration relationship among these variables cannot be rejected at level of 5% significance because the relevant ρ values are far greater than significance levels. Furthermore, such a long-run relationship of the independent and dependent variables is found by using both constant only and constant-trend models. In short, there is a cointegration relationship between the response variable (FDI) and explanatory variables referring to ρ values, and they move together over the long run.

3-6 Spatial Dependence Test

Before estimating the spatial spillover effects, examining the presence of spatial dependence is essential. To this end, we perform the Moran's I statistic. Moran's I test is based on the null hypothesis that the dependent variable (foreign direct investment) associated with different countries is spatially independent. In the case of negative Moran statistics values, negative spatial dispersion characterizes the data. In contrast, positive values of Moran statistics indicate positive spatial autocorrelation and spatial concentration. The following equations defines the global Moran's I index (Lee and Wong, 2001)[22].

$$Z(I)=\frac{I-E(I)}{\sqrt{var(I)}} \tag{12}$$

$$I=\frac{n\sum_{i=1}^{n}\sum_{j=1}^{n}W_{ij}(x_i-\bar{x})(x_j-\bar{x})}{\sum_{i=1}^{n}\sum_{j=1}^{n}W_{ij}\left(\sum_{i=1}^{n}(x_i-\bar{x})^2\right)} \tag{13}$$

Where E(I) is Moran's expectation I, var (I) is a variant of Moran's I, and W_{ij} denote the spatial weight matrix (W_q or W_d) between two countries i and j. n is number of countries.

Table (8): Spatial Dependency Testing Results of Variables Using Moran's I Test

Year	Matrix	Y		X_1		X_2		X_3		X_4		X_5	
		I	Z(I)	I	Z(I)	I	Z(I)	I	Z(I)	I	Z(I)	I	Z(I)
2015	W_q	0.284	7.172*	0.275	7.077*	0.293	8.05*	0.093	5.191*	0.059	3.446*	0.307	18.184*
	W_d	0.162	8.965*	0.103	5.842*	0.214	12.841*	0.079	6.283*	0.042	4.193*	0.273	19.208*
2016	W_q	0.406	10.238*	0.077	2.632*	0.297	7.939*	0.215	12.5*	0.392	10.135*	0.296	17.038*
	W_d	0.244	13.401*	0.034	2.691*	0.215	12.5*	0.201	13.251*	0.308	12.572*	0.229	19.011*
2017	W_q	0.373	9.447*	0.295	7.614*	0.364	9.421*	0.121	6.691*	0.075	4.323*	0.33	18.443*
	W_d	0.225	12.46*	0.102	5.854*	0.25	14.105*	0.097	8.573*	0.058	6.018*	0.308	20.802*
2018	W_q	0.377	9.547*	0.233	6.035*	0.393	10.136*	0.101	5.638*	0.055	3.229*	0.352	19.6*
	W_d	0.223	12.346*	0.068	3.924*	0.275	15.43*	0.095	8.301*	0.043	5.02*	0.292	21.23*
2019	W_q	0.381	9.637*	0.208	5.4*	0.436	11.147*	0.14	7.752*	0.049	2.873*	0.351	19.399*
	W_d	0.23	12.724*	0.061	3.545*	0.294	16.373*	0.106	9.024*	0.037	5.62*	0.305	22.318*
2020	W_q	0.353	9.009*	0.226	5.843*	0.417	10.729*	0.149	8.294*	0.047	2.779*	0.335	19.736*
	W_d	0.209	11.626*	0.065	3.768*	0.288	16.124*	0.092	10.211*	0.035	5.08*	0.327	20.01*
2021	W_q	0.361	9.112*	0.21	5.441*	0.385	9.907*	0.165	9.148*	0.036	2.17*	0.319	17.775*
	W_d	0.212	11.714*	0.06	3.482*	0.249	13.937*	0.103	12.138*	0.022	4.227*	0.285	19.215*
2022	W_q	0.29	7.324*	0.209	5.375*	0.383	9.861*	0.162	8.954*	0.036	2.145*	0.334	18.568*
	W_d	0.18	9.934*	0.065	3.789*	0.259	14.506*	0.109	10.103*	0.024	4.016*	0.311	19.13*
2023	W_q	0.32	8.063*	0.117	3.039*	0.391	10.617*	0.161	8.946*	0.028	2.037*	0.324	18.016*
	W_d	0.197	10.834*	0.053	3.111*	0.265	14.809*	0.108	10.225*	0.013	4.318*	0.215	20.129*
2024	W_q	0.134	3.873*	0.113	2.947*	0.349	9.013	0.175	9.638*	0.197	2.939*	0.325	18.113*
	W_d	0.083	5.288*	0.058	3.384*	0.252	14.142*	0.129	11.85*	0.093	4.671*	0.524	20.27*

(*) Significant at α = 5%, $Z_{0.025}$ = 1.96

(*) Significant at α = 5% , $Z_{0.025}$ = 1.96

Source: Author's calculation using the statistical package R.

Table (8) show the results of Moran's I test of all variables in 17 countries from 2015 to 2024, and their spatial autocorrelation was examined. As it is clear that the data of all years passed 5% significance test, rejecting the null hypothesis that "there is no spatial autocorrelation", indicating that there are significant positive spatial autocorrelations in variables over the years. As a result, the spatial correlation of the factors should be taken into account, otherwise the estimation will be biased or inconsistent. Therefore, the spatial econometrics model should be set up.

3-7 Statistical Tests For Model Form Selection

Before the spatial model estimation, it is necessary to apply the LM test and the Robust-LM test to determine whether the spatial correlation model is represented by the spatial autoregressive model (SAR) or spatial error model (SEM). The test results show that both the LM and the Robust-LM produce significant results at the 5% level (Table 9), but it is unclear whether (SAR) or (SEM) should be selected. By further adopting the Wald test and LR test, the null hypothesis is significantly rejected, indicating that the spatial Durbin model (SDM) cannot be simplified to (SAR) or (SEM).

Finally, the Hausman test determines whether to select a random or fixed effects (FE) model. As the results of Hausmat test indicated that fixed effects should be chosen. Therefore, this article chooses the (SDM) with spatial fixed effects to conduct the next analysis.

Table (9): The Identification of Spatial Panel Model

Test	Statistical value	p-value	Test	Statistical value	p-value
LM-lag	89.327*	0.001	Wald-lag	23.903*	0.008
Robust LM-lag	53.935*	0.001	LR-lag	38.625*	0.001
LM-error	77.089*	0.001	Wald-error	*19.731	0.005
Robust LM-error	32.493*	0.001	LR-error	47.629*	0.001
Hausman test	213.682*	0.001	-	-	-

(*) Significance at the 0.05 level.

Source: Author’s calculation using the statistical package R.

Table (10): Estimated Results of Pooled OLS, Fixed Effects (FE) and Spatial Econometric Models

Variables	Pooled OLS	FE	W_q				W_d			
			SAR-FE	SEM-FE	SDM-FE	DSDM-FE	SAR-FE	SEM-FE	SDM-FE	DSDM-FE
$\ln Y_{i,t-1}(\tau)$	-	-	-	-	-	0.593* (0.062)	-	-	-	0.603* (0.067)
$W\ln Yi,t\text{-}1(\eta)$	-	-	-	-	-	-0.482 (0.361)	-	-	-	-5.973 (3.874)
$W\ln Y_{i,t}(\rho)$	-	-	4.084* (1.293)	0.063 (0.061)	-5.981* (2.593)	0.109* (0.064)	5.108* (1.628)	0.072 (0.051)	-0.307* (0.981)	6.108* (2.682)
$\ln X_1$	1.903* (0.781)	2.614* (1.725)	1.768* (0.979)	1.624* (0.952)	1.429* (0.457)	0.173 (0.049)	2.375* (1.036)	1.927* (0.993)	1.232* (0.493)	0.180* (0.058)
$\ln X_2$	0.055 (0.724)	0.038 (0.701)	0.022 (0.971)	0.017 (0.973)	0.494 (0.429)	0.993* (0.527)	0.041 (0.715)	0.026 (0.963)	1.089* (0.564)	0.047 (0.179)
$\ln X_3$	-2.135* (1.013)	-3.151* (1.526)	-1.802* (0.991)	-1.812* (0.997)	-1.723* (0.763)	-0.301 (0.715)	-2.082* (1.253)	2.286* (1.308)	-0.537 (0.768)	-0.177 (0.158)
$\ln X_4$	0.068 (0.697)	0.131 (0.263)	0.043 (0.175)	0.039 (0.179)	0.081 (0.275)	0.405 (0.372)	0.066 (0.179)	0.115 (0.253)	0.281 (0.294)	0.022 (0.097)
$\ln X_5$	17.902* (5.502)	10.283* (3.157)	15.309* (4.607)	15.318* (4.615)	18.903* (5.712)	12.287* (3.802)	14.726* (4.401)	16.291* (4.728)	20.682* (6.975)	9.923* (3.309)
$W\ln X_1$	-	-	-	-	1.535* (0.658)	0.109* (0.063)	-	-	1.499* (0.649)	0.690* (0.353)
$W\ln X_2$	-	-	-	-	0.042 (0.072)	0.096 (0.138)	-	-	0.397 (0.388)	0.020 (0.060)
$W\ln X_3$	-	-	-	-	2.024* (1.160)	0.61 (0.062)	-	-	2.528* (1.182)	-0.043 (0.172)
$W\ln X_4$	-	-	-	-	0.030 (0.065)	0.083 (0.055)	-	-	0.032 (0.059)	0.048 (0.177)
$W\ln X_5$	-	-	-	-	0.847* (0.265)	5.083* (2.416)	-	-	0.928* (0.225)	4.331* (2.201)
$AdjR^2$	0.697	0.712	0.743	0.789	0.853	0.892	0.746	0.811	0.887	0.924
Log-likelihood	937.6	2852.1	3451.7	3450.2	3458.6	3691.4	3462.2	3470.8	3762.7	3851.6
AIC	6793.5	6521.3	6327.6	6372.4	5883.1	5495.7	6296.5	6319.4	5819.3	5327.2
BIC	6847.2	6579.3	6383.9	6433.7	5962.4	5563.1	6365.7	6385.9	5856.4	5406.3
$\tau+\eta+\rho$	-	-	-	-	-	0.619	-	-	-	0.738
No. of observations	170	170	170	170	170	170	170	170	170	170

Notes: W_q indicates the queen adjacency weight matrix,; W_d indicates weight matrix based on distance. Data in brackets are t value. AIC indicates Akaike information criterion; BIC indicates Bayesian information criterion.

(*) Significance at the 0.05 level.

The next step is to identify the best spatial model according to the described procedures. The results of four specifications of spatial models under queen adjacency weight matrix (W_q) and inverse distance weight matrix (W_d) are shown in Table (10). First, starting from the non-spatial linear model (pooled OLS), the coefficients of trade openness (X_1) and agglomeration (X_5) are both positive and significant, while the coefficient of population growth (X_3) is negative and significant. Second, since the Hausman test results show that fixed effects should be considered, the fixed effect model of the panel data is further estimated. In the non-spatial panel with the spatial fixed model, the parameter estimation coefficients of the three independent variables X_1, X_3, and X_5 are 2.614, -3.151 and 10.283, respectively, which are significant at the 0.05% level. if the spatial correlation is omitted in an econometric analysis, the estimation result of the model is likely to lead to bias (Yang et al., 2019)[41]. After further considering the spatial correlation, the (SDM) was found to be more appropriate. The Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) are employed to choose the appropriate model. The results indicate that the AIC and BIC in (DSDM) are lower than (SAC), (SEM) and (SDM), suggesting that (DSDM) is an appropriate specification. Third, considering the significance of the lagged dependent variable, (DSDM) is the best modeling approach for investigating the relationship among the variables. The main finding also shows that the Rho (ρ) coefficient is positive and significant, confirming that there are significant spatial effects between FDI of the countries. This mean a country's FDI also depends on the FDI in its neighboring countries. In addition, both the log-likelihood and R^2 increased, which reflects the need to construct the (DSDM). At the same time, the result of a robustness check from replacing a queen adjacency weight matrix (W_q) with the inverse distance weight matrix (W_d) is very similar to each other. Moreover, the calculation result $\tau+\eta+\rho=0.619$, 0.738 < 1 in the W_q and W_d weight matrix model respectively indicates that the (DSDM) turns out to be stable, and there is a spatial interaction of FDI inflows into the key economic regions of Arab countries. In general, from the pooled OLS to the (DSDM), due to the gradual consideration of fixed effects, spatial effects and the influence of factors other than regional FDI such as culture and institutions (represented by the first-order lag term of the dependent variable), the log-likelihood and R^2 of the models continue to increase, and the models gradually become more reasonable and reliable. Therefore, the (DSDM) is the most suitable model. Finally, the results are robust with both measures of the spatial weight matrix. However, compared with the (DSDM) with W_q, the (DSDM) with W_d has a larger log-likelihood and R^2, which further confirms that for this study, the distance weight matrix (W_d) has its advantage in model estimation. Therefore, the DSDM with W_d will be discussed as follows.

$$\ln Y_{i,t} = 0.603 \ln Y_{i,t-1} - 5.973 \, W\ln Y_{i,t-1} + 6.108 W\ln Y_{i,t} + 0.18 \ln X_{1,t} + 0.047 \ln X_{2,t} - 0.177 \ln X_{3,t} + 0.022 \ln X_{4,t} + 9.923 \ln X_{5,t} + 0.69 W\ln X_{1,t} + 0.02 \, W \ln X_{2,t} - 0.043 W\ln X_{3,t} + 0.048 \, W\ln X_{4,t} + 4.331 \, W\ln X_{5,t}$$

First, the FDI ($Y_{i,t}$) in different regions has significant time lag effects and spatial spillover effects. The coefficient of the time lag term of FDI ($\ln Y_{i,t-1}$) is significantly positive at the level of 5%, indicating that the development level of FDI in the previous period has a positive effect on the development of FDI in the current period. While the FDI spatial lag term ($W\ln Y_{i,t-1}$) coefficient is insignificantly negative at the 5% level.

In addition, the FDI spatial term ($W\ln Y_{i,t}$) coefficient is significantly positive at the 5% level, indicating that the FDI of a country will be positively affected by the FDI of neighboring countries.

Second, trade openness ($\ln X_1$) significantly affects FDI both directly and indirectly. The correlation between trade openness ($\ln X_1$) and FDI is significant and positive at the 5% level, according to the regression coefficient of $\ln X_1$. The spatial term coefficient of trade openness ($W\ln X_1$) is also significantly positive, indicating a positive effect on the FDI level of adjacent countries due to technology spillover and knowledge dissemination.

Third, the coefficients of agglomeration ($\ln X_5$) and spatial term are both significantly positive, indicating that an increase in agglomeration in this country will increase the FDI in this country and neighboring countries.

Fourth, the coefficients of the variables market size ($\ln X_2$), population growth ($\ln X_3$), infrastructure ($\ln X_4$) and their spatial terms ($W\ln X_2$, $W\ln X_3$ and $W\ln X_4$) were not significant.

3-8 Fit Test of (DSDM) Regression Model

To see the usability of the model as a whole, it is necessary to do a model compatibility test based on the following test:

H_0: $\rho=\beta_1=\beta_2=\beta_3=\beta_4=\beta_5=\theta_1=\theta_2=\theta_3=\theta_4=\theta_5=0$

H_1: $\rho\neq 0$ or there is at least one $\beta_j\neq 0$, $\theta_j\neq 0$, j = 1, 2, 3, 4, 5

Where it was found that the calculated value of F = 23.729 is more than that obtained from the $F_{(\alpha,k,n-k-1)}=F_{(0.05,5,164)}$ = 2.21 resulted from F table. Thus, it was decided that H_0 was rejected at the level of α =5%. Therefore, it is possible to say that the (DSDM) with five predictor variables is suitable to describe the proposed relationship.

It is also clear from table (10) that potential FDI flows are positively dependent on current FDI flows at the significant level of 5% (robustness check). This shows that there are short-term and long-term effects on factors affecting FDI flows in the host country of Arab countries. These impacts are identified through direct and indirect effects in the (DSDM) as follows.

Table(11): Results of the short-term and the long-term direct indirect, and total effects of the (DSDM) under two weight matrices

Matrix	Variable	Short-term effects			Long-term effects		
		Direct effect	Indirect effect	Total effect	Direct effect	Indirect effect	Total effect
W_q	lnX_1	0.7803* (0.4591)	-0.2517 (0.2013)	0.5286* (0.3792)	2.2815* (1.0123)	6.8059* (1.9805)	9.0874* (3.0258)
	lnX_2	0.0105 (0.0271)	2.3046* (0.9153)	2.3151* (1.0374)	0.0229 (0.0712)	5.3874* (1.8372)	5.4103 (1.9037)
	lnX_3	-0.3879 (0.6437)	-3.8279* (0.9757)	-4.2158* (1.7503)	-0.9443 (0.6192)	-5.2873* (1.4377)	-6.2316* (1.8179)
	lnX_4	0.0083 (0.3025)	-0.0051 (0.0729)	0.0032 (0.1103)	0.0322 (0.0962)	0.1244 (0.3912)	0.1566 (0.0137)
	lnX_5	16.7962* (4.6015)	-13.7142* (5.9307)	3.082* (0.9638)	29.0832* (8.3945)	-19.9357* (5.3841)	9.1475* (3.1527)
W_d	lnX_1	1.8358* (0.9913)	-0.9253 (1.0266)	0.9105* (0.3714)	4.3152* (1.9735)	-0.7359 (0.9372)	3.5793* (0.9932)
	lnX_2	0.0028 (0.0322)	0.8637 (1.0125)	0.8665* (0.7963)	0.0136 (0.2971)	1.9738 (1.5625)	1.9874* (0.9853)
	lnX_3	1.1839 (1.6201)	-7.0825* (2.0519)	-5.8986* (1.5827)	1.7392 (1.2035)	-12.8531* (4.3917)	-11.1139* (4.1372)
	lnX_4	0.0496 (0.8013)	-0.0039 (0.2081)	0.0457 (0.0122)	0.0939 (0.9105)	-0.0134 (0.1963)	0.0805 (0.0219)
	lnX_5	32.1032* (9.8371)	-24.8362* (7.0213)	7.267* (2.8361)	47.3055* (12.1392)	-38.2152* (10.6252)	9.0903* (3.0162)

(*) Significance at the 0.05 level, data in brackets are t value.

Table (11) reports the estimates of the short-term direct effect, short-term indirect effect, short-term total effect, long-term direct effect, long-term indirect effect and long-term total effect of the (DSDM) as follows:

1- The high trade openness (lnX_1) of a host country directly increases the FDI inflows in that country at a significant level of 5% and increases the FDI flows of the whole region at a significant level of 5% for both spatial weight matrices (W_q, W_d). it is also found that the high trade openness (lnX_1) in a country indirectly increases the FDI flows in neighboring countries at the 5% significance level at the long-term, but does not pass the significance test in short-term.
2- Although the results in Table (11) do not find any statistical evidence that the market size (lnX_2) in a host country has a direct effect on FDI attraction of itself country, it has an indirect effect on FDI in neighboring countries at the significant level of 5% when the spatial weight matrix is W_q. At the same time, the market size (lnX_2) affects the ability to attract FDI of the whole region at the significance level of 5% for both spatial weight matrices (W_q, W_d).
3- As it is clear from the results that increasing the population growth (lnX_3) of a host country indirectly reduces FDI flows in neighboring countries at a significant level of 5% and reduces FDI flows across the region by the 5% significance level for both spatial weight matrices (W_q, W_d).
4- While the direct, indirect and total coefficients of infrastructure (lnX_4) are not significant at the significant level of 5% for both spatial weight matrices (W_q, W_d).
5- The results also show that the Agglomeration (lnX_5) or concentration of enterprises in a country directly increases the FDI flows of that country at a significant level of 5%, and indirectly reduces FDI flows in neighboring countries at the 5% significance level. In terms of the total effect, enterprise agglomeration has a positive impact on FDI in flow for the whole region at the 5% significance level for both spatial weight matrices (W_q, W_d).

3-9 Robustness Checks

Robustness check is conducted to ensure the reliability of the spatial regression results in two ways. First, considering the importance of choice of the spatial weight matrix to estimate the spatial models, the (DSDM) was estimated with queen adjacency weight matrix (W_q) and inverse distance weight matrix (W_d), and the results are presented in Table (11). The results of the short-term and the long-term direct, indirect, and total effects using (W_q) are very similar with the estimated results using (W_d). Second, we estimate alternative models for analyzing the effect of main determinants of Foreign Direct Investment (FDI). The spatial effects are estimated using (DSDM) with only time-lagged dependent variable as in Table (12) and dynamic (SAR) (DSAR) as in Table (13).

Table (12): Results of the short-term and the long-term direct, indirect, and total effects of the (DSDM) with only time-lagged dependent variable under inverse distance weight matrix (W_d)

Variable	Short-term effects			Long-term effects		
	Direct effect	Indirect effect	Total effect	Direct effect	Indirect effect	Total effect
$lnY_{i,t-1}$	4.5027* (1.7031)	-3.1973* (1.8936)	1.3054* (0.4925)	6.0172* (2.0302)	-2.9302* (0.7163)	3.0872* (1.1358)
lnX_1	3.4916* (2.0138)	-0.0358 (0.0325)	3.4558* (2.0972)	6.8561* (2.2018)	-0.3008 (0.5529)	6.5553* (2.8731)
lnX_2	0.0383 (0.0157)	3.2851* (1.6205)	3.3234* (1.6683)	1.2261 (08673)	11.2572* (5.3807)	12.4833* (5.8825)
lnX_3	1.0293 (1.5382)	-7.4203 (1.6015)	-6.3911* (2.8739)	1.3629 (1.4028)	-13.7285 (8.6938)	-12.3656* (5.5831)
lnX_4	0.2123 (0.5538)	-0.1602 (1.1038)	0.0521 (0.0493)	0.1617 (0.5728)	-0.0452 (3.7625)	0.1165 (0.0451)
lnX_5	17.9305* (9.5823)	-14.3852* (8.1573)	3.5453* (1.7495)	33.0852* (10.6018)	-18.6291* (7.3826)	14.4561* (8.3103

(*) Significance at the 0.05 level, data in brackets are t value.

Table (13): Results of the short-term and the long-term direct, indirect, and total effects of the dynamic (SAR) (DSAR) under inverse distance weight matrix (W_d).

Variable	Short-term effects			Long-term effects		
	Direct effect	Indirect effect	Total effect	Direct effect	Indirect effect	Total effect
$lnY_{i,t-1}$	6.7928* (2.6725)	-0.2537* (0.1462)	6.5391* (3.1025)	9.2813* (2.8506)	-0.3029 (0.5023)	8.7984* (2.9135)
lnX_1	1.2038 (0.6135)	-0.0096 (0.0218)	1.1942* (0.6628)	0.3749 (0.4725)	-0.0173 (0.0252)	0.3576* (0.0973)
lnX_2	5.3029* (2.8372)	15.2638* (6.9274)	20.5667* (9.9722)	4.8205* (1.9791)	18.0294* (6.9029)	22.8499* (10.962)
lnX_3	0.0427 (0.0216)	-6.8619* (2.3015)	-6.8192* (2.3757)	1.8974* (0.7023)	-5.0602 (0.0281)	-3.1628* (1.3178)
lnX_4	0.3029 (0.5204)	-0.0429 (0.0351)	0.2601 (0.4735)	1.3258 (0.7733)	-1.0291 (0.6509)	0.2967 (0.0162)
lnX_5	12.3755* (6.2714)	-8.6501 (4.1291)	3.7254* (1.7375)	47.2352* (12.8935)	-32.3719* (10.3257)	14.8633* (8.8352)

(*) Significance at the 0.05 level, data in brackets are t value.

Source: Author's calculation using the statistical package R.

Tables (12) and (13) show the following:

1- The direct and total effect of the lagged dependent variable ($Y_{i,t-1}$) is positive and significant in both models, while its indirect effect is negative and significant only in (DSDM) with time-lagged dependent variable model.

2- The total effect of trade openness (lnX_1) is positive and significant in both models.

3- The indirect and total effect of the market size (lnX_2) is positive and significant in both models.

4- The total effect of the population growth (lnX_3) is negative and significant in both models.

5- The total effect of the infrastructure (lnX_4) is insignificant in both models.

6- The direct and total effect of the agglomeration ($\ln X_5$) is positive and significant in the two models, while its indirect effect is negative and significant in the two models as well.

All of these results are consistent with the results of (DSDM) applied in this study (Table 11).

4- Conclusion

1- This study analyses and exploits the long-run dynamics of foreign direct investment (FDI), trade openness($\ln X_1$), market size($\ln X_2$), population growth($\ln X_3$), infrastructure($\ln X_4$) and enterprise agglomeration ($\ln X_5$) for 17 Arab countries over the period of 2015-2024. As apart from a large part of previous literature, we test the possible presence of cross-sectional dependence in the panel time-series data, and accordingly, choose to employ second-generation unit root tests, namely, the CIPS and CADF panel unit root tests; a second-generation cointegration test, namely, the LM bootstrap panel cointegration test, all which are fine choices in the existence of cross-sectional dependence. In addition, this study estimates the long-run coefficients among variables using the (DSDM), which is superior to the (SAR), (SEM) and (SDM) for relatively small samples.

2- The CD test reports that all variables have cross- sectional dependence at 0.05 level of significance. The CIPS and CADF panel unit root tests show that all of the analyzed variables contain a unit root at their levels, but become stationary at their first differences at 0.05 level of significance. The LM bootstrap panel cointegration test indicates that the analyzed variables are cointegrated; alternatively, the variables have a long-run relationship. Thus, the coefficient estimates are assumed to be economically reliable and meaningful.

3- The results of Wald and LR tests reveal that the spatial Durbin model (SDM) cannot be simplified to (SAR) or (SEM).

In addition to the Hausman test results indicated that fixed effects model should be chosen .

4- This study provides an important and valuable contribution to the limited existing literature on FDI by achieving two objectives. Firstly, to examine the spatial dependence of both the lagged independent and dependent variables. Secondly, to determine the direct, indirect, and total effects of independent variables on FDI in the short-term and the long-term. For the first aim, we applied Moran's I test for FDI and independent variables. For the second aim, we used (DSDM) with fixed effects.

5- (DSDM) was the best model to handle this spatial effect because it had the smallest AIC and BIC value, and greatest Adj-R^2 and Log-likelihood value compared to Pooled OLS, FE, (SAR), (SEM) and (SDM). In the (DSDM) model, the factors that significantly affected the FDI in Arab countries were $\ln Y_{i,t-1}$, $W \ln Y_{i,t}$, $\ln X_1$, $\ln X_5$, $W \ln X_1$, $W \ln X_5$.

6- There is a direct effect of trade openness($\ln X_1$), which has a positive effect on FDI in the short and long terms, while there is an indirect effect of trade openness($\ln X_1$), which also has a positive effect FDI in the long term only.

7- There is an indirect effect of market size($\ln X_2$)and population growth($\ln X_3$)that has a positive and negative impact on FDI in the short and long terms, respectively.

8- There is no direct or indirect effect of the infrastructure($\ln X_4$) on FDI, whether in the short or long term.

9- There is a direct effect of enterprise agglomeration($\ln X_5$) that has a positive effect on FDI in the short and long terms, while there is an indirect effect of enterprise agglomeration ($\ln X_5$) that has a negative effect on FDI in both the short and long terms.

10- The reliability of the spatial regression results was confirmed in two ways. First, the (DSDM) was estimated with queen adjacency weight matrix(W_q) and inverse distance weight matrix (W_d).The results of the short-term and the long-term direct, indirect, and total effects using (W_q) are very similar with the estimated results using (W_d). Second, the spatial effects are estimated also using (DSDM) with only time-lagged dependent variable and dynamic (SAR) (DSAR), all results were also similar.

References

1- Anselin L 2001 A Companion to Theoretical Econometrics Spatial Econometrics (Oxford: Blackweel Publishing Ltd).

2- Anselin, Luc. 2007. Perspective on Spatial Data Analysis. Berlin: Springer.

3- Anselin, L. and S.J. Rey, 2010. Perspectives on Spatial Data Analysis. 1st Edn., Springer, Berlin, London,ISBN-10: 3642019757.

4- Ashari, A., Efendi, A., Pramoedyo, H., 2020. GSTARX-SUR modeling using inverse distance weighted matrix and queen contiguity weighted matrix for forecasting cocoa black pod attack in trenggalek regency. In: Proceedings of the Proceedings of the 13th International Interdisciplinary Studies Seminar, IISS 2019, 30-31 October 2019, Malang, Indonesia.

5- Arogundade, Sodiq, Mduduzi Biyase, and Santos Bila. 2022. Be Nice to Thy Neighbors: Spatial Impact of Foreign Direct Investment on Poverty in Africa. Economies 10: 128.

6- Baltagi, B.H. and G. Bresson, 2011. Maximum likelihood estimation and lagrange multiplier tests for panel seemingly unrelated regressions with spatial lag and spatial errors: An application to hedonic housing prices in Paris. J. Urban Econ., 69.

7- Debarsy, N., Ertur, C., and LeSage, J. P. (2012). Interpreting Dynamic Space-Time Panel Data Models. Stat. Methodol. 9, 158-171. doi: 10.1016/j.stamet.2011.02.002.

8- Elhorst, J.P., 2010. Applied spatial econometrics: raising the bar. Spat. Econ. Anal. 5 (1).

9- Elhorst J P 2010 Handbook of Applied Spatial Analysis (Berlin: Springer).

10- Elhorst, Paul, Eelco Zandberg, and Jakob De Haan. 2013. The Impact of Interaction Effects among Neighboring Countries on Financial Liberalization and Reform: A Dynamic Spatial Panel Data Approach. Spatial Economic Analysis 8.

11- Elhorst, J.P., 2014. Spatial Econometrics: From Cross-sectional Data to Spatial Panels, Vol. 479. Springer, Heidelberg.

12- Elhorst, J. P. (2014a). Matlab Software for Spatial Panels. Int. Reg. Sci. Rev. 37 (3).

13- Elhorst, J.P. (2014b). Spatial Econometrics: From Gross-Sectional Data to Spatial Panels. New York Dordrecht London: Springer. Heidelberg.

14- Feng Z and Chen W 2018 Environmental Regulation, Green Innovation, and Industrial Green Development Sustainability Journal 10 (1) 223.

15- Feng Y, Wang X, Du W, and Liu J 2018 Effects of Air Pollution Control on Urban Development Quality in Chinese Cities Based on Spatial Durbin Model Int. J. Environ. Res. Public Health 15(12) 2822.

16- Feng, Y., Wang, X., 2020. Effects of urban sprawl on haze pollution in China based on dynamic spatial Durbin model during 2003-2016. J. Clean. Prod. 242, 118368.

17- Jiang, L., Zhou, H.F., Bai, L., Zhou, P., 2018. Does foreign direct investment drive environmental degradation in China? An empirical study based on air quality index from a spatial perspective. J. Clean. Prod. 176, 864-872.

18- Lesage, J.P. and R.K. Pace, 2009. Introduction to Spatial Econometrics. 1st Edn., CRC Press, ISBN-10:142006424X.

19- LeSage, J. P., and Sheng, Y. (2014). A Spatial Econometric Panel Data Examination of Endogenous versus Exogenous Interaction in Chinese Province-Level Patenting. J. Geogr. Syst. 16, 233-262. doi: 10.1007/sl0109-014-0198-7.

20- LeSage, J. P. (2015). Software for Bayesian Cross Section and Panel Spatial Model Comparison. J. Geogr. Syst. 17 (4), 297-310. doi: 10.1007/sl0109-015-0217-3.

21- LeSage J P and Pace R K 2017 Spatial Econometrics Monte Carlo Studies: Raising The Bar Empirical Economics 55 (1).

22- Lee, J. and D.W.S. Wong, 2001. Statistical Analysis with ArcView GIS. 1st Edn., John i Wiley and Sons, New York, ISBN-10: 047143776X.

23- Lee, L.F., Yu, J., 2010. Estimation of spatial autoregressive panel data models with fixed effects. J. Econ. 154 (2).

24- Li, Mu-chen, Si-xian Feng, and Xing Xie. 2023. Spatial Effect of Digital Financial Inclusion on the Urban-Rural Income Gap in China-Analysis Based on Path Dependence. Economic Research-Ekonomska Istraživanja 36.

25- Millo G, and Miras P 2012 splm: Spatial Panel Data Models in R Journal of Statistical Software ,47/ 1.

26- Pesaran, M. H. (2021). General diagnostic tests for cross section dependence in panels. Empirical Economics 60(1).

27- Pesaran, M. H. (2007). A simple panel unit root test in the presence of cross-section dependence. Journal of Applied Econometrics, 22.

28- Pan, X., Guo, S., Li, M., & Song, J. (2021). The effect of technology infrastructure investment on technological innovation-A study based on spatial durbin model. Technovation, 107, 102315.

29- Paul, J., & Jadhav, P. (2019). Institutional determinants of foreign direct investment inflows: Evidence from emerging markets. International Journal of Emerging Markets, 15(2).

30- Putri Andayani Suaib, T. (2022). PENERAPAN SPATIAL DURBIN MODEL (SDM) PADA INDEKS PEMBANGUNAN GENDER DI PULAU SULAWESI. Majalah Ilmiah Matematika Dan Statistika, 22(2).

31- Ponce, P., Aguirre-Padilla, N., Oliveira, C., _Alvarez-Garcia, J., & del Rio-Rama, M. d. 1. C. (2020). The spatial externalities of tourism activities in poverty reduction. Sustainability, 12(15), 6138.

32- Quito, B., del Rio, M.D.L.C., 'Alvarez-Garcia, J., Bekun, F.V., 2023. Spatiotemporal influencing factors of energy efficiency in 43 European countries: a spatial econometric analysis. Renew. Sust. Energ. Rev. 182, 113340.

33- Triki, M.B. and S. Maktouf, 2012. Cross-country empirical studies of banking crisis: A spatial durbin model. Eur. J. Bus. Manage., 4.

34- Wang Z, Liang H Y, and Zhou Q 2013. Model for Effect of Spatial Weighted Matrix on Spatial Autocorrelation. International Journal of Applied Mathematics and Statistics. 44(14).

35- Wang, Q., Ge, Y., Li, R., 2023a. Does improving economic efficiency reduce ecological footprint? The role of financial development, renewable energy, and industrialization. Energy Environ.

36- Wang, Q., Hu, S., Li, R., 2023b. Could information and communication technology (ICT) reduce carbon emissions? The role of trade openness and financial development. Telecommun. Policy 102699.

37- Wang, Q., Sun, T., Li, R., 2023c. Does artificial intelligence promote green innovation? An assessment based on direct, indirect, spillover, and heterogeneity effects. Energy Environ.

38- Wang, Y., Kang, L., Wu, X., Xiao, Y., 2013. Estimating the environmental Kuznets curve for ecological footprint at the global level: a spatial econometric approach. Ecol. Indic. 34.

39- Westerlund, J., & Edgerton, D. L. (2007). A panel bootstrap cointegration test. Economics Letters, 97,185-190.

40- Wu, F., Yang, X., Chen, Y., 2023. Research on innovative mechanisms of financial agglomeration enabling green coordinated development in the Yangtze River Delta of China. Heliyon 9 (3), el4172.

41- Yang W, WangWand Ouyang S (2019) The influencing factors and spatial spillover effects of CO2 emissions from transportation in China. Science of The Total Environment 696: 133900.

42- Zambrano-Monserrate, M.A., Ruano, M.A., Orme˜no-Candelario, V., Sanchez-Loor, D.A., 2020. Global ecological footprint and spatial dependence between countries. J. Environ. Manag. 272, 111069.

43- Zhou, K., Yang, J., Yang, T., Ding, T., 2023. Spatial and temporal evolution characteristics and spillover effects of China’s regional carbon emissions. J. Environ. Manag. 325.

44- Zhao, M., Sun, T., 2022. Dynamic spatial spillover effect of new energy vehicle industry policies on carbon emission of transportation sector in China. Energy Policy 165, 112991.