Growth and Spatial Dependence: The Mankiw, Romer and Weil model revisited

Cem Ertur and Kalidou Thiaw

Laboratoire d'Economie et de Gestion (UMR 5118 CNRS)
Université de Bourgogne
Pôle d'Economie et de Gestion
B.P. 26611 – 21066 Dijon Cedex

Abstract

The aim of this paper is to analyze the theoretical and econometric implications of omitting spatial dependence in the Mankiw, Romer, and Weil model. Indeed, the international distribution of income levels and growth rates suggests the existence of large international disparities, and therefore the important role of location on economic performance. However, taking spatial dependence into account requires resorting to the methods of Spatial Econometrics, not only for a valid statistical inference, but also for revaluating the impact of the variables generally considered as crucial in the growth phenomenon and finding the processes underlying growth rates and income levels.

Keywords: *Economic growth, convergence, spatial econometrics*

JEL: C14, C31, O4

Introduction

In their 1992 paper entitled “A contribution to the empirics of economic growth”, Mankiw, Romer and Weil intend to show that, despite its shortcomings, the Solow model is a satisfactory enough framework for growth analysis. They show indeed that with a larger definition of capital that includes “human capital”, and the hypotheses of a technological progress and constant scale returns in production, one can explain a large part of international disparities in international income levels and per capita growth rates differences.

Therefore, Mankiw and *al.* revive the basic equations of the Solow model by adding a measure of human capital in the production function. However, as for most empirical studies in economic growth, they run cross-section regressions estimated by OLS, then omitting among others all space-related influences on income levels and growth rates. The observation of international data on growth though suggests the existence of a certain tendency to geographical pooling among rich nations, and the same is true of poor ones. This leads Temple [1999] to remind us that regional dummy variables have often been significant

Then, growth performance is probably not insensible to country location, even if those issues raised by the fact of taking space into account have very seldom inspired empirical studies in economic literature. Walter Isard already mentioned an “Anglo-Saxon bias” in the 50’s, regretting the absence of spatial dimensions in the analysis of economic phenomena. Therefore, it is only very recently that empirical studies started to explicitly integrate space effects on growth (Cf. Moreno et Trehan [1997], Easterly et Levine [1998], Fingleton [1999]).

However, it is a long time since problems resulting from the handling of spatial data have been highlighted in econometric models of regional science. The term “Spatial Econometrics” is precisely due to the latter discipline, more exactly to J. Paelinck, and is defined by Anselin (1988) as “the collection of techniques dealing with the peculiarities caused by space in the statistical analysis of regional science models”.

Spatial data actually have the special feature of bearing information not only on the observed value of a given variable, but also and above all, on the relative location of the observation unit. Henceforth, cross-section regressions usually implemented in empirical studies on growth no longer give satisfactory results because of the potential presence of two spatial effects: spatial autocorrelation and spatial heterogeneity.

However, while spatial heterogeneity can generally be treated by means of standard econometrics, the presence of spatial autocorrelation in the data substantially modifies statistical inference and requires the use of spatial econometric techniques. In particular, it reappraises a fundamental hypothesis of the Ordinary Least Squares method that is independence between observations, and then the estimations made by Mankiw and *al.* may turn out inconsistent or inefficient.

Therefore, on the basis of the standard specifications in the MRW model, the first step of our study will be to test the omission of spatial autocorrelation in the residuals of the least-squares estimation; the international distribution of income levels and growth rates indeed let us think that these equations are wrongly specified because geographic spillovers are not considered.

Then, we turn to finding the most convenient way of modelling relationships between countries when relative position in space is taken into account, and so resort to the different specifications and statistical tests suggested in spatial econometrics. We lastly come back to the conclusions of Mankiw and *al.* in order to assess the influence of location on the determinants of steady-state and growth performance.

1. The Mankiw, Romer and Weil Model

The Mankiw, Romer and Weil model relies on a production function that follows the traditional hypotheses of the Solow model, and henceforth verifies the inherent conditions of a neoclassical technology: decreasing and positive marginal productivities (for each factor of production), constant scale returns (for both factors) and Inada conditions.

However, the MRW model formally differs from the Solow model because of the insertion of a variable representative of “human capital” in the production function. Thus, we have:

$$Y(t) = F\left(K(t), H(t), A(t)L(t)\right) = K(t)^{\alpha}\, H(t)^{\beta}\left(A(t)L(t)\right)^{1-\alpha-\beta} \tag{1}$$

With $\alpha > 0,\ \ \beta > 0,\ \alpha + \beta > 1$; α and β are constant.

Y represents the flow of production; K, the physical capital stock; H, human capital; A, is the level of technological progress, and L labor supply. In addition, technology and labor supply grow exogenously, more precisely as following:

$$\begin{aligned} A(t) &= A(0)e^{gt} \\ L(t) &= L(0)e^{nt} \end{aligned} \tag{2}$$

Then it follows that the number of efficiency units in the economy grows at rate $n + g$. The weighting of the production function (using the constant returns hypothesis) by the inverse of the efficiency units $(1/AL)$ allows us to rewrite it in an intensive form. Thus, we obtain:

$$y = f(k, h) = k^{\alpha} h^{\beta} \tag{3}$$

Therefore, the flow of production per efficiency unit is a function of the physical and human capital stocks per efficiency unit. Moreover, every period, constant and exogenous parts of production (respectively s_k and s_h) are devoted to physical and human capital accumulation. In the meantime, capital obsolescence implies the vanishing of a constant part, δ, in the physical and human capital stocks. Thus, the effective depreciation rate of these stocks equals $n+g+\delta$, and then, the more the number of efficiency units grows, the more the capital stock (both physical and human) decreases. Henceforth, we can write the following system that formally describes the dynamics of the physical and human capital stocks per efficiency unit. We have:

$$\begin{cases} \dot{k} = s_k k^{\alpha} h^{\beta} - (n+g+\delta)k \\ \dot{h} = s_h k^{\alpha} h^{\beta} - (n+g+\delta)h \end{cases} \tag{4}$$

In addition, the steady-state of the economy may be characterized by equaling each of the above first-order differential equations to zero (at steady-state, the physical and human capital stocks no longer vary). The solution that ensues from this system then defines the long-term equilibrium of the economy, and relies on the following values for k and h:

$$\bar{k} = \left[\frac{s_k^{1-\alpha} s_h^{\alpha}}{n+g+\delta}\right]^{\frac{1}{1-\alpha-\beta}} \quad \text{et} \quad \bar{h} = \left[\frac{s_k^{1-\beta} s_h^{\beta}}{n+g+\delta}\right]^{\frac{1}{1-\alpha-\beta}} \tag{5}$$

From the expression of the intensive production function in a logarithmic form, and for the steady-state values $\bar{k}$ and $\bar{h}$, we derive the “income equation” which represents the first estimable equation in the Mankiw, Romer, and Weil model, that is:

$$\ln \hat{y} = a + \frac{\alpha}{1-\alpha-\beta}\ln s_k + \frac{\beta}{1-\alpha-\beta}\ln s_h - \frac{\alpha+\beta}{1-\alpha-\beta}\ln(n+g+\delta) + \varepsilon \tag{6}$$

This equation explains long-term per capita income level by the human and physical capital accumulation rates, and by the corresponding effective depreciation rate. Therefore, Mankiw and *al.* rely on this equation to show that the reaction of long-term per capita income to the saving and population growth rates is stronger when the definition of capital is expanded to integrate human capital; elasticities of production to the latter variables are indeed higher in the MRW model than in the Solow model, what leads Mankiw and *al.* to consider that their model is able to account for the large income disparities one can observe on an international scale. Their analysis then relies on the following “convergence equation”:

$$\ln\left[\frac{y(T)}{y(0)}\right] = a + \theta\frac{\alpha}{1-\alpha-\beta}\ln s_k + \theta\frac{\beta}{1-\alpha-\beta}\ln s_h - \theta\frac{\alpha+\beta}{1-\alpha-\beta}\ln(n+g+\delta) - \theta\ln y(0) + \varepsilon \tag{7}$$

Insofar as the income equation highlights the fact that economies will have as different long-term income levels as their saving and population growth rates are different, Mankiw and *al.* uphold that "absolute convergence" hypothesis, often adopted as a the conclusion in the Solow model, is no longer valid and that one should rather think of a "conditional convergence" when accounting for growth experiences throughout the world.

Mankiw, Romer and Weil show in addition that the convergence equation estimation results corroborate the predictions of the model. Indeed, with a value of $1/3$ for the α and β parameters, and a 1% rate for the population growth rate, one can expect a 2% estimation for the convergence rate (this result is very common in growth literature) whereas one would have obtained a 4% rate in the Solow model. Moreover, the estimation results for α and β parameters in the constrained version sensitively fit what the model suggests.

Therefore, our analysis will rely on these two estimable equations which definitely constitute the basis the MRW model. The latter indeed offers a relatively satisfactory framework for analyzing the growth process, even if it does not really provide any explanation for it as emphasized by endogenous growth theoreticians. Mankiw replies that what is at stake for the neoclassical growth model is rather being able to account for the large disparities in economic growth.

2. Spatial Autocorrelation

According to Anselin and Bera (1998), spatial autocorrelation "can be loosely defined as the coincidence of value similarity with locational similarity". Otherwise, it expresses the existence of a functional relationship between observations in different locations over the space considered. The potential presence of spatial autocorrelation is largely due to the bidimensional nature of spatial data and to the multidirectional feature of relations in space. As a matter of fact, for every observation unit, we have information both on the observed value of a certain variable and the location of the aforementioned unit. Henceforth, two different observation units may be correlated just because of their geographical position, and then, this happens in all directions.

Formally, the presence of spatial autocorrelation between two any observation units i and j can be expressed by a non-zero covariance between the values taken by the focus variable in the two corresponding locations. Thus, we obtain:

$$Cov\ (y_i y_j) = E(y_i y_j) - E(y_i)E(y_j) \neq 0 \quad ; \ avec \quad i \neq j \tag{8}$$

Here, y_i *and* y_j refer to the values of the focus variable, respectively in i and j locations. However, it is worth noting that the above covariance may have a real spatial meaning only when the distribution of the observation units can be interpreted in terms of a spatial structure, interaction or arrangement.

From a strictly econometric point of view, the zero covariance implies that there is no more independence between observations and then the hypothesis of spherical errors, which is fundamental to the OLS method, is no longer valid. It follows that when the error term is spatially autocorrelated, OLS estimators may turn out to be inconsistent and/or inefficient depending on the structure of spatial dependence that exists between the different locations in space.

Indeed, in spatial econometrics, this structure relies on the definition of a specific process, accounting for the distribution of spatial units and subsequently conditioning the functional form inter-individual covariances. Then, such an approach is quite different from the one adopted in geostatistics because, in the latter case, the structure of covariances is rather *a priori* imposed.

However, from an estimation standpoint, a problem of identification arises for the n observations available in the sample do not ensure for the estimation of n individual variance and $n(n-1)/2$ inter-individual covariance terms. Then, spatial econometrics offers such tools as spatial weights matrices and spatially lagged variables which allow dealing with such problems.

2.1 Weight matrices and spatially lagged variables

Weight matrices have a key role in Spatial Econometrics. Indeed, not only do they ensure the resolution of the estimation problems related to the bidimensional and multidirectional nature of spatial data, but they also allow the definition of a topology across the space under consideration (by defining the relative locations of spatial units) and the relative weight of the corresponding spatial units.

From the definition of spatial autocorrelation given by Anselin and Bera (1998), one may easily imagine the important role of notions such as “proximity” and “neighborhood” for outlining spatial patterns, and more generally, for the modelling of spatial autocorrelation. These notions may have different interpretations, however.

Two main conceptions, respectively based on contiguity and distance, have generally been adopted for defining spatial weights. The first measures of spatial dependence are due to

Moran (1948) and Geary (1950), and are based on the concept of binary contiguity. Besides, they lead to the use of "contiguity matrices" which rely on the sharing of a common border between spatial units. In other words, two locations will be said to be contiguous as far as they are neighbors.

Formally, a contiguity matrix represents each location of the spatial system in row and column. The "spatial weights" (*i.e.* the elements of the weight matrix) are supposed to be 1 whenever we have two contiguous locations and 0, should it be otherwise. In the meantime, a given location cannot be contiguous to itself. Then, for any region i , and J the set of its neighbors, the elements of the weight matrix $W = (w_{ij})$ are defined as it follows:

$$w_{ij} = \begin{cases} 1 \; for \quad j \in J \\ 0 \; for \quad j \notin J \end{cases} ; \quad w_{ii} = 0 \quad , \quad \forall i \tag{9}$$

The idea of neighborhood exclusively based on the notion of contiguity has a certain number of shortcomings. Indeed, binary contiguity describes the pattern of the spatial system only very roughly, and consequently, it does not allow for a true capturing of the strong dependence relationships which may exist between spatial units. Besides, the notion of contiguity is no longer obvious when one faces a regular spatial pattern.

On the other hand, the principle of distance matrices relies on the general idea of an interaction all the stronger (weaker) as the distance between two any spatial units is longer (shorter). Cliff and Ord (1981) originally provided this kind of specification for spatial weights, by combining a function of the inverse distance between two locations and the relative length of their common border. Thus, the elements of the corresponding matrix can be written as following:

$$w_{ij} = \begin{cases} (d_{ij})^{-a} (\beta_{ij})^{b} ; \quad \forall i \neq j \\ w_{ii} = 0 \end{cases} \tag{10}$$

Where d_{ij} represents the distance between two spatial units i and j ; β_{ij} , the relative share of the common border between spatial units i and j in the total perimeter of i ; a and b are fixed parameters.

However, the most common specifications implemented in empirical studies involve much simpler expressions for the spatial weights. In fact, weight matrices very often rely on a negative exponential function or the inverse distance between two any i and j spatial units.

Thus, we formally have:

$$\text{(a)} \begin{cases} w_{ij} = e^{-\alpha d_{ij}}, & \forall i \neq j \\ w_{ii} = 0 \end{cases} \quad \text{or} \quad \text{(b)} \begin{cases} w_{ij} = (d_{ij})^{-\beta}, & \forall i \neq j \\ w_{ii} = 0 \end{cases} \qquad (11)$$

Where, α and β are fixed parameters.

Moreover, these different kinds of spatial weights may be generalized by either setting a cut-off distance beyond which any interdependence disappears, or by restricting the neighborhood for each spatial unit to a certain number (k) of locations (there is no interaction beyond that space), and then, the corresponding matrix is called a "k-nearest neighbors".

The use of distance matrices precisely requires the choice of a distance criterion. The first criterion one may think of probably is the one of geographical distance (Euclidian distance, great circle distance ...) but other concepts, based on social or economic variables, have also been suggested in literature with the concern of better comprehending inter-individual relationships over space (Cf. Case [1993], Conley et Ligon [2001]). However, this type of distance crucially poses the question of exogeneity for the spatial weights.

In addition, spatial weight matrices are generally row-standardized so as to have an easier interpretation for the spatial weights at the end of estimation. Thus, each row i of a given spatial weight matrix, W, is typically divided by the sum of its j elements (w_{ij}) and then, the spatial weights can be written as it follows:

$$w_{ij}^{\ s} = \frac{w_{ij}}{\sum_j w_{ij}} \quad , \forall i \qquad (12)$$

The row-standardization of the weight matrix also offers the advantage of ensuring the comparison between spatial parameters resulting from different models insofar as the weights of a row-standardized matrix no longer express absolute values but relative ones.

The major role of spatial weights may also be appreciated through the concept of "spatial lag" which, for any location (i) and for any focus variable (y), relates back to the weighted (by spatial weights) average of the corresponding observations in neighboring locations (J). Then, it synthesizes the information relating to the neighborhood of each location, and is obtained pre-multiplying y (the vector of the values taken by the focus variable in each location) by the spatial weight matrix (W). Thus, it may be written as following:

$$[Wy]_i = \sum_{j \in J} w_{ij} y_j \qquad (13)$$

With, w_{ij} represents the j^{th} element in row i.

It is worth noting that spatial weights may be defined for higher orders of the weight matrix and then, they correspond to the weighted average of y-values for "neighbors of neighbors".

2.2 Moran's *I* Test

Moran's "I" test (1948, 1950) for the absence of spatial autocorrelation was the first ever specification test to be suggested in spatial econometrics and generally constitutes the first stage in searching for the spatial process that best matches the data. In fact, this test allows one to pronounce on the appropriateness of specifications that explicitly include spatial effects even if, in itself, it does not give any indication on the way spatial autocorrelation should be modeled[1]. The test statistic is the following[2]:

$$I = \frac{n}{s}\left(\frac{\varepsilon' W \varepsilon}{\varepsilon' \varepsilon}\right) \qquad (14)$$

The ε term represents the vector of residuals resulting from the OLS estimation of the non-spatial model (the basic statistical model); W is the spatial weights matrix, n is the sample size, and s a standardization factor corresponding to the sum of all the elements of the spatial weights matrix. Cliff and Ord (1981) showed that when the assumption of a normal distribution is made for the error term (ε), under the null hypothesis of no spatial dependence, then the mean and the variance of the (I) statistic can be written:

$$E(I) = \frac{n}{s} tr \frac{(MW)}{n-K} \qquad (15)$$

And,

$$V(I) = \left(\frac{n}{s}\right)^2 \frac{\left\{ tr(MWMW') + tr(MW)^2 + [tr(MW)]^2 \right\}}{(n-K)(N-K+2)} \qquad (16)$$

Thence, the asymptotic distribution of the I statistic can be drawn from these two moments, and the test relies on the following Z_I variable:

$$Z_I = \frac{I - E(I)}{\sqrt{V(I)}} \rightarrow \mathrm{N}(0,1) \qquad (17)$$

[1] Moran's « I » test does not offer an alternative to the hypothesis of no spatial autocorrelation.

[2] For a row-standardized weights matrix, $s = n$ and the test statistic is rewritten: $I = \varepsilon' W \varepsilon / \varepsilon' \varepsilon$. The shape of this I statistic makes it very similar to the Durbin-Watson (1950) autocorrelation test for Time-Series.

The null hypothesis of no global spatial autocorrelation is typically rejected when the OLS residuals lead to a Z_I-value that is higher than the critical value in the standard normal distribution[3].

2.3 Econometric Specifications

2.3.1 The Spatial Autoregressive Model

The spatial autoregressive model ensues from introducing a spatially lagged dependent variable among the regressors of the standard linear model. It is often implemented when one has to deal with a spatial interaction pattern resulting from a theoretical model (Cf. Case and $al.$[1993]; Moreno and Trehan[1997]). Formally, we have:

$$y = \rho Wy + X\beta + \varepsilon \tag{18}$$

Or, in a reduced form:

$$y = (I - \rho W)^{-1} X\beta + (I - \rho W)^{-1} \varepsilon \tag{19}$$

These expressions imply that spatial autocorrelation occurs through the correlation between the spatially lagged endogenous variable (Wy) and the error term (ε). In fact, contrary to the Time-Series case, for which the lag y_{t-1} is only correlated with the error term ε_t when the latter is autocorrelated itself, the correlation in the spatial case is independent from the distribution of ε. Indeed, when the vector ε is such that its components (ε_i) are $i.i.d(0,\sigma^2)$, the mean of the dependent variable (y) is given by:

$$E(y) = (I - \rho W)^{-1} X\beta \tag{20}$$

Then, the covariance matrix can be written:

$$V(y) = \sigma^2 \left[(I - \rho W)' (I - \rho W) \right]^{-1} \tag{21}$$

This matrix is full, what denotes the fact that all locations in the spatial system interact, and the presence of a spatial autocorrelation in the data[4]. Moreover, the full interaction pattern highlighted by the covariance matrix can be split into two separate effects; this is done by rewriting $(I - \rho W)^{-1}$ into an infinite form in equation (19). Thus, we obtain:

---

[3] Burridge (1980) shows that Moran's I test is asymptotically equivalent to the Lagrange Multiplier test, whereas King (1981) shows that it is a "Locally Best Invariant" test.

[4] In fact, this reduced expression of the model is only possible when the inverse matrix $(I - \rho W)$ is non singular $i.e.$ for $|I - \rho W| \neq 0$. This condition is confirmed for $|\rho| \neq 0$ and when $1/\rho$ is not an eigenvalue of the weights matrix W.

$$y = \left(I + \rho W + \rho^2 W^2 + \ldots\right) X\beta + \left(I + \rho W + \rho^2 W^2 + \ldots\right)\varepsilon \quad (22)$$

In this expression, the first term on the right side denotes a *spatial multiplier effect* which means that in every location, (y) depends not only on the observations in the same location, but also on the observations made in any other location of the spatial system. As for the second term on the right member, it represents a *spatial diffusion effect* so that an exogenous shock coming from a given spatial unit affects the dependent variable in this location, but stretches over all the other units in the spatial system. These two effects decrease in intensity as the neighbourhood order increases.

2.3.2 The Spatial Error Model

This specification is based on the rejection of the hypothesis of spherical errors in the standard linear model, the choice of an explicit spatial process for the error term (ε). In fact, several types of processes may be used but the spatial autoregressive specification is the most commonly used[5]. Thus, we have:

$$\begin{cases} y = X\beta + \varepsilon \\ \varepsilon = \lambda W \varepsilon + u \end{cases} \quad (23)$$

In this specification, λ represents the spatial autoregressive coefficient related to the spatially lagged error term $(W\varepsilon)$, and (u) is a vector of homoskedastic errors. The corresponding reduced form to this specification can be written as following[6]:

$$y = X\beta + \left(I - \lambda W\right)^{-1} u \quad (24)$$

When the error term (u) is such that its components are $i.i.d\left(0, \sigma^2\right)$, the mean for (y) is given by $E(y) = X\beta$ and the covariance matrix has the following expression:

$$V(y) = V(\varepsilon) = \sigma^2 \left[\left(I - \lambda W\right)' \left(I - \lambda W\right)\right]^{-1} \quad (25)$$

As for the SAR model, this matrix is full and the spatial interaction pattern it represents is global with the result that all the locations in the spatial system interact. However, this spatial

---

[5] Alternative specifications such as the Moving Average (Cliff et Ord [1981], Haining [1988, 1990]) and Kelejian et Robinson [1993,1995] processes have also been suggested for the error term. However, the use of these specifications remains relatively uncommon in empirical studies. They are generally obtained by breaking down the error term into two components, the first representing the specific shocks to each location while the second denotes a weighted average of errors in neighboring locations.

[6] As for the SAR specification, this reduced form only exists when the $(I - \lambda W)$ matrix is non singular , and λ subject to the same conditions as for ρ .

interdependence is only relies on a *spatial diffusion effect* since equation (24) can be rewritten:

$$y = X\beta + \left(I + \lambda W + \lambda^2 W^2 + ...\right)\varepsilon \qquad (26)$$

Then, an exogenous shock in a given spatial unit affects the dependent variable (y) in all the locations of the spatial system under consideration; however, this impact decreases when moving away from the same spatial unit.

2.3.3 Alternative Specifications

By analogy with first-order differentiation in Time-Series, the dependent variable can be *spatially filtered* in the two aforementioned models. This operation consists in isolating the spatial autocorrelation and leads to consistent and efficient OLS estimators.

In the case of the SAR model, transposing ρWy into the left side gives the following equation:

$$(I - \rho W)y = X\beta + \varepsilon \qquad (27)$$

Then, $(I - \rho W)y$ is the spatially filtered variable, whereas the right side of the equation is the same as the one in the standard linear model. On the other hand, the spatial error model is such that the spatial filter applies to the endogenous variable (y) as well as to the exogenous variables in the (X) matrix. Indeed, pre-multiplying both sides of (24) by $(I - \lambda W)$, we obtain the equation below:

$$(I - \lambda W)y = (I - \lambda W)X\beta + u \qquad (28)$$

It results that the spatial error model is equivalent to a standard linear model in which both the endogenous and exogenous variables are spatially filtered. In addition, this expression of the spatial error model can be rewritten into a "Spatial Durbin" specification (Cf. Anselin, 1988).

In fact, developing the equation above and shifting the spatial autoregressive term (λWy) into the right hand side, we have[7]:

$$y = \lambda Wy + X\beta - \lambda WX\beta + u \qquad (29)$$

Or[8]:

$$y = \lambda Wy + X\beta - WX\gamma + u \qquad (30)$$

[7] This "Spatial Durbin" model refers to the Durbin model usually implemented in Time-Series. It is also known as the "Common Factor" model.

[8] The equivalence between these two specifications is not so obvious and it requires testing for a certain number of nonlinear constraints. These constraints boil down to the following condition: $\lambda\beta = -\gamma$, and the corresponding test is called the "Common Factor" test.

The "Spatial Durbin" model represents a reduced form of the spatial error model, and equations (29) and (30) also show that it is an extension of the spatial autoregressive model, obtained by adding a set of spatially lagged exogenous variables (WX) to equation (18).

The spatial error model is generally implemented when the analysis is about accounting for the diffusion of shocks or disturbances over a given space. Moreover, it allows solving some problems related to the omission of decisive variables in the phenomenon under consideration, this is more particularly the case when these variables are spatially correlated[9].

The SAR and SEM specifications may be combined into a general spatial model (GSM) which formal expression is given by:

$$\begin{cases} y = \rho W_1 y + X\beta + \varepsilon \\ \varepsilon = \lambda W_2 \varepsilon + u \end{cases} \tag{31}$$

As in the previous cases, this model can be rewritten into a reduced form, thus we obtain:

$$y = \left(I - \rho W_1\right)^{-1} X\beta + \left(I - \rho W_1\right)^{-1}\left(I - \lambda W_2\right)^{-1} u \tag{32}$$

And the covariance matrix is given by:

$$V(y) = \sigma^2 \left[\left(I - \rho W_1' \right)^{-1} \left(I - \lambda W_2' \right) \left(I - \lambda W_2 \right) \left(I - \rho W_1 \right) \right] \tag{33}$$

Moreover, Equation can also be written in the form of an extended "Spatial Durbin" model. Indeed, one can show that:

$$y = \rho W_1 y + \lambda W_2 y - \rho\lambda W_2 W_1 y + X\beta - \lambda W_2 X\beta + u \tag{34}$$

For $W_1 = W_2 = W$, this equation becomes[10]:

$$y = (\rho + \lambda) Wy - \rho\lambda W^2 + X\beta - \lambda WX\beta + u \tag{35}$$

Although the general spatial model has been implemented in certain empirical studies such as those of Case [1991, 1992] related to demand analysis or to the diffusion of innovation, it remains rarely used in comparison with the specifications it generalizes[11].

In conclusion, several specifications may be used to model spatial autocorrelation. More complex processes such as Huang's SARMA (Spatial Autoregressive Moving Average)

---

[9] This model is notably implemented in studies on hedonic prices (Pace et Gilley [1998], Dubin [1998]) or conditional convergence (Fingleton [1999]).

[10] In this specific case, ρ and λ parameters are only identified when the (X) matrix contains at least one exogenous variable element (apart from the constant term). Moreover, nonlinear constraints must be imposed these two specifications so to ensure that spatial parameters will be unique and identified.

[11] According Anselin an Bera (1998), this type of processes often results from misspecification of the weights Matrix which entails, for example, the presence of spatial autocorrelation in a model with a spatially lagged dependent variable.

model (1984) have also been suggested in literature, but the SAR and SEM models offer the advantage of being relatively simple from a statistical inference standpoint.

3. Maximum likelihood estimation

The recourse to the maximum likelihood principle as an estimation method for spatial models is due to the original work of Cliff and Ord (1973) and Ord (1975) who applied it to the SAR (spatially lagged dependent variable) and SEM (spatially autocorrelated errors). However, when applied to spatial models, the maximum likelihood method requires that some particular conditions ensuring the consistency, and asymptotic normality and efficiency of estimators are verified[12].

As usual, the method relies on the log-likelihood function and on the normality hypothesis for the model residuals. Its application to the GSM model offers a global presentation from which one can easily draw the corresponding results for the SAR and SEM models. Indeed, starting with the GSM model and making the assumption of a normal joint distribution for the vector of error terms (*i.e.* $u \to N(0, \sigma^2 I)$), we can write the likelihood function as it follows:

$$L(u) = \left(2\pi\sigma^2\right)^{-n/2} e^{\left(-\frac{1}{2\sigma^2}u'u\right)} \qquad (36)$$

However, as the error term (u) cannot be observed, the likelihood function must be expressed in terms of the observations on the endogenous variable (y). Thus we use the Jacobian of the transformation (J), that is:

$$J = \det\left(\frac{\partial u}{\partial y}\right) = \left|I - \rho W_1\right|\left|I - \lambda W_2\right| \qquad (37)$$

Where $u = (I - \lambda W_2)(I - \rho W_1 y - X\beta)$, and we can write the log-likelihood function in terms of the endogenous variable (y). Indeed, we have[13]:

$$\ln L(y|\rho, \lambda, \beta) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln\left(\sigma^2\right) + \ln\left|I - \rho W_1\right| + \ln\left|I - \lambda W_2\right| - \frac{1}{2\sigma^2}u'u \qquad (38)$$

---

[12] These regularity conditions specified by Heijmans et Magnus (1986a, 1986b, 1986c) et Magnus (1978) and boil down to the existence of the log-likelihood function, and to continuous and differentiable elements for the corresponding Score and Hessian matrix. For the most common spatial models, these conditions come down to restrictions on the spatial weights and on the parameter space associated to each spatial coefficient.

[13] Then, the log-likelihood function only exists when the matrices forming the Jacobian, *i.e.* $(I - \rho W_1)$ et $(I - \lambda W_2)$ are non singular. Moreover, these matrices are not triangular (as this may be the case for Time-Series), which considerably complicates the computations related to the evaluation of the log-likelihood function. However, Ord (1975) showed that spatial jacobians can be written in terms of the eigenvalues of the corresponding weights matrix. Thus, we have: $\left|I - \rho W_1\right| = \prod_{i=1}^{n}\left(1 - \rho\omega_{1i}\right)$, and $\left|I - \lambda W_2\right| = \prod_{i=1}^{n}\left(1 - \lambda\omega_{2i}\right)$.

Furthermore, the log-likelihood function exists when J is strictly positive, now the determinants that make up the Jacobian are strictly positive when the following conditions are respectively verified:

$$\frac{1}{\omega_{1\ \min}} < \rho < \frac{1}{\omega_{1\ \max}}, \text{ and } \frac{1}{\omega_{2\ \min}} < \lambda < \frac{1}{\omega_{2\ \max}} \tag{39}$$

The terms $\omega_{\min}$ and $\omega_{\max}$ respectively represent the highest negative and positive eigenvalues (in absolute value) of the corresponding weights matrix (W_1 for ρ, W_2 for λ)[14].

Then, maximum likelihood estimators are obtained by resolving the following system: $S(\theta) = \frac{\partial \ln L}{\partial \theta} = 0$; with representing the Score vector (the first-order partial derivatives of the log-likelihood function and $\theta = [\beta, \rho, \lambda]$, the vector containing the parameters of the model.

One can show that these estimators are asymptotically efficient (provided that regularity conditions are confirmed), as the variance-covariance matrix equals the inverse of Fisher's Information Matrix (Cramer and Rao's Lower Bound); indeed we have:

$$V(\theta) = [I(\theta)]^{-1}, \text{ with: } I(\theta) = E\left[\frac{\partial^2 \ln L}{\partial \theta^2}\right]$$

On the other hand, some authors like Anselin (1980, 1988) showed that the Instrumental Variables Method may also be used in order to deal with the non-consistency of estimators resulting from the correlation between a spatially lagged dependent variable (Wy) and an error term (ε). The Generalized Method of Moments has also been suggested for estimating models with spatially autocorrelated errors insofar as it leads to consistent estimates for the spatial autoregressive coefficient.

4. Specification Tests

In the previous section, we showed that the spatial autocorrelation existing in the residuals of the standard linear model can be modelled in several ways. However, the choice between the presented specifications requires implementing a series of tests which alternative hypothesis offers an explicit spatial specification (contrary to Moran's "I" test). These tests procedures may be based on the Likelihood Ratio, Wald or Lagrange Multiplier principles; however, the Lagrange Multiplier principle offers the great advantage of only requiring estimation under

[14] When a moving average process is chosen for the error term (ε), the parameter space for the spatial autoregressive coefficient is given by the following interval: $\left]-1/w_{\max} ; -1/w_{\min}\right[$.

the null hypothesis, which most often boils down to the classical linear regression model, estimated by OLS, and so considerably facilitating statistical inference.

4.1 Test for an omitted spatial error autocorrelation

This test is based on the omission hypothesis of a spatial autoregressive process for the error term ε (*i.e.* $\varepsilon = \lambda W\varepsilon + u$) in the standard linear regression model and concerns the nullity of the coefficient $(H_0 : \lambda = 0)$. Then, the test statistic is the following (Cf. Burridge [1980])[15]:

$$LM_\lambda = \frac{\left[\hat{\varepsilon}'W\hat{\varepsilon} / \hat{\sigma}^2\right]^2}{T} \qquad (40)$$

In this expression, T represents the trace of matrix $\left(W'W + W^2\right)$, whereas $\hat{\varepsilon}$ and $\hat{\sigma}^2$ are the estimates for ε and σ^2 in the constrained model. Under the null hypothesis H_0, we have: $LM_\lambda \to \chi^2(1)$.

4.2 Test for an omitted spatially lagged endogenous variable

The corresponding statistic to this test was defined by Anselin [1988] and can be written as it follows, under the null hypothesis of $\rho = 0$:

$$LM_\rho = \frac{\left[\hat{\varepsilon}'Wy / \hat{\sigma}^2\right]^2}{\hat{T}_1} \qquad (41)$$

In the expression above, we have $\hat{T}_1 = \left[\left(WX\hat{\beta}\right)'\left(I - X\left(X'X\right)^{-1}X'\right)\left(WX\hat{\beta}\right) + T\hat{\sigma}^2\right] / \hat{\sigma}^2$ .Under H_0, LM_ρ also has a $\chi^2(1)$ distribution.

However, Anselin and Bera [1998] show that in the local presence of ρ (*resp.* de λ), when carrying out the LM_λ test (*resp.* LM_ρ), the corresponding test statistics are no more distributed with a $\chi^2(1)$[16]. Then, two different approaches can be used: the first one boils down to testing the joint hypothesis $H_0 : \lambda = \rho = 0$ in the general spatial specification, whereas the second is conditional insofar as it consists in testing the omission of a spatial

[15] It is shown that this test statistic is identical to the one for an error term (ε) that is subject to a spatial moving average process.

[16] For example, in the case of the LM_λ test, the null hypothesis might not be accepted, even for $\lambda = 0$.

error autocorrelation in a model containing a spatially lagged endogenous variable, and vice-versa.

4.3 Joint test for a spatially lagged dependent variable and a spatial error autocorrelation

This test relies on the null hypothesis in the general spatial model, *i.e.* $H_0 : \lambda = \rho = 0$. When this hypothesis is accepted, one finds the standard linear regression model again whereas otherwise, there is indication about the process underlining spatial dependence. The test statistic is the following:

$$LM_{\lambda\rho} = \hat{E}^{-1}\left[\left(\hat{d}_\lambda\right)^2 \frac{\hat{D}}{\hat{\sigma}^2} + \left(\hat{d}_\rho\right)^2 T_{22} - 2\hat{d}_\lambda \hat{d}_\rho T_{12}\right] \qquad (42)$$

Here, $E = \left(D/\sigma^2\right)T_{22} - \left(T_{12}\right)^2$ and : $T_{ij} = tr\left[W_i W_j + W_i' W_j\right]$; $D = \left(W_1 X\beta\right)' M\left(W_1 X\beta\right) + T_{11}\sigma^2$ and $T = \left[\left(W' + W\right)W\right]$. d_ρ and d_λ represent the score vectors, respectively for ρ and λ. Under H_0, this statistic is distributed with a $\chi^2(2)$.

4.4 Conditional Tests

This approach comes down to testing one of the basic specifications (spatially lagged endogenous variable or spatial error autocorrelation), supposing that the other one is already present. Then, we can test for the omission of a spatially lagged dependent variable in a SAR model with spatially autocorrelated errors estimated by maximum likelihood. Then, the test relies on the residuals and the statistic is written:

$$LM_{\lambda|\rho} = \frac{\tilde{d}_\rho^{\,2}}{T_{22} - \left(T_{21A}\right)^2 \tilde{V}\left(\tilde{\rho}\right)} \qquad (43)$$

With: $T_{21A} = tr\left[W_2 W_1 A^{-1} + W_2' W_1 A^{-1}\right]$, and $A = I - \tilde{\rho} W_1$. Under the $\left(H_0\right)$ null hypothesis, $LM_{\lambda|\rho}$ with an unbiased $\chi^2(1)$. The omission of a spatially lagged endogenous variable in the SEM model can also be tested with the help of the following statistic:

$$LM_{\rho|\lambda} = \frac{\left[\tilde{\varepsilon}' B' B W_1 y\right]^2}{H_\rho - H_{\theta\rho}\tilde{V}\left(\tilde{\theta}\right)H_{\theta\rho}'} \qquad (44)$$

In this statistic, $\hat{\varepsilon}$ represents the vector of residuals resulting from the model with spatial autoregressive errors estimated by maximum likelihood, $\theta = \left(\beta', \lambda, \sigma^2\right)$ and $B = I - \tilde{\lambda} W_2$. Moreover, we have:

$$H_\rho = tr\left(W_1^{\,2}\right) + tr\left(BW_1B^{-1}\right)'\left(BW_1B^{-1}\right) + \frac{\left(BW_1X\tilde{\beta}\right)'\left(BW_1X\tilde{\beta}\right)}{\tilde{\sigma}^2} \qquad (45)$$

And:

$$\begin{bmatrix} \dfrac{(BX)'\, BW_1X\tilde{\beta}}{\tilde{\sigma}^2} \\ tr\left(W_2B^{-1}\right)BW_1B^{-1} + trW_2W_1B^{-1} \\ 0 \end{bmatrix} \qquad (46)$$

Under $H_0 : \rho = 0$, it is shown that $LM_{\rho|\lambda} \to \chi^2(1)$.

4.5 Robust tests

On the basis of the framework defined by Bera and Yoon (1993), Anselin and *al.* (1996) implemented robust tests to a local misspecification which are adjusted versions of LM_λ and LM_ρ tests. In fact, they allow obtaining an unbiased $\chi^2(1)$ as asymptotic distribution under the null hypothesis, and this in the respective presence of ρ or λ.

The adjusted version of the LM_λ test under $H_0 : \lambda = 0$ is written as following:

$$LM_\lambda^{\ *} = \frac{\left[\hat{d}_\lambda - T_{12}\hat{\sigma}^2\hat{D}^{-1}\hat{d}_\rho\right]}{T_{22} - T_{12}^{\ 2}\hat{\sigma}^2\hat{D}} \text{ , or for } W_1 = W_2 = W \text{ , } \quad LM_\lambda^{\ *} = \frac{\left[\hat{d}_\lambda - T\hat{\sigma}\hat{D}^{-1}\hat{d}_\rho\right]}{T\left(1 - T\hat{\sigma}^2\hat{D}\right)} \qquad (47)$$

In the same vein, the robust test $LM_\rho^{\ *}$ is written as following under the null hypothesis $\left(H_0 : \rho = 0\right)$:

$$LM_\rho^{\ *} = \frac{\left[\hat{d}_\rho - T_{12}T_{22}^{\ -1}d_\lambda\right]}{\hat{\sigma}^{-2}\hat{D} - T_{12}^{\ 2}T_{22}^{\ -1}} \text{ ; or for } W_1 = W_2 = W \text{ , } LM_\rho^{\ *} = \frac{\left[\hat{d}_\rho - \hat{d}_\lambda\right]}{\hat{\sigma}^{-2}\hat{D} - T} \qquad (48)$$

Anselin and Rey (1991) suggest a combination of the tests presented above in order to choose the spatial model that best represents the data, when Moran's test concludes that spatial autocorrelation is present. This procedure originally relied on the significance levels for LM_λ and LM_ρ tests, but Anselin and Florax (1995) suggested the addition of the robust $LM^*{}_\lambda$ and $LM^*{}_\rho$ tests which allow refining the data.

The tools we have just presented complete the analytical framework that will lead us to the definition of the spatial process governing relationships between the countries of our sample. Then, spatial econometrics allows us to take into account relative location, and measure the impact it may have on economic performance. However, Moran's spatial autocorrelation test should more generally be implemented whenever cross-sectional data are concerned. Admittedly, the lack of data on location has been a problem for a long time, but the recent development of Geographic Information Systems allowed overcoming this obstacle.

5. The effects of spatial autocorrelation in the MRW model

5.1 Sample Data

The data we use in this paper come from the 6.0 version of the "Penn World Table" series initiated by Summers and Heston in 1988. Our study relies on the same variables as those used by Mankiw, Romer and Weil, but observed on the 1960-1995 period; then, the countries of the sample are chosen on the basis of data availability for the variables and the period under consideration. Furthermore, we follow Mankiw and *al.* by excluding oil-producing countries, what finally leads us to a sample of 89 countries.

Thus, $\ln gdp95$ and $grate6095$ stand for the endogenous variables in the income and convergence equations of the MRW model. Indeed, these two variables respectively represent real GDP per capita in 1995 and the average yearly growth rate of GDP per capita through the period under consideration, in each of the 89 countries. The level of GDP per capita is itself obtained by the ratio of real GDP to the number of workers of the corresponding year for each economy.

In addition, we approach the accumulation rate of physical capital by the share of investment in real GDP. The variable $\ln school$ stands for the share of the working-age population enrolled in secondary school through the period while $\ln pop$ is the effective depreciation rate of the human and physical capital stocks, with observations on the population growth rate or, more precisely, on the growth rate of the working-age population (people whose age is between 15 and 64), given that like Mankiw and *al.* we fix the sum of the technological and depreciation growth rates to 5%.

In our estimations, we use a single spatial weights matrix (W), which elements are obtained by taking the inverse distance between two any countries for quantifying the intensity of their economic relations. The latter distance is computed by locating each country with the help of the geographic coordinates of its national capital (latitude and longitude), and applying the great-circle distance criterion[17]. The inverse distance matrix is very common in spatial econometrics and has been used in studies like those of Moreno et Trehan (1997) or Florax and Nijkamp (2003) to formally express the spatial interaction degree between economies. Moreover, we row-standardize the spatial weights matrix in order to make the interpretation our results easier.

5.2 The search for spatial specification

As noted above, the first step of our analysis is based on testing for the absence of spatial error autocorrelation in the basic specifications of the MRW model, estimated by the Ordinary Least Squares method. Then, results stemming from the OLS estimation and from Moran's "I" test, respectively for the convergence and income models, are given below:

Table I : Ordinary Least Squares Estimation.

Convergence Model

Ordinary Least-squares Estimates
Dependent Variable = growth6095

R-squared	0.5810
Rbar-squared	0.5610
Std Dev- squared	0.1637
Nobs, Nvars	89,5

Variable	Coefficient	*t*-statistic	*p*-probability
const	3.424325	3.359787	0.001175
lngdp60	-0.437037	-5.007810	0.000003
lniony	0.520125	5.643363	0.000000
lnschool	0.350873	3.982224	0.000145
lnpop	-1.111355	-2.900470	0.004755

[17] The longitudes and latitudes of the country capitals come from the "Thesaurus of Geographic Names" which is available on the following website: www.getty.edu. The spherical distance formula (great circle distance) computed in Spacestat © is given by:

$$d_{ij} = 3959 * arc\ \cos\ \left\{ \cos\ \left| Y_i - Y_j \right| * \sin X_i * \sin X_J + \cos X_i * \cos X_j \right\}$$; the X and Y variables respectively stand for the latitudes (lat) and longitudes (lon) of the country capitals, transformed as following: X = $(90 - lat + \pi)/180$ et $Y = (lon + \pi/180)$.

Income Model

Ordinary Least-squares Estimates			
Dependent Variable = lngdp95			
R-squared			0.8086
Rbar-squared			0.8019
Std Dev- squared			0.2354
Nobs, Nvars			89,4
Variable	Coefficient	*t*-statistic	*p*-probability
const	5.821510	5.138242	0.000002
lniony	0.537594	4.871840	0.000005
lnschool	0.647379	7.205227	0.000000
lnpop	-2.346590	-5.889002	0.000000

Table II : Moran I-test for spatial correlation in residuals

	Convergence model	**Income model**
Moran I	0.04738847	0.03652254
Moran I-statistic	2.92649295	2.36077552
Marginal Probability (*p*)	0.00551039	0.02458622
mean	-0.01867508	-0.01749225
standard deviation	0.02257431	0.02288011

The results given by Moran's "I" test for both the "convergence model" and the "income model" lead to the rejection of the null hypothesis corresponding to the absence of spatial autocorrelation; indeed, the marginal probability (*p*) is lower than 5% in the two specifications. Therefore, the equations given by the MRW model are misspecified (for our sample) for the hypothesis of independent observations which underlies OLS estimation is no more valid, and so the resulting estimations are non-consistent and inefficient.

5.3 Specification Tests

Given that the presence of spatial autocorrelation is corroborated by Moran's test, the next step leads us to wondering about the functional form of the spatial processes which generate the data, and at this prospect, we resort to the procedure initially suggested by Anselin and Rey (1991) and enhanced by Florax, Folmer, and Rey (2002).

First, we carry out two simple hypothesis tests (based on the OLS residuals) which allow us to make a choice between specifications respectively involving a spatially lagged dependent variable and a spatial error autocorrelation, in order to take into account the spatial dependence found in the convergence and income equations. Therefore, we have:

Table III: Simple tests of a spatial error autocorrelation or a spatially lagged dependent variable omission in the convergence and income models

LM lag test for an omitted spatially lagged dependent variable

	Convergence model	**Income model**
LM value	5.50203643	6.73366539
Marginal Probability (*p*)	0.01899434	0.00946100
chi(1) .01 value	6.63500000	6.63500000

LM error test for spatial correlation in residuals

	Convergence model	**Income model**
LM value	2.61073135	1.55073935
Marginal Probability (*p*)	0.10614280	0.21302633
chi(1) .01 value	6.63500000	6.63500000

The results in Table III suggest *a priori* that the spatial autoregressive specification (spatially lagged endogenous variable) best matches the data generating process for the convergence model as well as the income model. In fact, the omission test of a spatially lagged endogenous variable is accepted for these two specifications (the marginal probability is lower than 5%) whereas the SEM specification is rejected.

However, as previously noted, these simple tests do not take into account the possible local presence of a spatially lagged dependent variable when an omission test for a spatial error autocorrelation is carried out, and vice versa. As a result, these tests may be biased and then, one should rather implement robust tests in order to deal with this flaw. The results corresponding to the latter are given in the table below:

Table IV: Robust tests of a spatial error autocorrelation or a spatially lagged dependent variable omission in the convergence and income models

Robust LM lag test for an omitted spatially lagged dependent variable

	Convergence model	**Income model**
LM value	2.94649641	5.19171438
Marginal Probability (*p*)	0.08606395	0.02269482
chi(1) .01 value	6.63500000	6.63500000

Robust LM error test for spatial correlation in residuals	Convergence model	Income model
LM value	0.05519133	0.00878833
Marginal Probability (*p*)	0.81426434	0.92531085
chi(1) .01 value	6.63500000	6.63500000

Table IV shows that the omission hypothesis of a spatial error autocorrelation is highly rejected as the marginal probability (*p*) is higher than 80% both for the convergence and income models. On the contrary, the robust tests suggest that the addition of a spatial lag, respectively for the average growth rate of income per capita and for long-tem income per capita (in logs), to the exogenous variables resulting from the theoretical model allows well capturing the spatial autocorrelation found following estimations in Table I; this also corroborates the results in Table III.

5.4 Estimation of the spatial specifications

Therefore, the SAR specification can be adopted as the data generating process both for the convergence and income models. Then, we can already estimate the "spatial convergence model" and the "spatial income model" by the maximum likelihood method, and so we obtain the following results:

Tableau IV: Estimation of spatial autoregressive specifications

Convergence Model

Spatial autoregressive Model Estimates
Dependent Variable = growth6095

R-squared	=	0.5970
Rbar-squared	=	0.5779
sigma^2	=	0.1443
Nobs, Nvars	=	89,5
log-likelihood	=	-9.9378068

Variable	Coefficient	Asymptot *t*-stat	*p*-probability
const	3.477497	3.627769	0.000286
lngdp60	-0.427146	-5.209702	0.000000
lniony	0.502146	5.781010	0.000000
lnschool	0.309672	3.692904	0.000222
lnpop	-0.891775	-2.404042	0.016215
rho	0.545988	2.674385	0.007487

Income Model

Spatial autoregressive Model Estimates			
Dependent Variable = lngdp95			
R-squared = 0.8198			
Rbar-squared = 0.8134			
Time for t-stats = 0.0310			
Nobs, Nvars = 89,4			
log-likelihood = -25.811205			
Variable	Coefficient	Asymptot *t*-stat	*p*-probability
const	2.748930	1.846952	0.064754
lniony	0.505757	4.844385	0.000001
lnschool	0.536354	5.868178	0.000000
lnpop	-1.808048	-4.223398	0.000024
rho	0.459989	2.814203	0.004890

Observing these results, we find that all the exogenous variables in the right side of the basic specifications (non spatial models) remain significant even after the insertion of a spatially lagged endogenous variable as an additional regressor. This is also the case for the spatial parameter (ρ) which is highly significant in the two spatial specifications respectively related to the convergence and income models.

Therefore, the specification search we have just carried out suggests that the spatial autoregressive specification offers the best representation of the data for the sample under consideration, and this is moreover confirmed by the following tests for the absence of spatial autocorrelation in the SAR specifications estimated in Table IV:

Tableau V*:* Tests for the absence of spatial autocorrelation in the residuals of the SAR specifications

Convergence Model

LM error tests for spatial correlation in SAR model residuals	
LM value	3.23582865
Marginal Probability (*p*)	0.07204385
chi(1) .01 value	6.63500000

Income Model

LM error tests for spatial correlation in SAR model residuals	
LM value	1.46114281
Marginal Probability (*p*)	0.22674837
chi(1) .01 value	6.63500000

Once the spatial data generating process is chosen, we move to the next step which consists in studying its relevance, *i.e.* assessing the impact of the spatially lagged dependent variable added to the convergence and income models on the econometric results and theoretical conclusions obtained by Mankiw, Romer, and Weil.

5.5 Results and Discussion

Income Model

The estimation results of a SAR specification for the income equation show that the spatial autoregressive parameter (ρ) is highly significant, with a marginal probability (*p*) which is about 0.5%. So, this corroborates the omission of a spatial lag for the endogenous variable in the standard specification (non spatial model) and the non-consistency and non-efficiency of the estimates obtained in Table I. In addition, the spatial parameter estimate shows that the spatial effects omitted in the standard specification are relatively important $(\tilde{\rho} = 0.46)$. As a matter of fact, for any given country, a 1% increase in the weighted average income per capita of its neighbors entails an increase of about 45% of its own long-term income per capita.

However, the addition of a spatially lagged endogenous variable to the standard specification does not considerably modify the effect of the accumulation rate of physical capital on the long-term per capita income level, even if it slightly decreases. This is also the case for the variables respectively standing for human capital and the population growth rate. Henceforth, these results comply with what is often found in literature, and highlight the fact that the most important part of the spatial autocorrelation found in the data is related to some omitted variables in the theoretical model; moreover, these variables are probably autocorrelated and have a significant impact on the long-run per capita income level.

Convergence Model

The estimation of the SAR specification in the case of the convergence model shows that all the coefficient estimates for the exogenous variables of the non-spatial model are significant, and the corresponding signs match what economic theory suggests. In particular, we have a negative and very significant coefficient for the log of initial income level, what corroborates the conditional convergence hypothesis. Moreover, as for the income equation, we notice that the coefficients resulting from the SAR estimation are very close to those for OLS, even if the impact of the corresponding variables is globally (in absolute value) higher in the standard estimation.

The results in table IV give a positive and significant estimation of the spatial autoregressive coefficient of about 0.55, which is again relatively important insofar as it the fact that the growth rate of an economy will *ceteris paribus react* to an 1% point increase in the weighted average growth rate of the other countries by a 0.55 percentage point increase. Thus, from the growth rate perspective, there exist important geographic spillover effects beyond THE economies of our sample.

In addition, the coefficient corresponding to the log of initial per capita income (*lngdp60*) in the SAR specification suggests an estimation of 1.6% for the speed of convergence, while the non-spatial model leads to an estimation of 1.5%.

In other words, the addition of a spatially lagged dependent variable to the basic specification does not significantly affect the estimation of the rate at which economies move towards their steady-states. Besides, this result is remarkably robust for several empirical studies on regional convergence lead to the same conclusion. More generally, the estimation of the speed of convergence given by the SAR specification remains close to the 1.4% rate obtained by Mankiw, Romer, and Weil for their sample of 98 countries, or to the 2% rate which is very recurrent in empirical studies on economic growth (Cf. Barro et Sala-i-Martin [1995]).

Conclusion

Finally, the results above have allowed us to show that the both the “convergence equation” and the “income equation” are misspecified and that neither the long-run income level, nor the growth rate escape from the effects of location and space. As a matter of fact, in both cases, Moran’s test leads us to strongly reject the null hypothesis and suggests the omission of a significant spatial autocorrelation in the specifications of the basic model; thus, we draw the conclusion that OLS estimations are non-consistent and non-efficient.

Henceforth, spatial econometric methods allow us to obtain a reliable statistical inference. In addition, when we implement the search procedure for spatial specifications suggested by Anselin and Rey (1991) and enhanced by Florax, Folmer, and Rey (2003), we are lead to adopt the SAR specification as the data-generating process, and the estimation of this specification for the convergence and income equations allows us to highlight the omission of important geographic spillover effects in the theoretical framework proposed by Mankiw, Romer, and Weil.

Indeed, the long-run per capita income and the growth rate of an economy will be all the higher as the weighted average (by inverse distance) of per capita incomes and growth rates of neighboring economies is high. In other words, the richer its neighbors will be (*resp.* poorer), the richer the economy under consideration will be (*resp.* poorer); the faster (*resp.* slower) its neighbors will grow, the faster (*resp.* slower) its per capita income growth will be.

Considering the specification of our spatial weights matrix, we also note that the interaction of a country with its neighbors will be all the weaker as the distance separating them is high. However, alternative specifications for the spatial weights based on explicit economic variables, however complex, could allow better capturing the intensity of economic relations in space but this type of matrices raises the issue of exogeneity for the spatial weights.

Even if the explicit consideration of spatial autocorrelation in the convergence and income equations does not change drastically the estimated coefficient values of the Mankiw, Romer, and Weil's model, the addition of a spatially lagged dependent variable in these equations allows highlighting the omission of some important variables which underlie the growth process and are strongly influenced by the geographic position of these economies.

Henceforth, several factors may explain the geographic spillovers we find in the convergence and income equations of the MRW model. As a matter of fact, capital mobility, international trade, and the technological transfers it generates, are so many vectors through which the spatial autocorrelation observed beyond per capita incomes and growth rates may happen.

In particular, the effects of technological diffusion absolutely match the MRW model which theoretical construction leads to explaining long-run growth through the only growth of technology. This is also the case for international trade which plays a fundamental role in popularizing technologies and probably considerably participates in making technical progress a public good.

Finally, even if the impact of the standard exogenous variables is not fundamentally changed, the consideration of spatial autocorrelation in the MRW model certainly allows obtaining more accurate results. In general, the addition of a spatially lagged endogenous variable in the convergence and income equations suggest that the effects of the exogenous variables are slightly overestimated and then, it notably follows a slight underestimation of the speed of convergence. This indicates the need to systematically test for the omission of spatial autocorrelation in the cross-section regressions traditionally implemented in empirical studies on growth.

References

Agénor, P.R. (2000), "*The Economics of Adjustment and Growth*", Academic Press.

Anselin, L., (1980), "*Estimation methods for spatial autoregressive structures*", Cornell University, Regional Science Dissertation and Monograph Series #8, Ithaca, NY.

Anselin, L., (1988), "Spatial Econometrics: Methods and Models", Kluwer Academic Publishers, Dordrecht.

Anselin, L., (2001a), "Rao's score test in spatial econometrics", *Journal of Statistical Planning and Inference*, 97, 113-139.

Anselin, L. (2001b), "*Spatial Econometrics*", in: Baltagi B. (ed.), Companion to Econometrics, Basil Blackwell, Oxford.

Anselin, L., Bera, A.K. (1998), "Spatial dependence in linear regression models with an introduction to spatial econometrics", in *Handbook of applied Economic Statistics* (A.Ullah, and D. Giles, Eds.). Marcel Dekker, New York.

Anselin, L., Florax, R.J.G.M. (1995), "Small sample properties of tests for spatial dependence in regression models, in: Anselin, L., Florax, R.J.G.M. (eds), *New Directions in Spatial Econometrics*, Springer, Berlin.

Anselin, L., Griffith, D.A. (1988), "Do spatial effects really matter in regression analysis?",
Papers of the Regional Science Association, 65, 11-34.

Anselin, L., Rey, S. (1991), "Properties of tests for spatial dependence in linear regression models, *Geographical Analysis*, 23, 112-131.

Aten, B. (1997), "Does space matter? International comparisons of the prices of tradables and
nontradables", *International Regional Science Review*, 20, 35-52.

Azariadis, C., Drazen, A. (1990), "Threshold externalities in economic development", *Quarterly Journal of Economics*, CV, 501-526.

Baumol, W.J. (1986), "Productivity growth, convergence and welfare: what the long run data show", *American Economic Review*, 78, 1155-1159.

Baumont, C., Ertur, C., Le Gallo, J. (2002), ‘‘Estimation des effets de proximité dans le processus de convergence régionale : une approche par l'économétrie spatiale sur 92 régions européennes (1980-1995), *Revue d'Economie Régionale et Urbaine*, 2.

Barro, R.J. (1991), "Economic growth in a cross-section of countries", *Quarterly Journal of Economics*, 106:2, 407-443.

Barro, R.J., Sala-i-Martin, X. (1995), "*Economic Growth*", New York, McGraw-Hill.

Benhabib, J., Spiegel, M.M. (1994), "The role of human capital in economic development: evidence from cross-country data", *Journal of Monetary Economics*, 34, 143-173.

Bera, A.K., Yoon, M. (1993), "Specification testing with locally misspecified alternatives", *Econometric theory*, 9, 649-658.

Bernard, A.B. and Durlauf, S.N. (1996), "Interpreting tests of the convergence hypothesis", *Journal of Econometrics*, 71, 161-173.

Can, A. (1990), "The measurement of neighborhood dynamics in urban house prices, *Economic Geography*, 66, 254-272.

Case, A.C. (1991), "Spatial patterns in household demand", *Econometrica*, 59, 953-965.

Case, A.C. (1992), "Neighborhood influence and technological change", *Regional Science and Urban Economics*, 22, 491-508.

Case, A.C., Rosen, H., S., Hines, J.R. (1993), "Budget spillovers and fiscal policy interdependence: Evidence from the states, *Journal of Public Economics*, 52, 285-307.

Caselli, F., Esquivel, G., Lefort, F. (1996), "Reopening the convergence debate: A new look at the cross-country growth empirics", *Journal of Economic Growth*, 1, 363-389.

Cass, D. (1965), "Optimum growth in a aggregative model of capital accumulation", *Review of Economic Studies*, 32, 223-240.

Chiang, A.C. (1984), "*Fundamental Methods of Mathematical Economics*", 2 ed. New York: McGraw-Hill Book Company.

Cho, D. (1996), "An alternative interpretation of conditional convergence results", *Journal of Money, Credit, and Banking*, 28, 669-681.

Cressie, N. (1993), "*Spatial Statistics for spatial data*", John Wiley, New York.

Cliff, A.D., Ord, J.K. (1972), "Testing for spatial autocorrelation among regression residuals", *Geographical Analysis*, 4, 267-284.

Cliff, A.D., Ord, J.K. (1973), "*Spatial Autocorrelation*", Pion, Londres.

Cliff, A.D., Ord, J.K. (1981), "*Spatial Processes: models and applications*", Pion, Londres.

Conley, T.G. (1996), "*Econometric modelling of cross sectional dependence*", Manuscript, Northwestern University.

De Long, J.B. (1988), "Productivity growth, convergence and welfare: Comment", *American Economic Review*, 78, 1138-1154.

Dubin, R.A. (1988), “Estimation of regression coefficients in the presence of spatially autocorrelated error terms, *Review of Economics and Statistics*, 70, 466-474.

Dubin, R.A. (1998), “Predicting house prices using multiple listings data”, *Journal of Real Estate Finance and Economics*, 17, 35-59.

Durlauf, S.N., Johnson, P.A. (1995), “Multiple regimes and cross-country growth behaviour”, *Journal of Applied Econometrics*, 10:4, 365-384.

Durlauf, S.N., Quah, D.T. (1999), “*The New Empirics of Economic Growth*” in Taylor, J.B., and Woodford, M., eds., Handbook of Macroeconomics Vol. 1A (North-Holland).

Easterly, W., Levine, R. (1998), “Troubles with the neighbours: Africa’s Problem, Africa’s Opportunity”, *Journal of African Economics*, 7:1, 120-142.

Ertur, C., Le Gallo, J. (2002), ‘‘An exploratory spatial data analysis of European regional disparities, 1980-1995, in: Fingleton B. (ed.), *European Regional Growth*, Springer, Berlin.

Fingleton, B. (1999), “Estimates of time to economic convergence: an analysis of regions of the European Union”, *International Regional Science Review*, 22, 5-34.

Florax R., Folmer H., Rey S. (2003) Specification searches in spatial econometrics: The relevance of Hendry’s methodology, forthcoming in *Regional Science and Urban Economics*, 33.

Florax, R.J.G.M. & Nijkamp, P. (2003), “*Misspecification in linear spatial regression models*”, Tinbergen Institute Discussion Papers 03-081/3, Tinbergen Institute.

Galor, O. (1996), “Convergence? Inferences from theoretical models”, *Economic Journal*, 106, 1056-1081.

Gandolfo, G. (1997), “*Economic Dynamics*”, 3rd Ed., Springer-Verlag, Berlin.

Grossman, G. M., Helpman, E. (1994), “Endogenous Innovation in the Theory of Growth”, *Journal of Economic Perspectives*, 8: 23 – 44.

Heijmans, R.D.H., Magnus, J.R. (1986a), “On the first-order efficiency and asymptotic normality of maximum likelihood estimators obtained from dependent observations”, *Statistica Neerlandica*, 40, 169-188.

Heijmans, R.D.H., Magnus, J.R. (1986b), “Asymptotic normality of maximum likelihood estimators obtained from normally distributed but dependent observations”, *Econometric Theory*, 2,374-412.

Heijmans, R.D.H., Magnus, J.R. (1986c), “Consistent maximum likelihood estimation with dependent observations: the general (non normal) case and the normal case, *Journal of Econometrics*, 32, 253-285.

Huang, J.S. (1984), ‘The autoregressive moving average model for spatial analysis”, *Australian Journal of Statistics*, 26, 169-178.

Inada, K. (1964), “Some structural characteristics of turnpike theorems”, *Review of Economic Studies*, 31 (January):43-58.

Magnus, J. (1978), “Maximum likelihood of the GLS model with unknown parameters in the disturbance covariance matrix, *Journal of Econometrics*, 7, 281-312 (Corrigenda, Journal of Econometrics, 10, 261).

Kelejian, H.H., Robinson, D.P. (1993), “A suggested method of estimation for spatial interdependent models with autocorrelated errors, and an application to a county expenditure model, *Papers in Regional Science*, 72, 297-312.

Kelejian, H.H., Robinson, D.P. (1995), “Spatial correlation: A suggested alternative to the autoregressive model”, in Anselin, L., Florax, R.J.G.M. (eds), *New Directions in Spatial Econometrics*, Springer, Berlin.

Kendrick, J.W. (1976), “*The formation and stocks of total capital*”, New York: Columbia University for NBER.

King, M.L. (1981), “A small sample property of the Cliff-Ord test for spatial correlation”, *Journal of the Royal Statistical Society*, B, 43, 263-264.

Klenow, P.J., Rodriguez-Clare, A. (1997), “The neoclassical revival in growth economics: has it gone too far?”, *NBER Macroeconomics Annual*, 1997, 73-102.

Koopmans, T. (1965), “*On the concept of optimal growth, The econometric approach to development planning*”, North-Holland, Amsterdam.

Lee, K., Pesaran, M., Smith, R. (1998), “Growth empirics: A panel data approach - a comment”, *Quarterly Journal of Economics*, vol. 113, Feb., 319-323.

Le Gallo, J. (2002), "*Disparités géographiques et convergence des régions européennes : une approche par l'économétrie spatiale*", Thèse de doctorat, LATEC, Université de Bourgogne.
Le Gallo, J., Ertur, C. (2002), ‘‘Exploratory spatial analysis of the distribution of regional per capita GDP in Europe, 1980-1995, *Papers in Regional Science*.
Le Gallo, J., Ertur, C., Baumont, C. (2002), "A spatial econometric analysis of convergence,
across European regions, 1980-1995, in: Fingleton B. (ed), *European Regional Growth*, Springer, Berlin.
Lucas, R.E. (1988), "On the mechanics of economic development", *Journal of Monetary Economics*, 22, 3-42.
Mankiw, N.G., Romer, D., Weil, D. (1992), "A contribution to the empirics of economic growth", *Quarterly Journal of Econometrics*, CVII, 407-437.
Mankiw, N.G. (1995), "The growth of nations", *Brooking Papers on Economic Activity*, 1, 275-310.
Moran, P. (1948), "The interpretation of statistical maps", *Journal of the Royal Statistical Society*, 10B, 243-251.
Moran, P. (1950), "A test for serial independence of residuals", *Biometrika*, 37, 178-181.
Moreno, R., Trehan, B. (1997), "Location and the growth of nations", *Journal of Economic Growth*, 2, 399-418.
Ord, J.K. (1975), "Estimation methods for models of spatial interaction", *Journal of the American Statistical Association*, 70, 120-126.
Pace, R. K., Gilley, Gilley, O.W.(1998), "Optimally Combining OLS and the Grid estimator," *Real Estate Economics*, Volume 26, Number 2, pp. 331-347.
Paelinck, J.H.P., Klaassen, L.H. (1979), "Spatial Econometrics", Saxon House, Farnborough. Pesaran, M.H., Smith, R. (1994), 'Estimating long-run relationships from dynamic heterogeneous panels", Working Paper, University of Cambridge and forthcoming, Journal of Econometrics.
Romer, P.M. (1986), "Increasing returns and long-run growth", *Journal of Political Economy*, 94, 1002-1037.
Rey, S.J., Montouri, B.D. (1999), "U.S. regional income convergence: a spatial econometric perspective", *Regional Studies*, 33, 145-156.
Solow, R. (1956), "A contribution to the theory of economic growth, *Quarterly Journal of Economics*, 70, 65-94.
Shone, R., (1997), "*Economic Dynamics: Phase Diagrams and Their Economic Applications*", Cambridge: Cambridge University Press.
Summers, R., Heston, A. (1988), "A new set of international comparisons of real product and prices: estimates for 130 countries", *Review of Income and Wealth*, 34, 1-26.
Summers, R., Heston, A. (1991), "The Penn World Table, Mark5: An expanded set of international comparisons, 1950-1988", *Quarterly Journal of Economics*, 106:2, 327-368.
Temple, J. (1999), "The new growth evidence", *Journal of Economic Literature*, 37, 112-156.