Convergence, Human Capital and International Spillovers

Cem ERTUR and 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
tel: +33-380-39-35-23
fax: +33-380-39-54-43
21066 Dijon Cedex, France
cem.ertur@u-bourgogne.fr
kochwilfriedfr@aol.com

First Draft: March 5, 2006

Abstract

This paper develops a growth model with physical and human capital externalities together with technological interdependence between economies. It leads to a spatial autoregressive reduced form for the convergence equation characterized by parameter heterogeneity. A locally linear spatial autoregressive specification is then estimated providing a different convergence speed estimate for each country in a sample of 89 countries over the period 1960-1995. Finally, counterfactual density estimates show that our model better fits the observed income distribution than the well known augmented neoclassical growth model.

KEYWORDS: Conditional convergence, spatial externalities, spatial autocorrelation, bayesian estimation, parameter heterogeneity, locally linear estimation

JEL: C14, C21, O41

*The authors wish to thank K. Behrens and A. Desdoigts for their valuable comments and J. LeSage for the invaluable help he provided for the implementation of the Bayesian MCMC estimation methodology through his Econometric Toolbox for Matlab. Wilfried Koch acknowledges financial support of the CNRS-GIP-ANR "young researchers" program. The usual disclaimer applies.

1 Introduction

Since the seminal paper of Mankiw, Romer and Weil (1992, henceforth MRW), human capital is commonly introduced as a production factor along with the physical capital in growth regressions. In this framework, countries are most of the time considered as independent observations with no interactions between them. Growth regressions are then estimated by OLS under the standard assumptions. However, evidence in favor of cross-sectional correlation, i.e. spatial autocorrelation, is now well documented in the empirical literature (Conley and Ligon, 2002; Ertur *et al.*, 2006; Moreno and Trehan, 1997) and can no more be neglected in growth modeling. Moreover, technological progress and its diffusion channels: trade, foreign direct investment or geographical proximity among others are motivating, from a theoretical perspective, the observed world-wide interdependence (Keller, 2004).

Ertur and Koch (2005) propose therefore a growth model with technological interdependence and physical capital externalities that yields a spatial autoregressive convergence equation as a reduced form. Their results show the importance of technological interdependence in growth and development processes and underline the fact that the traditional results obtained in empirical growth regressions may be severely biased. Moreover, they show that the capital share is close to one third even without including human capital as in MRW (1992), once technological interdependence is modeled.

The question is then: what is the effect of human capital in such a framework? Therefore, in the first section, we extend the model proposed by Ertur and Koch (2005) by adding human capital as a production factor as MRW (1992). Human capital externalities are modeled along the lines of Lucas (1988) and physical capital externalities along the lines of Romer (1986). Technological interdependence is modeled in the form of spatial externalities in order to take account of the world-wide diffusion of knowledge across borders. Our model also yields a spatial autoregressive conditional convergence equation including both spatial autocorrelation and parameter heterogeneity as a reduced form.

In order to estimate our model, we use recent developments of spatial econometrics. Indeed, the simplest version of our model that we use as a benchmark, obtained by imposing identical convergence rates, leads to a spatial Durbin model including both spatial lags of exogenous and endogenous variables to be compared to the standard MRW model (1992). First, we use the quasi-maximum likelihood estimation method (Lee, 2004). However, since the empirical growth literature underlines the importance of heteroscedasticity and outliers (Temple, 1998), we also use the Bayesian heteroskedastic estimation approach developed by LeSage (1997). This approach handles outliers and addresses robustness concerns in the context of spatial modeling using the Markov Chain Monte Carlo methodology generalizing the approach of Geweke (1993) in the classical linear model. Finally, we use the locally linear spatial autoregressive estimation method developed by LeSage and Pace (2004) in order to estimate our full reduced form model characterized by varying convergence rates leading to complete parameter heterogeneity.

Our results show the importance of the spatial autocorrelation process and therefore of the underlying technological interdependence phenomenon. They show also that the coefficient of human capital is low and not significant when it is used as a simple production factor. This result is consistent with the human capital puzzle raised in the literature (Benhabib and Spiegel, 1994; Bils and Kleenow, 2000; Pritchett, 2001). Moreover, counterfactual density estimates show that the MRW (1992) model is not able to capture the well-known twin peaks characterizing the international per worker income distribution whereas our model indeed capture this feature. These results cast doubt on the MRW (1992) treatment of human capital as a simple production factor in the neoclassical growth model.

2 The model

Let us consider an aggregate Cobb Douglas production function for country i at time t exhibiting constant returns to scale in labor and reproducible physical and human capital:

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

with $Y_i(t)$ the output, $K_i(t)$ the level of reproducible physical capital, $H_i(t)$ the level of reproducible human capital, $L_i(t)$ the level of raw labor and $A_i(t)$ the aggregate level of technology:

$$A_i(t) = \Omega(t)k_i(t)^{\phi_K}h_i(t)^{\phi_H}\prod_{j\neq i}^{N} A_j(t)^{\gamma w_{ij}} \tag{2}$$

This function describing the level of technology in a country i depends of three terms. First, as the Solow (1956) and MRW (1992) models, we suppose that a part of technological progress depends on exogenous technical progress defined by: $\Omega(t) = \Omega(0)e^{\mu t}$ where μ is the exogenous technical progress growth rate and $\Omega(0)$ is its initial level. Second, we suppose that the technical progress increases with accumulated factors. It increases with per worker physical capital $k_i(t) = K_i(t)/L_i(t)$ reflecting the learning by doing process underlined by Arrow (1962) and Romer (1986) and with per worker human capital $h_i(t) = H_i(t)/L_i(t)$ reflecting the effect of human capital externalities as underlined by Lucas (1988). Parameters ϕ_K and ϕ_H reflect respectively the strength of physical capital externalities and human capital externalities. Finally, we suppose as Ertur and Koch (2005), that the technical progress of a country i depends positively on the technical progress of other countries. The last term represents technological interdependence between countries since they are all connected by parameters w_{ij} reflecting the interactions between countries i and j. These terms are assumed non negative, non stochastic and finite. We suppose also that: $0 \leq w_{ij} \leq 1$, $w_{ij} = 0$ if $j = i$ and $\sum_{j\neq i}^{N} w_{ij} = 1$ for $i = 1, ..., N$. In matrix form, these terms are to be collected in a connectivity matrix also called the spatial weight matrix denoted by W. The parameter γ reflects the degree of international technological interdependence.

Resolving equation (2) for A_i and replacing in the production function per worker, we obtain the following expression:

$$y_i(t) = \Omega(t)^{\frac{1}{1-\gamma}}k_i(t)^{u_{ii}}h_i(t)^{v_{ii}}\prod_{j\neq i}^{N} k_j(t)^{u_{ij}}h_j(t)^{v_{ij}} \tag{3}$$

where $y_i(t) = Y_i(t)/L_i(t)$ is per worker income, $u_{ii} = \alpha + \phi_K\left(1 + \sum_{r=1}^{\infty}\gamma^r w_{ii}^{(r)}\right)$, $v_{ii} = \beta + \phi_H\left(1 + \sum_{r=1}^{\infty}\gamma^r w_{ii}^{(r)}\right)$, $u_{ij} = \phi_K\left(1 + \sum_{r=1}^{\infty}\gamma^r w_{ij}^{(r)}\right)$ and $v_{ij} = \phi_H\left(1 + \sum_{r=1}^{\infty}\gamma^r w_{ij}^{(r)}\right)$. Where

$w_{ij}^{(r)}$ is the element in row i and column j of the connectivity matrix W to the power of r. Note that our model implies parameter heterogeneity in the production function. Finally, we suppose that there are social decreasing returns where all physical and human capital increase in the same level:

$$\sum_{j=1}^{N}(u_{ij}+v_{ij})=\alpha+\beta+\frac{\phi_K+\phi_H}{1-\gamma}<1 \tag{4}$$

As in the MRW (1992) model, we suppose that an exogenous part of the income is saved and invested in the physical and human capital respectively $s_{K,i}$ and $s_{H,i}$. We suppose also that each capital depreciates at the same level δ and the population growth grows exogenously at the rate n_i. The fundamental dynamic equation of Solow is written as follows:

$$\dot{k}_i(t)=s_{K,i}y_i(t)-(n_i+\delta)k_i(t) \tag{5}$$

for the physical capital and:

$$\dot{h}_i(t)=s_{H,i}y_i(t)-(n_i+\delta)h_i(t) \tag{6}$$

for the human capital. At steady state, physical and human capital grow at the same rate denoted by g, and defined by : $g=\frac{\mu}{(1-\gamma)(1-\alpha-\beta)-\phi_K-\phi_H}$. Since the production function per worker is characterized by decreasing returns, equations (5) and (6) imply that the physical capital-output and human capital-output ratios of country i, for $i=1,...,N$, are constant so that: $k_i^\star/y_i^\star=s_{K,i}/(n_i+g+\delta)$ for physical capital and $h_i^\star/y_i^\star=s_{H,i}/(n_i+g+\delta)$ for human capital. Replacing these expressions in the production function written per worker, we obtain per worker income at steady state as follows:

$$\begin{aligned}
\ln y_i^\star &= \frac{\alpha+\phi_K}{1-\alpha-\beta-\phi_K-\phi_H}\ln\left[\frac{s_{K,i}}{n_i+g+\delta}\right]+\frac{\beta+\phi_H}{1-\alpha-\beta-\phi_K-\phi_H}\ln\left[\frac{s_{H,i}}{n_i+g+\delta}\right]\\
&- \frac{\alpha\gamma}{1-\alpha-\beta-\phi_K-\phi_H}\sum_{j\neq i}^{N}w_{ij}\ln\left[\frac{s_{K,j}}{n_j+g+\delta}\right]\\
&- \frac{\beta\gamma}{1-\alpha-\beta-\phi_K-\phi_H}\sum_{j\neq i}^{N}w_{ij}\ln\left[\frac{s_{H,j}}{n_j+g+\delta}\right]\\
&+ \frac{(1-\alpha-\beta)\gamma}{1-\alpha-\beta-\phi_K-\phi_H}\sum_{j\neq i}^{N}w_{ij}\ln y_j^\star
\end{aligned} \tag{7}$$

Taking log-linearization of equations (5) and (6) around steady states and using the same technic as the one developed in Ertur and Koch (2005), we obtain the following differential equation for each country:

$$\frac{d\ln y_i(t)}{dt}=\frac{\mu}{1-\gamma}-\lambda_i[\ln y_i(t)-\ln y_i^\star] \tag{8}$$

with:

$$\lambda_i=\Lambda_i-\sum_{j=1}^{N}(u_{ij}+v_{ij})\frac{1}{\Theta_j}(n_j+g+\delta) \tag{9}$$

the speed of convergence. This expression contains two scale parameters specific to each country.

First, the scale parameter Λ_i represents the fact that each country has a different population growth rate n_i. If all countries have the same population rate of growth n, this scale parameter reduces to $(n+g+\delta)$. It is defined as follows:

$$\Lambda_i = \frac{\sum_{j=1}^{N}\frac{1}{\theta_j}(u_{ij}+v_{ij})(n_j+g+\delta)}{\sum_{j=1}^{N}\frac{1}{\theta_j}(u_{ij}+v_{ij})} \tag{10}$$

with : $\ln k_i(t) - \ln k_i^\star = \theta_j[\ln k_j(t) - \ln k_j^\star]$. The second scale parameter Θ_j represents the relations between the gaps of countries in respect to their own steady states in the following form:

$$\ln y_i(t) - \ln y_i^\star = \Theta_j[\ln y_j(t) - \ln y_j^\star] \quad \text{for} \quad j=1,...,N \tag{11}$$

When this parameter is equal to 1, countries i and j are in the same distance in respect to their steady states. If this parameter is higher than 1 (respectively lower than 1) the country i is farther (respectively closer) to its own steady state than the country j. It represents the effects of technological interdependence between countries. In fact, the speed of convergence depends positively on this parameter since: $\frac{\partial\lambda_i}{\partial\Theta_j} = \frac{(u_{ij}+v_{ij})(n_j+g+\delta)}{\Theta_j^2} > 0$. Therefore, when a country i is far from its own steady state, its speed of convergence is high which is a traditional result. When the neighbors of country i are close to their own steady states, the convergence speed of country i is high. The last relation may reflect the fact that the closer its neighbors are to their own steady states, the more they can spill on the technological progress of country i. Moreover, we can note that the speed of convergence is specific to each country, so the model implies complete parameter heterogeneity since λ_i is a function of the parameters w_{ij} reflecting the links between countries. Therefore, in this model, heterogeneity comes from technological interdependence. Note that, when there are no physical and human capital externalities (ϕ_K and ϕ_H), the speed of convergence reduces to that of the MRW model: $\lambda_i = -(1-\alpha-\beta)(n_i+g+\delta)$.
Finally, after resolving the differential equation (8) and some algebraical manipulations, we can write the econometric specification of the spatial conditional convergence equation in the following form:

$$\begin{aligned}\frac{[\ln y_i(t) - \ln y_i(0)]}{T} &= \beta_{0i} + \beta_{1i}\ln y_i(0) + \beta_{2i}\ln s_{K,i} + \beta_{3i}\ln s_{H,i} + \beta_{4i}\ln(n_i+g+\delta)\\ &+ \theta_{1i}\sum_{j\neq i}^{N} w_{ij}\ln y_j(0) + \theta_{2i}\sum_{j\neq i}^{N} w_{ij}\ln s_{K,j} + \theta_{3i}\sum_{j\neq i}^{N} w_{ij}\ln s_{H,j}\\ &+ \theta_{4i}\sum_{j\neq i}^{N} w_{ij}\ln(n_j+g+\delta) + \rho_i\sum_{j\neq i}^{N} w_{ij}\frac{[\ln y_j(t) - \ln y_j(0)]}{T}\end{aligned} \tag{12}$$

with $\beta_{0i} = \frac{(1-e^{-\lambda_i T})}{T}(\frac{g}{1-\gamma}\frac{1}{\lambda_i} + \frac{1}{1-\alpha-\beta-\phi_K-\phi_H}\Omega(T))$, $\beta_{1i} = -\frac{(1-e^{-\lambda_i T})}{T}$,
$\beta_{2i} + \beta_{3i} = -\beta_{4i} = \frac{(1-e^{-\lambda_i T})}{T}\frac{\alpha+\beta+\phi_K+\phi_H}{1-\alpha-\beta-\phi_K-\phi_H}$, $\theta_{1i} = \frac{(1-e^{-\lambda_i T})}{T}\frac{1-\alpha-\beta}{1-\alpha-\beta-\phi_K-\phi_H}\gamma$,
$\theta_{4i} = -(\theta_{2i}+\theta_{3i}) = \frac{(1-e^{-\lambda_i T})}{T}\frac{\alpha+\beta}{1-\alpha-\beta-\phi_K-\phi_H}\gamma$ and $\rho_i = \frac{(1-e^{-\lambda_i T})}{\Gamma_i}\frac{1-\alpha-\beta}{1-\alpha-\beta-\phi_K-\phi_H}\gamma$.
The term Γ_i is a scale parameter reflecting the effects of the speeds of convergence in the neighboring countries and T represents the time lag between the current period and the initial period. The growth rate of real income per worker is a negative function of the initial level of income per worker reflecting the convergence process. We also control for the determinants

of the steady state. More specifically, the growth rate of real income depends positively on saving rate in physical and human capital ($s_{K,i}$ and $s_{H,i}$) and negatively on the effective rate of capital depreciation ($n_i + g + \delta$). It depends also in the same direction on the saving rate in physical and human capital in the neighboring countries, and negatively on their effective capital depreciation rate.[1] The rate of growth also depends positively on the initial income per worker in the neighboring countries. Finally, the last term of equation (12) represents the rate of growth in the neighboring countries reflecting the spatial autocorrelation process implied by technological interdependence.

3 The econometric specification

Rewriting equation (12) in matrix form, we obtain:

$$y = DX\beta + DWX\theta + \rho D\Gamma Wy + \varepsilon \tag{13}$$

where y is the $(N \times 1)$ vector of the logarithms of real income per worker, X the $(N \times 4)$ matrix of the explanatory variables, including the constant term, the vector of the logarithms of the investment rates in physical and human capital and the vector of the logarithms of the effective rate of capital depreciation. W is the row standardized $(N \times N)$ spatial weight matrix, WX is the $(N \times 3)$ matrix of the spatially lagged exogenous variables and Wy is the endogenous spatial lag variable. $\beta' = [\beta_0, \beta_1, \beta_2, \beta_3]$, $\theta' = [\theta_1, \theta_2, \theta_3]$ are the parameters defined in equation (12) without the terms $(1 - e^{\lambda_i T})/T$, which are included in the diagonal matrix D reflecting the specific effects of the convergence speed in each country. Γ is also a diagonal matrix containing the scale parameters Γ_i. Finally, ρD is the spatial autoregressive parameter specific to each country. ε is the $(N \times 1)$ vector of independently and identically distributed errors with mean zero and variance σ^2.

This kind of specification, including the spatial lags of exogenous variables in addition to the lag of the endogenous variable, is referred to as the spatial Durbin model (SDM) in the spatial econometric literature. More specifically, we refer to this model as the global spatial Durbin model when all speeds of convergence are identical and then all parameters are homogenous. In contrast, when the speeds of convergence are specific to each country, we refer to the equation (12) as the local spatial Durbin model with heterogenous parameters.

Under the hypothesis of normality of the error term, we estimate the global spatial Durbin model assuming that the speed of convergence is homogenous and so identical for all countries: $\lambda_i = \lambda$ for $i = 1, ..., N$ in equation (12) using a maximum likelihood estimation framework (see Anselin and Bera, 1998 for details). 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 quasi-maximum likelihood estimators of the SDM model can also be considered if the disturbance in the model are not truly normally distributed (Lee, 2004). The heterogenous local SDM model in equations (12)

[1]We note that there are negative signs associated with the saving rates and a positive sign with the rate of depreciation in the neighboring countries in equation (12). This seems in contradiction with our claim in the text, however equation(12) is written in implicit form and the net effect must be considered in its explicit form. See Ertur and Koch (2005) for more details.

and (13) is in turn estimated by the recursive spatial maximum likelihood approach developed by LeSage and Pace (2004).

In addition to maximum likelihood, the method of instrumental variables (Kelejian and Prucha, 1998) may also be applied to estimate SDM models as well as the Bayesian estimation method (LeSage, 1997).

Furthermore, as outlined in Temple (1998, 1999), heteroskedasticity and outliers are well known problems in the empirical growth regression literature.[2] Use of non-parametric and robust estimation methods have thus been advocated (for instance least trimmed squares and reweighed least squares in Temple, 1998). LeSage (1997) recently proposed an alternative Bayesian heteroskedastic estimation approach which handles ouliers and addresses robustness concerns in the context of spatial modeling using the Markov Chain Monte Carlo methodology (Gelfand and Smith, 1990). This approach extends to spatial autoregressive models the Bayesian treatment of heteroskedasticity suggested by Geweke (1993) in the linear model. As proved by this author, the bayesian approach to modeling heteroskedastic disturbances in the linear model is equivalent to a specification that assumes an independent Student-t distribution for the errors. This type of leptokurtic distributions has frequently been used to deal with sample data containing outliers (Lange *et al.*, 1989).

These models allow the disturbances to take the form $\varepsilon \sim N(0, \sigma^2 V)$, where V is a diagonal matrix containing variance scalars $v_1, v_2, \ldots, v_n$, estimated using Markov Chain Monte Carlo (MCMC) methods. Prior information regarding the variance scalars v_i takes the form of a set of N independent, identically distributed, $\chi^2(r)/r$ distributions, where r represents the single degree of freedom parameter of the χ^2 distribution. This allows us to estimate the additional N non-zero variance scaling parameters v_i by adding only a single parameter r, to the model.

The specifics regarding the prior assigned to the v_i terms can be motivated by considering that the mean equals unity and the variance of the prior is $2/r$. This implies that as r becomes very large, the terms v_i will all approach unity, resulting in the non-zero variance scalars taking the form $V = I_N$, the traditional assumption of constant variance across space. On the other hand, small values of r lead to a skewed distribution permitting large values of v_i that deviate greatly from the prior mean of unity. The role of these large v_i values is to accommodate outliers or observations containing large variances by down-weighting these observations. In practice, one can assign an informative prior for the parameter r based on the Gamma distribution with parameters m and k. This distribution has a mean of m/k and a variance of m/k^2, so using $m = 8$ and $k = 2$ would assign a prior to r centered on a small value equal to 4 with a variance of 2. It is also possible to treat r as a hyperparameter in the model, set to a small value, for example $r = 4$. An extended approach would be to estimate to degree of freedom parameter of the Student-t distribution for the error terms along with other parameters in an equivalent specification by using a tractable prior distribution, for example an exponential prior distribution as suggested by Geweke (1993).

[2]see Durlauf et al. (2005) for a general discussion on problems raised with regard to this literature.

4 Data

Following the literature on empirical growth, we draw our basic data from the Heston, Summers and Aten (2002) Penn World Tables (PWT version 6.1), which contain information on real income, investment and population (among many other variables) for a large number of countries. In this paper, we use a sample of 89 countries over the period 1960-1995. These countries are those of the MRW (1992) non-oil sample, for which PWT 6.1 provides data.

We mesure n as the average growth of the working-age population (ages 15 to 64). Therefore, we compute the number of workers as: $RGDPCH \times POP/RGDPW$, where $RGDPCH$ is real GDP per capita computed by the chain method, $RGDPW$ is real-chain GDP per worker, and POP is the total population. Real income per worker is measured by the real GDP computed by the chain method, divided by the number of workers. The saving rate s_K is the average share of gross investment in GDP as in MRW (1992). Finally, the variable s_H, labeled SCHOOL in MRW (1992), is the average percentage of a country's working-age population in secondary school. More specifically, MRW (1992) define SCHOOL as the percentage of school-age population (12-17) attending secondary school times the percentage of the working-age population that is of secondary-school age (15-19). In this paper, we use data provided by Bernanke and Gürkaynak (2003). Their data on working-age population and its components are from the World Bank's *World Tables*, the World Bank's *World Development Indicators* 2000 CD-ROM and the UN World Population Prospects.

The spatial weight matrix W is the matrix commonly used in spatial econometrics to model spatial interdependence between regions or countries (Anselin and Bera 1998). More precisely, each country is connected to a set of neighboring countries by means of a purely spatial pattern introduced exogenously in W. Elements w_{ii} on the main diagonal are set to zero by convention whereas elements w_{ij} indicate the way country i is spatially connected to 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. Lee (2004) presents some technical properties for the W matrix. It is important to stress that the friction terms w_{ij} should be exogenous to the model to avoid the identification problems raised by Manski (1993) in the social sciences. This is why we consider pure geographical distance, more precisely great-circle distance between capitals, which is indeed strictly exogenous. Geographical distance has also been considered by Eaton and Kortum (1996) or Klenow and Rodriguez-Clare (2005) among others.[3] The functional form we consider is simply the inverse of squared distance, which can be interpreted as reflecting a gravity function. The general term of this matrix $W1$ is defined as follows, in standardized form $[w1_{ij}]$:

$$w1_{ij} = w1^*_{ij} / \sum_j w1^*_{ij} \quad \text{with} \quad w1^*_{ij} = \begin{cases} 0 & \text{if} \quad i = j \\ d_{ij}^{-2} & \text{otherwise} \end{cases} \tag{14}$$

where d_{ij} is the great-circle distance between country capitals.

[3] Klenow and Rodriguez-Clare (2005) suggest that use of pure geographical distance could capture trade and FDI related spillovers.

5 Results

In this section, we assess the predictions for conditional convergence of the MRW model compared to our model, which is considered in two polar cases.[4] First, we suppose, for simplicity and comparability that the speed of convergence is identical for all countries and we refer to this case as the global homogenous SDM, which is our benchmark model. Second, we estimate a model with complete parameter heterogeneity and we refer to this case as the local heterogenous SDM, which is the full econometric specification of our theoretical model.

5.1 The MRW model

We first test the convergence predictions of the MRW (1992) model in Table 1 using OLS-White in the first column and heteroscedastic bayesian MCMC in the second column. Our results are essentially identical to those of MRW (1992).

Table 1: OLS and Bayesian heteroscedastic estimation results of the MRW convergence model.

Model	MRW convergence OLS-White	MRW convergence Bayesian het	
Dependent variable	$\frac{\ln y_i(1995)-\ln y_i(1960)}{35} \times 100$	$\frac{\ln y_i(1995)-\ln y_i(1960)}{35} \times 100$	95% HPDI
Obs. / Weight matrix	89	89	
const.	9.367	9.885	$[4.742, 15.181]$
	(2.978) (0.002)	(3.165)	
$\ln y_i(1960)$	-1.320	-1.291	$[-1.636, -0.959]$
	(0.200) (0.000)	(0.207)	
$\ln s_{K,i}$	1.423	1.340	$[0.864, 1.838]$
	(0.321) (0.000)	(0.296)	
$\ln s_{H,i}$	1.334	1.338	$[0.889, 1.782]$
	(0.266) (0.000)	(0.271)	
$\ln(n_i + 0.05)$	-4.008	-3.661	$[-5.205, -2.142]$
	(0.798) (0.000)	(0.926)	
Implied λ	0.017	0.017	
Half-life	53	53	
Moran's I test ($W1$)	0.257		
	(0.000)		
Restricted regression			
constant	12.296	12.471	$[9.433, 15.597]$
	(1.896) (0.000)	(1.864)	
$\ln y_i(1960)$	-1.282	-1.295	$[-1.641, -0.959]$
	(0.206) (0.000)	(0.206)	
$\ln s_{K,i} - \ln(n_i + 0.05)$	1.496	1.446	$[1.002, 1.907]$
	(0.312) (0.000)	(0.276)	
$\ln s_{H,i} - \ln(n_i + 0.05)$	1.323	1.352	$[0.903, 1.790]$
	(0.274) (0.000)	(0.270)	
Moran's I test ($W1$)	0.253		
	(0.000)		
Test of restriction	1.313		
	(0.255)		
Implied λ	1.7% (0.000)	1.7%	
Half-life	53	53	
Implied α	0.364 (0.000)	0.354	
Implied β	0.323 (0.000)	0.329	

Notes: coefficient estimates and standard deviations as well as p-values (in parentheses) are reported for OLS using White heteroskedasticity consistent covariance matrix estimator.
p-values for the implied parameters are computed using the delta method.
Posterior means and standard deviations (in parentheses) are reported for Bayesian heteroskedastic estimation using MCMC. Associated 95% Highest Posterior Density Intervals are reported in brackets in the 3rd column.

The coefficient on the initial level of income is significant and negative; that is, there is strong evidence of convergence when we control for investment rates in physical and human capital and

[4] Actually the MRW model is a particular case of our SDM model without externalities when $\phi_K = \phi_H = \gamma = 0$.

for the growth rate of working-age population. The results also support the predicted signs of each variables.

Estimation results for the equation imposing the restriction that the coefficients sum to zero are presented in the bottom part of Table 1. We find that this restriction is not rejected (the $p-$value of the test is 0.255). The implied values of α and β are close to the predicited values that is one third. As also underlined by MRW (1992), these regressions give a somewhat larger weight to physical capital and a somewhat smaller weight to human capital.

The implied value of λ, the parameter governing the speed of convergence, is derived from the coefficient on initial per worker income level. The value of the speed of convergence, $\lambda = 1.7$ % is close to that obtained by MRW (1992) and implies a half-life of about 53 years. Bayesian heteroskedastic MCMC estimation results support all the maximum likelihood results despite the presence of some potential outliers in both the constrained and unconstained MRW convergence models as illustrated in the upper part of Figure 1: Botswana, Hong-Kong, Jamaica, Zaire and Zambia have posterior mean v_i values exceeding 4 in both cases.

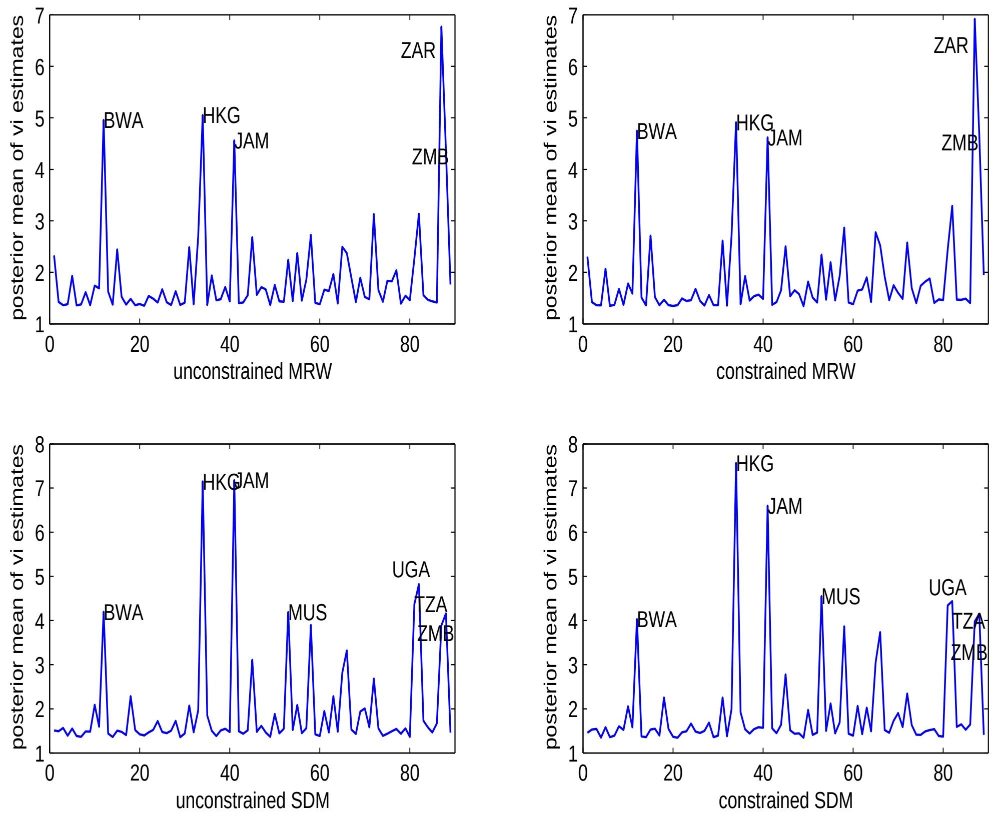

Figure 1: Posterior mean of v_i estimates for the MRW convergence model and homogenous global SDM.

Note that the MRW (1992) model is misspecified since it omits variables due to technological interdependence and physical and human capital externalities. Therefore, the error terms of MRW (1992) contain omitted information and are spatially autocorrelated as also indicated by Moran's I (Anselin and Bera, 1998) tests in Table 1.

Table 2: ML and Bayesian heteroscedastic estimation results of the homogenous global SDM.

Model	SDM-convergence QML	SDM-convergence Bayesian het	
Dependent variable	$\frac{\ln y_i(1995)-\ln y_i(1960)}{35} \times 100$	$\frac{\ln y_i(1995)-\ln y_i(1960)}{35} \times 100$	95% HPDI
Obs. / Weight matrix	89 / ($W1$)	89 / ($W1$)	
constant	1.005	3.645	$[-3.438, 10.775]$
	(3.931) (0.798)	(4.338)	
$\ln y_i(1960)$	−1.718	−1.667	$[-1.999, -1.331]$
	(0.214) (0.000)	(0.203)	
$\ln s_{K,i}$	1.431	1.544	$[1.124, 1.955]$
	(0.232) (0.000)	(0.253)	
$\ln s_{H,i}$	1.196	1.034	$[0.626, 1.456]$
	(0.253) (0.000)	(0.252)	
$\ln(n_i + 0.05)$	−3.790	−3.707	$[-5.433, -2.017]$
	(1.129) (0.000)	(1.037)	
$W \ln y_i(1960)$	1.346	1.426	$[0.708, 2.087]$
	(0.358) (0.000)	(0.420)	
$W \ln s_{K,j}$	−1.461	−1.277	$[-2.326, -0.225]$
	(0.580) (0.012)	(0.638)	
$W \ln S_{H,j}$	−0.256	−0.314	$[-1.260, 0.662]$
	(0.601) (0.670)	(0.585)	
$W \ln(n_j + 0.05)$	1.590	2.966	$[0.031, 5.828]$
	(1.882) (0.398)	(1.764)	
$W\left(\frac{\ln y_j(1995)-\ln y_j(1960)}{35} \times 100\right)$	0.504	0.517	$[0.320, 0.824]$
	(0.115) (0.000)	(0.114)	
Common factor test	7.987 (LR)	1.000 PMP unrest.	
SDM vs. SEM	(0.092)	0.000 PMP rest.	
Restricted regression			
constant	3.110	3.505	$[-1.608, 8.869]$
	(2.760) (0.260)	(3.195)	
$\ln y_i(1960)$	−1.713	−1.695	$[-2.020, -1.369]$
	(0.214) (0.000)	(0.199)	
$\ln s_{K,i} - \ln(n_i + 0.05)$	1.481	1.547	$[1.146, 1.940]$
	(0.228) (0.000)	(0.242)	
$\ln s_{H,i} - \ln(n_i + 0.05)$	1.200	1.067	$[0.669, 1.481]$
	(0.252) (0.000)	(0.248)	
$W \ln y_i(1960)$	1.461	1.401	$[0.700, 2.059]$
	(0.342) (0.000)	(0.415)	
$W[\ln s_{K,j} - \ln(n_j + 0.05)]$	−1.208	−1.323	$[-2.190, -0.414]$
	(0.498) (0.010)	(0.541)	
$W[\ln s_{H,j} - \ln(n_j + 0.05)]$	−0.472	−0.280	$[-1.162, 0.618]$
	(0.558) (0.398)	(0.546)	
$W\left(\frac{\ln y_j(1995)-\ln y_j(1960)}{35} \times 100\right)$	0.528	0.528	$[0.331, 0.817]$
	(0.111) (0.000)	(0.114)	
Test of restriction	1.544 (LR)	0.029 PMP unrest.	
	(0.819)	0.971 PMP rest.	
Common factor test	7.877 (LR)	1.000 PMP unrest.	
SDM vs. SEM	(0.049)	0.000 PMP rest.	
Implied λ	2.61% (0.000)	2.57%	
Half-life	40.5	40.0	
Implied Γ	1.324 (0.000)	1.333	
Implied α	0.398 (0.014)	0.440	
Implied β	0.147 (0.342)	0.093	
Implied ϕ_K	−0.061 (0.697)	-0.081	
Implied ϕ_H	0.126 (0.388)	0.154	
Implied γ	0.731 (0.000)	0.697	
Implied scale returns	0.787 (0.000)	0.774	

Notes: coefficient estimates and standard deviations as well as p-values (in parentheses) are reported for quasi maximum likelihood estimation (QML). p-values for the implied parameters are computed using the delta method. Posterior means and standard deviations (in parentheses) are reported for Bayesian heteroskedastic estimation using MCMC. PMP stands for posterior model probability. Associated 95% Highest Posterior Density Intervals are reported in brackets in the last column.

5.2 The homogenous model

In this section we assume that the speed of convergence is homogenous and so identical for all countries: $\lambda_i = \lambda$ for $i = 1, ..., N$ in equation (12) and estimate the homogenous SDM by quasi-maximum likelihood estimation method and Bayesian MCMC estimation method. Note that the latter is robust with regard to heteroscedasticity and potential outliers.

In Table 2, we estimate the conditional convergence equation. Many aspects of the results support our model. First, the spatial autocorrelation coefficient ρ is positive and significant which shows the impact of global technological interdependence on the growth of countries. Second, all coefficients have the predicted signs and are significant except the spatial lags of human capital investment and working-age population growth rate when considering quasi-maximum likelihood estimation. However, we can note that the latter coefficient becomes positive as predicted and significant when estimated by Bayesian heteroskedastic MCMC. Posterior mean v_i estimates are displayed in the lower left of Figure 1 for the unconstrained MRW convergence model showing some evidence of non-constant variance or outliers. We note that Botswana, Hong-Kong, Jamaica, Mauritius, Uganda, Tanzania and Zambia present large values exceeding 4: these observations may therefore be interpreted as outliers. The downward bias of the quasi-maximum likelihood estimate of the coefficient of the spatial lag of population growth rate is represented in Figure 2.

Third, the coefficient on the initial level of income is negative and significant, so there is strong evidence of convergence after controlling for those variables that are determining the steady state in the SDM homogenous convergence model. Fourth, the linear constraints implied by the model are not rejected since the $p-$value of the LR test is 0.819 and the posterior model probability (PMP) is 0.971 for the constrained model against 0.029 for the unconstrained model. Note that the same potential outliers are detected in the constrained model (lower right of Figure 1) and down-weighted in the MCMC estimation procedure. However, the Bayesian heteroskedastic estimates remain close to those obtained by quasi-maximum likelihood. Significance test results are not affected.

The implied value of the convergence speed λ is higher than that found by MRW (1992) because of technological interdependence. The value is 2.6% and the half-life is 40.5. The implied value of the capital share in income is close to 0.4, the upper bound generally admitted in the literature. The implied value of the human capital parameter is close to 0.15, a very low value indeed and moreover is not significant. Moreover, the physical and human capital externalities are not individually significant. We can also jointly test the absence of physical and human capital externalities using the common factor test. More precisely, the null hypothesis of this test is jointly: $\phi_K = \phi_H = 0$ and $\Gamma = 1$, that is there is no physical and human capital externalities and the scale parameter is equal to 1. However, in the homogenous global SDM convergence equation, the scale parameter is theoretically equal to 1. Therefore, the common factor tests allow to reject the absence of externalities in the model, i.e. $\phi_K \neq 0$ or $\phi_H \neq 0$. In fact, the $p-$values of the LR test range from 0.049 to 0.092 and the $PMPs$ strongly play in favor of the unconstrained model. The implied scale returns are significant and significantly below 1 (p-value of 0.065).

The value of the coefficient for global technological interdependence is significant and close to 0.7 confirming its crucial role in growth and convergence processes.

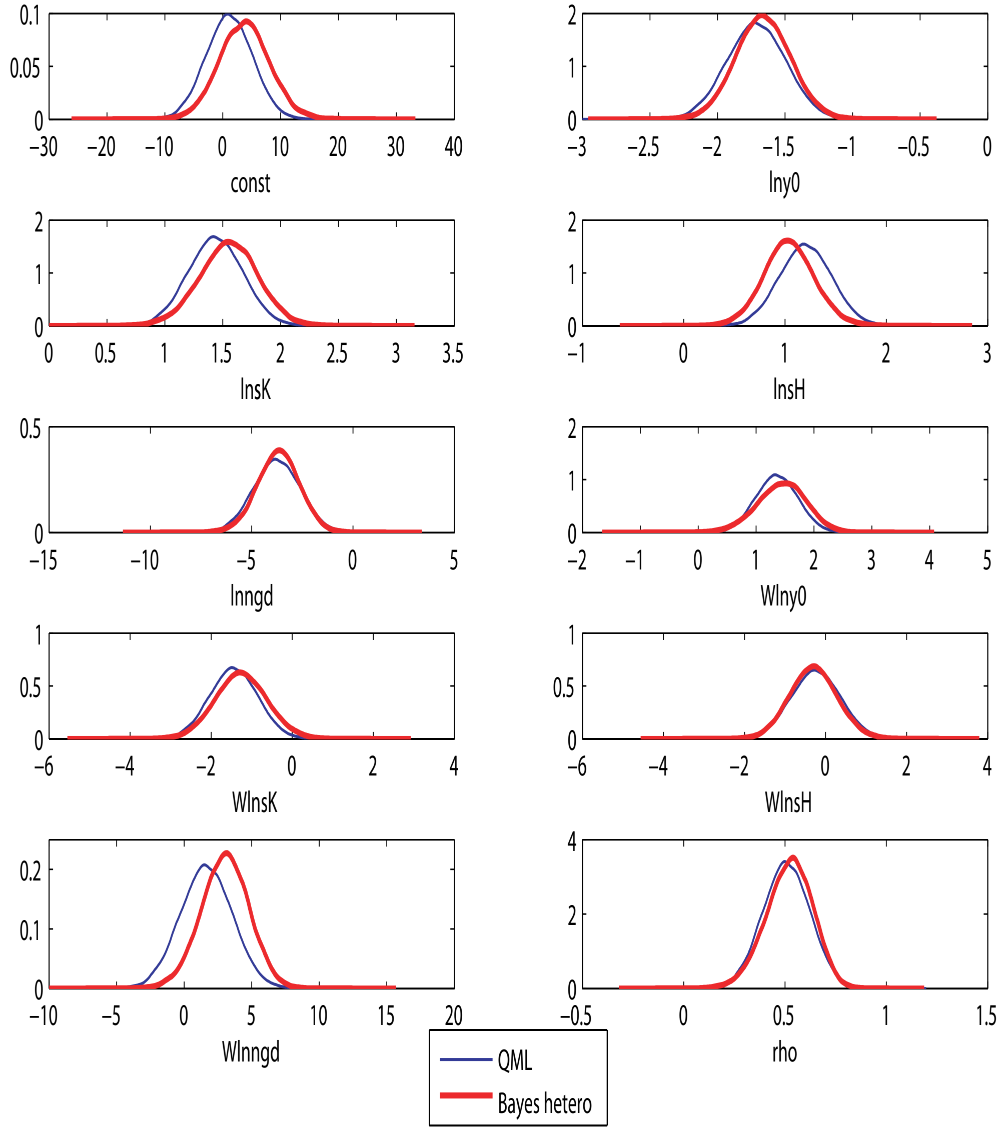

Figure 2: Kernel density estimates of Bayesian posterior and simulated quasi-maximum likelihood distributions of the parameters for the unconstrained homogenous global SDM

5.3 The heterogenous model

In recent papers, Durlauf (2001) and Brock and Durlauf (2001) draw attention to the assumption of parameter homogeneity imposed in cross-section growth regressions. Indeed, it is unlikely to assume that the parameters describing growth are identical across countries. Moreover, evidence of parameter heterogeneity has been found using different statistical methodologies such as in Canova (2004), Desdoigts (1999), Durlauf and Johnson (1995) and Durlauf et al. (2001). Each of these studies suggests that the assumption of a single linear statistical growth model applying to all countries is incorrect.

From the econometric methodology perspective, Islam (1995) or Evans (1998) among others have suggested the use of panel data to address this problem, but this approach is of limited use in empirical growth contexts, because variation in the time dimension is typically small. Some variables as for example political regime do not vary by nature over high frequencies and some other variables are simply not measured over such high frequencies. Moreover high frequency data will contain business cycle factors that are presumably irrelevant for long run output movements. The use of long run averages in cross sectional analysis has still a powerful justification for identifying growth as opposed to cyclical factors. Durlauf and Quah (1999) underline also that it might appear to to be a proliferation of free parameters not directly motivated by economic theory.

The empirical methodology we propose takes into account the heterogeneity embodied in equation (12). To accommodate both spatial dependence and heterogeneity, we produce estimates using N-models, where N represents the number of cross-sectional sample observations, using the locally linear spatial autoregressive specification. The original methodology was proposed by LeSage and Pace (2004) and labeled spatial autoregressive local estimation (SALE). We implement here this specification and label it the local SDM model:

$$U(i)y = U(i)X\beta_i + U(i)WX\theta_i + \rho_i U(i)Wy + U(i)\varepsilon \tag{15}$$

where $U(i)$ represents an $N \times N$ diagonal matrix containing distance-based weights for observation i that assign weights of one to the m nearest neighbors to observation i and weights of zero to all other observations. This results in the product $U(i)y$ representing an $m \times 1$ sub-sample of observed GDP growth rates associated with the m observations nearest in location to observation i (using great-circle distance). Similarly, the product $U(i)X$ extracts a sub-sample of explanatory variable information based on m nearest neighbors and so on. The local SDM model assumes $\varepsilon_i \sim N(0, \sigma_i^2 U(i) I_N)$. The model is estimated by the recursive spatial maximum likelihood approach developed by LeSage and Pace (2004).

The scalar parameter ρ_i measures the influence of the variable, $U(i)Wy$ on $U(i)y$. We note that as $m \to N$, $U(i) \to I_N$ and these estimates approach the global estimates based on all N observations that would arise from the global SDM model. The suggested range for sub-sample size is $\frac{N}{4} < m < \frac{3N}{4}$ (LeSage and Pace, 2004).

The local SDM model in the context of convergence analysis means that each country converges to its own steady state at its own rate represented by the parameter λ_i. Therefore, heterogeneity in both level of steady state and transitional growth rates toward this steady state is allowed. We implement the estimation procedure for $m = 30, 45$ and 60.

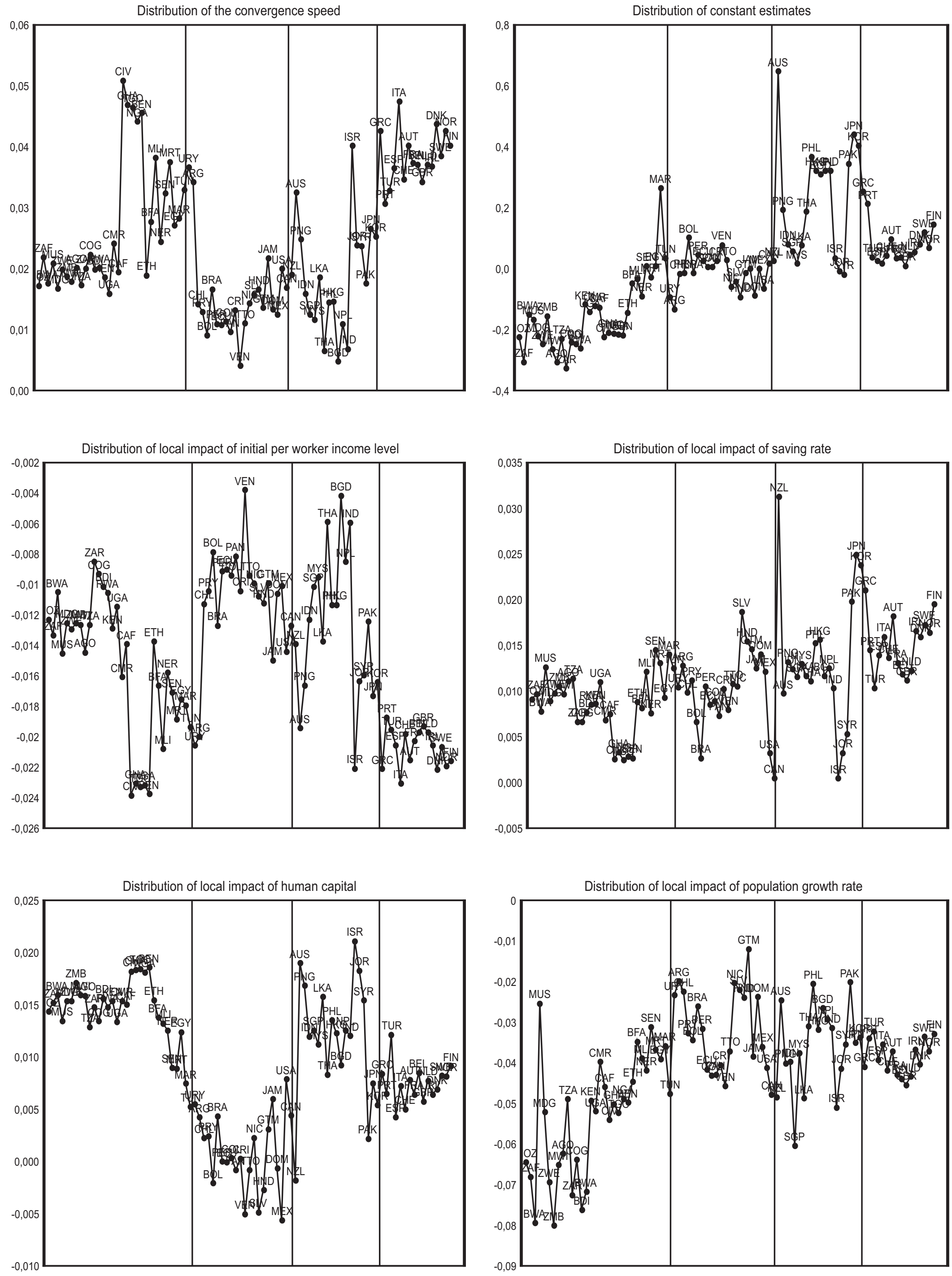

Figure 3: Distributions of local impacts of variables of interest in the heterogenous local SDM with $m = 30$.

Due to lack of space, estimation results are presented in Figure 3 only for $m = 30$.[5] Countries are ordered by continent (Africa, America, Asia and Europe) and increasing latitude in each continent.

We note strong evidence for parameter heterogenity as in Durlauf et al. (2001) or Ertur and Koch (2005). This heterogeneity is furthermore linked to the geographical location of the observations. For sub-sample size of $m = 30$, in Figure 3, we report the local impact of each variable after considering the explicit form of the SDM heterogenous convergence model. We note that the convergence speed is higher in European and Northern African countries than American and Asian countries. Note that the convergence speed for USA is just above 2% whereas for Northern European countries it is double, close to 4%. Japan is in between those convergence speeds. We also note that the distribution of the impact of the initial level of income per worker is symmetrical with regard to the distribution of the convergence speed. The local impact of the investment rate in physical capital accumulation seems more important in South East Asian countries as Japan and South Korea. This result is convergent with those of Young (1995) about the importance of factor accumulation in their development process whereas it is very low for some Central African countries. The local impact of human capital seems very low in American countries with respect to the rest of the world. However, it seems to play an important role in African countries since these countries are characterized by a lack of human capital together with low growth rates. Finally, we see that the local impact of the population growth rate is higher for African countries than for the rest of the world. We also note that all these results are globally robust to the choice of the sub-sample size.

5.4 Counterfactual income kernel density estimates

The effect of the different theoretical frameworks on the world income dynamics are estimated by applying kernel density methods. Practical application of kernel density estimation is crucially dependent on the choice of the smoothing parameter. In the following analysis, we use the plug-in method of Sheather and Jones (1991) as bandwidth selector for the gaussian kernel that is also chosen by Di Nardo, Fortin and Lemieux (1996) and Desdoigts (2004). The results are presented in Figure 4.

In the upper left box of Figure 4, univariate density estimates of the world real output per worker is displayed. The solid line represents the density estimates in 1960 and the dashed line represents the density estimates in 1995. We can note that the dashed line shows more structure than the solid line. In fact, the dynamics of the cross-section distribution of countries exhibit polarization and the so-called phenomenon of twin peaks distribution dynamics across countries is at work over the period 1960-1995 (Quah, 1996).

In order to evaluate the local impact of models in the world income distribution, we compare the observed per worker income distribution (dashed line) with the distributions implied by the MRW (1992) model, our homogenous and heterogenous models (solid lines). First, in the upper right box of Figure 4, we can note that the MRW (1992) model does not capture the structure of the observed income distribution. In fact, there is only one mode whereas there are three modes in the observed distribution. Our model captures two of three modes in the observed income

[5]Results for $m = 45$ and $m = 60$ are available from the authors upon request.

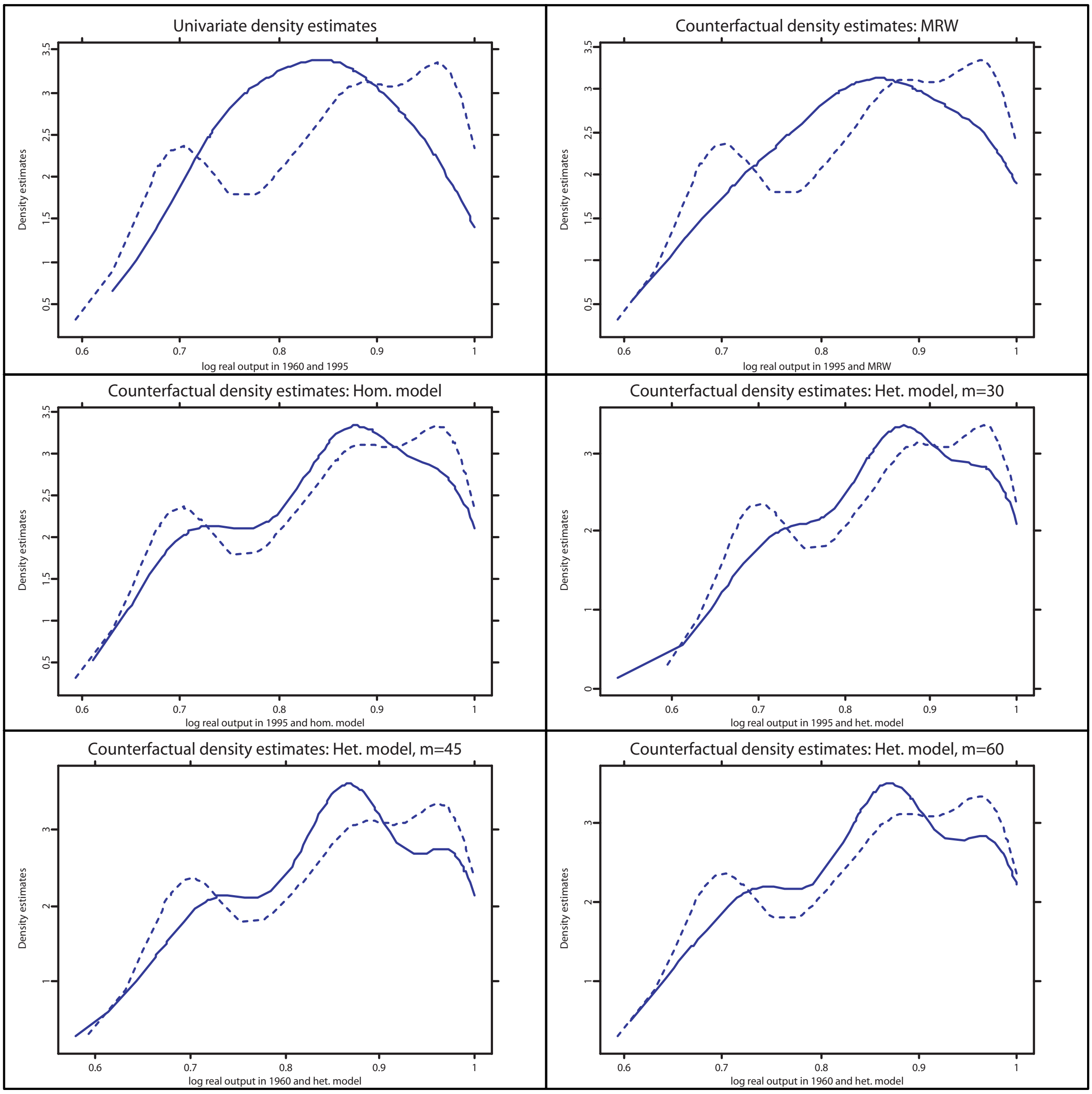

Figure 4: Counterfactual kernel density estimates

distribution. However, when we consider the local estimation of our heterogenous model for all sub-sample sizes considered ($m = 30, 45$ and 60), we capture the three modes of the observed distribution. The latter model seems to perform the best.

6 Conclusion

In this paper we develop a spatially augmented growth model with physical and human capital externalities together with technological interdependence between countries. Implied econometric specifications are estimated by spatial econometric technics: quasi-maximum likelihood as well as robust bayesian heteroskedastic MCMC estimators are used in order to take into account heteroskedasticity and potential outliers. The results support our model for different reasons. First, as also underlined by Benhabib and Spiegel (1994) and Pritchett (2001), we find, in contrast to the MRW (1992) model, that human capital does not influence growth as a simple production factor. This result is consistent with the human capital puzzle often underlined in the literature. Second, spatial autocorrelation is highly positively significant showing the importance of global technological interdependence from both the theoretical and empirical perspectives. Third, our model implies a specification characterized by parameter heterogeneity estimated using spatial autoregressive local estimation. We present evidence with regard to the local differentiated impacts of the investment rates in physical capital and human capital in each country of our sample. Finally, counterfactual density estimates show that our heterogenous convergence model better fits the observed income distribution, reproducing more accurately the modes that characterize it in 1995, than the well known MRW (1992) model. Further research is oriented towards the development of a more general bisectorial growth model with two sectors using different production functions in order to investigate more deeply the role played by human capital in growth and convergence processes once technological interdependence is introduced. Moreover, the diffusion channels of technological progress should also be investigated in this framework.

References

[1] Anselin L, Bera AK. 1998. Spatial dependence in linear regression models with an introduction to spatial econometrics, In *Handbook of Applied Economics Statistics*, Ullah A, Giles DEA (eds). Marcel Dekker: New York.

[2] Arrow K. 1962. The economic implications of learning by doing. *Review of Economic Studies* **29:** 155-173.

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

[4] Bernanke BS, Gurkaynak RS. 2003. Is growth exogeneous? taking Mankiw, Romer and Weil seriously, In *NBER Macroeconomics Annual*, Bernanke BS, Rogoff K (eds). MIT Press: Cambridge.

[5] Bils M, Klenow PJ. 2000. Does schooling cause growth? *American Economic Review* **90:** 1160-1183.

[6] Brock WA, Durlauf SN. 2001. Growth empirics and reality. *The World Bank Economic Review* **15:** 229-272.

[7] Canova F. 2004. Testing for convergence clubs in income per capita: A predictive density approach. *International Economic Review*, **45:** 49-77.

[8] Conley TG, Ligon E. 2002. Economic distance and cross-country spillovers. *Journal of Economic Growth* **7:** 157-187.

[9] Desdoigts, A. 1999. Patterns of economic development and the formation of clubs. *Journal of Economic Growth* **4:** 305-330.

[10] Desdoigts, A. 2004. Neoclassical convergence versus technological catch-up: a contribution for reaching a consensus. *Problems and Perspectives in Management* **3:** 15-42.

[11] Di nardo J., Fortin NM, Lemieux T. 1996. Labor market institutions and the distribution of wages, 1973-1992: a semi-parametric approach. *Econometrica* **64:** 1001-1044.

[12] Durlauf SN, Johnson P, Temple J. 2005. Growth empirics. In *The Handbook of Economic Growth*, Aghion P and Durlauf S (eds). North-Holland: Elsevier.

[13] Durlauf SN. 2001. Manifesto for a Growth Econometrics. *Journal of Econometrics* **100:** 65-69.

[14] Durlauf SN, Johnson P. 1995. Multiple regimes and cross-country growth behavior. *Journal of Applied Econometrics* **10:** 365-384.

[15] Durlauf SN, Quah D. 1999. The new empirics of economic growth, In *Handbook of Macroeconomics*, Taylor J, Woodford M (eds). North-Holland: Elsevier Science.

[16] Durlauf SN, Kourtellos A, Minkin A. 2001. The local Solow growth model. *European Economic Review* **45:** 928-940.

[17] Eaton J, Kortum S. 1996. Trade in ideas patenting and productivity in the OECD. *Journal of International Economics* **40:** 251-276.

[18] Ertur C, Koch W. 2005. Growth, technological interdependence and spatial externalities: theory and evidence, Regional Economics Applications Laboratory-Technical Series, 05-T-10, University of Illinois at Urbana-Champaign, presented at the Econometric Society World Congress, August 2005.

[19] Ertur C, Le Gallo J, Baumont C. 2006. The European regional convergence process, 1980-1995: do spatial regimes and spatial dependence matter? *International Regional Science Review* **29:** 3-24.

[20] Evans P. 1998. Using panel data to evaluate growth theories. *International Economic Review* **39:** 295-306.

[21] Gelfand AE, and Smith AFM. 1990. Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association* **85:** 398-409.

[22] Geweke J. 1993. Bayesian treatment of the independent student t linear model. *Journal of Applied Econometrics* **8:** 19-40.

[23] Heston A, Summers R, Aten B. 2002. *Penn World Tables Version 6.1.* Downloadable dataset. Center for International Comparisons, University of Pennsylvania.

[24] Islam N. 1995. Growth Empirics: A panel data approach. *Quarterly Journal of Economics* **110:** 1127-1170.

[25] Kelejian HH, Prucha IR. 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.

[26] Keller W. 2004. International technology diffusion. *Journal of Economic Literature* **42:** 752-782.

[27] Klenow PJ, Rodriguez-Clare A. 2005. Externalities and growth. In *The Handbook of Economic Growth*, Aghion P and Durlauf S (eds). North-Holland: Elsevier.

[28] Lange KL, Little RJA, Taylor JMG. 1989. Robust statistical modeling using the t distribution. *Journal of the American Statistical Association* **84:** 881-896.

[29] Lee LF. 2004. Asymptotic distributions of quasi-maximum likelihood estimators for spatial autoregressive models. *Econometrica* **72:** 1899-1925.

[30] LeSage JP. 1997. Bayesian estimation of spatial autoregressive models. *International Regional Science Review* **20:** 113-129.

[31] LeSage J, Pace RK. 2004. Spatial auroregressive local estimation. In *Spatial Econometrics and Spatial Statistics*, Getis A, Mur J, Zoller H (eds). New York: Palgrave MacMillan.

[32] Lucas RE. 1988. On the Mechanics of Economic Development. *Journal of Monetary Economics* **22:** 3-42.

[33] Mankiw NG, Romer D, Weil DN. 1992. A contribution to the empirics of economic growth. *Quarterly Journal of Economics* **107:** 407-437.

[34] Manski CF. 1993. Identification of endogenous social effects: the reflection problem. *Review of Economic Studies* **60:** 531-542.

[35] Moreno R, Trehan B. 1997. Location and growth of nations. *Journal of Economic Growth* **2:** 399-418.

[36] Pritchett L. 2001. Where has all education gone? *World Bank Economic Review* **15:** 367-391.

[37] Quah D. 1996. Twin peaks: growth and convergence in models of distribution dynamics. *Economic Journal* **106:** 1045-1055.

[38] Romer PM. 1986. Increasing returns and long run growth. *Journal of Political Economy* **94:** 1002-1037.

[39] Solow RM. 1956. A contribution to the theory of economic growth. *Quarterly Journal of Economics* **70:** 65-94.

[40] Sheather SJ, Jones MC. 1991. A reliable data-based bandwith selection method for kernel density estimation. *Journal of the Royal Statistical Society, Series B* **53:** 683-690.

[41] Temple J. 1998. Robusness tests of the augmented Solow model. *Journal of Applied Econometrics* **13:** 361-375.

[42] Temple J. 1999. The new growth evidence. *Journal of Economic Literature* **37:** 112-156.

[43] Young A. 1995. The tyranny of numbers: confronting the statistical realities of the East Asian growth experience. *Quarterly Journal of Economics* **110:** 641-680.