Total Factor Productivity equation and development accounting: New evidence

Wilfried KOCH*
Laboratoire d'Economie et de Gestion
UMR CNRS 5118 - Université de Bourgogne
Pôle d'Economie et Gestion, B.P. 26611,
21066 Dijon Cedex, France
kochwilfriedfr@aol.com

First Draft: April 20, 2006

Abstract

The aim of this paper is reconciliating spatial autocorrelation often observed in the world distribution of wealth and its effects on development gap. Neglecting spatial autocorrelation potentially biased our vision on the role of physical capital in development process. We therefore show that the total contribution of physical capital account almost 90% of differences between developing countries and the most richest countries.

KEYWORDS: Development accounting, Total Factor Productivity, Spatial Autocorrelation, Technological interdependence.

JEL: C21, O33, O40

*I would like to thank, Cem Ertur for valuable comments and suggestions. The usual disclaimer applies.

1 Introduction

Many papers evaluate critically the roles of investment and physical capital accumulation in economic growth and development. The differences in physical capital accumulated between countries seem not sufficient to explain the development gap in the world repartition of wealth. Growth and development accounting provide a powerful and simple tool to study this enormous cross-country differences in income per worker level and growth.

Growth accounting provides a breakdown of observed economic growth into components associated with changes in factor inputs and a residual that reflects technological progress and other elements (Barro, 1999). The basics of growth accounting were presented in Solow (1957), Kendrick (1961), Denison (1962, 1967), Maddison (1982), Jorgenson and Griliches (1967), and Jorgenson (1995a, 1995b). The residual component is usually named the Solow residual in the literature since the seminal paper of Solow (1957) or our "measure of ignorance" on the causes of growth (Denison, 1962).

In this paper, we focus explicitly on levels instead of rates of growth for several reasons as also underline by Hall and Jones (1999). First, levels capture the differences in long-run economic performance that are most directly relevant to welfare as measured by the consumption of goods and services. Second, many papers highlights the empirical relevance of levels differences studies instead of a growth rates study (Easterly et al. (1993), Hall and Jones (1999), Jones (1995)). Moreover, a number of recent models of idea flows across countries such as Barro and Sala-i-Martin (1997), Coe and Helpman (1995), Eaton and Kortum (1996), Howitt (2000), Keller (2002), Klenow and Rodriguez-Clare (2005), Parente and Prescott (1994) imply that all countries will grow at a common rate in the long run. We use this literature to challenge an augmented development accounting equation involving technological interdependence implied by this international idea flow. Moreover, it is a stylized fact that the wealth and the poverty are concentrated in space. This fact have an important consequence on the estimation of the contribution of the physical capital in the development gap between countries. In fact, the accounting equation specification will then suffers of a spatial autocorrelation problem. We show that the literature underestimates the role of the physical capital in explaining development gap because of technological interdependence generated by geographical diffusion of knowledge.

Using a very simple accounting equation, the contribution of physical capital input is evaluated and some proprieties of spatial autocorrelation are considered. The spatial repartition of wealth

and input, exhibiting strong spatial autocorrelation, plays an important role in the development gap between countries. The global technological interdependence in our equation shows also the importance of transitory technological shock diffusion in explaining transitory differences between countries.

The paper is organized as following. In the next section, we present the equation used in this paper in order to evaluate the contribution of the inputs and the Total Factor Productivity in the developing gap using a development accounting approach. Development accounting uses cross-country data on output and inputs, at point in time, to assess the relative contribution of differences in factor quantities, and differences in the efficiency with which those factors are used, to these vast differences in per-worker incomes (Caselli, 2005). Conceptually, development accounting can be thought of as quantifying the relationship:

$$Income = F(Inputs, TFP) \tag{1}$$

If one found that inputs are able to account for most of the differences, then development economies could focus explaining low rates of factor accumulation. Instead, should one find that Total Factor Productivity differences play a large role, then one would have to confront the additional task of explaining why some countries extract more output than others from their factors of production. The consensus view in development accounting is that Total Factor Productivity plays a very large role since it accounts for at least 50 % of differences in per capita income.

Among Caselli (2005), one potentially promising research strategy is to try to "chip away" at this residual by improving on these two steps that is by looking at alternative functional forms, and by attempting a more sophisticated measurement of Income and Factors. In other word, he proposes to find the missing input. The missing input we propose is "Space". More precisely, we suppose that the income of a country i depends on its accumulated inputs, its Total Factor Productivity and on Total Factor Productivity of other countries j:

$$Income_i = F(Inputs_i, TFP_i, TFP_j) \tag{2}$$

With a Cobb Douglas production function we develop a specification which includes explicitly this technological interdependence. The model will be estimated with a spatial econometric

specification and the structural parameter will be recovered.

In section 3, we present the database constructed with the last version of the Penn World Table 6.1 (Heston, Summers and Aten, 2005). An evidence of global and local spatial autocorrelation is displaying using Exploratory Spatial Data Analysis (ESDA) tools (Anselin, 1995). We estimate the model using spatial econometric specification and evaluate its predictions in section 4. The results support our model and permit us to evaluate the contribution of the physical capital in the development gap using the method proposed by Denison (1967) King and Levine (1994) and Easterly and Levine (2001). For a group of nine countries, Denison asked the development accounting question: what part of cross-country differences in income per capita is accounted for by differences in physical capital per capita? Denison (1962, 1967) and others[1] found that differences in the level of physical capital per worker accounted for very little of the differences in income per capita across countries. Our results leads to consider the contribution of physical capital as very important for the most poorest countries (higher than 90 %) in the development gap.

2 Theoretical and empirical model

2.1 The missing input: "the space"

Let us consider N countries indexes by $i = 1, ..., N$. Each country is characterized by a Cobb-Douglas production function exhibiting constant return to scale:

$$Y_i(t) = A_i(t)K_i(t)^{\alpha}L_i(t)^{1-\alpha} \tag{3}$$

with the standards notations: Y_i is the output, K_i the physical capital, L_i the labor and A_i the Total Factor Productivity or the Solow residual. However, we suppose that this residual is contaminated by physical capital externalities because of Learning by Doing process (Arrow 1962, Romer 1986) and by technological interdependence between countries. As in the Romer (1986) model, each firm behaves competitively, taking as given the aggregate capital stock $K_i(t)$ and the stock of knowledge of all countries in the world. However, they gain better on the knowledge of the foreign countries if there exists an higher proximity between the countries.

[1] King and Levine (1994) and Easterly and Levine (2001) using the same methodology than Denison found that these differences account for between 28 and 59 percent of the gap development between the most richer countries and the most poorer countries.

The Solow residual is written now:

$$A_i(t) = \Omega(t) k_i(t)^{\phi} \prod_{j \neq i}^{N} A_j(t)^{\gamma w_{ij}} \tag{4}$$

with $k_i(t) = K_i(t)/L_i(t)$ the physical capital per worker, $0 < \phi < 1$ the level of physical capital externalities reflecting the Learning by Doing process. The technological interdependence is represented by the parameter $0 < \gamma < 1$, and we assume that the international diffusion of knowledge underlying this technological interdependence is not perfect because there exist frictions between the home country i and foreign countries $j \neq i$, $j = 1, ..., N$ represented by w_{ij}. These terms are supposed positive, lower by 1, fixed, and $\sum_{j \neq i}^{N} w_{ij} = 1$. $\Omega(t)$ is the new residual obtained after considering these elements which model the relationship between countries. More lower are these elements and more higher the firms of the home countries profit on the knowledge of foreign countries. These parameters are the material of knowledge diffusion. For instance, Coe and Helpman (1995) view these parameters as the trade of commodities between countries. The theory behind this interpretation follows Grossman and Helpman (1991) who claims that "*international trade in tangible commodities facilitates the exchange of intangible ideas*" (Grossman and Helpman (1991), Chapter 6, p.166). In fact, the larger the volume of international trade, the greater presumably will be the number of personal contacts between domestic and foreign individuals. Another interpretation of these parameters should be the technological diffusion between countries in a quality lader model (Howitt (2000), Barro and Sala-i-Martin (1997)), transport cost in inputs trade (Keller 2002), FDI flows between countries (Caves, 1996) or the importance of geographic localization effects (Ciccone and Hall (1996)).[2] In this paper, we follow Klenow and Rodriguez-Clare (2005) focusing on the simplest approach, that assuming that the parameters w_{ij} are completely determined by distance. As also underlined by Klenow and Rodriguez-Clare (2005), this would capture trade and FDI that are related to distance. It is also well known that the transmission of idea is better with direct contact because of tacit information. These meetings are more easy if the distance between country is low and the traditional gravity equation show the importance of physical distance in trade. Moreover, Keller (2002) suggests that the international diffusion of technology is geographically localized, in the sense that the productivity effects of R&D decline with the geographic distance between countries.

[2] Keller (2004) reviews the recent literature on technological change and diffusion.

These considerations along with the stylized fact that there exists strong spatial autocorrelation in the repartition of wealth and inputs must be seriously taken into account in our specification. Taking the logarithms of this expression and rewriting it in matrix form, we obtain:

$$A = \Omega + \phi k + \gamma W A \tag{5}$$

We can resolve this linear system for A, if $\gamma \neq 0$ and if $1/\gamma$ is not an eigenvalue of W:[3]

$$A = (I - \gamma W)^{-1}\Omega + \phi(I - \gamma W)^{-1}k \tag{6}$$

The stock of knowledge contained in the Solow residual is function of capital accumulated in all countries. Replacing this expression in the production function written in logarithms and in matrix form, $y = A + \alpha k$, we obtain:

$$y = (I - \gamma W)^{-1}\Omega + \phi(I - \gamma W)^{-1}k + \alpha k \tag{7}$$

Pre-multiplying both sides by $(I - \gamma W)$ we have finally:

$$y = \Omega + (\alpha + \phi)k - \alpha\gamma W k + \gamma W y \tag{8}$$

Therefore, the level of development of a given country depends positively of its own production factors accumulated as the traditional model. It also depends positively on the same factors in foreign countries and on their development level. This equation shows that the development level is high if there exists a strong concentration of production factors in space. Moreover, the spatial distribution of income influences the development level with the last term. The net effect of accumulated production factor depends on the location of the countries as the following proposition:

Proposition 1 *We can evaluate the elasticities of the level of output per worker of a country i in respect to its own level of physical capital per worker and in respect to the level of physical*

[3] Actually $(I - \gamma W)^{-1}$ exists if and only if $|I - \gamma W| \neq 0$. This condition is equivalent to: $|\gamma||W - (1/\gamma)I| \neq 0$ where $|\gamma| \neq 0$ and $|W - (1/\gamma)I| \neq 0$.

capital per worker of the foreign countries:

$$\Xi = \alpha I + \phi(I - \gamma W)^{-1} \tag{9}$$

Proof:

We can rewrite equation (7) in the following form:

$$y = (I - \gamma W)^{-1}\Omega + (\alpha I + \phi(I - \gamma W)^{-1})k \tag{10}$$

The expression of elasticities follows easily. QED

Thus, these effects of inputs on level of output per worker are stronger than the same effects in the traditional residual equation. In fact, there exists a spatial multiplier effect in accumulation of factors because of physical capital externalities and technological interdependence. These externalities increase the effect of capital on development level in respect to the neoclassical model. Hence, the coefficient α in the traditional equation understates by $\phi(I - \gamma W)^{-1}$ the direct contribution of k. This understatement arises because assumption of physical capital externalities and technological interdependence of knowledge, the social marginal product of capital[4] $\left(\alpha + \frac{\phi}{1-\gamma}\right) Y/K$ exceeds the private marginal product $\alpha Y/K$. The social marginal product is defined here as the increase of output consecutive on an increase of all inputs in the world. In fact, because of technological interdependence generated by knowledge diffusion, the foreign inputs can be viewed used by the firms of the home country (see the reduced form of the production function (8) where the level of per worker income depends on all per worker physical capital level.). We can test the possibility of endogenous growth linked with a constant return for the social capital testing $\alpha + \frac{\phi}{1-\gamma} = 1$. Therefore, it is possible to have constant social marginal product of physical capital even if $\alpha + \phi < 1$ since the parameter γ is assumed lower by 1. Romer (1987), for instance, find a value comprising between 0.7 and 1 for $\alpha + \phi$, and Ertur and Koch (2005) using this production function in the context of the Solow growth model find that this value is close to $0.6 - 0.7$ for $\alpha + \frac{\phi}{1-\gamma}$. We also note that the absence of physical capital externalities reduces the expression of social marginal product to α, the private marginal

[4] Its expression follows directly with the Markov line propriety of the matrix W. In fact, if we take the sum of a line i of the expression (9), for $i = 1, ..., N$, we have: $\alpha + \phi(1 + \gamma + \gamma^2 + ...)$, since: $\sum_{j=i}^{N} w_{ij} = \sum_{j=i}^{N} w_{ij}^{(2)} = ... = 1$ for $i = 1, ..., N$ where $w_{ij}^{(r)}$ is the element of the W matrix to the power of r, for $r = 1, ..., \infty$. Because of $|\gamma| < 1$, the expression in the text follows directly.

product of physical and only the residual contain technological interdependence with random shocks. Therefore, we can claim the following proposition:

Proposition 2 *The Total Factor Productivity variable of the Solow equation is contaminated by physical capital externalities and by the spatial repartition of this factor in the world:*

$$TFP_{Solow} = \Omega + \phi k + \gamma W TFP_{Solow} = (I - \gamma W)^{-1}\Omega + \phi(I - \gamma W)^{-1}k \tag{11}$$

where $TFP_{Solow} = y - \alpha k$.

Therefore, Total Factor Productivity of the Solow accounting equation contains spatial autocorrelation du to technological interdependence and physical capital externalities. This global technological interdependence implies cross-sectional correlation in the errors terms and the appropriate empirical specification is given by spatial econometric literature.

2.2 Spatial econometric specification

The Total Factor Productivity variable of the spatially augmented accounting equation can be decomposed on two part: a constant identical for each country and a country specific shock so that $\ln \Omega(t) = \beta_0 + \varepsilon_i$. The empirical specification of equation (8) is written in the following form for a country i:

$$\ln Y_i = \beta_0 + \beta_1 \ln K_i + \beta_2 \ln L_i + \theta_1 \sum_{j \neq i}^{N} w_{ij} \ln K_j + \theta_2 \sum_{j \neq i}^{N} w_{ij} \ln L_j + \gamma \sum_{j \neq i}^{N} w_{ij} \ln Y_j + \varepsilon_i \tag{12}$$

with β_0 the constant, ε_i an error term independently and identically distributed representing the country specific shock. We have also the following theoretical constraint in the parameters: $\beta_1 + \beta_2 = 1$ and $\theta_1 + \theta_2 = \gamma$ where γ is the coefficient of spatial autocorrelation. We can write this model in matrix form:

$$y = X\beta + WX\theta + \gamma W y + \varepsilon \tag{13}$$

In the spatial econometrics literature, this kind of specification, including the spatial lags of both endogenous and exogenous variables, is referred to as the spatial Durbin model (see Anselin and Bera, 1998). Here y is the $(N \times 1)$ vector of logarithms of real income, X the $(N \times 3)$ matrix

with the constant,[5] the vectors of logarithms of the aggregate level of physical capital and the logarithms of the aggregate level of worker, W the $(N \times N)$ spatial weight matrix, $\beta = [\beta_0\ \beta_1\ \beta_2]$, $\theta = [\theta_1\ \theta_2]$. ε is the $(N \times 1)$ vector of errors supposed normally and identically distributed so that $\varepsilon \rightsquigarrow N(0, \sigma^2 I)$. We can easily see that model (12) reduces to the textbook Solow model when $\phi = \gamma = 0$.

This model can be viewed as an equilibrium model. Assuming that $\gamma \neq 0$ and the inverse of γ is not an eigenvalue of W, we can write the equilibrium vector y as following:

$$y = (I - \gamma W)^{-1} X\beta + (I - \gamma W)^{-1} WX\theta + (I - \gamma W)^{-1}\varepsilon \tag{14}$$

This reduced form of the spatial Durbin model have two very interesting proprieties. Note first that the inverse matrix $(I - \gamma W)^{-1}$ is a full matrix which implies an infinite series for the explicative variable and for the errors terms associated with all locations.[6]

The first propriety of this model is the spatial multiplier effect. The equilibrium value of y of a country i is explaining by the explicative variables of this country and by these same variables in all others countries $j = 1, ..., N,\ j \neq i$ throughout the inverse matrix $(I - \gamma W)^{-1}$. This spatial multiplier effect decreases because of spatial friction parameters w_{ij}. In the context of our model, this propriety is identical to the claim of Proposition 1. When a country increases its physical capital, it increases its knowledge with the Learning by Doing process and spill over foreign countries'firms which gain on this new knowledge.

The second effect of is on interest is the spatial diffusion process contained in the error terms. This specification implies that a random shock affecting one country propagates to all the other locations of the sample through the spatial transformation $(I - \gamma W)^{-1}$. This propriety can be viewed in our model as a technologic shock affecting one country and the diffusion of it throughout the world. We can model the effect of a transitory shock on the technology of one country in our sample.[7]

Therefore, the traditional accounting equation neglects these effects and the error term contains omitted variable and spatial autocorrelation phenomenon:

[5] In practice, the spatially lagged constant is not included in WX, since there is an identification problem for row-standardized W (the spatial lag of a constant is the same as the original variable.).

[6] This propriety of this inverse matrix is equivalent to those of Leontieff matrix with $\gamma = 1$.

[7] Le Gallo et al. (2005) use this spatial diffusion effect to simulate random shock on distribution of growth in Europe.

Proposition 3 *If there exists physical capital externalities and spatial externalities, the error term of the traditional Total Factor Productivity equation contains omitted information and spatial autocorrelation. We can write this error term in the following form:*

$$\varepsilon_{SOLOW} = \phi(I - \gamma W)^{-1}k + (I - \gamma W)^{-1}\varepsilon \tag{15}$$

These proprieties implies that usual estimation methods, as OLS, are mistake. The appropriate estimation method in this case is the maximum likelihood estimator (Anselin and Bera, 1998; Lee, 2004) or the GMM estimator (Kelejian and Prucha, 1999).[8]

A last propriety of this model is linked with the presence or not of physical capital externalities in the model:

Proposition 4 *If there is no physical capital externalities, that is: $\phi = 0$, the econometric model is a spatial Durbin model in constrained form which is formally identical to a spatial error model. Therefore, it is possible to test the absence of physical capital externalities with the following non linear constrains: $\gamma\beta + \theta = 0$.*

Proof:

Replacing these non linear constrains $\theta = -\gamma\beta$ in the expression (13), subtracting by γWy and premultiplying by $(I - \gamma W)^{-1}$ in both sides, we have:

$$y = X\beta - \gamma WX\beta + \gamma Wy + \varepsilon = X\beta + (I - \gamma W)^{-1}\varepsilon \tag{16}$$

QED

This propriety can be us to test the presence or not of physical capital externalites. We can challenge the claim of Benhabib and Jovanovic (1991) of the absence of these type of externalities is the growth accounting equation. In fact, in their paper, they show that there is no support for the assertion that capital-related externalities are present.[9]

[8] In this paper, we use the maximum likelihood estimator with the spatial econometric toolbox of James Lesage: spatial.econometric.com. In appendix, we describe more precisely the estimation method and the consequence on OLS of these proprieties.

[9] They show instead that the variation in countries'growth rates is consistent with each country having the same constant return to scale production function and with a stochastic process for technical change that is the same across different countries but starts from different initial positions.

3 Evidence for spatial autocorrelation

3.1 Database and spatial weight matrix

The basic data set used in this paper come from the version 6.1 of the Penn World Tables (PWT6.1, Heston, Summers and Aten, 2002), that is the latest version of the Summers and Heston (1991) data set. From PWT6.1, we extract output, physical capital stock and the number of workers. We measure Y/L as real GDP per worker in PPP (call RGDPW in the original data set). We generate estimates of the physical capital stock, K, using to the perpetual inventory equation:

$$K(t) = I(t) + \delta K(t-1) \tag{17}$$

where $I(t)$ is investment and δ is the depreciation rate. We measure $I(t)$ from PWT6.1 as real aggregate investment in PPP, computed as $RGDPL \times POP \times KI$, where $RDGPL$ is real income per capita obtained with the Laspeyres method, POP is the population, and KI is the investment share in total income. Following Griliches (1980), we compute the initial capital stock $K(0)$ using its expression at steady state as $I(0)/(g+\delta)$, $I(0)$ is the value of the investment series between the first year it is available, and g is the geometric growth rate for the investment series between the first year with available data and 1970. Following the literature, we set δ to 0.06, and also consider $\delta = 0.04$ and $\delta = 0.10$ to evaluate robustness of results with reasonable change in the depreciation rate. Finally, the number of workers is obtained as $RGDPCH \times POP/RGDPW$, where $RGDPCH$ is real GDP per capita computed with the chain method. The sample consists on 91 countries in 1995. These countries are those of the Mankiw et al. (1992) non-oil sample, for which PWT 6.1 provide data.

In order to model spatial interdependence between all countries we must specify a spatial weight matrix W. More precisely, each country is connected to a set of neighboring countries by means of a purely spatial pattern introduced exogenously in W. The elements $w_{ii} = 0$ on the diagonal are set to zero whereas the elements w_{ij} indicate the way the country i is spatially connected to the country j. In order to normalize the outside influence upon each country, the weight matrix is standardized such that the elements of a row sum up to one. For the variable x, this transformation means that the expression Wx, called the spatial lag variable, is simply the weighted average of the neighboring observations. Various matrices are considered in the literature: a simple binary contiguity matrix, a binary spatial weight matrix with a distance-

based critical cut-off, above which spatial interactions are assumed negligible, more sophisticated generalized distance-based spatial weight matrices with or without a critical cut-off. The notion of distance is quite general and different functional forms based on distance decay can be used (for example inverse distance, inverse squared distance, negative exponential etc.). The critical cut-off can be the same for all countries or can be defined to be specific to each country leading in the latter case, for example, to $k-$nearest neighbors weight matrices when the critical cut-off for each country is determined so that each country has the same number of neighbors. It is important to stress that the connectivity terms w_{ij} should be exogenous to the model to avoid the identification problems raised by Manski (1993) in social sciences. This is the reason why we consider pure geographical distance, more precisely great circle distance between capitals, which is indeed strictly exogenous; the functional forms we consider are simply the inverse of squared distance denoted by $W1$, which can be interpreted as reflecting a gravity function and the negative squared exponential distance denoted by $W2$. The general terms of the spatial weight matrices $W1$ and $W2$ are defined as follows in standardized form $[w_{ij}]$: $w_{ij} = w^*_{ij}/\sum_{j\neq i}^{N} w^*_{ij}$ with $w1^*_{ij} = d_{ij}^{-2}$ or 0 if $i = j$ for $W1$ and $w2^*_{ij} = e^{-2\times d_{ij}}$ or 0 if $i = j$ for $W2$. The term d_{ij} is the great circle distance between country capitals.[10]

3.2 Evidence for global and local spatial autocorrelation

Spatial dependence or autocorrelation can be defined as the coincidence of value similarity with locational similarity (Anselin, 2001). Therefore, there is positive spatial autocorrelation when high or low values of a random variable tend to cluster in space and there is negative spatial autocorrelation when geographical areas tend to be surrounded by neighbors with very dissimilar values. Spatial interactions between regions can be evaluated using Exploratory Spatial Data Analysis (ESDA), which is a set of techniques aimed at describing and visualizing spatial distributions, at identifying atypical localizations or spatial outliers, at detecting patterns of spatial association, clusters or hot spots, and at suggesting spatial regimes or other forms of spatial heterogeneity. These methods provide measures of global and local spatial autocorrelation. The

[10]The great-circle distance is the shortest distance between any two points on the surface of a sphere measured along a path on the surface of the sphere (as opposed to going through the sphere's interior). It is computed using the equation:

$$d_{ij} = radius \times \cos^{-1}[\cos|long_i - long_j| \cos lat_i \cos lat_j + \sin lat_i \sin lat_j]$$

where radius is the Earth's radius, lat and $long$ are respectively latitude and longitude for i and j.

measurement of global spatial autocorrelation is usually based on Moran's I statistic (Cliff and Ord, 1981). This statistic is written in the following matrix form:

$$I = \frac{N}{S_0} \frac{z' W z}{z' z} \tag{18}$$

where z is the vector of the N observations in deviation from the mean. W is the spatial weight matrix. S_0 is a scaling factor equal to the sum of all the elements of W. For row-standardized spatial weights $S_0 = N$ and this expression consequently simplifies. Moran's I statistic gives a formal indication of the degree of linear association between the vector of z observed values and the vector Wz of spatially weighted averages of neighboring values, called the spatially lagged vector. Values of I larger (respectively smaller) than the expected value $E(I) = -1/(N-1)$ indicate positive (respectively negative) spatial autocorrelation. Statistical inference is based on the permutation approach with 10000 permutations (Anselin, 1995). Since Moran's I yields a single result for the entire data set, it cannot discriminate between a spatial clustering of high values and a spatial clustering of low values in the case of a global positive spatial autocorrelation. We can study the local spatial instability by means of the Moran scatterplot (Anselin, 1996), which plots the spatial lag Wz against the original values z. The four different quadrants of the scatterplot correspond to the four types of local spatial association between a country and its neighbors: HH a country with a high value surrounded by countries with high values, LH a country with a low value surrounded by countries with high values, etc. Quadrants HH and LL (respectively LH and HL) refer to positive (respectively negative) spatial autocorrelation indicating spatial clustering of similar (respectively dissimilar) values. The Moran scatterplot may thus be used to visualize atypical localizations, i.e. countries in quadrant LH or HL. The global spatial autocorrelation may also be visualized on this graph since, from its expression Moran's I is formally equivalent to the slope coefficient of the linear regression of on using a row-standardized weight matrix. However, despite the information given by the Moran scatterplots, we do not have any indication about the statistical significance of spatial associations. We must therefore compute the Local Indicators or Spatial Association (LISA) proposed by Anselin (1995) and defined as any statistics satisfying two criteria: first, the LISA for each observation gives an indication of significant spatial clustering of similar values around that observation; second, the sum of the LISA for all observations is proportional to a global indicator of spatial

association. The local version of Moran's I statistic for each country i is written as:

$$I_i = \frac{x(i) - \mu}{m_0} \sum_{j \neq i}^{N} w_{ij}(x(j) - \mu) \quad \text{with} \quad m_0 = \sum_{j \neq i}^{N} (x(j) - \mu)^2 / N \tag{19}$$

where $x(i)$ is the observation in country i, μ is the mean of the observations across countries and where the summation over j is such that only neighboring values of j are included. A positive value for I_i, indicates spatial clustering of similar values (high or low) whereas a negative value indicates spatial clustering of dissimilar values between a country and its neighbors. In presence of global spatial autocorrelation, inference must be based on the conditional permutation approach. This approach is conditional in the sense that the value x_i at location i is held fixed, while the remaining values are randomly permuted over all locations. 10000 permutations were used here to compute the empirical distribution function which provides the basis for statistical inference. The $p-$values obtained for the local Moran's statistics are then pseudo-significance levels (Anselin, 1995, p. 96 and p. 99-100).

[Figure 1 around here]

Figure 1 displays the types of spatial association prevailing across countries in our samples. They represent the Moran scatterplot and the Moran's I representing respectively the local and global autocorrelation in the inputs and in the income.[11]

It first appears that per capita GDP are positively spatially autocorrelated since the Moran's I statistics are strongly significant with $p = 0.000$ ($I = 0.673$). This result suggests that the distribution of per capita GDP is by nature clustered in the world. In other word, the countries with relatively high per capita GDP (respectively low) are localized close to other countries with relatively high per capita GDP (respectively low). We also observe in quadrant HH numerous countries in Europe, North and South America, and South-Est Asia. However, we note that only European countries, United States and Canada have significant LISA statistics. In quadrant LL consisting of poor countries surrounded by poor countries, we observe mainly Sub Saharan African and South Asian (India, Pakistan, Nepal, Sri Lanka) countries. Only African countries are significant at 5% in this quadrant. In quadrant HL, we observe four countries which are richer than their surrounding countries: South Africa, Mauritius, Costa Rica and El Salvador. Only

[11] The results are given for $W1$ and $\delta = 6\%$. They are robust for $\delta = 4\%$ and $\delta = 10\%$ or the other spatial weight matrix $W2$. They are available upon request.

Mauritius has significant LISA statistic. Finally, in the quadrant LH, we have some countries in south East Asia as Indonesia, Papua New Guinea, Philippines and in the Central America as Bolivia, Jamaica, Honduras and Nicaragua. Only Asian countries are significantly more poorer than their neighboring countries.

It also appears that inputs are positively spatially autocorrelated corroborating the stylized fact of Easterly and Levine (2001) about the concentration of inputs in space implying some externalities. In fact, physical capital and per worker physical capital exhibit strong positive spatial autocorrelation (respectively $I = 0.481$ and $I = 0.644$ with a $p-$value $= 0.000$). Labor input doesn't exhibit significant global spatial autocorrelation ($I = 0.072$ with $p-$value $= 0.108$). The structure of local spatial autocorrelation for per worker physical capital is almost identical to the one of per worker income. These results show the strong spatial autocorrelation in the repartition of wealth and physical capital input. These stylized facts must be taken into account in the model and empirical study.

4 Econometrics Results

4.1 Estimation of the Total Factor Productivity equation

The estimation of the Total Factor Productivity equation in level, with $\phi = \gamma = 0$, is in the table 1. Each column represents the results for different depreciation rate ($\delta = 3\%$, 6% and 10% respectively in the first, the second and the third columns).

[Table 1 around here]

The coefficients are strongly significant and have the good signs. The constant is also strongly significant showing the importance of Total Factor Productivity in explaining of the differences of GDP and per capita GDP between countries. This result has been interpreted in the literature as a reject of the neoclassical explanation of development gap with factors differences. The linear restriction implies by the constant return to scale is not rejected. Finally, the coefficients associated with the inputs, physical capital and labor, are positive and highly significant. The capital share in the income, α, is too high and close to 0.7. However, this equation ignore spatial autocorrelation in GDP and in inputs. This difference can also be explained by physical capital externalities, technological interdependence and spatial repartition of wealth. In fact, Moran's I test of residuals shows that we can't reject the spatial autocorrelation hypothesis as suggested

in the proposition 3. Therefore, we estimate our spatially augmented version of this equation including both physical capital externalities and technological interdependence (equation 12). The results of the unconstrained Spatial Durbin Model are displaying in the Table 2. We present the results for each spatial weight matrix ($W1$ for the one to third columns and $W2$ for the last third columns) and each rate of depreciation. We interpret essentially the results for the inverse squared distance spatial weight matrix ($W1$) and $\delta = 6\%$.

[Table 2 around here]

The coefficients are significants and have the good signs in the constrained and the unconstrained models for all cases. The linear constrains implied by constant returns to scale are not rejected (the $p-$value of the likelihood test is $p = 0.560$). The α value, representing the capital share in the income, is close to 0.47, the high bound value generally admitted in the literature and more closer to expected than the precedent regression where the value is $\alpha = 0.71$. The value of the physical capital externalities is equal to 0.18, an identical value obtained by Ertur and Koch (2005) in the neoclassical growth model context.[12] The coefficient of spatial autocorrelation is highly significant in the constraint and the unconstraint models. This last result confirms the strong spatial autocorrelation in the sample which we can explain by the presence of technological interdependence between countries. The parameter γ, equal to the spatial autocorrelation parameter, is highly significant with a value close to 0.3 - 0.4 in both models. Finally, the social returns to scale are significantly close to the expected value 0.7, find by Romer (1987), where γ = 0, and Ertur and Koch (2005).

[Table 3 around here]

We note however, that the value of the parameter representing the coefficient of physical capital externalities ($\phi = 0$), is not significant. The absence of physical capital externalities can be tested by a common factor test as shown in Proposition 4. To implement this test, we first estimate the Spatial Error Model implied when $\phi = 0$ and test the non-linear restriction with the likelihood ratio test in the table 3. The test reject the absence of physical capital externalities in our sample (the $p-$value of the likelihood test is $p = 0.068$ for the unconstrained model

[12] This result contrast with those of Mankiw *et al.* (1992) and Benhabib and Spiegel (1994) where their results differ significantly using a neoclassical growth specification or a Total Factor Productivity equation.

and $p = 0.030$ for the constrained model). We can't reject the presence of physical capital externalities. However, the test do not give clear-cut results and seems in contradiction with the individual significant test of ϕ. More caveats are necessary to control for omitted variables, as human capital for example, to conclude for presence or absence of physical capital externalities. These results show the importance of a missing factor: the space, in explaining development gap. The Solow residual term of the traditional Total Factor Productivity equation contains omitted variables and spatial autocorrelation.

4.2 Estimation of the contribution of the physical capital to the development gap

In order to give an estimation of the influence of physical capital on development gap, we use the method developed by Denison (1962) and recently by King and Levine (1994) and Easterly and Levine (2001).

Proposition 5 *The fraction of differences in national output levels due to capital equals the following ratio in matrix form:*

$$\Upsilon_k = \alpha 1 + D^{-1}\Xi\kappa^\star + \phi D^{-1}(I - \gamma W)^{-1}y^\star \tag{20}$$

where $\ln \kappa_i^\star = \ln(\kappa_i/\overline{\kappa})$ *letting* $\kappa_i = k_i/y_i$ *the capital-output ratio for the country* i *and* $\overline{\kappa}$ *represents the capital-output ratio for a reference country and* D *is a diagonal matrix with the following terms on its main diagonal:* $\ln(y_i/\overline{y})$ *where* $\overline{y}$ *is the output per worker for a reference country.*

Proof:

Take equation (7) writing in respect to a reference country:

$$y^\star = (I - \gamma W)^{-1}\Omega^\star + (\alpha I + \phi(I - \gamma W)^{-1})k^\star \tag{21}$$

Take: $\ln \kappa_i = \ln k_i - \ln y_i$, the capital-output ratio. In matrix form and in respect to a reference country: $k^\star = \kappa^\star + y^\star$. Then, we obtain:

$$y^\star = (I - \gamma W)^{-1}\Omega^\star + \alpha(\kappa^\star + y^\star) + \phi(I - \gamma W)^{-1}(\kappa^\star + y^\star) \tag{22}$$

Premultiplying both sides by the inverse of the following diagonal matrix: $D = diag(\ln y_1^\star, ..., \ln y_i^\star, ..., \ln y_N^\star)$, we obtain:

$$D^{-1}y^\star = D^{-1}(I - \gamma W)^{-1}\Omega^\star + \alpha D^{-1}(\kappa^\star + y^\star) + \phi D^{-1}(I - \gamma W)^{-1}(\kappa^\star + y^\star) \tag{23}$$

Therefore, we obtain the contribution of the capital to the gap of the development:

$$\Upsilon_k = \alpha 1 + (\alpha D^{-1} + \phi D^{-1}(I - \gamma W)^{-1})\kappa^\star + \phi D^{-1}(I - \gamma W)^{-1}y^\star \tag{24}$$

Replacing by equation (9), we obtain (20). *QED*

The influence of physical capital on development gap depends on three terms. First, it depends on the capital share in the income (α). Second, the higher the capital-output ratio is the higher the role of capital is in the development gap. A country with a strong difference of this ratio in a reference country will have a strong difference in the development. This contribution of capital-output ratio is increasing with the elasticities of income in respect to the physical capital (Ξ), its depends, therefore, on technological interdependence. Finally, the contribution of capital to the level of development depends on the spatial distribution on wealth. More a country has riches neighbors more its capital contributes to its development level. The importance of this relative spatial distribution of wealth shows the importance of space and neighborhood in the development process. A country with poor neighbors cannot beneficiates on externalities generated by their physical capital stock (see equation 9). We note that this expression reduces to those of Kind and Levine (1994) if there is no physical externalites ($\phi = 0$).

[Figure 2 around here]

We use the most richest country as reference that is the United States in our sample. We compute the percentage shortfall in output for country i relative to a reference country j: $P_i = 100 \times (y_j - y_i)/y_j$. The gap between the most poorest country in respect to the United States is around 99% for the Republic Democratic of Congo and around 9% for Hong Kong, the richest country after the United States.[13] The thick line displays this gap in the figure 2. In this figure, we can show also the contribution of the physical capital per worker in the textbook Total Factor Productivity equation of Solow in fine line and in our spatially augmented Total

[13] Recall that we work with the per worker GDP.

Factor Productivity model in dotted line. Figure 2 shows that the total contribution of capital go from around 60% for developing countries to negatives figures for the richest countries. Considering the spatially augmented Total Factor Productivity model, the total contribution of physical capital increases dramatically. The physical capital explains almost the totality of development gap since the dotted line in the figure 1, representing this contribution, is almost identic to the thick line.

[Figure 3 around here]

In the Total Factor Productivity model, with $\phi = \gamma = 0$, the contribution of physical capital to the development gap reduces to: $\Upsilon_k = \alpha 1 + \alpha D^{-1}\kappa^{\star}$. This expression, those of King and Levine (1994) and Easterly and Levine (2001), contains two components: a capital-share component (αP_i) and a capital-output ratio component $\alpha(D^{-1}\kappa^{\star})P_i$. The figure 3 (in the left) displays the two components of this case. For the capital-share component (in the dotted line), we use the value of $\alpha = 0.468$.[14] With this value, capital account directly for 45% of the development gap for the developing country whereas it is weak for the developed countries. Considering differences between the capital-output ratio component (in fine line), there is an additional of less by 10% attributed to capital and an subtracting up to 10% for the most richest countries. This result is due to the fact that the developed countries have a capital-output ratio higher than those of United States. Therefore, the total contribution of capital in a traditional accounting equation up to 60% for the most poorest countries. Therefore, the capital accounts for a greater share of the gap in respect to the United States in poorer countries. However, this gap is not explained by the physical capital for the most developed countries. The results are convergent with Denison (1962), King and Levine (1994) and Easterly and Levine (2001): the physical capital explains only the half to 60% of the development gap between United States and the most poorest countries.

Considering the spatially augmented Total Factor Productivity model, the components of the contribution of capital are displaying in the figure 3 (on the right). The first component is the capital share in the income (dotted line). This component accounts for 45% of the develop-

[14] This value is close to those of King and Levine (1994) who take a conservative value (0.4) in order to critic the potential contribution of capital accumulation to output. We use the implied α in our estimation result in the spatially augmented Total Factor Productivity model. The results using $\alpha = 0.33$ are identical and available upon request.

ment gap for the developing countries. The second component is the ratio capital-output (fine line). This component accounts for 20% for developing countries, around 10% for intermediate countries and is negative for developed countries. These negatives figures for the most richest countries are explaining as preceding: the capital-output ratio is weak in Unites States relative to other developed countries. The third component is the relative distribution of wealth (dashed line). This component explains around 25% of the contribution of capital in the shortfall in development for the poorest countries and around 10% for the most developed countries. Therefore the contribution of capital in the development gap is underestimating in the traditional developing accounting equation. Taking into account technological interdependence and physical capital externalities, this contribution increases dramatically for the developing countries.

5 Conclusion

In this paper, we reconciliate spatial autocorrelation often observed in world distribution of wealth and its effects on development gap. We show that, using appropriates tools, per worker income and inputs exhibit strong global and local spatial autocorrelation. Modeling technological interdependence in Total Factor Productivity equation, we show that traditional results obtained in the literature are biased and the role of physical capital is underestimated when we neglect the global interdependence between countries.

However, we don't want to reject the importance of Total Factor Productivity component in theoretical and empirical studies. Taking into account global interdependence in Total Factor Productivity component potentially modify dramatically our vision about growth and development process and must be taken into account both in theory and empirical studies.

References

[1] Anselin L. (1988), "*Spatial econometric: Methods and Model*", Kluwer Academic Publishers, Dordrecht.

[2] Anselin L. (1995), "Local indicators of spatial association-LISA", *Geographical Analysis*, 27, 93-115.

[3] Anselin L. (1996), "The Moran scatterplot as an ESDA tool to assess local instability in spatial association" in Fisher M., Scholten H.J., Unwin D. (eds), *Spatial analytical perspectives on GIS*. London, Taylor & Francis.

[4] Anselin L. (2001), "Spatial econometrics", In: Baltagi B. (eds) *Companion to econometrics*. Oxford, Basil Blackwell.

[5] Anselin L., Bera A.K. (1998), "Spatial Dependence in Linear Regression Models with an Introduction to Spatial Econometrics", In: *Handbook of Applied Economics Statistics*, A. Ullah and D.E.A. Giles (eds). New York: Marcel Dekker.

[6] Arrow K. (1962), "The economic implications of learning by doing", *Review of Economic Studies*, 29, 155-173.

[7] Barro R.J. (1999), "Notes on Growth Accounting", *Journal of Economic Growth*, 4, 119-137.

[8] Barro R.J., Sala-i-Martin X. (1997), "Technological Diffusion, Convergence and Growth", *Journal of Economic Growth*, 1, 1-26.

[9] Benhabib J., Jovanovic B. (1991), "Externalities and Growth Accounting", *American Economic Review*, 81, 82-113.

[10] Benhabib J., Spiegel M.M. (1994), "The role of human capital in economic development: evidence from aggregate cross-country data", *Journal of Monetary Economics*, 34, 2, 143-174.

[11] Caselli F. (2004), "Accounting for cross-country income differences", In: *Handbook of Economic Growth*, Aghion P. And Durlauf S. (eds), North-Holland: Elsevier.

[12] Caves R.E. (1996), "Multinational Enterprise and Economic Analysis", second edition, Cambridge University Press, New York.

[13] Ciccone A., Hall R.E. (1996), "Productivity and the Density of Economic Activity", *American Economic Review*, 86, 54-70.

[14] Cliff A.D., Ord J.K. (1981), "*Spatial Processes: Models and Applications*", Pion: London.

[15] Coe D., Helpman E. (1995), "International R & D spillovers", *European Economic Review*, 39, 859-897.

[16] Denison E.F. (1962), "*The sources of Economic Growth in the United States and the Alternatives Before Us*" Washington, DC: Committee for Economic Development.

[17] Denison E.F. (1967), "*Why Growth Rates Differ*", Washington, D.C.: The Brookings Institution.

[18] Easterly W., Kremer M., Pritchett L., Summers L. (1993), "Good Policy or Good Luck? Growth Performance and Temporary Shocks", *Journal of Monetary Economics*, 32, 459-84.

[19] Easterly W., Levine R. (2001), "Its Not Factor Accumulation: Stylised Facts and Growth Models", *The World Bank Economic Review*, 15, 177-219.

[20] Eaton J., Kortum S. (1996), "Trade in ideas Patenting and Productivity in the OECD", *Journal of International Economics*, 40, 251-276.

[21] Ertur C., Koch W. (2005), "Growth, Technological Interdependence and Spatial Externalities: Theory and Evidence", Regional Economics Applications Laboratory, Discussion papers, REAL 05-T-10, University of Illinois at Urbana-Champaign. Document de travail, Laboratoire d'Economie et de Gestion (LEG), n^o2005-03, Université de Bourgogne.

[22] Griliches Z. (1980), "R&D and the Productivity Slowdown", *American Economic Review*, 70, 343-348.

[23] Grossman G., Helpman E. (1991), "*Innovation and growth in the global economy*", Cambridge, MA: MIT Press.

[24] Hall R.E., Jones C. (1999), "Why Do Some Countries Produce So Much More Output per Worker than Others?", *Quarterly Journal of Economics*, 114, 83-116.

[25] Heston A., Summers R., Aten B. (2002), Penn World Tables Version 6.1. Downloadable dataset. Center for International Comparisons at the University of Pennsylvania.

[26] Howitt P. (2000), "Endogeous Growth and Cross-Country Income Differences", *American Economic Review*, 90, 4, 829-846.

[27] Jones C.I. (1995), "R&D-based models of economic growth", *Journal of Political Economy*, 103, 759-784.

[28] Jorgenson D.W. (1995a), "*Productivity. Volume 1: Postwar U.S. Economic Growth*", MIT Press.

[29] Jorgenson D.W. (1995b), "*Productivity. Volume 2: International Comparisons of Economic Growth*", MIT Press.

[30] Jorgenson D.W., Griliches Z. (1967), "The Explanation of Productivity Change", *Review of Economic Studies*, 34, 249-283.

[31] Kelejian H.H., Prucha I.R. (1998), "A Generalized Spatial Two-Stages Least Squares Procedure for Estimating a Spatial Autoregressive Model with Autoregressive Disturbances", *Journal of Real Estate Finance and Economics*, 17, 99-121.

[32] Kelejian H.H., Prucha I.R. (1999), "A Generalized Moments Estimator for the Autoregressive Parameter in a Spatial Model", *International Economic Review*, 40, 509-33.

[33] Keller W. (2002), "Geographic Localization of International Technology Diffusion", *American Economic Review*, 92, 120-142.

[34] Keller W. (2004), "International technology diffusion", *Journal of Economic Literature*, 42, 752-782.

[35] Kendrick J.W. (1976), *The formation and stocks of total capital*, New-York, Columbia University for NBER.

[36] King R.G., Levine R. (1994), "Capital Fundamentalism, Economic Development and Economic Growth", *Carnegie-Rochester Conference Series on Public Policy*, 40, 259-292.

[37] Klenow P.J., Rodriguez-Clare A. (2005), "Externalities and Growth", In: *Handbook of Economic Growth*, Aghion P. And Durlauf S. (eds.), North-Holland: Elsevier.

[38] Lee L.F. (2003), "Best Spatial Two-Stages Least Squares Estimation for a Spatial Autoregressive Model with Autoregressive Disturbances", *Econometric Reviews*, 22, 307-335.

[39] Lee L.F. (2004), "Asymptotic Distributions of Quasi-Maximum Likelihood Estimators for Spatial Autoregressive Models", *Econometrica*, 72, 6, 1899-1925.

[40] Le Gallo J., Baumont C., Dall'erba S., Ertur C. (2005), "On the Property of Diffusion in the Spatial Error Model", *Applied Economics Letters*, 12, 533-536.

[41] Maddison A. (1982), "Phases of Capitalist Development", Oxford: Oxford University Press.

[42] Mankiw N.G., Romer D. et Weil D. N. (1992), "A contribution to the empirics of economic growth", *Quarterly Journal of Economics*, 107, 407-437.

[43] Manski C.F. (1993), "Identification of endogenous social effects: the reflection problem", *Review of Economic Studies*, 60, 531-542.

[44] Ord J.K. (1975), "Estimation methods for models of spatial interaction", *Journal of American Statistical Association*, 70, 120-126.

[45] Parente S.L., Prescott E.C. (1994), "Barriers to Technology Adoption and Development", *Journal of Political Economy*, 102, 298-321.

[46] Romer P.M. (1986), "Increasing returns and long run growth", *Journal of Political Economy*, 94, 1002-1037.

[47] Romer P.M. (1987), "*Crazy Explanations for the productivity slowdown*", NBER Macroeconomics Annual, MIT Press, 163-202.

[48] Solow R.M. (1957), "Technical Change and the Aggregate Production Function", *Review of Economics and Statistics*, 39, 312-320.

[49] Summers R., Heston A. (1991), "The Penn World Table (Mark 5): An Expanded Set of International Comparisons 1950-88", *Quarterly Journal of Economics*, 106, 327-68.

Table 1: Estimation results. Textbook TFP equation

	$\delta = 4\%$	$\delta = 6\%$	$\delta = 10\%$
Unrestricted regression			
$\ln A$	2.235	2.484	2.862
	(0.000)	(0.000)	(0.000)
$\ln K$	0.718	0.715	0.709
	(0.000)	(0.000)	(0.000)
$\ln L$	0.273	0.268	0.269
	(0.000)	(0.000)	(0.000)
Moran's I test ($W1$)	0.181	0.152	0.125
	(0.004)	(0.013)	(0.035)
Moran's I test ($W2$)	0.137	0.116	0.099
	(0.038)	(0.067)	(0.101)
Restricted regression			
$\ln A$	2.151	2.246	2.674
	(0.000)	(0.000)	(0.000)
$\ln k$	0.717	0.715	0.708
	(0.000)	(0.000)	(0.000)
Test of restriction	0.175	0.534	1.129
	(0.677)	(0.467)	(0.291)
implied α	0.718	0.715	0.709
	(0.000)	(0.000)	(0.000)
Moran's I test ($W1$)	0.182	0.156	0.137
	(0.004)	(0.012)	(0.025)
Moran's I test ($W2$)	0.137	0.121	0.116
	(0.041)	(0.063)	(0.072)

Notes: p-values are in parentheses; p-values for the implied parameters are computed using the delta method. Tests of restriction are realized with Wald test. $W1$ is the inverse squared matrix and $W2$ is the negative squared exponential matrix.

Table 2: Estimation results. Spatially augmented TFP equation (SDM)

Spatial weight matrix	$W1$			$W2$		
	$\delta = 4\%$	$\delta = 6\%$	$\delta = 10\%$	$\delta = 4\%$	$\delta = 6\%$	$\delta = 10\%$
$\ln A$	1.067	1.469	2.341	1.509	1.847	2.341
	(0.029)	(0.004)	(0.000)	(0.000)	(0.000)	(0.000)
$\ln K$	0.637	0.647	0.660	0.644	0.655	0.660
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$\ln L$	0.345	0.332	0.317	0.335	0.323	0.317
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$W \ln K$	-1.191	-0.162	-0.104	-0.138	-0.123	-0.104
	(0.053)	(0.116)	(0.277)	(0.133)	(0.192)	(0.277)
$W \ln L$	-0.223	-0.200	-0.145	-0.169	-0.154	-0.145
	(0.000)	(0.001)	(0.008)	(0.003)	(0.006)	(0.008)
$W \ln Y$	0.421	0.359	0.229	0.311	0.271	0.229
	(0.001)	(0.007)	(0.076)	(0.011)	(0.031)	(0.076)
Restricted regression						
$\ln A$	0.981	1.270	1.631	1.383	1.612	1.890
	(0.003)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$\ln k$	0.640	0.650	0.653	0.649	0.659	0.662
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$W \ln k$	-0.194	-0.170	-0.154	-0.142	-0.135	-0.142
	(0.050)	(0.098)	(0.139)	(0.124)	(0.149)	(0.126)
$W \ln y$	0.417	0.363	0.323	0.308	0.279	0.273
	(0.001)	(0.007)	(0.019)	(0.012)	(0.026)	(0.030)
Test of restriction	0.779	1.159	1.934	0.831	1.183	1.968
	(0.677)	(0.560)	(0.380)	(0.660)	(0.554)	(0.374)
implied α	0.464	0.468	0.476	0.460	0.484	0.519
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)
implied ϕ	0.175	0.181	0.177	0.189	0.175	0.143
	(0.094)	(0.139)	(0.186)	(0.146)	(0.196)	(0.240)
implied γ	0.418	0.363	0.323	0.308	0.279	0.273
	(0.001)	(0.007)	(0.019)	(0.012)	(0.026)	(0.030)
implied $\alpha + \frac{\phi}{1-\gamma}$	0.766	0.753	0.738	0.733	0.727	0.716
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Notes: p-values are in parentheses; p-values for the implied parameters are computed using the delta method. Tests of restriction are realized with the likelihood ratio test. $W1$ is the inverse squared matrix and $W2$ is the negative squared exponential matrix.

Table 3: Estimation results. Spatially augmented TFP equation (SEM)

Spatial weight matrix	$W1$			$W2$		
	$\delta = 4\%$	$\delta = 6\%$	$\delta = 10\%$	$\delta = 4\%$	$\delta = 6\%$	$\delta = 10\%$
$\ln A$	2.641	2.732	2.986	2.508	2.626	2.918
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$\ln K$	0.679	0.688	0.690	0.691	0.698	0.697
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$\ln L$	0.310	0.298	0.294	0.298	0.288	0.288
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
γ	0.530	0.456	0.391	0.359	0.294	0.250
	(0.000)	(0.000)	(0.002)	(0.001)	(0.014)	(0.044)
Common factor test	6.447	5.385	4.986	5.625	4.489	4.017
	(0.040)	(0.068)	(0.083)	(0.060)	(0.106)	(0.134)
Restricted regression						
$\ln A$	2.515	2.598	2.848	2.393	2.512	2.785
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$\ln k$	0.682	0.690	0.690	0.693	0.698	0.696
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
γ	0.512	0.448	0.393	0.346	0.303	0.263
	(0.000)	(0.000)	(0.001)	(0.000)	(0.011)	(0.032)
Common factor test	5.951	4.684	3.677	5.036	3.691	2.612
	(0.015)	(0.030)	(0.055)	(0.025)	(0.055)	(0.106)
Test of restriction	0.282	0.458	0.625	0.242	0.384	0.564
	(0.595)	(0.498)	(0.429)	(0.623)	(0.535)	(0.453)
implied α	0.682	0.690	0.690	0.693	0.698	0.696
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
implied γ	0.512	0.448	0.393	0.359	0.303	0.263
	(0.000)	(0.000)	(0.001)	(0.001)	(0.011)	(0.032)

Notes: p-values are in parentheses; p-values for the implied parameters are computed using the delta method. Tests of restriction and Common Factor Tests are realized with the likelihood ratio test. $W1$ is the inverse squared matrix and $W2$ is the negative squared exponential matrix.

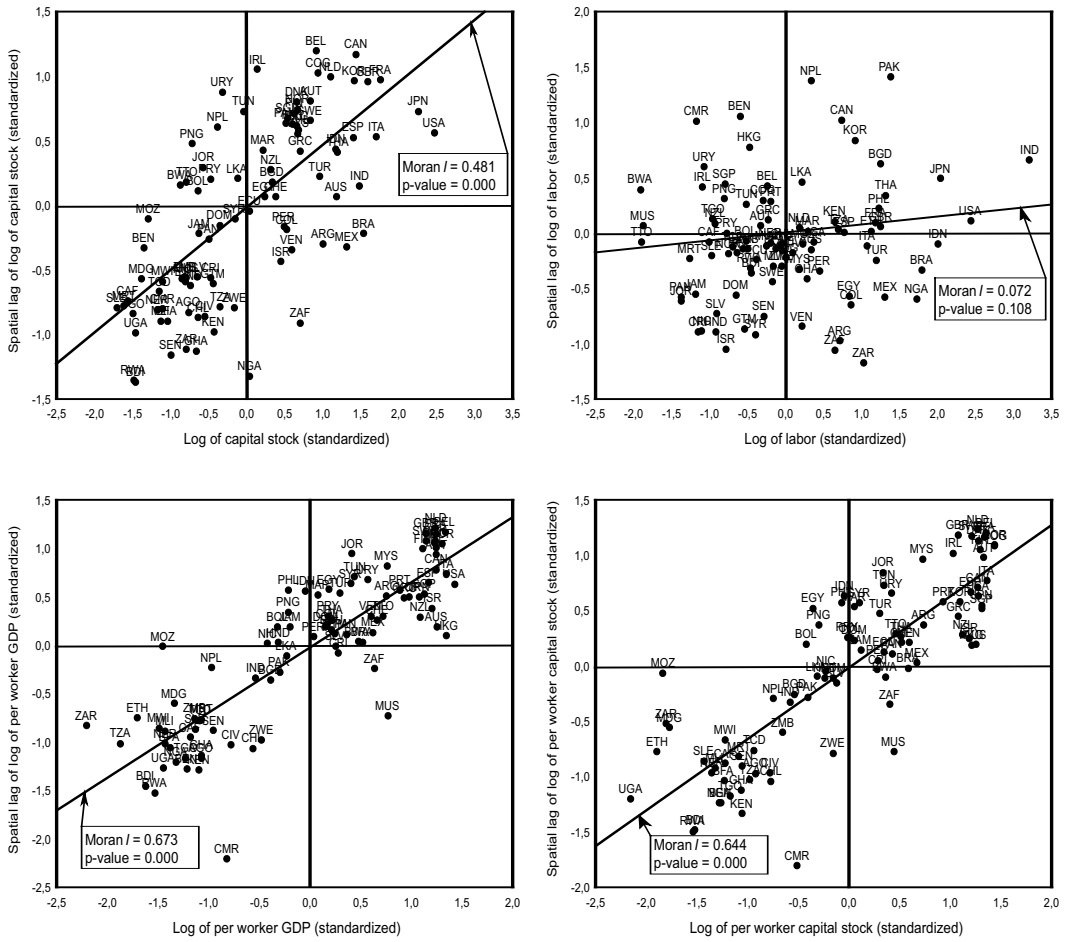

Figure 1: Moran scatterplot for log of physical capital $\ln K$ (upper left), log of labor $\ln L$ (upper right), log of per worker income $\ln y$ (bottom left) and log of per worker physical capital $\ln k$ (bottom right).

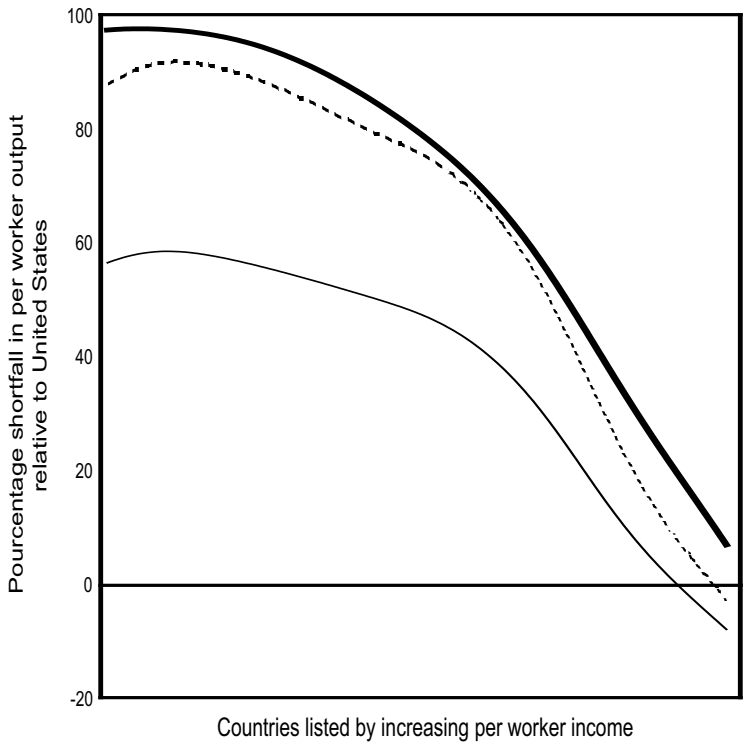

Figure 2: Development accounting for 91 countries with the textbook Total Factor Productivity equation (fine line) and the spatially augmented Total Factor Productivity equation (dotted line). The thick line represents the percentage shortfall in per worker output relative to United States.

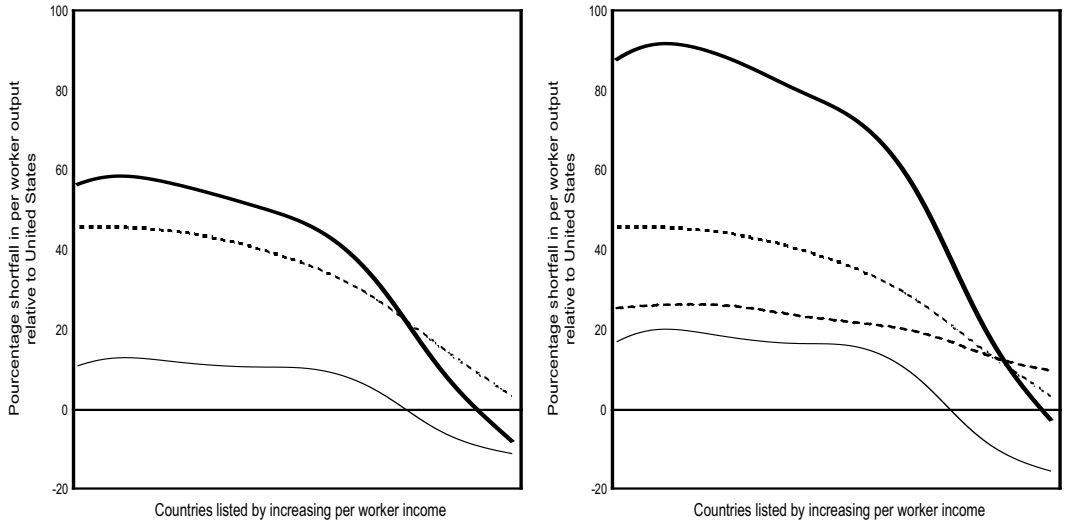

Figure 3: Components of physical capital contribution in the shortfall in per worker output relative to United States with the textbook Total Factor Productivity equation (on the left) and the spatially augmented Total Factor Productivity equation (on the right). The thick line represents the total contribution, the dotted line represents the direct contribution of capital share α, the fine line the contribution of relative capital-output ratio and the dashed line the contribution of spatial repartition of wealth.

Appendix : Estimation method

For ease of exposition, equation (13) may also be written as follows:

$$y = \widetilde{X}b + \rho Wy + \varepsilon \tag{25}$$

with $\widetilde{X} = [X \quad WX]$ and $b = (\beta', \theta')'$. If $\rho \neq 0$ and if $1/\rho$ is not an eigenvalue of W, $(I - \rho W)$ is nonsingular and we have, in reduced form:

$$y = (I - \rho W)^{-1}\widetilde{X}b + (I - \rho W)^{-1}\varepsilon \tag{26}$$

Since for row-standardized spatial weight matrices $|\rho| < 1$ and $w_{ij} < 1$, the inverse matrix in equation (26) can be expanded into an infinite series as:

$$(I - \rho W)^{-1} = (I + \rho W + \rho^2 W^2 + ...) \tag{27}$$

The reduced form has two important implications. First, in conditional mean, real income per worker in a location i will not only be affected by an increase of the physical capital input in i, but also by those in all other locations through the inverse spatial transformation $(I - \rho W)^{-1}$. This is the so-called spatial multiplier effect or global interaction effect. Second, a random shock in a specific location i does not only affect the real income per worker in i, but also has an impact on the real income per worker in all other locations through the same inverse spatial transformation. This is the so-called spatial diffusion process of random shocks.

The variance-covariance matrix for y is easily seen to be equal to:

$$Var(y) = \sigma^2 (I - \rho W)^{-1}(I - \rho W')^{-1} \tag{28}$$

The structure of this variance-covariance matrix is such that every location is correlated with every other location in the system, but closer location more so. It is also interesting to note that the diagonal elements in equation (28), the variance at each location, are related to the neighborhood structure and therefore are not constant, leading to heteroskedasticity even though the initial process (25) is not heteroskedastic.

It also follows from the reduced form (26) that the spatially lagged variable Wy is correlated

with the error term since:

$$E(Wy\varepsilon') = \sigma^2 W(I - \rho W)^{-1} \neq 0 \tag{29}$$

Therefore OLS estimators will be biased and inconsistent. The simultaneity embedded in the Wy term must be explicitly accounted for in a maximum likelihood estimation framework as first outlined by Ord (1975).[15] More recently, Lee (2004) presents a comprehensive investigation of the asymptotic properties of the maximum likelihood estimators of SAR models.

Under the hypothesis of normality of the error term, the log-likelihood function for the SAR model (25) is given by:

$$\begin{aligned} \ln L(\beta',\, \rho,\, \sigma^2) &= -\frac{N}{2}\ln(2\pi) - \frac{N}{2}\ln(\sigma^2) + \ln|I - \rho W| \\ &\quad -\frac{1}{2\sigma^2}\left[(I-\rho W)y - \widetilde{X}b\right]'\left[(I-\rho W)y - \widetilde{X}b\right] \end{aligned} \tag{30}$$

An important aspect of this log-likelihood function is the Jacobian of the transformation, which is the determinant of the $(N \times N)$ full matrix $I - \rho W$ for our model. This could complicate the computation of the ML estimators which involves the repeated evaluation of this determinant. However Ord (1975) suggested that it can be expressed as a function of the eigenvalues ω_i of the spatial weight matrix as:

$$|I - \rho W| = \prod_{i=1}^{N}(1 - \rho\omega_i) \quad \Longrightarrow \quad \ln|I - \rho W| = \sum_{i=1}^{N}\ln(1 - \rho\omega_i) \tag{31}$$

This expression simplifies considerably the computations since the eigenvalues of W only have to be computed once at the outset of the numerical optimization procedure.

From the usual first-order conditions, the maximum likelihood estimators of β and σ^2, given ρ, are obtained as:

$$\hat{\beta}_{ML}(\rho) = (\widetilde{X}'\widetilde{X})^{-1}\widetilde{X}'(I - \rho W)y \tag{32}$$

$$\hat{\sigma}^2_{ML}(\rho) = \frac{1}{N}\left[(I-\rho W)y - \widetilde{X}\hat{\beta}_{ML}(\rho)\right]'\left[(I-\rho W)y - \widetilde{X}\hat{\beta}_{ML}(\rho)\right] \tag{33}$$

Note that, for convenience, $\hat{\beta}_{ML}(\rho) = \hat{\beta}_O - \rho\hat{\beta}_L$ where $\hat{\beta}_O = (\widetilde{X}'\widetilde{X})^{-1}\widetilde{X}'y$ and $\hat{\beta}_L = (\widetilde{X}'\widetilde{X})^{-1}\widetilde{X}'Wy$. Define $\hat{e}_O = y - \widetilde{X}\hat{\beta}_O$ and $\hat{e}_L = y - \widetilde{X}\hat{\beta}_L$, it can be then easily seen that $\hat{\sigma}^2_{ML}(\rho) = \left[\frac{(\hat{e}_O - \rho\hat{e}_L)'(\hat{e}_O - \rho\hat{e}_L)}{N}\right]$.

[15] In addition to the maximum likelihood method, the method of instrumental variables (Anselin 1988, Kelejian and Prucha 1998, Lee 2003) may also be applied to estimate SAR models.

Substitution of (30) and (31) in the log-likelihood function (28) yields a concentrated log-likelihood as a non-linear function of a single parameter ρ:

$$\begin{aligned} \ln L(\rho) &= -\frac{N}{2}\left[\ln(2\pi)+1\right]+\sum_{i=1}^{N}\ln(I-\rho\omega_i) \\ &\quad -\frac{N}{2}\ln\left[\frac{(\hat{e}_O-\rho\hat{e}_L)'(\hat{e}_O-\rho\hat{e}_L)}{N}\right] \end{aligned} \tag{34}$$

where $\hat{e}_O$ and $\hat{e}_L$ are the estimated residuals in a regression of y on X and Wy on X, respectively. A maximum likelihood estimate for ρ is obtained from a numerical optimization of the concentrated log-likelihood function (34).[16] Under the regularity conditions described for instance in Lee (2004), it can be shown that the maximum likelihood estimators have the usual asymptotic properties, including consistency, normality, and asymptotic efficiency. The asymptotic variance-covariance matrix follows as the inverse of the information matrix, defining $W_A = W(I-\rho W)^{-1}$ to simplify notation, we have:

$$\text{AsyVar}[b',\rho,\sigma^2] = \begin{bmatrix} \frac{1}{\sigma^2}\widetilde{X}'\widetilde{X} & \frac{1}{\sigma^2}(\widetilde{X}'W_A\tilde{X}b)' & 0 \\ \frac{1}{\sigma^2}\widetilde{X}'W_A\tilde{X}b & tr\left[(W_A+W_A')W_A\right]+\frac{1}{\sigma^2}(W_A\widetilde{X}b)'(W_A\widetilde{X}b) & \frac{1}{\sigma^2}trW_A \\ 0 & \frac{1}{\sigma^2}trW_A & \frac{N}{2\sigma^4} \end{bmatrix}^{-1} \tag{35}$$

[16] The quasi-maximum likelihood estimators of the SAR model can also be considered if the disturbance in the model are not truly normally distributed (Lee 2004).